

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ЦИВІЛЬНОГО ЗАХИСТУ УКРАЇНИ

Кафедра психології діяльності в особливих умовах

В. Ф. Боснюк, Г. С. Грибенюк

ЕКСПЕРИМЕНТАЛЬНА ПСИХОЛОГІЯ

Навчальний посібник

Черкаси
2025

УДК 159.9(075.8)
Б 85

*Рекомендовано до друку вченою радою
Національного університету цивільного захисту України,
(протокол від 27.11.2025. № 4)*

Рецензенти:

О.М. Кокун, доктор психологічних наук, професор, дійсний член (академік)
НАПН України, директор Інституту психології імені Г.С. Костюка
Національної Академії педагогічних наук України;
Т.С. Яценко, доктор психологічних наук, професор, дійсний член (академік)
НАПН України, професор кафедри психології діяльності в особливих умовах
навчально-наукового інституту оперативно-рятувальних сил
Національного університету цивільного захисту України.

Боснюк В. Ф., Грибенюк Г.С.

Б 85 Експериментальна психологія: навчальний посібник. Черкаси : НУЦЗ України, 2025.
247 с.

Навчальний посібник «Експериментальна психологія» призначений для здобувачів вищої освіти психологічних спеціальностей та спрямований на формування системного розуміння науково-дослідницького підходу в психології, а також практичних умінь щодо планування, проведення та аналізу експериментальних і кореляційних досліджень.

У посібнику послідовно розглядаються методологічні основи психології як науки, принципи наукової достовірності, особливості вимірювання психічних явищ, логіка організації експерименту, різновиди експериментальних і неекспериментальних планів, а також методи контролю побічних впливів. Значну увагу приділено проблемі реплікаційної кризи, питанням академічної доброчесності, прозорості дослідницьких практик та сучасним вимогам доказової психології. Окремо висвітлено психометричні аспекти оцінювання надійності та валідності вимірювальних інструментів.

Практична складова посібника включає приклади застосування різних методів дослідження, завдання для самостійної роботи, рекомендації щодо статистичного аналізу та представлення результатів. Такий підхід дозволяє здобувачам освіти не лише засвоїти теоретичні положення, а й сформувати навички, необхідні для самостійного проведення наукових досліджень.

Посібник орієнтований на здобувачів вищої освіти, викладачів, молодих науковців і практичних психологів. Він буде корисним усім, хто займається проведенням психологічних досліджень, прагне оволодіти сучасними науковими методами й у галузі психології.

УДК 159.9(075.8)

ЗМІСТ

ПЕРЕДМОВА.....	6
РОЗДІЛ 1. ПСИХОЛОГІЯ ЯК НАУКА ПРО ЛЮДИНУ.....	8
1.1. Предмет вивчення експериментальної психології.....	8
1.2. Закономірності розвитку науки в психології.....	12
1.3. Еволюція розвитку наукової парадигми в психології	17
1.4. Комплекс методологічної неповноцінності у психології.....	24
1.5. Методологічні принципи психології як науки про людину	29
1.6. Природничо-наукова та гуманітарна парадигми до дослідження особистості в психології.....	31
Завдання на самостійну підготовку.....	36
РОЗДІЛ 2. РЕПЛІКАЦІЙНА КРИЗА В ПСИХОЛОГІЇ ТА ШЛЯХИ ЇЇ ПОДОЛАННЯ.....	37
2.1. Суть та основні прояви реплікаційної кризи в психології.....	37
2.2. Сумнівні дослідницькі практики як причини реплікаційної кризи...	40
2.3. Проблема публікаційного здвигу і селективності наукових журналів.....	42
2.4. Статистична потужність як фундаментальний чинник реплікаційної кризи	46
2.5. Роль традиційної парадигми перевірки статистичної значущості нульових гіпотез у реплікаційній кризі.....	49
2.6. Проблема недостатньої прозорості досліджень та відсутність відкритих даних.....	52
2.7. Шляхи подолання реплікаційної кризи в психологічних дослідженнях.....	55
Завдання на самостійну підготовку.....	58
РОЗДІЛ 3. ПЛАНУВАННЯ НАУКОВОГО ДОСЛІДЖЕННЯ	60
3.1. Етапи наукового дослідження.....	60
3.2. Роль і значення огляду літератури в сучасній психологічній науці...	66
3.3. Етика наукового дослідження.....	69
3.4. Академічна доброчесність у плануванні наукового дослідження.....	71
Завдання на самостійну підготовку.....	74
РОЗДІЛ 4. ОСОБЛИВОСТІ ВИМІРЮВАННЯ В ПСИХОЛОГІЇ.....	76
4.1. Проблема вимірювання у психології	76
4.2. Методи та види вимірювань у психології.....	79
4.3. Види та характеристики шкал відповідей на запитання в опитувальниках.....	83
4.4. Психометрична надійність тесту.....	87
4.5. Валідність як психометрична властивість тесту	90
4.6. Тестові норми в психометричному вимірюванні.....	95
4.7. Етапи розробки вимірювального інструменту.....	98
4.8. Психосемантична парадигма конструювання психодіагностичних методик.....	102

4.9 Сучасні тенденції та виклики в психометрії.....	106
Завдання на самостійну підготовку.....	112
РОЗДІЛ 5. НЕЕКСПЕРИМЕНТАЛЬНІ МЕТОДИ ДОСЛІДЖЕННЯ ПСИХІКИ.....	113
5.1. Основні методи збору даних у психологічних дослідженнях	113
5.2. Метод спостереження.....	117
5.3. Анкетування як метод дослідження психіки.....	121
5.4. Метод інтерв'ю (бесіди).....	124
5.5. Архівний метод в психології.....	127
5.6. Метод контент-аналізу в психологічних дослідженнях.....	130
Завдання на самостійну підготовку.....	134
РОЗДІЛ 6. ХАРАКТЕРИСТИКА МЕТОДУ ПСИХОЛОГІЧНОГО ЕКСПЕРИМЕНТУ.....	135
6.1. Проблема вивчення причинно-наслідкових зв'язків у психології....	135
6.2. Характеристика та логіка психологічного експерименту.....	137
6.3. Види змінних у психологічному експерименті.....	141
6.4. Психологічні та ситуативні джерела побічних впливів в експерименті.....	144
6.5. Методи контролю побічних змінних в експерименті.....	147
6.6. Валідність психологічного експерименту як критерій наукової достовірності.....	150
6.7. Загрози достовірності висновків психологічного експерименту.....	153
6.8. Етапи реалізації психологічного експерименту.....	156
Завдання на самостійну підготовку.....	159
РОЗДІЛ 7. ДИЗАЙНИ ПСИХОЛОГІЧНОГО ЕКСПЕРИМЕНТУ...	161
7.1. Основні вимоги до експериментального дизайну.....	161
7.2. Види експериментальних дизайнів.....	164
7.3. Характеристика доекспериментальних планів.....	166
7.4. Характеристика істинних експериментальних планів для однієї незалежної змінної.....	169
7.5. Декілька незалежних змінних: повні факторні плани.....	172
7.6. Неповні, змішані та інші експериментальні плани.....	176
7.7. Характеристика квазіекспериментальних планів.....	179
7.8. Дизайн досліджень для маленьких вибірок.....	182
7.9. Експериментальні плани для одного досліджуваного.....	184
Завдання на самостійну підготовку.....	187
РОЗДІЛ 8. ПЛАНИ КОРЕЛЯЦІЙНИХ ДОСЛІДЖЕНЬ В ПСИХОЛОГІЇ.....	189
8.1. Теоретико-методологічні засади кореляційних досліджень у психології.....	189
8.2. Коефіцієнт кореляції як показник статистичних взаємозв'язків.....	192
8.3. Основні джерела похибок у кореляційних дослідженнях.....	194
8.4. Типологія та конструктивні варіації планів кореляційних досліджень.....	197

8.5. Психогенетичні та крос-культурні дослідження як специфічні види кореляційних досліджень.....	200
Завдання на самостійну підготовку.....	203
РОЗДІЛ 9. БАЗОВІ МЕТОДИ СТАТИСТИЧНОГО АНАЛІЗУ РЕЗУЛЬТАТІВ ПСИХОЛОГІЧНИХ ДОСЛІДЖЕНЬ.....	205
9.1. Особливості підготовки результатів дослідження до статистичного аналізу.....	205
9.2. Алгоритм вибору статистичного критерію.....	210
9.3. Статистичні критерії порівняння емпіричного розподілу з очікуваним.....	212
9.4. Статистичні критерії порівняння двох емпіричних розподілів.....	215
9.5. Статистичні критерії порівняння трьох і більше емпіричних розподілів.....	217
9.6. Статистичні критерії для виявлення взаємозв'язків між змінними...	220
9.7. Оцінка розміру ефекту в психологічних дослідженнях.....	223
Завдання на самостійну підготовку.....	226
РОЗДІЛ 10. ПРЕДСТАВЛЕННЯ РЕЗУЛЬТАТІВ НАУКОВОГО ДОСЛІДЖЕННЯ.....	228
10.1. Значення візуалізації даних в науковому дослідженні.....	228
10.2. Принципи та правила якісного графічного представлення даних...	232
10.3. Графічне представлення даних залежно від типу шкали вимірювання.....	235
10.4. Особливості представлення результатів дослідження в таблицях...	241
Завдання на самостійну підготовку.....	245
СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ.....	246

ПЕРЕДМОВА

У сучасному науковому просторі психологія дедалі більше утверджується як доказова, емпірично орієнтована дисципліна, що поєднує гуманітарне розуміння людини з природничо-науковими підходами до вивчення її психічних процесів. Тенденція до підвищення стандартів наукової достовірності, необхідність критичної оцінки отриманих результатів, а також зростання вимог до відтворюваності досліджень зумовлюють потребу у формуванні у здобувачів вищої освіти високого рівня дослідницької компетентності. В цих умовах експериментальна психологія виступає основою наукового стилю мислення, що забезпечує можливість перевірки гіпотез, побудови моделей, дослідження причинно-наслідкових зв'язків та формування якісної емпіричної бази.

Експеримент як метод посідає ключове місце в психологічній науці, адже він дозволяє не лише фіксувати прояви психічних явищ, а й виявляти механізми їх виникнення та функціонування. На відміну від описових або кореляційних підходів, експеримент забезпечує контроль змінних, можливість відокремити суттєві фактори від несуттєвих і оцінити реальний вплив дослідницьких умов на поведінку та психічні стани. Проте ефективне використання цього методу неможливе без знання ширшої методологічної бази, що включає розуміння розвитку наукових парадигм, особливостей вимірювання у психології, психометричних характеристик інструментів, етичних норм дослідницької діяльності та сучасних стандартів академічної доброчесності.

Метою цього навчального посібника є формування у здобувачів системного й цілісного уявлення про експериментальну психологію як основу наукового дослідження психіки. У межах посібника подано інтегрований матеріал, що охоплює методологічні принципи сучасної науки, логіку побудови емпіричного дослідження, особливості вимірювання психологічних конструктів, принципи конструювання експериментальних і кореляційних планів, а також методи аналізу та представлення результатів. Такий підхід дозволяє здобувачу вищої освіти побачити експеримент не окремо, а в контексті цілісного наукового процесу – від теоретичного обґрунтування проблеми до формулювання висновків і презентації даних.

Важливою особливістю посібника є акцент на сучасних викликах психологічної науки. Реплікаційна криза, проблема сумнівних дослідницьких практик, питання статистичної потужності досліджень, вимоги до прозорості отриманих даних – усі ці аспекти є невід'ємною частиною сьогоденної наукової реальності. Саме тому значна увага приділяється формуванню вміння критично оцінювати дані, дотримуватися етичних та методологічних стандартів, запобігати системним помилкам і відповідально ставитися до інтерпретації результатів. Здобувачі мають розуміти, що якість психологічного дослідження визначається не лише технікою проведення

експерименту, але й коректністю вимірювання, обґрунтованістю статистичного аналізу та чесністю дослідника щодо отриманих результатів.

Матеріал укладено таким чином, щоб забезпечити поступове й усвідомлене засвоєння знань: від основ наукової методології до практичних аспектів експериментального дизайну, аналізу даних і презентації результатів. Така послідовність сприяє цілісному розумінню наукового процесу та формуванню практичних навичок, необхідних для самостійного проведення дослідження.

Посібник адресовано здобувачам вищої освіти, молодим науковцям, викладачам та практичним психологам. Він може бути корисним як під час вивчення дисциплін методологічного блоку, так і у процесі підготовки кваліфікаційних робіт.

Автори висловлюють сподівання, що це видання сприятиме розвитку дослідницького мислення, підвищенню культури наукового пізнання та формуванню відповідального ставлення до організації дослідження, збору, аналізу й інтерпретації психологічних даних. Адже саме вміння працювати з достовірною інформацією, критично оцінювати результати та дотримуватися принципів науковості є основою професійної компетентності психолога та важливим кроком до розвитку сучасної доказової психології в Україні.

РОЗДІЛ 1

ПСИХОЛОГІЯ ЯК НАУКА ПРО ЛЮДИНУ

1.1. Предмет вивчення експериментальної психології

Питання про предмет вивчення експериментальної психології здається на перший погляд простим: начебто йдеться про «психічні явища, досліджувані в умовах експерименту». Проте уважніша методологічна рефлексія показує, що тут перетинаються щонайменше три лінії дискусії: що саме вважати психічною реальністю; чи можна звести її до поведінки, фізіології або інформаційних процесів; і чи є експериментальна психологія окремою галуззю з власним предметом чи радше методологічним «каркасом» для всієї науки про психіку. Тому розгляд предмета експериментальної психології неминуче пов'язаний із різними концептуальними підходами, які по-різному окреслюють те, що саме має бути центром наукового вивчення.

У найзагальнішому розумінні експериментальну психологію часто визначають як галузь, що вивчає базові психічні процеси, такі як сприймання, увагу, пам'ять, мислення, емоції, мотивацію, прийняття рішень, навчання тощо за допомогою спеціально організованих експериментальних процедур. У цьому підході предметом виступають універсальні закономірності функціонування психіки, які проявляються незалежно від конкретного життєвого контексту, соціальної ролі чи індивідуальної біографії людини. Такий погляд тяжіє до уявлення про експериментальну психологію як фундаментальну дисципліну, що створює «базову карту» психічних механізмів, на які потім спираються інші, прикладні напрями – клінічна, освітня, організаційна, юридична психологія тощо. Саме в цьому сенсі говорять про експериментальну психологію як «фундамент» емпіричної науки про психіку: вона не обмежується окремою сферою практики, а прагне виявити загальні закони, що діють у різних контекстах.

Інша лінія розуміння предмета експериментальної психології пов'язана з тим, наскільки до нього включаються власне внутрішні стани й переживання. У рамках поведінкових підходів ХХ століття було висунуто тезу, що справжнім об'єктом наукового дослідження може бути лише спостережувана поведінка, оскільки вона піддається точному опису, фіксації й контролю. У цьому випадку предмет експериментальної психології звужується до закономірностей формування, зміни й підтримання поведінкових реакцій під впливом стимулів, підкріплень, покарань та інших факторів середовища. Психіка в традиційному розумінні розглядається або як «чорна скринька», або як сукупність проміжних змінних, які не мають самостійного онтологічного статусу. В такій парадигмі експериментальна психологія – це наука про поведінкові закони, що вивчаються в строго контрольованих умовах.

Когнітивний та нейронауковий підходи, навпаки, розширюють предмет експериментальної психології, включаючи до нього не лише зовнішню поведінку, а й внутрішні процеси переробки інформації та нейронні

механізми. Тут у центрі уваги опиняються ментальні репрезентації, схеми, стратегії, процеси кодування й зберігання інформації, а також їхні мозкові кореляти. Психологічний експеримент у такому розумінні спрямований на дослідження того, як людина сприймає, інтерпретує, організує й використовує інформацію, як структурується свідомий і несвідомий досвід на рівні когнітивних операцій. Предметом стає «архітектура» психічних процесів – від мікрорівня сенсорної обробки сигналів до складних форм міркування, планування чи морального вибору. Поведінка тут виступає важливим, але не єдиним показником; аналіз доповнюється записами часу реакцій, точності виконання завдань, даними нейровізуалізації, показниками руху очей, психофізіологічними параметрами.

Існує також позиція, згідно з якою експериментальна психологія не має «свого» предмета в сенсі окремої ділянки реальності, а радше описує певний методологічний рівень дослідження. З цього погляду експериментальний підхід може застосовуватися у будь-якій галузі психології: у клінічній – для перевірки ефективності терапевтичних інтервенцій; у педагогічній – для дослідження результативності методик навчання; у соціальній – для вивчення впливу групових норм і стереотипів; в організаційній – для аналізу мотиваційних схем чи стилів лідерства. Тоді предмет експериментальної психології описують як «структурні та динамічні закономірності психічних процесів, виявлені шляхом навмисної маніпуляції змінними й контролю ситуації». Іншими словами, експериментальна психологія – це спосіб теоретико-емпіричного мислення, спрямованого на встановлення причинних зв'язків у будь-якій сфері психічної реальності.

У межах діяльнісного та соціокультурного підходів предмет експериментальної психології трактується ширше, ніж сукупність окремих функцій психіки. Тут акцент робиться на структурах діяльності, взаємодії й спілкування, у яких психічне реалізується як регуляція відносин «людина – світ» і «людина – людина». Експеримент, відповідно, покликаний моделювати не ізольовані реакції, а фрагменти цілісної діяльності: навчальної, ігрової, професійної, комунікативної. Предметом стають ті психологічні механізми, які забезпечують побудову дії, прийняття рішень у контексті цілей і мотивів, узгодження індивідуальних і соціальних смислів. У таких дослідженнях експериментальна ситуація може бути наближена до реального середовища (польові, природні експерименти), а змінними виступають не лише стимульні параметри, а й організація діяльності, форма зворотного зв'язку, структура взаємодії в групі.

Попри відмінності, усі ці підходи сходяться в тому, що експериментальна психологія вивчає не просто психічні явища, а закономірності змін психіки й поведінки під впливом керованих чинників. Саме ця настанова на виявлення причинно-наслідкових залежностей відрізняє експериментальний підхід від описових, кореляційних або суто клінічних методів. Там, де спостереження лише фіксує співіснування явищ, експеримент намагається встановити, чи призводить маніпуляція однією змінною до

систематичної зміни іншої. Тому як би не формулювали предмет – як базові когнітивні процеси, поведінкові реакції, структури діяльності чи інформаційні потоки, – ядром експериментальної психології лишається вивчення детермінованості психічного: що, за яких умов, у кого й чому змінюється.

Із цього розуміння випливають і основні завдання експериментальної психології. Насамперед це теоретичне завдання – побудова й уточнення моделей психічних процесів, які дозволяють робити чіткі, формалізовані передбачення щодо поведінки й переживань людей у різних ситуаціях. Експериментальна психологія покликана не просто фіксувати емпіричні зв'язки, а перевіряти гіпотези про механізми, які ці зв'язки породжують. Наприклад, у галузі дослідження уваги йдеться не лише про опис обсягу чи стійкості, а про виявлення принципів організації уваги в умовах конкуренції стимулів, подвійних завдань, емоційного навантаження. Такі моделі мають бути достатньо точними, щоб їх можна було спростувати експериментальними даними, але водночас досить узагальненими, щоб пояснювати різні феномени.

Другою групою завдань є методологічні й інструментальні. Експериментальна психологія розробляє, стандартизує та валідизує експериментальні парадигми, методи реєстрації й аналізу даних, підходи до контролю побічних змінних. Йдеться як про класичні методи (реакція вибору, часові пороги, парадигми запам'ятовування), так і про сучасні комп'ютеризовані протоколи, віртуальні середовища, інтерактивні завдання. У цьому сенсі експериментальна психологія виступає «лабораторією» наукової інфраструктури: від точності інструкцій, конструювання стимулів, параметрів рандомізації до процедур пілотажу і перевірки надійності залежить якість усієї емпіричної психології, незалежно від її змістового спрямування.

Третє важливе завдання стосується опису меж функціонування психіки, дослідження граничних умов, у яких звичайні закономірності порушуються або набувають іншого характеру. Експериментальна психологія тут звертається до ситуацій перевантаження, стресу, дефіциту інформації, багатозадачності, незвичного сенсорного середовища, змінених станів свідомості тощо. Дослідження, наприклад, уваги й прийняття рішень у умовах ризику, втоми або часової нестачі дозволяють зрозуміти, де саме починають діяти «обмеження» когнітивної системи. Ці дані мають прямі наслідки для проектування інтерфейсів, організації праці, безпеки транспорту, клінічних втручань.

Четверте завдання пов'язане з аналізом варіативності й індивідуальних відмінностей у межах експериментальних ефектів. Якщо класична експериментальна психологія часто була зорієнтована на «усередненого» суб'єкта, то сучасні підходи дедалі більше враховують, що один і той самий вплив може по-різному діяти на людей з різними рисами особистості, мотивацією, когнітивним стилем, культурним досвідом. Тому завданням експериментальної психології стає виявлення не лише загальних закономірностей, а й модераторів і медіаторів ефектів: які фактори посилюють або послаблюють вплив, через які проміжні ланки він реалізується. Це

наближає експериментальну психологію до персоналізованих підходів у клініці, освіті, організаційному консультуванні.

П'яте завдання – критично-перевірочне. Експериментальна психологія має бути «механізмом самокорекції» науки, перевіряючи стійкість раніше виявлених ефектів, узагальнень, теоретичних висновків. Реплікаційні дослідження, метааналізи, порівняння ефектів у різних вибірках, культурах та умовах є невід'ємною частиною сучасної експериментальної роботи. Тут експериментальна психологія виконує роль своєрідного «методологічного суду присяжних», оцінюючи, які твердження дійсно витримують перевірку часом і контекстом.

Окрема дискусійна площина стосується співвідношення експериментальної психології з прикладними галузями та теоретичною психологією. Частина авторів наполягає, що предмет експериментальної психології – переважно «елементарні» процеси (сенсорика, увага, пам'ять тощо), тоді як особистісні, соціальні, культурні виміри належать іншим галузям. Інші, навпаки, підкреслюють, що немає принципової межі між «елементарним» і «вищим», а експеримент можна застосувати до будь-якого рівня психічної організації, якщо вдається його коректно операціоналізувати. Звідси виникають різні уявлення про те, чи має експериментальна психологія власний «змістовий» предмет, чи це радше певний стиль дослідження, який пронизує всю науку про психіку.

Сучасні тенденції в розвитку експериментальної психології суттєво розширюють і поглиблюють її предмет. По-перше, відбувається інтенсивна інтеграція з нейронауками: експериментальні парадигми поєднуються з методами реєстрації мозкової. Унаслідок цього предмет експериментальної психології дедалі більше описується як багаторівнева система – від нейронних мереж до усвідомлених рішень і соціальних дій. Дослідження пам'яті, уваги, мови, прийняття рішень доповнюються картографією відповідних нейронних контурів, аналізом пластичності мозку, впливу навчання й досвіду на його організацію.

По-друге, посилюється роль комп'ютерного моделювання й обчислювальних підходів. Експериментальна психологія вже не лише фіксує кореляції між змінними, а й конструє формальні моделі – наприклад, моделі процесів прийняття рішень, розпізнавання образів, контролю дій, – які потім перевіряються на експериментальних даних. Це дозволяє говорити про предмет експериментальної психології як про систему інформаційних перетворень, що реалізуються в мозку та поведінці. Зі свого боку, штучний інтелект і машинне навчання пропонують нові інструменти для аналізу великих масивів експериментальних даних, виявлення прихованих патернів, кластерів, траєкторій змін.

По-третє, зростає увага до екологічної валідності експериментальних досліджень. Якщо класична лабораторна експериментальна психологія часто критикувалася за «штучність» умов і відірваність від реального життя, то нині дедалі ширше застосовуються польові експерименти, мобільні додатки,

віртуальна і доповнена реальність, які дозволяють моделювати життєві ситуації з високим ступенем контролю. Предметом стають закономірності психічного функціонування в умовах реального середовища: взаємодія з цифровими інтерфейсами, поведінка в соціальних мережах, реагування на багатоканальні потоки інформації, прийняття рішень у складних, динамічних контекстах.

По-четверте, важливим трендом є орієнтація на відкриту науку та підвищення прозорості експериментальних досліджень. Пререєстрація гіпотез та плану аналізу, відкриті бази даних, спільні протоколи, мультилабораторні проекти змінюють не лише організацію роботи, а й саме розуміння предмета експериментальної психології. Він дедалі менше пов'язується з «одиначним експериментом» і дедалі більше – з накопиченням і зіставленням великої кількості досліджень, що дозволяє уточнювати ефекти, виявляти модератори, перевіряти узагальнюваність результатів.

По-п'яте, експериментальна психологія дедалі активніше виходить на рівень складних соціальних явищ. Експериментальні підходи застосовуються до дослідження довіри, кооперації, агресії, упереджень, моральних суджень, політичних рішень. Предмет розширюється від індивідуальних процесів до групових і масових, але зберігається ключовий принцип – контроль і маніпуляція змінними. Це ставить перед експериментальною психологією нові етичні й методологічні виклики, але водночас демонструє її здатність працювати з найбільш актуальними проблемами сучасних суспільств.

Таким чином, предмет вивчення експериментальної психології не можна звести до однієї-єдиної формули. Для одних дослідників це насамперед базові ментальні процеси, для інших – закономірності поведінки, для третіх – структури діяльності й взаємодії, для четвертих – інформаційні й нейронні механізми, що лежать в основі свідомості. Проте всі ці підходи пов'язані спільною інтенцією: виявити, у чому полягає закономірна, відтворювана детермінованість психіки та поведінки, якщо штучно, але контрольовано змінювати умови функціонування людини. Саме ця каузально-експериментальна перспектива й визначає обличчя експериментальної психології в сучасній науці, водночас зберігаючи історичну спадкоємність і відкритість до нових концептуальних рішень.

1.2. Закономірності розвитку науки в психології

Розвиток психології як науки – це складний шлях переходу від споглядальних філософських міркувань про душу до системи емпірично перевірюваних знань. Еволюція психологічної думки демонструє, як змінювалися не лише уявлення про психіку, а й критерії науковості самої психології.

Початки наукового мислення в психології сягають філософських традицій раціоналізму та емпіризму. Раціоналісти, зокрема Рене Декарт, наголошували на ролі розуму як основного джерела достовірного знання. Його

теза «*Cogito, ergo sum*» («Я мислю, отже, існую») заклала уявлення про свідомість як первинну реальність і про дуалізм душі та тіла. Для психології це стало поштовхом до пошуку механізмів взаємодії між психічним і фізіологічним.

Протилежну позицію займали емпірики Френсіс Бекон і Джон Локк. Вони вважали, що свідомість людини є «чистою дошкою» (*tabula rasa*), на якій досвід залишає свої сліди. Бекон сформулював метод емпіричного пізнання, що спирається на спостереження, експеримент і індукцію. Його ідеї про очищення розуму від хибних переконань і упереджень, що стали основою для розвитку об'єктивної науки.

Раціоналізм і емпіризм утворили діалектичну пару, що визначила подальший розвиток психології. Раціоналізм привніс логічну структуру і системність, тоді як емпіризм забезпечив перевірюваність і опору на досвід. Їхнє поєднання стало підґрунтям експериментальної методології Вундта, який прагнув поєднати внутрішнє спостереження з об'єктивним вимірюванням реакцій.

Після Декарта і Локка ключову роль у формуванні наукової психології відіграли ідеї трьох мислителів – Лейбніца, Канта й Юма. Готфрід Вільгельм Лейбніц висунув припущення про ступені усвідомлення: за його думкою, психічні процеси можуть бути як свідомими, так і несвідомими. Саме він запровадив поняття «апперцепції» як активного акту спрямування свідомості на певний зміст. Ця ідея пізніше стала основою для вчень Вундта і Гербарта про внутрішню динаміку свідомості.

Іммануїл Кант, у свою чергу, висловив сумнів у можливості перетворення психології на точну науку, оскільки, на його думку, психічні явища не піддаються кількісному вимірюванню. Проте саме його категоріальний апарат (поняття простору, часу, причинності) вплинув на подальшу побудову психологічних теорій сприйняття і мислення. Філософ довів, що свідомість не просто відображає світ, а активно його структурує, що згодом стало підґрунтям для гештальтпсихології та когнітивізму.

Девід Юм продовжив емпіричну лінію, обґрунтувавши асоціативну теорію мислення. Він вважав, що ідеї у свідомості поєднуються за принципом схожості, суміжності у просторі й часі або причинно-наслідкового зв'язку. Цей принцип асоціацій пізніше стане центральним у біхевіористських моделях навчання, де поведінка розглядається як система зв'язків між стимулами та реакціями.

Таким чином, ще до виникнення власне експериментальної психології у філософії сформувався інтелектуальний фундамент для її становлення. Вебер і Фехнер у ХІХ столітті зробили наступний крок, створивши психофізику – першу дисципліну, що досліджувала кількісні співвідношення між фізичним подразником і психічною реакцією. Формула Вебера–Фехнера показала, що суб'єктивні відчуття підкоряються об'єктивним законам. Відтепер психіка перестала бути «чорною скринькою» і перетворилася на феномен, доступний вимірюванню.

Психофізика стала методологічним містком між філософськими роздумами й експериментальною наукою. Вона не лише дала технічні інструменти – визначення абсолютного та диференціального порогів, методи середніх помилок і постійних стимулів, а й утвердила ідею кількісного опису психічного. Саме ця традиція підготувала ґрунт для відкриття Вундта, який зумів поєднати емпіричні методи вимірювання з теоретичним аналізом структури свідомості.

У ХІХ столітті під впливом філософії позитивізму психологія почала набувати сучасних рис природничої науки. Огюст Конт, засновник позитивізму, вважав, що справжнє знання можливе лише через емпіричне спостереження фактів і виявлення закономірностей між ними. Філософія мала відмовитися від метафізики й займатися тим, що піддається перевірці.

Позитивізм відіграв надзвичайно важливу роль у становленні психології як експериментальної науки. Його головна теза – наука має спиратися лише на факти, які можна спостерігати, виміряти і перевірити. У цьому полягала радикальна відмова від метафізики, тобто спроб пояснювати явища через недоступні спостереженню сутності. Для психології це означало відхід від абстрактних понять на кшталт «душі» чи «свідомості» у їхньому філософському значенні – на користь конкретних процесів, що піддаються емпіричному аналізу: відчуттів, емоцій, реакцій, поведінки.

Філософи логічного позитивізму (Рудольф Карнап, Моріц Шлік, Отто Нейрат) продовжили розвивати ці ідеї у ХХ столітті. Вони наголошували, що справжня наука потребує уніфікованої мови, вільної від двозначностей.

Логічний позитивізм сприяв також розвиткові математичної статистики в психології. Ідея, що наукове твердження має бути верифікованим, тобто підтвердженим досвідом, призвела до активного використання ймовірнісних моделей, аналізу варіацій, перевірки гіпотез. З'явилися концепції надійності та валідності тестів, що зробили психологічні вимірювання стандартизованими.

Водночас у 1930-х роках почали лунаати й критичні голоси. Деякі науковці зауважували, що абсолютна вимога «фактичності» призводить до ігнорування складних ідейних структур, які не піддаються прямому спостереженню. Так народилася потреба у більш гнучкому підході до наукового пізнання – підході, який дозволяє визнавати гіпотези, навіть якщо вони ще не перевірені експериментально, але мають пояснювальну силу. Цей інтелектуальний рух і підготував ґрунт для фальсифікаційної теорії Карла Поппера, яка визначила сучасне розуміння науковості.

Значний вплив на методологічну культуру психології справив Карл Поппер, який критикував позитивізм за надмірну віру у верифікацію. Він запропонував принцип фальсифікації як головний критерій науковості: теорія є науковою лише тоді, коли її можна спростувати досвідом. Це означало, що жодна кількість підтверджень не робить теорію остаточно істинною, але один переконливий контрприклад може її зруйнувати.

Психологічна наука, прийнявши цей підхід, змінила спосіб мислення: гіпотеза перестала сприйматися як «вічна істина» і почала розглядатися як тимчасова модель, що підлягає перевірці. Така позиція відкрила шлях до методологічної самокритичності, яка і сьогодні є етичним стандартом наукового дослідження.

Попперів підхід привів до розвитку гіпотетико-дедуктивної моделі, що стала основою сучасного експериментального мислення. Відповідно до неї, дослідження починається з теоретичного припущення – гіпотези, з якої дедуктивно виводяться передбачення, що підлягають перевірці. Якщо результати експерименту суперечать передбаченням, гіпотезу потрібно змінити або відкинути.

Цей цикл – гіпотеза → експеримент → перевірка → корекція теорії – став серцем наукового пізнання у психології. Саме тому наукове дослідження в психології не може бути зведене лише до спостереження: воно передбачає пояснення, побудову моделей і постійну перевірку їхньої адекватності.

Додатковим внеском позитивістської методології стало впровадження операціоналізації понять. Психологічні конструкти – інтелект, мотивація, тривожність, креативність – отримали конкретні операційні визначення, що дозволило вимірювати їх емпірично.

Ідея операціоналізації мала далекосяжні наслідки для всієї психологічної практики. Уперше вона була сформульована фізиком Персі Бріджменом у 1927 році: будь-яке поняття є науковим лише тоді, коли воно визначене через конкретні операції спостереження або вимірювання. Психологи швидко прийняли цей принцип, бо він дозволив перевести складні психічні явища у вимірювані показники.

Наприклад, інтелект почали описувати через коефіцієнт IQ, який визначається співвідношенням розумового віку до хронологічного. У цьому випадку «розумовий вік» – це вже операціоналізація: його вимірюють за допомогою стандартизованих завдань на логіку, пам'ять, увагу. Подібно, тривожність вимірюють через бали за тестом Спілбергера, самооцінку – за шкалою Розенберга, депресію – за шкалою Бека.

Завдяки такій стандартизації психологія отримала інструмент порівняння результатів між різними групами і культурами. Але одночасно виникла проблема культурної валідності. Те, що вважається «нормальним рівнем тривожності» в одній культурі, може мати зовсім інше значення в іншій. Дослідження, проведені в Японії, Індії та країнах Європи, показали, що культурні норми стриманості чи відкритості емоцій суттєво впливають на результати опитувальників.

Це поставило перед наукою нове завдання – створювати тести, адаптовані до культурного контексту, а не універсальні шаблони. Так з'явилися методики перекладу й адаптації психодіагностичних інструментів, стандарти яких нині затверджуються на рівні Міжнародної тестової комісії (ITS). Сучасна психологія дедалі частіше говорить не про універсальні норми,

а про контекстуальну надійність – відповідність інструменту культурі, у якій він застосовується.

Операціоналізація також викликала методологічну дискусію: чи не втрачає психологія через надмірну формалізацію свою гуманітарну глибину. Прихильники феноменології вважали, що суб'єктивний досвід не можна звести до цифр без спотворення сенсу. Сьогодні більшість учених схилиються до інтегративної позиції: кількісні та якісні методи мають співіснувати, утворюючи дві сторони одного пізнавального процесу – об'єктивного вимірювання і смислового тлумачення.

Таким чином, операціоналізація стала одним із головних інструментів перетворення психології на точну науку, але водночас вона нагадала про необхідність зберігати зв'язок із реальним людським досвідом. Адже мета психології – не лише виміряти людину, а й зрозуміти її.

Справжня наукова революція в психології відбулася завдяки впровадженню експерименту як головного методу. Він дозволив переходити від опису явищ до виявлення причинно-наслідкових зв'язків. У цьому контексті важливим є поняття контролю змінних – тобто ізоляції впливу досліджуваного фактора від усіх інших.

Як зазначає Б. Скіннер, лише за умов системного контролю можна робити достовірні висновки про поведінкові закономірності. Наприклад, у класичних дослідженнях І. Павлова стимул (дзвінок) виконував роль незалежної змінної, а слиновиділення – залежної. Контроль усіх сторонніх факторів (час доби, інтенсивність стимулу, стан тварини) забезпечував чистоту експерименту.

Сучасна експериментальна психологія використовує більш складні дизайни: міжгрупові, внутрішньогрупові, кросфакторні, що дозволяють аналізувати не лише прості, а й комбіновані впливи змінних. Наприклад, можна досліджувати взаємодію між рівнем стресу і когнітивною гнучкістю в умовах вирішення завдань.

Попри домінування позитивістського підходу, у середині XX століття в науці відбувся гуманістичний зсув. З'явився інтерес до суб'єктивного досвіду, особистісного сенсу, творчості та самореалізації. Представники гуманістичної психології поставили в центр уваги не поведінку, а людину як цілісну істоту з власною системою цінностей і смислів.

У цьому контексті почали активно розвиватися якісні методи дослідження: феноменологічне спостереження, аналіз життєвих історій, глибинні інтерв'ю, наративний аналіз. Їхня мета – не узагальнення закономірностей, а розкриття унікального досвіду конкретної людини.

Як зазначає К. Роджерс, наука, яка ігнорує внутрішній досвід людини, втрачає зв'язок із реальністю людського існування. Тому гуманістична традиція доповнила експериментальну психологію розумінням контексту, смислу і свободи вибору.

Поступово стало очевидно, що жоден метод – ані суто кількісний, ані суто якісний – не може повністю відобразити складність психічної реальності.

Це призвело до методологічного плюралізму – ідеї про необхідність поєднання різних підходів.

Сучасна психологія активно інтегрує експериментальні, клінічні та соціокультурні методи. Наприклад, дослідження депресії сьогодні включає одночасно нейробіологічні вимірювання (рівень серотоніну, активність певних ділянок мозку), опитувальники самооцінки та аналіз особистісних наративів. Це дозволяє описати явище на кількох рівнях – біологічному, психологічному й соціальному.

Поєднання кількісної та якісної методологій створює «методологічну симфонію», де кожен підхід компенсує слабкі сторони іншого. Саме такий інтегративний характер робить сучасну психологію дійсно наукою про людину – не лише як об'єкт, а як суб'єкт, носія досвіду, цінностей і вибору.

1.3. Еволюція розвитку наукової парадигми в психології

Психологія, як і будь-яка наука, розвивається не лінійно, а через зміну парадигм – систем фундаментальних уявлень, що визначають, які явища вважаються науковими, які методи вважаються допустимими і які питання – важливими. Кожна парадигма не зникає безслідно: вона залишає після себе ідеї, що інтегруються в наступні підходи.

Першою науковою парадигмою психології став структуралізм. Його засновником вважають Вільгельма Вундта, який 1879 року створив першу лабораторію експериментальної психології в Лейпцигу. Метою Вундта було описати структуру свідомості, подібно до того, як хімія описує будову речовин. Він вважав, що складні психічні явища можна розкласти на елементи (відчуття, уявлення, почуття тощо) і дослідити їх взаємозв'язок.

Основним методом структуралізму була інтроспекція – самоспостереження. Досліджуваного просили описати свої переживання у відповідь на стимул (наприклад, звук чи зображення), фіксуючи відчуття та емоційні реакції. Попри новаторство, метод інтроспекції мав серйозні обмеження: він залежав від суб'єктивності учасника і не дозволяв перевірити результати незалежно.

Проте історичне значення структуралізму величезне – він започаткував наукову лабораторну традицію в психології, сформулював принцип вимірюваності психічних процесів і поставив питання про системність свідомості.

На зміну структуралізму прийшов функціоналізм, що сформувався у США під впливом еволюційної теорії Чарльза Дарвіна. Вільям Джеймс, Джон Дьюї, Гарві Керр як його головні представники зосередили увагу не на структурі свідомості, а на її функціях. Свідомість, на їхню думку, є інструментом адаптації до середовища: вона допомагає людині виживати, навчатися, приймати рішення.

Функціоналісти досліджували, як психічні процеси допомагають організму ефективно діяти. Цей підхід змістив фокус із опису на пояснення, з

«що» на «навіщо». Функціоналізм вперше поєднав експериментальні методи з вивченням поведінки у природних умовах, що стало передумовою розвитку прикладних напрямів – педагогічної, клінічної та індустріальної психології.

Функціоналізм відіграв ключову роль у формуванні прикладної психології. Його головна ідея про корисність психічних процесів підштовхнула дослідників до пошуку практичних способів їх використання. Джон Дьюї та Джеймс Енджел створили фундамент шкільної психології у США. Вони довели, що ефективне навчання не можна звести до механічного заучування, а слід розглядати як динамічний процес адаптації дитини до середовища.

Під впливом функціоналізму розвинулася психометрія – наука про вимірювання індивідуальних відмінностей. З'явилися перші тести здібностей, уваги, пам'яті, що використовувалися для відбору працівників у промисловості та армії. Саме завдяки цьому напрямові психологія отримала вихід за межі лабораторій і стала соціально необхідною дисципліною.

Функціоналісти також уперше розглянули емоції як адаптивний механізм, що допомагає організму реагувати на середовище. Роботи Вільяма Джеймса дали початок психофізіологічним дослідженням емоцій, які продовжують існувати й сьогодні.

Цей напрям зумів поєднати філософію прагматизму з експериментальним методом, відкривши шлях до біхевіоризму, який зробив поведінку головним предметом наукового спостереження.

Радикальний злам відбувся на початку XX століття з появою біхевіоризму. Його засновник Джон Б. Уотсон у маніфесті 1913 року проголосив: «Психологія повинна вивчати не свідомість, а поведінку». Він вважав, що внутрішні стани неможливо об'єктивно виміряти, тому єдиним предметом науки має бути зовнішня реакція на стимул.

Біхевіоризм можна вважати найпослідовнішим утіленням позитивістських принципів у психології. Його метою було зробити психологію такою ж об'єктивною, як фізика чи біологія. Якщо структуралісти вивчали зміст свідомості, то біхевіористи відкинули його як «суб'єктивну ілюзію», зосередившись виключно на спостережуваній поведінці.

Ранній біхевіоризм, представлений Джоном Уотсоном, ґрунтувався на простій схемі S–R (стимул – реакція). Людина розглядалася як система, яка реагує на подразники середовища, і цими реакціями можна керувати, змінюючи стимули. Відомий «експеримент із маленьким Альбертом», у якому дитину привчили боятися білого пацюка, демонструє, як емоції можуть формуватися через асоціативне навчання.

Пізніше з'явився нео-біхевіоризм, який намагався пояснити поведінку складніше. Кларк Халл і Едвард Толмен запровадили проміжні змінні – такі, як потреби, установки, когнітивні карти, – які хоч і не спостерігаються безпосередньо, але опосередковують зв'язок між стимулом і реакцією. Саме Толмен увів поняття когнітивної карти – внутрішнього образу простору, що формується у процесі навчання, довівши, що навіть тварини не лише реагують, а й «мислять» під час орієнтації в лабіринті.

Беррес Скіннер зробив наступний крок, створивши теорію оперантного навчання. На відміну від класичного умовного рефлексу Павлова, оперантне навчання описує формування поведінки через наслідки. Якщо дія супроводжується позитивним підкріпленням (нагородою), вона повторюється; якщо негативним – згасає. Ця модель стала основою для методів поведінкової терапії, навчання і навіть маркетингу.

Сильні сторони біхевіоризму – експериментальна точність і контрольованість умов. Його підходи лягли в основу терапевтичних методів, таких як поведінкова терапія, і педагогічних стратегій навчання. Однак він виявився надто редуccionістським: заперечуючи роль свідомості, емоцій і мотивації, біхевіоризм звів людину до автомату реакцій. При цьому біхевіоризм мав і практичні успіхи: саме він створив підґрунтя для прикладного аналізу поведінки, який сьогодні використовується у корекції поведінки дітей з аутизмом. Однак його головна слабкість полягала у надмірному редуccionізмі. Ігноруючи внутрішній досвід, він не міг пояснити мотивацію, креативність чи моральний вибір. Ця вада зумовила появу когнітивної революції, що повернула свідомість у центр психологічної науки.

Паралельно з біхевіоризмом розвивався психоаналіз, започаткований Зигмундом Фройдом. На відміну від експериментаторів, Фройд працював із клінічними випадками, прагнучи пояснити внутрішні конфлікти людини. Він вважав, що поведінка значною мірою визначається несвідомими потягами, дитячими переживаннями та механізмами захисту «Я».

Психоаналіз, на відміну від біхевіоризму, спирався не на спостереження за поведінкою, а на аналіз внутрішніх конфліктів особистості. Психоаналіз дав психології нову категорію «несвідоме» і методи його вивчення: вільні асоціації, аналіз сновидінь, тлумачення помилкових дій. Зигмунд Фройд побудував свою теорію на трьох головних елементах: Ід (Воно) – джерело несвідомих потягів; Его (Я) – посередник між інстинктами і реальністю; Супер-Его (Над-Я) – внутрішній моральний цензор. Взаємодія цих сил визначає поведінку людини, а невідповідність між ними породжує неврози.

Фройд уперше ввів поняття «механізмів психологічного захисту» – витіснення, проєкції, раціоналізації, сублімації. Вони описують, як психіка захищає себе від надмірного напруження, приховуючи неприйнятні імпульси у несвідомому. Ця концепція згодом стала базою для клінічної психотерапії.

Важливо, що психоаналіз став не просто теорією, а методологічним підходом до вивчення особистості. Його техніки вільних асоціацій, аналіз сновидінь, тлумачення помилкових дій була спробою зробити несвідоме доступним пізнанню. Хоча цей метод важко піддавався перевірці, він відкрив новий вимір психологічної науки – глибинну мотивацію поведінки.

Подальші напрями психоаналізу розвинули ідеї Фрейда. Альфред Адлер заснував індивідуальну психологію, де головну роль відіграє прагнення до переваги і почуття спільності. Карл Юнг запропонував концепцію колективного несвідомого – спадкових архетипів, які проявляються в міфах,

снах і культурних символах. Еріх Фромм, у свою чергу, з'єднав психоаналіз із соціальною філософією, розглядаючи особистість як продукт культури.

Сучасна психодинамічна психологія вже не сприймає Фрейда догматично, але зберігає головний принцип – поведінка має приховані мотиви, які можна зрозуміти лише через аналіз внутрішніх суперечностей. Психоаналітична традиція суттєво вплинула на мистецтво, літературу, культурологію, а також на розвиток психотерапевтичних практик. Саме завдяки їй психологія остаточно вийшла за межі лабораторії та набула гуманітарного виміру. Попри критику за ненауковість, цей підхід зробив неocenний внесок у розуміння особистості, мотивації, емоцій і психотерапії.

У Німеччині на початку ХХ століття виникла гештальтпсихологія – ще один напрям, що протистояв редукціонізму біхевіоризму. Макс Вертгеймер, Курт Коффка і Вольфганг Кьолер довели, що психічні явища не можна розкласти на елементи без втрати змісту: ціле більше, ніж сума частин.

Дослідження сприйняття показали, що людина спонтанно організовує стимули у структуровані цілісності (гештальти). Закони близькості, схожості, замикання і фігури-фону стали класикою психології сприйняття. Гештальтпсихологи також вивчали мислення – наприклад, феномен інсайту (раптового розв'язання проблеми). Їхні ідеї вплинули на подальші когнітивні теорії.

У середині ХХ століття психологія зіткнулася з так званою методологічною кризою. Її суть полягала в усвідомленні, що жодна з існуючих шкіл – ні біхевіоризм, ні психоаналіз, ні гештальтизм – не може повністю пояснити людську психіку. Кожна мала сильні сторони, але жодна не давала цілісної картини.

Цю кризу підсилювало зростання спеціалізації: з'явилися десятки піддисциплін (соціальна, клінічна, когнітивна, нейропсихологія), що втрачали спільну методологічну мову. Психологи почали усвідомлювати, що необхідно подолати розрив між природничо-науковими й гуманітарними підходами, між експериментом і тлумаченням.

Відповіддю на кризу став новий етап – когнітивна революція 1950–1960-х років. У цей період психологія знову звернулася до вивчення внутрішніх процесів – мислення, пам'яті, мови, уваги.

Когнітивна психологія повернула свідомість у центр науки, але на новому рівні – інформаційному. Вона використовувала експерименти, моделювання, штучні нейронні мережі, поєднуючи точність біхевіоризму з внутрішнім змістом феноменології.

Саме когнітивна парадигма дала старт нейронаукам, когнітивній лінгвістиці, когнітивній соціології. Вона показала, що свідомість – це не «чорна скринька», а складна система перетворення інформації, що може бути досліджена об'єктивно.

Когнітивна революція не лише повернула поняття свідомості у психологію, а й сформувала зовсім нову модель пояснення психічних процесів. Якщо біхевіоризм розглядав людину як пасивного реагента, то

когнітивна психологія побачила її як активного переробника інформації. Центральним метафоричним образом стала інформаційна система, яка отримує, зберігає, обробляє і видає дані – подібно до комп'ютера.

Ця парадигма дозволила пояснити внутрішню структуру мислення через етапи:

1. Сприйняття стимулу;
2. Кодування інформації;
3. Зберігання у короткочасній або довготривалій пам'яті;
4. Відтворення або прийняття рішення.

Джордж Міллер у класичній статті «The Magical Number Seven, Plus or Minus Two» показав, що обсяг короткочасної пам'яті людини обмежений приблизно сімома одиницями інформації. Це відкриття стало емпіричною основою для всієї когнітивної науки.

Ноам Чомскі, лінгвіст і філософ, розгромив поведінкову модель мови, показавши, що діти не просто наслідують почуті фрази, а створюють нові – тобто володіють вродженою граматикою. Цей аргумент остаточно поховав біхевіоризм як універсальну теорію навчання і зробив когнітивну науку головним напрямом психологічних досліджень другої половини XX століття.

Поступово когнітивізм розширився за межі власне психології. Він охопив когнітивну антропологію, когнітивну економіку, когнітивну соціологію – галузі, які вивчають, як людина пізнає світ у різних культурних і соціальних контекстах. З'явився термін когнітивна карта світу – внутрішня модель реальності, яку люди створюють, щоб орієнтуватися у складному середовищі.

Когнітивна парадигма також започаткувала розвиток когнітивної терапії, заснованої на роботах Аарона Бека та Альберта Елліса. Вони довели, що психічні розлади, зокрема депресія, пов'язані не лише з емоціями, а й із спотвореннями мислення – автоматичними негативними думками. Це дало змогу перейти від пояснення «несвідомих потягів» до практичної корекції мисленнєвих схем.

У результаті когнітивна революція створила новий образ людини як системи, що постійно інтерпретує, моделює і перетворює реальність. Саме тому сучасна психологія дедалі частіше говорить про когнітивну екологію – середовище ідей, знань і значень, у якому формується свідомість людини XXI століття.

Паралельно з когнітивним напрямом продовжує розвиватися гуманістична психологія. Абрахам Маслоу, Карл Роджерс, Віктор Франкл повернули у науку ідею людяності, свободи й сенсу. Вони стверджували, що людина – це не об'єкт для дослідження, а суб'єкт, який сам визначає напрям свого життя.

Карл Роджерс створив клієнт-центровану терапію, засновану на трьох принципах: безумовному прийнятті, емпатії та автентичності. Такий підхід не нав'язує людині «правильних» рішень, а створює умови для саморозкриття й самоцілення.

Абрахам Маслоу запропонував модель ієрархії потреб, на вершині якої – самоактуалізація, прагнення людини реалізувати власний потенціал. Його роботи започаткували «позитивний» погляд на психіку: не як на систему розладів, а як на механізм зростання.

Віктор Франкл, творець логотерапії, зосередився на пошуку сенсу як ключовому чиннику психічного здоров'я. Його дослідження досвіду концтаборів показали, що навіть у найекстремальніших умовах людина здатна зберігати гідність, якщо знаходить сенс у стражданні.

Гуманістична та екзистенційна традиції стали філософською протиположністю механістичним моделям психіки, нагадуючи, що будь-яка наука про людину має залишатися етичною.

Наприкінці ХХ століття з'явилася позитивна психологія – напрям, що вивчає не патологію, а ресурси людини: щастя, оптимізм, креативність, стійкість, альтруїзм. Її засновник Мартін Селігман запропонував концепцію PERMA – п'яти складових добробуту: позитивні емоції (Positive emotions), залученість (Engagement), стосунки (Relationships), сенс (Meaning), досягнення (Accomplishment).

Позитивна психологія базується на експериментальних і лонгітюдних дослідженнях. Наприклад, багато робіт демонструють, що практика вдячності підвищує рівень психологічне благополуччя, а емпатійне спілкування зменшує симптоми депресії. Таким чином, позитивна психологія об'єднує гуманістичні цінності з доказовими методами.

Популярність позитивної психології пояснюється не лише її «оптимістичним» тоном, а й намаганням зробити щастя об'єктом системного наукового вивчення. На відміну від філософських роздумів про благополуччя, цей напрям базується на емпіричних дослідженнях та довготривалих спостереженнях.

Одним із найвідоміших таких проектів став Harvard Study of Adult Development – найдовше лонгітюдне дослідження благополуччя, яке триває з 1938 року. Учені спостерігали за життям понад 700 людей протягом восьми десятиліть, щоб з'ясувати, що саме визначає щастя. Висновок виявився простим, але глибоким: найважливішим чинником є якість міжособистісних стосунків, а не матеріальний достаток чи соціальний статус.

Інший масштабний проект – World Happiness Report, що публікується ООН щорічно з 2012 року, – використовує великі вибірки для оцінки рівня задоволеності життям у різних країнах. Аналіз цих даних показує, що національний рівень благополуччя найбільше залежить від довіри, соціальної підтримки, очікуваної тривалості життя й свободи вибору. Це свідчить про те, що психологічне благополуччя є не лише індивідуальною, а й соціальною категорією.

Сучасна психологія визнає, що людська психіка не існує поза контекстом культури. Соціокультурна психологія, спираючись на ідеї Льва Виготського, вивчає, як мова, традиції та цінності формують мислення й

поведінку. Вона показує, що пізнання – це не індивідуальний процес, а спільна діяльність, укорінена в культурному середовищі.

У цьому ж руслі розвивається екопсихологія, яка досліджує взаємозв'язок людини й природи. Психологічний добробут, згідно з її концепцією, залежить від екологічної гармонії, а руйнування довкілля спричиняє не лише матеріальні, а й психічні наслідки – тривожність, апатію, втрату сенсу.

Сьогодні психологія існує в умовах методологічного плюралізму. Вона визнає, що жодна теорія не має монополії на істину, а кожна відображає лише один із рівнів психічної реальності. Наприклад, страх може бути одночасно фізіологічною реакцією (підвищення адреналіну), психологічним станом (відчуття тривоги) і соціальним явищем (колективна паніка). Тому наукове пояснення має бути мультирівневим. Сформована нова мета-парадигма передбачає поєднання кількісних і якісних методів. Для дослідження, скажімо, стресу, використовують одночасно опитувальники, біохімічні вимірювання та глибинні інтерв'ю. Це називається триангуляцією – перехресною перевіркою результатів із різних джерел.

Такий підхід дозволяє створювати цілісні моделі психіки, у яких біологічне, соціальне й особистісне не протиставляються, а взаємодіють. Сучасна психологія дедалі більше стає інтегративною наукою, що прагне описати людину в усій складності її природи – як біологічної істоти, соціального суб'єкта і носія свідомості.

У ХХІ столітті перед психологією постали нові виклики. Штучний інтелект, соціальні мережі, цифрова залежність, віртуальна реальність – усе це створює безпрецедентне середовище для людської психіки. З'явилися нові галузі: кіберпсихологія, що досліджує поведінку в онлайн-просторі, та нейроетика, яка аналізує моральні наслідки технологічного втручання в мозок.

Психологічні дослідження показують, що надмірне користування соціальними мережами може призводити до зниження самооцінки та відчуття ізоляції, тоді як помірковане використання – навпаки, підтримує соціальні зв'язки. Таким чином, психологія стає ключовим інструментом адаптації людини до технологічного світу.

Психологія ХХІ століття дедалі чіткіше усвідомлює свою місію: не просто пояснювати поведінку, а допомагати людині жити усвідомлено, гідно й цілісно. Сучасна наука про психіку повертає до життя головне – ідею, що науковість і людяність не суперечать одне одному.

Кінець ХХ – початок ХХІ століття став періодом інтеграції. Якщо раніше парадигми психології існували як конкуруючі табори, то тепер усе частіше говорять про плюралістичну єдність. Науковці визнають, що складність психіки вимагає багаторівневого пояснення: від нейронної активності до смислових структур свідомості.

Психологія перестала бути розділеною між «гуманітарним» і «природничим» полюсами. Сучасні дослідження дедалі частіше об'єднують кількісні вимірювання (експерименти, статистику) з якісними методами

(інтерв'ю, аналіз досвіду, інтерпретація поведінки). Цей підхід закріпився у понятті методологічного інтегралізму, який передбачає використання різних методів залежно від цілей дослідження, а не від ідеології школи.

Міждисциплінарність стала головним трендом: психологія активно взаємодіє з нейронауками, генетикою, соціологією, культурологією, економікою та навіть штучним інтелектом. У цьому полягає її сучасна сила: психологічне знання стає універсальним інструментом розуміння людини у світі технологій, соціальних змін і криз.

1.4. Комплекс методологічної неповноцінності у психології

Психологія як наука пройшла довгий і складний шлях розвитку – від античних спекуляцій про природу душі до сучасних експериментальних, когнітивних і нейронаукових підходів, які поєднують методи біології, інформатики, лінгвістики та соціології. Варто підкреслити, що сучасна психологія не існує у вакуумі – вона дедалі активніше взаємодіє з іншими дисциплінами, створюючи міждисциплінарні поля знання. Так, когнітивна наука інтегрує підходи філософії свідомості, штучного інтелекту та нейрофізіології; соціальна нейронаука вивчає, як соціальні процеси виявляються у мозкових структурах; психолінгвістика досліджує, як мова формує мислення. Цей перехід психології свідчить про її зрілість: лише через співпрацю з іншими галузями наука про психіку може пояснити складність людського буття.

Вона стала однією з найдинамічніших галузей наукового знання XX–XXI століть. Проте, попри стрімке накопичення емпіричних даних, поява тисяч досліджень щороку і розвиток потужних статистичних інструментів, психологія досі залишається наукою з нестійким методологічним підґрунтям. Це пов'язано з тим, що психіка людини є надзвичайно складним і багатовимірним феноменом. На відміну від фізичних об'єктів, її явища неможливо виміряти безпосередньо; вони завжди виявляються через поведінку, мовлення, невербальні прояви, що створює додатковий рівень невизначеності для дослідника. Саме тому навіть високотехнологічні методи, такі як нейровізуалізація чи експериментальні моделі, лише частково відображають реальність психічного життя. Цю ситуацію дослідники дедалі частіше описують терміном «комплекс методологічної неповноцінності». Це поняття є своєрідним узагальненням численних критичних зауваг до методологічного фундаменту психології як науки, її внутрішньої цілісності, суворості процедур, обґрунтованості інструментів вимірювання і реплікаційності досліджень.

Поняття цього комплексу з'явилося у наукових колах у середині XX століття, коли стало очевидно, що навіть найточніші експериментальні процедури не гарантують істинності висновків у гуманітарних науках. Своєю чергою, психологія виявилася особливо вразливою через подвійний статус – вона прагне точності природничих наук, але водночас має справу з живим,

мінливим людським досвідом. Дослідники, такі як П. Клайн і Д. Кемпбелл, підкреслювали, що ця «методологічна тривожність» є свідченням розвитку, а не слабкості: тільки наука, здатна сумніватися у власних основах, може оновлюватися. Окремо варто зазначити, що сучасна психологія не лише вивчає поведінку та психічні процеси, а й розробляє інструменти для їх прогнозування та модифікації, що робить її прикладною наукою.

Цей комплекс відображає низку глибинних суперечностей у теоретичній, емпіричній, інструментальній і соціальній сферах психології. Йдеться не лише про технічні труднощі чи нестачу об'єктивних показників, а про саму природу психологічного знання – його межі, критерії істинності, репрезентативність і відтворюваність. Як слушно зауважував К. Поппер, будь-яка наука проходить етапи становлення через сумнів у власних методах, і психологія сьогодні переживає саме такий етап.

На теоретичному рівні головною проблемою є відсутність єдиної інтегративної парадигми. Психологія складається з безлічі теоретичних напрямів, що часто не лише не співпрацюють, а й конкурують між собою. Психоаналіз З. Фрейда, біхевіоризм Дж. Вотсона і Б. Скіннера, гештальтпсихологія М. Вертгеймера, гуманістична школа К. Роджерса і А. Маслоу, когнітивна психологія Дж. Брунера і У. Найссера – усі ці напрями мають власні теоретичні моделі, методи і категоріальний апарат. Вони описують різні аспекти психіки, але не інтегруються в єдину систему знання. У результаті психологічна наука радше нагадує мозаїку або конгломерат часткових теорій, ніж монолітну дисципліну на кшталт фізики чи біології. Ця відсутність уніфікації призводить до того, що дослідження часто проводяться у вузьких рамках конкретної школи або підходу, що ускладнює формування загальної картини психічних процесів. Така ситуація породжує явище, яке деякі методологи називають «плюралізмом без діалогу». Кожна школа психології накопичує великий обсяг знань, проте між цими знаннями майже не відбувається обміну. Це призводить до того, що психологічна теорія залишається фрагментованою: когнітивна модель пояснює процеси мислення, але ігнорує емоційний вимір; психоаналіз заглиблюється у внутрішню динаміку особистості, проте часто нехтує емпіричною перевіркою; соціальна психологія фокусується на групових закономірностях, однак втрачає індивідуальний контекст. У результаті виникає методологічна «вежа з різних мов», де кожен говорить про людину, але різними категоріями. Наприклад, біхевіористичні методи можуть показувати закономірності поведінки, але вони не пояснюють мотиваційні та смислові аспекти, які вивчає гуманістична психологія. Така мозаїчність знання робить необхідним розвиток міждисциплінарних методів і моделей, здатних об'єднати різні рівні аналізу.

Додаткову складність створює множинність термінів і визначень, які вживаються у різних школах по-різному. Наприклад, поняття «особистість» може означати сукупність ролей (у соціальній психології), систему рис (у психометрії), структуру самості (в гуманістичній традиції) або внутрішню динаміку несвідомого (у психоаналізі). Це ускладнює порівняння результатів

різних досліджень і формування загальних висновків. Спроби створити метатеорії, як-от культурно-історичний підхід Л. Виготського чи системна психологія У. Бронфенбреннера, лише частково розв'язують проблему – на місце старої розрізненості приходить нова, складніша структура. Сучасні дослідники намагаються стандартизувати поняття через розробку так званих «конструктів» і теоретичних моделей, що дозволяє більш чітко визначати предмет дослідження.

На емпіричному рівні ключовим проявом комплексу методологічної неповноцінності є так звана реплікаційна криза. Реплікаційна криза стала не лише технічною проблемою, а й філософським викликом: вона поставила під сумнів уявлення про психологію як про кумулятивну науку, де знання поступово накопичується. Коли навіть класичні ефекти не відтворюються, виникає питання – що саме ми досліджуємо: властивості психіки чи особливості експериментальної ситуації? Це призвело до нового етапу саморефлексії науки, в якому точність поступається місцем прозорості, а кількість експериментів – їхній якості. Психологія почала визнавати власну контекстуальність: результати – не абсолютні істини, а моделі поведінки, що діють за певних умов.

Вона почалася приблизно у 2010-х роках, коли низка незалежних дослідницьких груп спробували повторити класичні експерименти соціальної та когнітивної психології. Результати виявилися тривожними: за різними оцінками, лише 35–40% експериментів дали подібні ефекти. Навіть фундаментальні ефекти, як-от «ефект сили волі» (ego depletion) чи «ефект первинності», не завжди відтворюються. Це поставило під сумнів надійність методології: малі вибірки, недостатня статистична потужність, «p-hacking» (маніпулювання даними для отримання значущих результатів) та відсутність попередньої реєстрації гіпотез стали предметом гострої критики. У відповідь на реплікаційну кризу зростає інтерес до методів пререєстрації плану дослідження та відкритого доступу до даних, що сприяє підвищенню надійності досліджень.

Проблема реплікації не обмежується соціальною психологією – вона торкнулася клінічної, дитячої, освітньої та навіть нейропсихологічної галузей. Це виявило глибшу проблему – так званий «кризовий реалізм»: дослідники змушені визнати, що експеримент у психології ніколи не є чистим, а суб'єкти дослідження – не абстрактні одиниці, а соціальні істоти, що інтерпретують сам факт участі у досліді. Тому сучасна методологія поступово переходить до природних експериментів, польових спостережень, лонгітудних досліджень. Такий зсув від контрольованого до життєвого контексту дозволяє зменшити штучність даних і краще відобразити динаміку реальної поведінки. Нерідко виявляється, що ефекти, отримані в одній культурі, не повторюються в іншій. Це означає, що психологічні феномени можуть мати культурну контекстуальність, тобто залежать від соціальних норм, мови та системи цінностей, у яких живе людина. Таким чином, універсальність психологічних законів виявляється ілюзорною. Дослідники підкреслюють необхідність

кроскультурних досліджень та адаптації методів для різних соціальних і мовних середовищ.

На інструментальному рівні труднощі зводяться до проблеми вимірювання психічних явищ. Психологічні конструкти – емоції, інтелект, мотивація, стрес, щастя – не мають прямого фізичного аналога. Ми не можемо «зважити» чи «виміряти» їх у традиційному сенсі. У психології доводиться використовувати опосередковані методи – тести, шкали, спостереження, інтерв'ю. Проте будь-який тест є лише модельним відображенням реальності, яке базується на припущеннях про зв'язок між поведінкою і психічними процесами. У цьому сенсі вимірювання у психології завжди має гіпотетичний характер. Ми не фіксуємо сам феномен, а лише його зовнішні прояви – вербальні відповіді, поведінкові реакції, біометричні сигнали. Тому між «даними» і «психікою» завжди залишається інтерпретаційний розрив. Цей розрив заповнюється теорією – саме вона надає сенс результатам вимірювань. Без теоретичного каркасу будь-який тест втрачає значення. Саме тому провідні методологи (А. Анастасі, Р. Стернберг) наголошують: жодне вимірювання не може бути нейтральним, воно завжди закорінене у певній науковій парадигмі. Як зазначав Л. Кронбах, «будь-яке вимірювання у психології є лише частковим наближенням до істини». Проблема полягає в тому, що надмірна стандартизація тестів може спотворювати індивідуальну картину, а недостатня – унеможливорює порівняння результатів. Тому дедалі популярнішими стають комбіновані методи: поєднання якісного аналізу з кількісними показниками, що підвищує достовірність висновків.

Крім того, кроскультурна невідповідність методик призводить до того, що тест, розроблений у США, може давати інші результати в Японії чи Україні, навіть після адаптації. Такі спотворення пояснюються як перекладними нюансами, так і культурними відмінностями у сприйнятті шкал оцінювання (наприклад, схильність до середніх або крайніх відповідей). Тому дедалі більшої уваги набувають концепції еквівалентності вимірювання (measurement invariance) та валідності конструкту, які закликають перевіряти, чи справді тест вимірює те саме у різних групах респондентів.

На соціальному рівні комплекс методологічної неповноцінності підриває авторитет психології як науки у суспільстві. З одного боку, поп-культура активно експлуатує психологічну термінологію: у медіа побутують поняття «токсичні стосунки», «аб'юзери», «тригери», які часто вживаються без фахового розуміння. З іншого – процвітають псевдонаукові напрями, такі як «нейрографіка», «астропсихологія» чи «хіромантия», які подаються як альтернативна психологія. Відсутність жорсткого професійного контролю, як у медицині чи юриспруденції, дозволяє некваліфікованим «спеціалістам» діяти під маркою психологів. Це підсилює скептицизм суспільства і формує уявлення про психологію як «м'яку» або «ненаукову» дисципліну. Окрім поп-культури та псевдонаукових практик, на сприйняття психології впливає і медійне висвітлення наукових результатів. Часто складні статистичні дані спрощуються до заголовків типу «вчені довели, що...», що формує у

громадськості відчуття суперечливості або непослідовності науки. Це створює парадокс: з одного боку, психологія стає все більш науково підкованою, з іншого – її реальні результати сприймаються як «непевні» або «сумнівні».

Історично цей стан має глибоке коріння. З часу, коли В. Вундт заснував психологічну лабораторію, психологія балансувала між двома полюсами – природничим і гуманітарним. Вона прагнула бути точною, як фізика, але зберігала інтерес до суб'єктивного досвіду, властивого філософії. Цей дуалізм і породив методологічну гетерогенність: одні школи тяжіли до вимірювань і статистики, інші – до розуміння і тлумачення смислів. Суперечка між прибічниками кількісних і якісних методів не вщухає й досі. Як підкреслював Дж. Брунер, «психологія розщепилася на дві культури: експериментальну і наративну».

Попри все, у ХХІ столітті з'являються тенденції, які свідчать про поступ у подоланні методологічної фрагментації. Однією з них є рух за відкриту науку, що вимагає прозорості даних, відкритого доступу до протоколів і публікації «негативних результатів» (тобто тих, що не підтвердили гіпотезу). Це сприяє зменшенню публікаційного упередження і підвищує довіру до емпіричних даних. Розвиваються також міждисциплінарні галузі – когнітивна нейронаука, психогенетика, афективна нейропсихологія, які поєднують суб'єктивні методи з об'єктивними нейрофізіологічними показниками. Таким чином, психологія поступово формує гібридну методологію, здатну інтегрувати природничо-наукові й гуманітарні принципи.

Важливу роль у зміцненні методологічної культури відіграє освіта. Сучасні програми підготовки психологів включають не лише класичні дисципліни, а й курси з наукової доброчесності, статистичного аналізу, етики досліджень. Формується нове покоління дослідників, орієнтованих на доказову практику і критичне мислення. Цей процес відображає загальну тенденцію до «методологічного реалізму» – усвідомлення, що психологічне знання не може бути абсолютно точним, але може бути максимально обґрунтованим. Освітні програми нового типу навчають не лише застосовувати статистику, а й розуміти її філософію: що означає значущість, як відрізнити причинність від кореляції, як уникати підтверджувальних упереджень. У результаті формується культура відкритої дискусії, де сумнів – не слабкість, а критерій науковості. Це дозволяє розглядати комплекс методологічної неповноцінності не як ознаку слабкості, а як точку росту науки, що саморефлексує і самовдосконалюється. Він вказує на важливість інтеграції, уточнення методології, підвищення стандартів освіти та практики. Якщо ці виклики будуть усвідомлені і прийняті як передумова до зростання, психологія зможе подолати внутрішні суперечності і зміцнити свій статус як наука, що досліджує одну з найскладніших і найважливіших сфер – людську психіку.

Щоб глибше зрозуміти природу психології як науки, варто звернутися до її методологічних принципів, що формуються на фоні подолання комплексу методологічної неповноцінності.

1.5. Методологічні принципи психології як науки про людину

Методологічні принципи – це те, що визначає наукову зрілість будь-якої дисципліни. У психології вони виконують роль концептуальної основи, без якої неможливо забезпечити логічну цілісність, доказовість і відтворюваність знань. Психологія як наука про людину унікальна тим, що поєднує природничі, соціальні й гуманітарні виміри. Тому її методологічні принципи повинні не лише описувати об'єктивні закономірності психічних процесів, а й враховувати суб'єктивність, контекст, історичність і моральну складову людського буття.

Психіка – це не ізольована структура, а живий процес, що розвивається у взаємодії з середовищем. Вона не підкоряється механічним законам природи, хоча має біологічну основу. Таке розуміння перегукується з ідеєю екологічного підходу Дж. Гібсона, який розглядав сприйняття як функцію взаємодії організму і середовища. Людина не просто реагує на стимули – вона постійно вибудовує власну «екологію сенсів», де об'єкти мають значення залежно від її потреб і цілей. Ця динаміка робить психіку відкритою системою, у якій біологічні, когнітивні та соціальні рівні постійно взаємодіють. Саме у такій інтеграції проявляється справжня складність психологічного об'єкта. Саме тому психологія виробила власну систему принципів, що поєднує об'єктивність наукового підходу і розуміння людського досвіду. Серед них ключовими вважаються принцип системності, принцип розвитку та принцип активності. Вони є не лише методологічними орієнтирами, а й філософськими засадами, що визначають спосіб мислення психолога.

Ідея системності увійшла в психологію під впливом загальної теорії систем Л. фон Берталанфі. Цей принцип виходить з того, що психіка – не сума окремих функцій, а цілісна організація, де всі процеси від відчуття до свідомості пов'язані між собою. Кожен психічний акт має багаторівневу природу: нейрофізіологічний, когнітивний, емоційний, мотиваційний і соціальний рівні функціонують у тісній взаємодії.

Системний підхід дозволяє уникнути редукціонізму, тобто спрощеного пояснення психіки лише через один рівень – біологічний, соціальний чи когнітивний. Наприклад, агресію можна трактувати як результат гормональних змін, як соціальний конструкт або як наслідок фрустрації. Лише системний аналіз дозволяє побачити, що всі ці чинники діють одночасно. У цьому сенсі психіка – це динамічна система, що саморегулюється, адаптується і підтримує рівновагу між внутрішнім і зовнішнім світом.

Важливим наслідком принципу системності є визнання взаємозв'язку між різними рівнями організації психіки – від мікропроцесів у мозку до макрорівня соціальних відносин. Це означає, що будь-яке психологічне явище слід аналізувати в контексті – не лише біологічному чи поведінковому, а й соціокультурному. Як підкреслював Л. Виготський, психічні функції не існують поза середовищем, у якому вони формуються. Саме середовище є джерелом розвитку, а психіка – засобом його освоєння.

Принцип розвитку в психології означає визнання того, що всі психічні явища перебувають у русі, а будь-який психічний стан є лише моментом у безперервній лінії еволюції особистості. Цей принцип має глибоке філософське підґрунтя – від діалектики Гегеля до культурно-історичної теорії Виготського, який підкреслював, що розвиток психіки людини відбувається через засвоєння суспільного досвіду у спільній діяльності.

У сучасній психології розвиток розглядається не лише як вікова зміна, а як універсальний процес, який триває протягом усього життя – від дитинства до старості. Людина постійно формує нові способи мислення, моделі поведінки, ціннісні орієнтири. Розвиток може мати поступовий характер, але часто супроводжується кризами, переломними моментами, які відкривають шлях до якісних змін. Як зазначав Е. Еріксон, саме кризи і внутрішні конфлікти є рушієм особистісного зростання.

Принцип розвитку вимагає від психолога враховувати часову перспективу дослідження: будь-яке явище потрібно розглядати у контексті його становлення, а не як фіксований стан. Це стосується як онтогенезу (індивідуального розвитку), так і філогенезу (еволюції психіки виду). Застосування цього принципу дозволяє зрозуміти, чому психологічні закономірності не можуть бути універсальними: те, що вірно для дитини, не обов'язково відповідає дорослій людині; те, що працює в одній культурі, може не діяти в іншій.

Важливо, що принцип розвитку передбачає ідею саморозвитку – здатність людини активно змінювати себе. Це підводить до наступного фундаментального принципу – активності.

Принцип активності стверджує, що людина – це не пасивний об'єкт, що реагує на зовнішні впливи, а активний суб'єкт власної життєдіяльності. Людина сама формує свій досвід, ставлення до подій, активно впливає на середовище, перетворює його та себе. Цей принцип протистоїть поглядам, згідно з якими психіка – це лише набір реакцій на стимули (як у класичному біхевіоризмі). Навпаки, людина має свободу волі, здатність до вибору, саморефлексії, відповідальності за свої вчинки. У психологічній практиці цей принцип означає, що кожна особистість здатна до змін, навіть за складних умов. Активність проявляється не лише у зовнішній поведінці, але й у внутрішній роботі – прийнятті рішень, подоланні внутрішніх конфліктів, постановці цілей. Наприклад, у психотерапії саме опора на активну позицію клієнта дозволяє досягати стійких позитивних змін.

Активність проявляється у всіх рівнях психічного життя: від елементарних актів уваги до складних форм саморефлексії. Навіть сприйняття – це не механічна реакція, а вибіркова, спрямована діяльність, у якій людина відбирає значущу інформацію. Як наголошував Рубінштейн, у психіці «відображення і діяльність – дві сторони одного процесу». Без активності немає розвитку, адже саме через власну дію, через випробування і помилки людина пізнає себе та світ.

Психологічна практика підтверджує значення цього принципу. У психотерапії, наприклад, клієнт не є пасивним реципієнтом допомоги, а активним учасником змін. Психолог лише створює умови, за яких людина може мобілізувати власні ресурси. У навчанні – це також основа сучасних освітніх технологій: знання не передаються, а конструюються самим учнем у процесі діяльності.

Принцип активності має також етичний вимір. Він передбачає, що людина відповідальна за власні рішення і за той вплив, який чинить на світ. Це протиставляє психологічний підхід фаталізму та біологічному детермінізму, наголошуючи на свободі вибору і творчому потенціалі особистості.

Системність, розвиток і активність – не ізольовані ідеї, а три сторони єдиного цілого. Системність описує структуру психіки, розвиток – її динаміку, активність – механізм змін. Разом вони формують цілісний підхід до розуміння людини як суб'єкта, що живе у взаємозв'язку з природою, суспільством і культурою.

Ці принципи допомагають психології зберігати баланс між науковою об'єктивністю й гуманістичним змістом. Цей баланс є центральною умовою зрілості будь-якої науки про людину. Якщо психологія надмірно тяжіє до природничого підходу, вона втрачає здатність розуміти внутрішній світ; якщо ж переважає гуманітарний компонент – зникає перевіряємість. Саме поєднання цих двох вимірів забезпечує методологічну гнучкість і глибину. Як показують сучасні дослідження з когнітивної нейробиології, наука і гуманістика не повинні конкурувати – вони описують різні, але взаємопов'язані сторони людського існування. Завдяки їм психологія не зводиться до статистики, а залишається наукою про сенси, вибір і внутрішню свободу. Водночас вони створюють основу для розвитку доказової практики – оскільки будь-яке психологічне втручання має спиратися на розуміння системності процесів, динаміки розвитку і активної ролі людини.

Методологічні принципи є не лише теоретичними положеннями, а й орієнтирами для практики. Вони визначають, як психолог ставиться до людини, як інтерпретує її поведінку, як вибудовує дослідження. Через них психологія поступово долає свій «комплекс методологічної неповноцінності», наближаючись до рівня інтегрованої, цілісної науки, що здатна одночасно вимірювати і розуміти.

Дослідження методологічних принципів психології логічно підводить нас до розгляду двох ключових наукових підходів у її розвитку – «природничо-наукової» та «гуманітарної» парадигм, які відображають різні способи розуміння природи людини та психіки.

1.6. Природничо-наукова та гуманітарна парадигми до дослідження особистості в психології

Проблема двох парадигм – природничо-наукової та гуманітарної – є однією з найглибших у методології психології. Вона відображає не просто

різницю у методах чи предметі дослідження, а глибинну суперечність між двома способами пізнання людини: через вимірювання і через розуміння. Ця дихотомія існує від самого зародження психології як науки, а її наслідки визначають обличчя всієї дисципліни до сьогодні. Питання двох парадигм – це не лише історичний факт, а й жива проблема сучасності. У ХХІ столітті воно проявляється в дебатах між прихильниками нейронаук і прибічниками феноменологічних методів. Одні вважають, що мозок – головний ключ до свідомості; інші стверджують, що свідомість не можна звести до нейронних процесів, бо вона завжди переживається зсередини. Ця суперечність є джерелом інтелектуальної напруги, що стимулює розвиток науки. Адже саме на межі між вимірюванням і розумінням виникають найцікавіші відкриття. Ці парадигми формують каркас, на якому ґрунтуються теорії, методи, критерії істинності та навіть роль психолога в суспільстві. Розуміння відмінностей між ними дозволяє краще орієнтуватися в різних напрямках психології та об'єктивно оцінювати їх можливості й обмеження.

Природничо-наукова парадигма психології бере свій початок із ідеалів позитивізму та механістичного світогляду ХІХ століття, які утвердились у фізиці, біології, хімії та медицині. Основна мета цього підходу – виявити об'єктивні закономірності психічної діяльності, які можна вивчати, вимірювати, узагальнювати й використовувати для передбачення поведінки. У межах цієї парадигми психіку розглядають як явище, що підпорядковується загальним законам природи. Людина в цьому підході – біопсихологічний об'єкт, складна система, яка реагує на стимули й чинники середовища. Найважливіші риси цієї парадигми – експериментальність, об'єктивність, статистична перевірка гіпотез, використання методів природничих наук. Прикладом є біхевіоризм, когнітивна психологія, нейропсихологія, психофізіологія, де дослідник шукає чіткі причинно-наслідкові зв'язки між психічними станами й поведінкою або між мозковими процесами й свідомістю. Ця парадигма вважає, що психіку можна досліджувати незалежно від її носія, тобто суб'єктивний досвід відходить на другий план.

Таке уявлення ґрунтувалося на біологічному редукціонізмі, який був типовим для наукової думки ХІХ століття. Психічні явища намагалися пояснити виключно через фізіологічні процеси. Ідеї Гельмгольца, Фехнера та Мюллера заклали основи психофізики – науки, що прагнула кількісно описати зв'язок між інтенсивністю стимулу і суб'єктивним відчуттям. Зокрема, закон Вебера–Фехнера став першою спробою «математизувати» досвід людини. Відтоді психологія прагнула наслідувати точні науки, створюючи власні стандарти вимірювання і відтворення результатів.

Саме тому у фокусі опинився експеримент – спосіб контролювати змінні, встановлювати причинно-наслідкові зв'язки і виключати «людський фактор». Під впливом природничо-наукової методології психологія почала розвивати поняття валідності, надійності, об'єктивності – тобто критерії, які раніше вважалися винятковою ознакою фізики чи хімії. Такий підхід дозволив

виробити мову кількісного опису психічного, але водночас відсунув суб'єктивний досвід на периферію.

Перші спроби зробити психологію «точною наукою» належать Вільгельму Вундту. Він прагнув вивчати психічні процеси так само, як фізіологи вивчають нервову систему – через контрольовані експерименти, вимірювання реакцій, часів сприйняття, інтенсивності відчуттів. З цього моменту психологія почала наслідувати модель природничих наук, запозичуючи їхній інструментарій – точність, статистику, гіпотетико-дедуктивний метод. Однак це наслідування мало не лише методологічні, а й філософські наслідки. Психологія опинилася під впливом позитивізму Огюста Конта, який вважав, що наука повинна займатися лише тим, що можна спостерігати і вимірювати. Це призвело до поступового витіснення понять «душа», «сене», «свідомість» – їх замінили більш «наукові» категорії: реакція, стимул, функція, процес. Завдяки цьому психологія отримала статус експериментальної дисципліни, але втратила частину гуманітарної сутності. Наукова раціональність перетворила психіку на об'єкт контролю, а людину – на механізм, який реагує на подразники. Такий підхід мав свої переваги – він створював простір для емпіричної перевірки гіпотез, але й породжував етичні питання: чи можна «вимірювати» внутрішній світ за тими самими законами, що й температуру чи тиск?

У ХХ столітті ця парадигма особливо розкрилася у біхевіоризмі. Джон Вотсон, а згодом Беррес Скіннер, стверджували, що психологія повинна вивчати лише те, що можна спостерігати – поведінку. Вони заперечували внутрішній досвід як об'єкт науки, оскільки він не піддається вимірюванню. Людина, з цієї точки зору, – це система «стимул → реакція». Ця модель, попри свою простоту, виявилася надзвичайно впливовою. Вона дозволила створити прогнозовані алгоритми поведінки, які використовувалися в освіті, рекламі, менеджменті, навіть у військовій підготовці. Експерименти Івана Павлова з умовними рефlekсами та роботи Едварда Торндайка про ефект закону результату стали фундаментом «поведінкової інженерії». Згодом цю модель уточнили когнітивісти, які ввели у неї поняття «обробки інформації», але зберегли природничо-наукову логіку: свідомість стала розглядатися як механізм, подібний до комп'ютера.

Перевага цього підходу – у його точності та відтворюваності. Природничо-наукова психологія дозволяє створювати стандартизовані методики діагностики, прогнозувати поведінкові реакції, розробляти інтервенції. Вона дає змогу інтегрувати психологію з медициною, генетикою, інформатикою. Проте її обмеження стають очевидними, коли йдеться про унікальний, неповторний досвід людини – про любов, віру, сене, творчість, страждання. Там, де починається людська індивідуальність, чиста об'єктивність перестає працювати. Саме ця межа стала точкою народження гуманітарної парадигми.

Гуманітарна парадигма психології виникла як реакція на редукціонізм природничих підходів. Вона спирається на філософські традиції

феноменології (Е. Гусерль), екзистенціалізму (М. Гайдеггер, Ж.-П. Сартр), герменевтики (В. Дільтей, Г. Гадамер). Тут основна увага приділяється унікальності людського досвіду, внутрішньому світу, суб'єктивності, особистісному змісту життя. Людина розглядається не як об'єкт, а як суб'єкт, який усвідомлює себе, шукає сенси, ставить цілі, приймає рішення. Головна мета дослідження – не встановити універсальні закони, а зрозуміти людину у її конкретному існуванні, в її культурному, соціальному, моральному контексті. У цій парадигмі немає жорсткого розділення між дослідником і об'єктом – дослідник сам включений у процес розуміння. Методи – якісні: інтерв'ю, аналіз біографій, наративів, спостереження, опис життєвих ситуацій, інтерпретація. Гуманістична, екзистенційна, наративна, культурно-історична психологія – яскраві представники цієї парадигми. Вона дозволяє вивчати складні й унікальні аспекти особистості, які неможливо звести до чисел і формул: любов, віру, творчість, відчай, відповідальність. Психологія тут – це наука про розуміння, а не про пояснення.

Таке протиставлення вперше чітко сформулював Вільгельм Дільтей, який розмежував природничі та духовні науки. Якщо перші прагнуть пояснювати події через закони, то другі – зрозуміти їх через смисл. Цей підхід знайшов потужне відображення у феноменологічній традиції Едмунда Гусерля, який закликав дослідника «повернутися до самих речей» – тобто до безпосереднього досвіду.

Феноменологічна психологія не заперечує емпірію, але наголошує: будь-яке знання про людину починається з її переживань. Це стало філософською основою для гуманістичної та екзистенційної психології ХХ століття, які зробили акцент на унікальності, свободі та відповідальності особистості.

Гуманітарна парадигма розглядає людину як істоту, що створює власну реальність через інтерпретацію досвіду. Її психіка – не механізм, а система смислів, що формуються у культурі. Відповідно, завдання психолога – не лише виміряти, а зрозуміти, як людина надає значення своїм діям і подіям.

Класичним прикладом гуманітарної парадигми є гуманістична психологія Карла Роджерса і Абрахама Маслоу. Вони наголошували, що людина – не об'єкт спостереження, а суб'єкт, здатний до саморозвитку, творчості, пошуку сенсу. Психологічна допомога, за Роджерсом, – це не маніпулювання поведінкою, а створення простору прийняття, у якому клієнт сам знаходить рішення.

У цьому ж руслі розвивалася екзистенційна психологія (В. Франкл, Р. Мей), яка бачить головне завдання психотерапії у допомозі людині знайти сенс свого існування навіть у межових умовах. Франкл, переживши концтабір, писав, що навіть коли всі зовнішні свободи відібрані, у людини залишається свобода ставлення до ситуації. Ця ідея демонструє, наскільки гуманітарна парадигма зосереджена на внутрішньому світі, духовності, унікальності особистого досвіду. Підхід Франкла й Мея дав людині не лише пояснення її страждання, а й шлях до подолання. Франкл підкреслював, що сенс не можна

нав'язати – його можна лише відкрити, і саме це відкриття є актом свободи. На відміну від біхевіористів, які прагнули контролювати поведінку, екзистенційні психологи шукали способи допомогти людині стати автором власного життя. Ролло Мей писав, що справжня зрілість полягає не у відсутності тривоги, а у здатності витримувати її, перетворюючи на творчу силу. Це радикально гуманітарна ідея: психічне здоров'я не вимірюється відсутністю симптомів, а глибиною осмислення буття.

Гуманітарна психологія також тяжіє до міждисциплінарності. Вона використовує досягнення антропології, культурології, літературознавства. Так, у нарративній психології Д. Брунера людське життя розглядається як історія, яку кожен розповідає сам про себе. Аналіз наративів дозволяє зрозуміти, як особистість конструює свою ідентичність і знаходить смисл у подіях.

Слабким місцем гуманітарного підходу є його обмежена верифікованість: результати часто суб'єктивні, не завжди відтворювані іншими дослідниками. Але це не недолік, а риса самої природи людського пізнання: смисли не можна зважити чи підрахувати. Тут критерій істини – не відтворюваність, а глибина розуміння.

Природничо-наукова і гуманітарна парадигми не є взаємовиключними – вони взаємодоповнюють одна одну. Перша забезпечує точність, об'єктивність і можливість передбачення; друга – глибину, контекст і моральний вимір. Без природничої парадигми психологія ризикує перетворитися на філософію без доказів; без гуманітарної – на технічну дисципліну без душі. Саме тому дедалі більше психологів виступають за інтегративний підхід, який долає стару дихотомію «об'єктивного» та «суб'єктивного». У сучасних дослідженнях дедалі частіше застосовують «змішані методи», що поєднують кількісні експерименти та якісний аналіз досвіду. Наприклад, у клінічних програмах вивчення депресії нейрофізіологічні показники аналізуються разом із наративами пацієнтів. Такий підхід створює більш повну картину психічного життя, де дані не витісняють сенси, а сенси уточнюють дані. Водночас це вимагає нової культури мислення – відкритої, міждисциплінарної, етично вмотивованої.

Сьогодні дедалі більше дослідників говорять про необхідність інтеграції цих підходів. Так виникає інтегративна парадигма – спроба поєднати експериментальну точність із феноменологічною глибиною. Наприклад, у сучасній нейропсихології вивчення мозкових процесів поєднується з аналізом суб'єктивних переживань (нейрофеноменологія, яку розвиває Ф. Варела). У психотерапії з'являються еклектичні моделі, де когнітивно-поведінкові методи поєднуються з гуманістичними підходами.

Ідея інтеграції парадигм також пов'язана з розвитком постнекласичної науки. Вона визнає, що повна об'єктивність неможлива, а дослідник завжди є частиною досліджуваної системи. Це особливо важливо у психології, де межа між суб'єктом і об'єктом тонка: психолог, спостерігаючи людину, впливає на неї, а отже, бере участь у самому феномені.

Завдання на самостійну підготовку

1. У чому полягає складність визначення предмета експериментальної психології?
2. Як різні психологічні підходи (поведінковий, когнітивний) по-різному визначають предмет експериментальної психології?
3. Які основні завдання виконує експериментальна психологія в сучасній науці?
4. Чому в експериментальній психології пріоритет має виявлення причинно-наслідкових зв'язків, а не описових закономірностей?
5. Як раціоналізм і емпіризм вплинули на становлення психології як наукової дисципліни?
6. У чому полягає внесок психофізики Фехнера та Вебера в становлення психології як науки?
7. Яку роль відіграв позитивізм у формуванні критеріїв науковості психології?
8. У чому полягає сутність фальсифікаціонізму Поппера та як він вплинув на сучасну психологічну методологію?
9. Чому операціоналізація стала ключовим інструментом розвитку психодіагностики та експериментальної психології?
10. Чому сучасна психологія характеризується методологічним плюралізмом і поєднанням кількісних і якісних методів?
11. У чому проявляється «комплекс методологічної неповноцінності» та які чинники його зумовлюють?
12. Як проявляється реплікаційна криза та чому вона стала важливим методологічним викликом для психології?
13. Які інструментальні труднощі пов'язані з вимірюванням психологічних конструктів (інтелекту, мотивації, тривожності тощо)?
14. Яким чином поп-культура, псевдонаука та медіа впливають на авторитет психології як наукової дисципліни?
15. Як рух за відкриту науку сприяє подоланню методологічних проблем сучасної психології?
16. У чому полягає зміст принципу системності психіки та як він запобігає редукціонізму?
17. Як принцип розвитку пояснює динамічність психічних явищ та їх залежність від контексту?
18. У чому полягає принцип активності та чому він є ключовим для розуміння людини як суб'єкта?
19. Які основні відмінності між природничо-науковою та гуманітарною парадигмами дослідження психіки?
20. Чому сучасна психологія потребує інтегративної парадигми та як поєднання експериментальної точності та феноменологічної глибини формує нові підходи до дослідження людини?

РОЗДІЛ 2

РЕПЛІКАЦІЙНА КРИЗА В ПСИХОЛОГІЇ ТА ШЛЯХИ ЇЇ ПОДОЛАННЯ

2.1. Суть та основні прояви реплікаційної кризи в психології

Реплікаційна криза в психології стала одним із наймасштабніших і найважливіших викликів для сучасної науки про людину, бо вона поставила під сумнів здатність психології продукувати стабільні, узагальнювані й надійні знання. Її суть полягає у невідповідності між результатами, опублікованими в наукових журналах, і тими даними, які отримуються при незалежному повторному проведенні експериментів. Тобто значна частина ефектів, описаних у класичних і сучасних дослідженнях, виявляється нестабільною або взагалі не відтворюється, коли інші дослідники точно відтворюють методи, процедури та умови первинного дослідження. Сам факт того, що результати, які тривалий час розглядалися як фундаментальні компоненти психологічного знання, не повторюються у незалежних реплікаціях, засвідчує глибинні проблеми в структурі емпіричної бази психології і спонукає до переосмислення її методологічних засад.

Найбільш показовими для розуміння суті реплікаційної кризи стали масштабні міжнародні реплікаційні проєкти. Найвідомішим серед них є проєкт Open Science Collaboration, у якому 270 науковців впродовж 2012–2015 років намагалися відтворити 100 експериментів, опублікованих у провідних психологічних журналах. Незважаючи на те, що 97 із цих досліджень містили статистично значущі результати, підтвердити їх вдалося лише у 36 випадках. Ці дані засвідчили, що відтворюваність ефектів, опублікованих у престижних журналах, не перевищує третини. Таке різке розходження між первинними та повторними результатами спровокувало глибоку дискусію щодо точності та стійкості знань, які психологія вважала перевіреними та емпірично підтвердженими.

Суть реплікаційної кризи полягає не лише в окремих випадках невдалих повторних досліджень, а у системному характері цього явища. Реплікації не повторюють результати не тому, що їх виконують непрофесійно або з порушенням процедур. Навпаки, багато реплікаційних проєктів були виконані з максимальною точністю, із використанням більших вибірок, кращого інструментарію та сучасних аналітичних методів. Незважаючи на це, величина ефектів у повторних дослідженнях часто виявлялася значно нижчою, ніж у первинних роботах, а інколи ефекти взагалі зникали. Ця ситуація демонструє ключову рису реплікаційної кризи: психологічні ефекти, що здавалися переконливими й широко цитованими, виявилися залежними від випадкових варіацій, локальних умов або специфічних характеристик окремих вибірок.

Одним із найважливіших проявів реплікаційної кризи є суттєве зниження величини ефектів у повторних дослідженнях. У великій кількості випадків реплікації демонстрували той самий напрямок ефекту, але його сила

була значно меншою, ніж зазначено в оригінальній роботі. Це свідчить про те, що первинні дослідження могли завищувати величину ефекту, а психологічні феномени, які здавалися сильними, насправді мають набагато слабший вплив. Такий прояв кризи є особливо важливим, адже саме величина ефекту визначає практичну значущість психологічних закономірностей та їхній внесок у формування теоретичних моделей.

Іншим проявом реплікаційної кризи є те, що велика кількість ефектів виявляється залежною від контексту. Психологічні процеси часто мають культурну, соціальну або ситуативну специфіку. Ефект, зафіксований у певній лабораторії, у певній країні або серед певної групи учасників, може не проявлятися за інших умов. Масові реплікаційні проєкти продемонстрували, що ефекти, які раніше вважалися універсальними, зникали або істотно змінювалися при проведенні досліджень у інших культурних контекстах чи з різними категоріями учасників. Це свідчить про те, що частина психологічних знань має значно обмеженішу сферу застосування, ніж передбачалося першопочатково, і потребує глибшого аналізу модераторів, які визначають стабільність або варіативність ефектів.

Ще одним характерним проявом кризи є відсутність повторюваності ефектів, які тривалий час вважалися класичними. Численні експерименти в галузях соціальної та когнітивної психології, що стали основою підручників, теорій і практичних рекомендацій, не вдалося підтвердити у сучасних реплікаційних проєктах. Наприклад, низка ефектів соціального праймінгу, що свого часу вважалися одними з найбільш переконливих доказів впливу неусвідомлюваних стимулів на поведінку, не витримали повторної перевірки в багатоцентричних дослідженнях. Ці невдачі не свідчать про недоліки оригінальних дослідників, але вони підкреслюють, що багато експериментальних парадигм були занадто чутливими до незначних змін у процедурі й не мали достатньої внутрішньої стійкості.

Важливим проявом кризи є також суттєві розбіжності між результатами, отриманими різними лабораторіями. Навіть коли дослідницькі групи використовують ідентичний протокол, однакові інструкції, ті самі матеріали та інструменти, результати можуть суттєво варіюватися. Це свідчить про наявність прихованих факторів, які впливають на поведінку учасників і не можуть бути повністю стандартизовані. Такі фактори можуть включати культурні відмінності, індивідуальні особливості вибірки, соціально-психологічний контекст лабораторії або навіть особистісні характеристики експериментатора. Така варіативність підкреслює складність психологічних процесів і демонструє, що багато ефектів є менш універсальними та значно більш залежними від ситуаційних факторів, ніж вважалося.

Суттєвим проявом реплікаційної кризи є також низька узгодженість між даними первинних і повторних досліджень. У деяких випадках оригінальні дослідження показували сильні й чіткі ефекти, тоді як повторні дослідження демонстрували повну відсутність ефекту або протилежний напрямок. Така ситуація є особливо проблемною, адже вона свідчить, що первинні висновки

не мають стабільного емпіричного підґрунтя. Для науки це означає, що значна частина теоретичних моделей, побудованих на таких результатах, повинна бути переглянута або переосмислена.

Не менш важливим проявом є те, що реплікаційна криза виявила розрив між офіційно опублікованими дослідженнями і реальним станом емпіричних даних у лабораторіях. У багатьох сферах дослідники знали про низьку відтворюваність певних ефектів, але ці відомості залишалися непублічними. Реплікаційна криза відкрила доступ до цих даних, зробивши проблему очевидною для широкої наукової спільноти. Вона оголила те, що психологічна література містила значну кількість результатів, які ніколи не були перевірені незалежними дослідниками, і що механізми наукової самокорекції діяли недостатньо ефективно.

Таким чином, суть реплікаційної кризи полягає у системному порушенні відтворюваності психологічних результатів, а її основними проявами є зниження величини ефектів у повторних дослідженнях, повна відсутність відтворюваності в окремих випадках, залежність результатів від контексту, нестабільність даних між лабораторіями, «зникнення» класичних ефектів при їх повторному тестуванні та зростання недовіри до емпіричної основи окремих теоретичних моделей. Реплікаційна криза не свідчить про неспроможність психології як науки, але вона демонструє необхідність системного оновлення методологічних стандартів, перегляду експериментальних парадигм та впровадження нових механізмів контролю за якістю й достовірністю наукових доказів. У цьому сенсі криза є не тільки проблемою, а й важливим етапом еволюції психологічної науки, який відкриває шлях до більшої прозорості, точності та наукової відповідальності.

Реплікаційна криза стала дзеркалом, у якому відобразилися структурні вади психології як науки: недостатня увага до статистичної потужності, неправильне використання *p-value*, домінування публікаційної орієнтації, непрозорість методології, відсутність практик відкритої науки, нестача системних реплікаційних досліджень і некритичне накопичення теоретичних моделей, що ґрунтувалися на поодиноких і безсистемних емпіричних ефектах. Важливо усвідомити, що криза не є свідченням неспроможності психології; навпаки, вона стала важливим етапом розвитку, який стимулював перегляд наукових стандартів. Історія науки показує, що кризи часто є передумовою прогресу: вони руйнують старі неефективні структури й відкривають шлях для формування нових, більш якісних методологічних засад. Саме такою кризою стала реплікаційна криза в психології – сигналом до глибоких структурних реформ, які спрямовані на підвищення надійності, відкритості та валідності психологічного знання. Усвідомлення причин кризи, їх критичне осмислення та системна перебудова практик дослідження є ключовими передумовами того, щоб психологія зберегла свій статус як емпірична дисципліна, здатна виробляти стабільні та узагальнювані знання про складну і багатовимірну природу людської поведінки.

2.2. Сумнівні дослідницькі практики як причини реплікаційної кризи

Серед численних причин реплікаційної кризи в психології особливе місце посідають так звані сумнівні дослідницькі практики, які стали однією з найглибших і найменш контрольованих проблем сучасної науки. Ці практики не завжди є результатом свідомого обману чи навмисної маніпуляції, навпаки, їхня поширеність пов'язана з тим, що вони поступово сформувалися в умовах академічної системи, яка довгий час стимулювала не стільки якісні, скільки «успішні» й статистично значущі результати. Сумнівні методологічні стратегії проникли у різні етапи наукового процесу – від формування гіпотез до вибору статистичних моделей, інтерпретації висновків і способів подання результатів. Усі вони в тій чи іншій мірі сприяли утворенню помилкових, випадкових або нестабільних ефектів, які потім ставали частиною психологічної літератури і сприймалися як емпірично доведені факти. Саме тому розуміння природи сумнівних дослідницьких практик є ключовим елементом аналізу причин реплікаційної кризи.

Найхарактернішою та однією з найбільш поширених сумнівних практик є *p-hacking* – маніпулювання аналітичними процедурами з метою досягнення статистично значущого результату. У цій практиці дослідник змінює модель аналізу, додає або виключає змінні, пробує різні варіанти очищення даних, переглядає критерії відбору учасників, змінює порядок контролювання коваріат або розглядає додаткові статистичні показники, поки значення *p* не стане нижчим за умовний поріг 0,05. Основна проблема полягає в тому, що, проводячи багаторазові тестування одних і тих самих даних, дослідник неминуче збільшує ймовірність випадкового виявлення ефекту – те, що з погляду статистики є закономірним, але з погляду наукової інтерпретації є хибною знахідкою. Такий ефект може бути цілком випадковим, але при цьому він отримує статус «наукового відкриття», оскільки задовольнив формальний критерій значущості. У психології, де традиційно використовуються вибірки невеликого обсягу, *p-hacking* стає особливо небезпечним, оскільки нестабільність даних підсилює вплив кожного статистичного рішення, збільшуючи ймовірність хибних висновків.

Іншою формою сумнівної практики є *HARKing* (*Hypothesizing After the Results are Known*) – формулювання гіпотез після того, як результати вже отримані, з поданням їх як первинних. Ця практика виникає з прагнення посилити теоретичну узгодженість дослідження, коли гіпотеза вибудовується з урахуванням реальних результатів, а не навпаки. Хоча в описовому дослідженні така практика може бути корисною для формування нових гіпотез, у контексті підтверджувального експерименту вона є науково неприйнятною, оскільки створює ілюзію передбачуваності й пояснювальної сили моделі. Гіпотеза, що виникла *postfactum*, часто виглядає переконливо, але не має підтверджувального статусу; вона лише оптимізує інтерпретацію наявних даних, а не передбачає нові факти. Внаслідок *HARKing* результати

можуть виглядати переконливіше, але вони позбавлені справжньої прогностичної сили, яка є фундаментальною ознакою наукової теорії. У психології ця практика набула широкого поширення через орієнтацію на публікацію, де дослідження з «непередбачуваними» або хаотичними результатами вважалися менш цінними.

Важливою складовою сумнівних практик є вибіркове повідомлення результатів (*selective reporting*). Дослідник може проводити багато аналізів, але в публікацію включає лише ті, що показали значущі або «очікувані» результати. Усе інше – слабкі, нульові, суперечливі ефекти – залишається поза межами статті. Хоча така практика може бути несвідомою і продиктованою стандартами наукового письма, вона суттєво спотворює наукову літературу. Вона створює ілюзію стабільності ефектів, тоді як у реальності дослідник спостерігав значну варіативність або відсутність будь-якої закономірності. Публікаційний здвиг, який полягає у відсутності публікацій із нульовими результатами, лише підсилює цей ефект, адже дослідник апріорі знає, що «негативне» дослідження має значно менші шанси бути опублікованим, а отже, несвідомо віддає перевагу повідомленню лише значущих фрагментів.

Іншою поширеною сумнівною практикою є довільне зупинення або продовження збору даних. У традиційній моделі перевірки статистичної значущості нульової гіпотези (NHST) збільшення обсягу вибірки підвищує статистичну потужність і може вплинути на значущість ефекту. Тому дослідник може, усвідомлено або ні, зупинити збір даних, коли отримано бажане значення *p-value*, або продовжити його, коли результат «не дотягує» до 0,05. Це порушує принцип рандомізації та робить висновки недостовірними, оскільки критерій зупинки повинен бути визначений до початку експерименту. У реальній практиці такі маніпуляції часто не фіксуються у публікаціях, що ускладнює оцінку достовірності результатів.

До сумнівних практик належить і маніпуляція змінними або спостереженнями. Вона може проявлятися у виключенні «незручних» учасників, зміні операціоналізації конструкта або включенні додаткових змінних у модель, якщо це покращує статистичний результат. У багатьох випадках дослідник має певні підстави для таких змін: наприклад, респондент не виконав інструкції або дані містять артефакти. Але за відсутності прозорої звітності межа між коректним очищенням даних і маніпуляцією стає розмитою. У результаті читач не може оцінити, чи були зміни виправданими, і чи могли саме вони спричинити появу значущого ефекту.

Не менш небезпечним проявом є використання надмірно гнучких статистичних моделей, що дозволяють отримати значущі результати навіть за відсутності справжнього ефекту. Психологічні дані часто містять значну кількість шуму, і дослідник може, використовуючи гнучкі моделі, «підігнати» дані під очікувану закономірність. У ситуації, коли вибірка мала, навіть незначні зміни в моделі здатні кардинально змінити результат. Така гнучкість аналізу зробила психологічні дослідження особливо вразливими до хибнопозитивних висновків.

Інституційні механізми, зокрема публікаційні вимоги журналів, також сприяли поширенню сумнівних практик. Журнали часто відхиляли роботи без значущих ефектів, що стимулювало дослідників «удосконалювати» аналіз так, щоб отримати бажані результати. У системі рецензування недостатньо уваги приділялося методологічній прозорості, а акцент робився на кінцевих висновках. Унаслідок цього наукова спільнота отримувала привабливо оформлені, але методологічно сумнівні роботи.

Ще однією причиною сумнівних практик є недостатній рівень статистичної грамотності багатьох дослідників. Невірне розуміння p -value, статистичної потужності, індексів розміру ефекту, корекції на множинні порівняння та інших аспектів аналізу призводить до неправильного застосування статистичних методів. Дослідники можуть щиро вважати результати значущими, не усвідомлюючи, що отримані значення є наслідком випадковості або неправильної моделі. У поєднанні з малою вибіркою це створює передумови для виникнення ілюзії ефекту, що зникає при повторному дослідженні.

Сумнівні дослідницькі практики не є ізольованими діями окремих учених – це системний феномен, укорінений у структурі наукового процесу. Вони виникають у контексті тиску на публікацію, конкуренції за ресурси, вимог до новизни та нестачі інституційної підтримки якісних досліджень. Саме тому для подолання реплікаційної кризи недостатньо закликати дослідників до етичної поведінки чи методологічної сумлінності, потрібна системна перебудова академічної культури. Вона включає реформу статистичної освіти, запровадження попередньої реєстрації, обов'язкове розкриття даних, підтримку реплікаційних досліджень, зміну редакційної політики та створення системи стимулів, які заохочуватимуть не «значущі» результати, а методологічно коректну, прозору і відтворювану науку.

Таким чином, сумнівні дослідницькі практики є однією з ключових причин реплікаційної кризи, оскільки саме вони формують підґрунтя для появи хибних, випадкових і нестабільних ефектів, які потім поширюються в науковій літературі й набувають статусу емпірично підтверджених закономірностей. Усуваючи ці практики шляхом реформування наукової культури, психологія має змогу повернути собі статус дисципліни, що ґрунтується на достовірних, перевірених і відтворюваних знаннях, створюючи фундамент для теоретичних моделей, які справді відображають складну природу людської психіки та поведінки.

2.3. Проблема публікаційного здвигу і селективності наукових журналів

Публікаційний здвиг і селективність наукових журналів становлять одну з ключових і найглибших причин реплікаційної кризи, оскільки вони визначають структуру того, що стає доступним для наукової спільноти та використовується як підґрунтя для подальших теоретичних узагальнень.

Природа цієї проблеми полягає в тому, що у процесі рецензування та відбору до публікацій надається значна перевага дослідженням, які містять статистично значущі результати, тоді як роботи з нульовими, слабкими або суперечливими ефектами рідко знаходять місце у провідних виданнях. Така форма селекції створює системне викривлення наукової літератури: у ній накопичуються лише ті ефекти, які, можливо, виникли випадково або були отримані за рахунок сумнівних методологічних прийомів. У той час як дослідження, що не підтверджують певні гіпотези або не виявляють очікуваного ефекту, часто залишаються неопублікованими й невідомими для інших науковців. Унаслідок цього література відображає не об'єктивний стан досліджуваної проблеми, а лише її вибірккову, «успішну» частину, що суттєво знижує можливість формування адекватних теоретичних моделей. Саме публікаційний здвиг став одним із головних каталізаторів реплікаційної кризи, оскільки створив середовище, у якому наукові результати оцінювалися не за якістю методології, а за відповідністю критерію статистичної значущості.

Публікаційний здвиг нерозривно пов'язаний із феноменом «культури значущості», що формувався протягом десятиліть у психологічній науці. Журналам, щоб підтримувати високий імпакт-фактор і наукову «видимість», було вигідніше публікувати роботи з новими, сильними та «сенсаційними» ефектами. Дослідження, які не демонстрували статистично значущих результатів, вважалися менш цікавими, менш цитованими і тому «непридатними» для видання. У результаті рецензенти і редактори фактично формували не просто фільтр академічної якості, а фільтр статистичної значущості. Така тенденція сприяла накопиченню великої кількості досліджень, які пройшли в наукову літературу лише тому, що містили формально значущі ефекти, незалежно від їхньої реальної наукової цінності або практичної важливості. У той же час сильні за методологією, але «негативні» дослідження не отримували можливості бути почутими. Таким чином, у науковому просторі формувався своєрідний «ефект айсберга», коли тільки верхівка – найбільш привабливі, але не обов'язково надійні результати – ставала доступною, тоді як велика частина емпіричних даних залишалася прихованою.

Ця селективність створює серйозні проблеми для накопичення наукового знання. Відсутність інформації про нульові та негативні результати спотворює уявлення дослідників про реальну силу ефектів і спричиняє перебільшення значущості тих, що були опубліковані. Коли в літературі домінують лише дослідження зі значущими результатами, метааналітичні оцінки ефектів стають систематично завищеними. Це призводить до того, що психологічні явища починають виглядати значно більш сильними, стабільними та універсальними, ніж вони є насправді. У своїй суті публікаційний здвиг руйнує можливість отримати адекватну картину наукової реальності, бо дослідники більше не можуть розрізняти, які ефекти є справді стійкими, а які виникли випадково. Такий стан речей не тільки ускладнює

побудову теорій, але й сприяє появі «ефектів-привидів», які тривалий час циркулюють у науці без жодного реального емпіричного підґрунтя.

Однією з найважливіших причин публікаційного здвигу стала інституційна система, яка винагороджує продуктивність і новизну вище, ніж методологічну строгість. Університети, наукові фонди й академічні комітети оцінюють дослідників за кількістю публікацій у журналах з високим імпакт-фактором. У такому середовищі дослідник усвідомлює, що його кар'єра залежить не від того, наскільки точним і ретельно спланованим є його дослідження, а від того, чи може він продемонструвати значущі результати. Цей тиск може бути непрямим, але його наслідки відчутні: дослідники несвідомо віддають перевагу стратегіям, які підвищують шанси отримати статистично значущі висновки. Поступово формується культура, у якій «негативні» дослідження не лише не цінуються, але й вважаються ознакою невдачі. Це, у свою чергу, підштовхує дослідників до проведення множинних аналізів, гнучкого тлумачення методів, вибіркового подання даних – і всі ці практики посилюють *p-hacking*, *HARKing* та інші форми сумнівної дослідницької поведінки, які вже були окреслені як одні з головних причин реплікаційної кризи.

Селективність наукових журналів також має наслідки для структури наукової дискусії. Коли до публікації потрапляють лише роботи зі значущими висновками, дискусія будується на неповній інформації. Оригінальні дослідження рідко перевіряються незалежними групами, але при цьому стають підґрунтям для десятків теоретичних публікацій, оглядів, методичних рекомендацій та навчальних посібників. Ідеї, що не мають стабільного емпіричного підтвердження, перетворюються на наукові «факти» лише тому, що вони були опубліковані та активно тиражувалися. У цей спосіб публікаційний здвиг працює як механізм самопідсилення: опубліковані «значущі» ефекти отримують наукову вагу, навіть якщо вони нестабільні; натомість непублікація негативних результатів не дозволяє виправити помилкові уявлення.

Ситуацію ускладнює і те, що багато психологічних журналів традиційно віддавали перевагу новим, оригінальним та концептуально цікавим результатам, нехтуючи дослідженнями реплікаційного характеру. Реплікації вважалися «несенсаційними» й такими, що не сприяють зростанню імпакт-фактору. Унаслідок цього навіть добре виконані реплікаційні проєкти часто не отримували можливості бути опублікованими, а значна кількість негативних реплікацій залишалася у «сірій зоні» – існувала лише в особистих архівах дослідників або у вигляді неопублікованих звітів. Саме тому реплікаційна криза стала настільки глибокою: вона була прихованою десятиліттями через вибіркковість публікацій та орієнтацію журналів на ефекти, що підкріплюють популярні теорії. Таким чином, селективність журналів діяла як механізм, що посилює всі інші форми методологічних спотворень.

Публікаційний здвиг має і важливі статистичні наслідки. Наявність у літературі переважно значущих результатів призводить до систематичного

завищення оцінок ефектів у метааналізах, що далі підсилює помилкові уявлення про силу і стійкість психологічних впливів. Коли метааналізи виконуються на основі лише опублікованих даних, вони неминуче відтворюють і підсилюють упередження, властиві початковому корпусу літератури. У деяких галузях психології метааналітичні оцінки ефектів були завищені в кілька разів порівняно з реальними показниками, отриманими у масових реплікаційних проєктах. Це виявляє глибоку системну проблему: метааналіз, який колись вважався «золотим стандартом» накопичення знань, стає ненадійним інструментом, якщо дані для нього вибираються вибірково.

У контексті реплікаційної кризи особливо важливо враховувати взаємодію публікаційного здвигу з іншими системними чинниками. Наприклад, сумнівні дослідницькі практики стають значно більш поширеними тоді, коли дослідник знає, що журнал не прийме його дослідження без значущих результатів. Аналітична гнучкість стає способом «адаптації» до очікувань редакторів. Низька статистична потужність досліджень, поєднана з селективною публікацією, створює середовище, у якому випадкові ефекти мають непропорційно великі шанси потрапити у літературу. Цей цикл самопідсилення сприяє накопиченню псевдонаукових результатів і створенню ілюзії емпіричної підтримки теорій, що не мають стабільного підґрунтя.

Важливим етапом подолання публікаційного здвигу стала поява нових редакційних підходів, зокрема формату попередньої реєстрації плану дослідження (Registered Reports), у якому рішення про публікацію ухвалюється до проведення дослідження – лише на основі теоретичного обґрунтування та якості методології. Такий формат дозволяє усунути прив'язаність публікації до значущості результатів і кардинально зменшує можливості для р-hacking та інших сумнівних практик. Крім того, усе більше журналів зобов'язують авторів надавати дані, матеріали та аналітичні скрипти, що є кроком до подолання інформаційної закритості. Однак ці реформи поки що охоплюють лише частину академічного середовища, і повне подолання публікаційного здвигу вимагатиме глибших, системніших змін на всіх рівнях наукової системи.

Таким чином, публікаційний здвиг і селективність наукових журналів становлять фундаментальний механізм, що протягом десятиліть формувал викривлену картину психологічних досліджень і став однією з ключових причин реплікаційної кризи. Вони створили умови, у яких значущість стала важливішою за достовірність, новизна – важливішою за відтворюваність, а поодинокі випадкові ефекти – важливішими за систематичні наукові перевірки. Подолання цього феномена вимагатиме докорінної трансформації редакційної політики, перегляду критеріїв наукової цінності та переходу до культури відкритості й відтворюваності, яка дозволить психології повернути собі статус дисципліни, заснованої на надійному та перевіреному знанні.

2.4. Статистична потужність як фундаментальний чинник реплікаційної кризи

Одним із найглибших і найчастіше недооцінених чинників реплікаційної кризи у психології є проблема недостатньої статистичної потужності досліджень, яка протягом десятиліть непомітно формувала середовище, сприятливе для появи випадкових, нестабільних і невідтворюваних результатів. Статистична потужність (power) – це ймовірність того, що дослідження виявить реальний ефект, якщо він справді існує. Іншими словами, це здатність експериментального дизайну відрізнити сигнал від шуму. Дослідження з низькою потужністю не лише з великою ймовірністю пропускають справжні ефекти (високий ризик помилки другого роду), але й водночас суттєво підвищують імовірність появи хибнопозитивних висновків. Низька потужність створює ситуацію, у якій навіть випадкові коливання в даних можуть виглядати як статистично значущі закономірності. Саме тому проблема статистичної потужності є одним з основних джерел нестабільності психологічних результатів та їх невідтворюваності у незалежних дослідженнях.

Психологія, як дисципліна, тривалий час характеризувалася використанням малих вибірок, що зумовлено як історичними особливостями, так і специфікою дослідницького процесу. Багато експериментів проводилися на зручних вибірках студентів, кількість яких обмежувалася ресурсами конкретної лабораторії. У той час як у біологічних або медичних науках вибірка вважається недостатньою, якщо вона не дозволяє виявити основні ефекти, у психології довгий час домінувала практика, за якої 20–30 учасників вважалися прийнятним обсягом вибірки для серйозного емпіричного дослідження. Однак з погляду статистики такі вибірки рідко забезпечують потужність, що перевищує 30–40 %, тобто дослідження з самого початку має більшу ймовірність не виявити реального ефекту, ніж знайти його. Проте ще більш небезпечним є те, що низька потужність збільшує ймовірність того, що якщо дослідження і виявить значущість, то цей ефект буде випадковим або значно завищеним порівняно з реальним значенням.

Проблема низької статистичної потужності проявляється не лише у пропусканні реальних ефектів, але й у формуванні хибних уявлень про величину й стабільність виявлених ефектів. Невеликі вибірки створюють умови, у яких результати стають надзвичайно чутливими до випадкових флуктуацій. Навіть невеликі відхилення в поведінці учасників або методологічні нюанси можуть спричинити різкі зміни у показниках. Унаслідок цього дослідження з низькою потужністю можуть давати виключно позитивні або виключно негативні результати, які не відповідають реальній величині ефекту.

Проблема низької потужності має системний характер і підсилюється іншими факторами, зокрема селективністю публікацій та сумнівними дослідницькими практиками. Дослідники, які працюють із малими вибірками,

часто вдаються до аналітичних стратегій, що дозволяють «підсилити» ефект: вони можуть вибирати ті змінні або ті підвибірки, де результати здаються більш значущими. Така практика створює ілюзію, що навіть малі вибірки здатні демонструвати «сильні» ефекти, хоча в реальності ці ефекти є нестабільними й випадковими. Це своєю чергою підсилює публікаційний здвиг: журнали публікують лише значущі результати, не враховуючи їхню реальну статистичну надійність. Унаслідок цього психологічна література стає насиченою ефектами, які виявилися «позитивними» лише тому, що були отримані в умовах низької потужності.

Ще однією важливою проблемою є те, що дослідники часто не проводять попереднього аналізу потужності (power analysis), який дозволяє визначити необхідний розмір вибірки. Це зумовлено як відсутністю достатньої статистичної підготовки, так і ресурсними обмеженнями. У багатьох випадках дослідники самостійно вирішують, яка вибірка є «достатньою», керуючись аналогіями з попередніми дослідженнями або рекомендаціями наставників, але не математичними обґрунтуваннями. Навіть коли аналіз потужності проводиться, він часто базується на завищених очікуваннях щодо величини ефекту. Якщо дослідник припускає, що ефект є сильнішим, ніж він є насправді, це призводить до недооцінки розміру вибірки. Внаслідок цього дослідження виявляється нерепрезентативним і не здатним виявити ефект належним чином.

Принциповою особливістю психологічних досліджень є висока варіативність поведінкових реакцій і вплив різноманітних контекстуальних факторів, що роблять результати більш шумними порівняно з багатьма природничими науками. Ця шумність повинна компенсуватися збільшенням вибірки або застосуванням більш точних методів вимірювання, але в реальності цього не відбувається. У результаті велика частина психологічних досліджень існує у зоні малої потужності, де шанс виявити стабільний ефект є невисоким. У такій ситуації навіть незначні зміни в експериментальних умовах можуть істотно вплинути на результат, що робить дослідження чутливими до найменших варіацій. Саме тому багато класичних ефектів у психології не витримали перевірки у масових реплікаційних проєктах – їхня стабільність була ілюзорною і виникала лише в умовах слабких, нестійких даних.

Особливе значення має проблема завищеної величини ефектів, що спостерігається у дослідженнях з низькою потужністю. Це явище відоме як «ефект помилкового підсилення» (inflation bias). Коли ефект виявляється значущим у дослідженні з малою вибіркою, це, як правило, означає, що реальна величина ефекту або була завищена випадковим чином, або отримана внаслідок шумних даних. У той же час дослідження з низькою потужністю значно рідше публікують нульові результати, тому спільнота бачить лише завищені ефекти й не бачить безлічі досліджень, у яких ефект не був виявлений. Така ситуація сприяє формуванню викривленої картини, у якій психологічні впливи здаються сильнішими, ніж вони є насправді.

Проблема низької статистичної потужності особливо гостро проявляється у дослідженнях, де очікувані ефекти є невеликими або середніми за величиною. Багато психологічних явищ справді мають слабкий ефект, оскільки на поведінку людини впливає велика кількість незалежних чинників. У такій ситуації дослідження потребують дуже великих вибірок, щоб виявити ефект із достатньою точністю. Проте, у психології довгий час недооцінювалася необхідність великих вибірок, що призводило до втрати значної кількості інформації.

Реплікаційні дослідження, такі як Many Labs Projects, продемонстрували, що ефекти, які вважалися «сильними» на основі малих вибірок, часто зникають або суттєво зменшуються при аналізі на великих вибірках. Це свідчить про те, що багато опублікованих результатів були наслідком низької потужності, а не реальних психологічних закономірностей. Дослідження з малою потужністю не лише сприяють появі хибних висновків, але й унеможлиблюють точну оцінку величини ефекту. Теоретичні моделі, які будуються на основі таких результатів, неминуче виявляються слабкими і складно витримують перевірку.

Проблема статистичної потужності має також інституційний вимір. Дослідницькі інститути, грантові комісії та університети рідко встановлювали вимогу щодо попереднього аналізу потужності, і дослідники часто планували свої експерименти виходячи не з наукових, а з ресурсних можливостей. У таких умовах навіть найбільш добросовісні дослідники не завжди можуть дозволити собі великі вибірки. Особливо це стосується галузей, де збір даних є складним або дорогим – наприклад, у клінічній психології, розвитку або когнітивній нейронауці. Таким чином, проблема потужності не є суто методологічною – вона також пов'язана з організаційними й економічними аспектами наукової практики.

Щоб подолати проблему низької статистичної потужності, необхідно здійснити системну реформу методологічних стандартів. Передусім, дослідники повинні проводити попередній аналіз потужності, визначаючи розмір вибірки відповідно до очікуваної величини ефекту. Не менш важливим є впровадження відкритих практик, таких як попередня реєстрація, яка дозволяє запобігти маніпуляціям із вибіркою та аналізом. Крім того, важливо заохочувати створення мультилабораторних проєктів та досліджень із великими вибірками, які здатні забезпечити стабільність і точність оцінок. Журнали повинні вимагати від авторів інформацію про статистичну потужність, а також підтримувати публікацію досліджень з нульовими результатами, щоб уникнути селективності публікацій.

Отже, статистична потужність є одним із фундаментальних чинників реплікаційної кризи, оскільки вона визначає якість і стабільність емпіричних висновків. Низька потужність сприяє виникненню хибних результатів, їх завищенню та нестабільності, що унеможлиблює відтворюваність. Подолання цієї проблеми потребує глибокого перегляду методологічних стандартів, зміни інституційних стимулів та формування нової культури дослідження, яка

орієнтується не на швидкі публікації, а на достовірність, точність і відтворюваність наукових висновків.

2.5. Роль традиційної парадигми перевірки статистичної значущості нульових гіпотез у реплікаційній кризі

Одним із найбільш суттєвих і водночас найглибших методологічних чинників реплікаційної кризи в психології стало неправильне тлумачення та некритичне використання p -value у межах традиційної парадигми перевірки статистичної значущості нульових гіпотез (Null Hypothesis Significance Testing – NHST). Протягом багатьох десятиліть NHST залишалася основним статистичним інструментом психологічних досліджень, що значною мірою визначав спосіб мислення, прийняття рішень та інтерпретацію результатів у науковій спільноті. Однак поступово саме цей інструмент став одним із центральних джерел системних помилок, що сприяли накопиченню нестабільних, випадкових та невідтворюваних ефектів. Неправильне розуміння p -value, надмірна довіра до його значення, формальне ототожнення статистичної значущості з науковою важливістю, а також численні методологічні спотворення у застосуванні NHST призвели до того, що психологічна література виявилася насиченою результатами, які не витримують перевірки у повторних дослідженнях. Саме тому аналіз ролі NHST та інтерпретацій p -value є критично важливим у розумінні причин реплікаційної кризи.

У традиційній моделі NHST p -value визначається як ймовірність отримати такі або більш екстремальні результати за умови, що нульова гіпотеза є істинною. Тобто p -value не є ні ймовірністю істинності гіпотези, ні мірою сили ефекту, і тим більше не є гарантією того, що результат можна відтворити. Проте в психологічній практиці p -value протягом багатьох років трактувалося саме таким чином. Дослідники часто інтерпретували низьке p як доказ того, що альтернативна гіпотеза є істинною або що нульова гіпотеза є хибною. Таке тлумачення є фундаментально помилковим, адже p -value не може говорити про істинність гіпотези – воно лише інформує про відповідність даних певній статистичній моделі. Проте ця модель є надзвичайно загальною і залежить від численних припущень, які в реальних експериментальних умовах рідко дотримуються повністю.

Одним із найбільш поширених методологічних хибнотлумачень є сприйняття значущості як доказу існування ефекту й незмінності величини цього ефекту. У численних психологічних дослідженнях статистична значущість почала виступати частиною наукової риторики, яка «легітимує» висновки автора. Результат із $p < 0,05$ ставав необхідною умовою для публікації, а часто – і для визнання дослідження науково значущим. У такій ситуації дослідники були схильні приділяти увагу не теоретичному обґрунтуванню, не якості дизайну і не точності вимірювання, а саме

досягненню статистичної значущості. Це спричинило появу численних практик, орієнтованих на «оптимізацію» p -value.

P -value стало не просто статистичним показником, а фактично символом наукового успіху, що визначав долю дослідження. Саме тому багато наукових висновків виявилися слабкими та невідтворюваними: вони існували лише тому, що дослідник знайшов спосіб отримати значення p трохи нижче за умовний поріг 0,05.

Поріг значущості $\alpha = 0,05$, запропонований Фішером у середині ХХ століття як умовна межа для прийняття рішень, із часом перетворився з евристичного орієнтира на жорстку вимогу. У межах психологічної літератури значення p стало розглядатися як бінарний маркер: результат або «значущий», або «незначущий». Такий підхід суттєво спростив логіку інтерпретації, але призвів до низки серйозних методологічних помилок. По-перше, він ігнорує величину ефекту: навіть надзвичайно слабкий ефект у великій вибірці може дати $p < 0,05$, при цьому не маючи жодного практичного значення. По-друге, він створює хибну ілюзію певності: дослідники схильні вважати, що значущий результат є «доказом», у той час як статистичний критерій насправді ніколи не дає достовірного підтвердження чи спростування гіпотези. Така бінарна логіка суперечить природі психологічних явищ, які рідко бувають однозначними, із чіткими межами та стабільними ефектами. Проте саме така логіка формувала науковий дискурс упродовж десятиліть.

Неправильне тлумачення p -value нерозривно пов'язане, як уже говорили, з низькою статистичною потужністю психологічних досліджень. У дослідженнях із малою вибіркою навіть слабкі або випадкові ефекти можуть стати значущими, якщо випадкові коливання даних співпадуть у потрібному напрямі. У таких умовах p -value стає надзвичайно непередбачуваним показником: воно часто опускається нижче порогу значущості не тому, що існує справжній ефект, а тому, що дослідник випадково потрапив у статистичну комбінацію, яка створила ілюзію закономірності. За відсутності повторних досліджень ці випадкові ефекти починають сприйматися як реальні та використовуватися для формування теоретичних висновків. Унаслідок цього психологічна література накопичує значну кількість «ефектів», які є лише статистичним артефактом.

Не менш важливою проблемою є те, що p -value не забезпечує інформації про відтворюваність ефекту. Два дослідження можуть мати однакові значення p , але при цьому мати абсолютно різні величини ефекту або різний рівень варіативності у вибірках. Значущість одного експерименту нічого не говорить про стабільність результату у повторному аналізі. Проте протягом тривалого часу p -value сприймалося як універсальний критерій «наукової істини», що призвело до того, що науковці часто некритично приймали значущі результати як підтвердження гіпотез. Саме ця особливість моделі NHST спричинила накопичення величезної кількості нестабільних ефектів, які не повторюються у реплікаційних дослідженнях.

Роль NHST у реплікаційній кризі полягає не лише у неправильному тлумаченні p -value, але й у структурних обмеженнях самої моделі. NHST передбачає перевірку гіпотези про відсутність ефекту, але не дає можливості отримати оцінку ймовірності альтернативної гіпотези. Тобто система побудована так, що вона не може сказати, наскільки великою є ймовірність того, що ефект існує – вона лише оцінює ступінь невідповідності даних нульовій моделі. Крім того, NHST не передбачає інтерпретації щодо практичної значущості або величини ефекту, і тому дослідники часто зосереджуються на показнику значущості, ігноруючи параметри, які мають більший науковий зміст. Система NHST також стимулює формування дослідницьких стратегій, орієнтованих на досягнення формальної значущості, а не на побудову теоретично і методологічно обґрунтованих моделей.

Ще однією проблемою є те, що NHST не враховує множинних порівнянь, які часто проводяться у психологічних дослідженнях. Коли дослідник виконує багато статистичних тестів одночасно, ймовірність отримати хоча б один хибнопозитивний результат суттєво зростає. Однак у багатьох психологічних роботах ця проблема не враховувалася, що призводило до того, що навіть при відсутності реального ефекту дослідники отримували значущі висновки. Такий підхід робив p -value ще більш ненадійним, особливо в умовах великих наборів даних, де дослідники мали змогу тестувати безліч змінних та взаємодій.

У межах реплікаційної кризи особливо важливою стала проблема «мислення за порогом». Коли дослідники орієнтуються лише на те, чи опустилося p нижче 0,05, вони часто ігнорують важливу інформацію, таку як довірчі інтервали, розміри ефектів і точність вимірювань. Така практика формує спрощену, бінарну модель наукового мислення, яка не відповідає складності психологічних явищ. Вона також створює враження, що ефекти є стабільними та універсальними, якщо вони «пройшли тест на значущість». Проте реплікаційні проєкти довели, що значущі ефекти часто не повторюються, а незліченна кількість «непозитивних» досліджень залишилася непоміченою.

З огляду на масштаб проблем, пов'язаних із NHST і неправильним тлумаченням p -value, сучасні наукові спільноти активно пропонують альтернативи. Однією з таких альтернатив є використання оцінок величини ефекту та довірчих інтервалів – підходу, який дозволяє побачити не лише факт наявності ефекту, а й його силу та точність. Іншим напрямом є баєсові методи статистики, що дозволяють отримувати ймовірнісні оцінки для гіпотез. Сучасні журнали дедалі частіше вимагають від авторів не лише p -value, а й повний набір статистичних показників, що робить результати більш інформативними. Однак ці реформи поки що охоплюють лише частину науки, тоді як структурні проблеми NHST продовжують впливати на реплікаційність результатів.

Таким чином, неправильне тлумачення p -value та некритичне використання NHST стали одними з ключових чинників реплікаційної кризи в психології. Система, що спочатку була створена як інструмент оцінки статистичної відповідності, перетворилася на жорстку парадигму, що визначає логіку наукових публікацій. Унаслідок цього психологічна література наповнилася значущими, але нестабільними результатами, які не витримують перевірки. Подолання цієї проблеми вимагає кардинальної зміни підходів до статистичного аналізу, переосмислення ролі p -value, впровадження альтернативних методів оцінки даних та формування нової культури дослідження, що зосереджується не на значущості, а на достовірності, точності та відтворюваності наукових висновків.

2.6. Проблема недостатньої прозорості досліджень та відсутність відкритих даних

Недостатня прозорість досліджень та обмежений доступ до вихідних даних є одним із найважливіших і водночас найменш усвідомлених чинників, які сприяли розгортанню реплікаційної кризи в психології. Цей аспект проблеми виявився особливо критичним, коли наукова спільнота отримала можливість порівнювати результати первинних досліджень із тими висновками, які отримували незалежні групи, що намагалися відтворити ті самі експериментальні процедури. Виявилося, що відсутність доступу до первинних даних, аналітичних алгоритмів, матеріалів дослідження, протоколів процедур та критеріїв обробки інформації значною мірою унеможлиблювала точне відтворення досліджень, навіть коли реплікатори прагнули виконати роботу максимально коректно. У цьому контексті прозорість досліджень перестає бути простою етичною вимогою і перетворюється на фундаментальну умову наукової достовірності та відтворюваності результатів.

Упродовж десятиліть у психології домінувала практика, за якої опублікована стаття містила лише підсумкові статистичні показники, короткий опис вибірки та загальні процедури проведення експерименту. У статтях рідко подавалися детальні інструкції, стимульні матеріали, контрольні змінні, інформація про виключення учасників, процедури очищення даних, опис проміжних аналізів чи критерії прийняття рішень щодо включення або виключення даних. У результаті реплікаційні спроби часто виявлялися утрудненими через те, що дослідники не мали доступу до детального протоколу експерименту, а багато критично важливих рішень залишалися прихованими. Дослідницька практика передбачала широке використання «лабораторного знання», яке існувало лише у формі внутрішніх домовленостей, неформальних традицій чи інтуїтивних рішень дослідника. Таке знання не документувалося, не публікувалося і не було доступним для тих, хто намагався повторити експеримент. Унаслідок цього реплікатори часто створювали процедури

заново, ґрунтуючись на власних інтерпретаціях тексту статті, що призводило до розбіжностей у результатах досліджень.

Відсутність відкритих даних стала ще одним критично важливим чинником, який значно ускладнив перевірку достовірності психологічних висновків. У більшості опублікованих досліджень вихідні дані ніколи не були доступними громадськості. Це створювало ситуацію, у якій незалежні дослідники не могли перевірити правильність статистичних розрахунків, адекватність застосованих моделей, обґрунтованість виключення окремих учасників або правильність вибору аналітичних процедур. У науковій літературі накопичувалася велика кількість результатів, які не могли бути незалежно перевірені, а отже, їхня стійкість залишалася невідомою. В умовах, коли статистичні методи часто є складними, а аналітичні рішення можуть суттєво впливати на підсумкові висновки, відсутність доступу до даних перетворювала кожне дослідження на своєрідну «чорну скриньку», зміст якої був недоступний іншим науковцям.

Під час великих реплікаційних проєктів стало очевидно, що навіть мінімальні розбіжності у процедурі обробки даних можуть суттєво вплинути на результати. Наприклад, різні методи очищення даних, способи обробки пропущених значень, об'єднання змінних, фільтрація за критеріями швидкості реакції чи виключення «екстремальних» значень можуть призводити до істотно різних статистичних висновків. Якщо автори первинного дослідження не описували ці процедури детально, а дані залишалися недоступними, то реплікатори опинялися в ситуації, коли відтворення результатів ставало практично неможливим, навіть якщо вони використовували той самий експериментальний дизайн. У цій ситуації відсутність прозорості фактично не дозволяла перевірити, чи були результати первинного дослідження стійкими, чи вони були наслідком специфічних рішень щодо обробки даних, які інші дослідники не могли повторити.

Недостатня прозорість виявилася пов'язаною не лише з відсутністю відкритих даних, але й із відсутністю доступу до аналітичних скриптів. Сучасні статистичні аналізи часто включають складні процедури, які можуть реалізовуватися у спеціалізованих програмних середовищах. Навіть незначні зміни у реалізації алгоритмів можуть призвести до відмінностей у результатах. Коли аналітичні скрипти залишаються недоступними, реплікація стає складним завданням, що потребує від дослідників не лише повторення методів, але й реконструкції всього аналітичного шляху. Водночас навіть наявність вихідних даних не гарантує можливості точно повторити аналіз без доступу до кодів, які визначають логіку обробки. Відсутність відкритих аналітичних скриптів унеможливорює незалежну перевірку правильності статистичних процедур та робить науковий результат вразливим до прихованих методологічних похибок.

Ще одним важливим аспектом недостатньої прозорості стало обмеження доступу до матеріалів дослідження. У багатьох експериментальних

парадигмах стимульні матеріали, інструкції, візуальні або текстові стимули, а також алгоритми взаємодії з учасником є критично важливими компонентами дизайну. Зміни навіть у дрібницях – у формулюванні інструкції, кольорі стимулу, розміщенні елементів на екрані чи часі експозиції – можуть суттєво змінити поведінку учасників. Коли ці матеріали недоступні, реплікатори змушені створювати власні версії, що неминуче призводить до варіацій, які можуть пояснити розбіжності у результатах. У такій ситуації навіть добросовісна реплікація може не дати результатів, які збігаються з оригіналом, не через те, що ефект не існує, а через те, що дослідники використовують матеріали, які не повністю відповідають первинним.

Недостатня прозорість також стосується неповного висвітлення процедур. Часто дослідники описують лише основні етапи експерименту, опускаючи деталі, які вважають другорядними або очевидними. Однак саме ці деталі можуть відігравати ключову роль у формуванні результатів. Наприклад, спосіб рекрутингу учасників, особливості інструктажу, умови проведення експерименту, особливості поведінки експериментатора, попередній досвід учасників – всі ці фактори можуть впливати на результати, але рідко описуються у статтях. У сучасних реплікаційних проєктах стало очевидно, що відтворення результатів залежить від повної документації всіх процедур, навіть тих, які раніше здавалися незначущими.

Ще одним проявом недостатньої прозорості стало те, що дослідницькі групи рідко повідомляли про всі спроби аналізу або про проміжні дослідження, які передували опублікованим результатам. У статтях зазвичай відображався лише «успішний» аналіз, тоді як невдалі або непослідовні результати залишалися поза увагою. Це створювало викривлену картину: результати виглядали більш переконливими, ніж вони є насправді, а реплікатори не мали інформації про те, які саме умови були необхідними для виявлення ефекту. Відсутність доступу до повної історії дослідження спотворювала уявлення про його стійкість і робила реплікацію надзвичайно складною.

У підсумку недостатня прозорість досліджень та відсутність відкритих даних стали фундаментальними структурними чинниками, які значною мірою сприяли розгортанню реплікаційної кризи. Вони створили середовище, у якому наукова спільнота не могла ефективно перевіряти, відтворювати або уточнювати результати. Реплікаційна криза показала, що прозорість є не додатковою, а необхідною умовою наукової надійності. Без доступу до даних, матеріалів, процедур і аналітичних скриптів наука перетворюється на діяльність, що не піддається незалежній перевірці, а отже – втрачає здатність до самокорекції. Сучасні реформи відкритої науки демонструють, що прозорість може стати ключем до подолання кризи й формування нової культури дослідження, у якій надійність і відтворюваність стають центральними критеріями наукової якості.

2.7. Шляхи подолання реплікаційної кризи в психологічних дослідженнях

Шляхи подолання реплікаційної кризи в психології формують складний багаторівневий процес, який передбачає глибоку трансформацію наукової культури, стандартизацію методологічних підходів, реформування публікаційних практик, запровадження механізмів відкритої науки та переосмислення принципів статистичної перевірки гіпотез. Ця криза виявила не лише технічні недоліки у способах збору та оброблення даних, а й фундаментальні проблеми організації наукової діяльності: від структури стимулів до системи мотивації дослідників. Тому її подолання не може бути зведене до впровадження окремих процедур або технік; воно потребує комплексної реформи, що охоплює як індивідуальний рівень (підготовка дослідника та його методологічна компетентність), так і рівень наукових інституцій, журналів, професійних спільнот і міжнародних регулятивних практик. Центральним елементом цієї трансформації є розвиток відкритої науки, яка забезпечує прозорість усіх етапів дослідження, починаючи від планування і закінчуючи публікацією та архівуванням даних. Відкрита наука передбачає доступ до матеріалів, протоколів, аналітичних скриптів і первинних даних, що дозволяє іншим дослідникам перевіряти, аналізувати та відтворювати результати незалежно від авторів оригінального експерименту. Така практика формує новий рівень наукової відповідальності, коли автор не лише презентує результат, але й демонструє повний шлях його отримання.

Запровадження відкритих даних та відкритих матеріалів створює низку важливих переваг. По-перше, це підвищує якість дослідження: коли матеріали доступні для перегляду та перевірки, дослідник більш відповідально підходить до складання протоколу. По-друге, це стимулює розвиток кумулятивної науки: інші дослідники можуть використовувати вже створені матеріали, порівнювати результати між собою, проводити метааналізи, будувати складніші експериментальні дизайни. По-третє, це зменшує ризик прихованих методологічних спотворень, оскільки відкритість створює механізм зовнішнього контролю, який знижує імовірність вибіркового подання або маніпуляції даними. Нарешті, відкритість даних сприяє формуванню нових етичних стандартів, коли наукова спільнота визнає відповідальність не лише за створення, а й за перевірку знань.

Важливим засобом подолання реплікаційної кризи є також поширення практики попередньої реєстрації досліджень. Пререєстрація передбачає, що дослідник фіксує гіпотези, план аналізу, методологію, розмір вибірки та основні змінні ще до початку збору даних. Це дозволяє чітко розмежувати підтверджувальні й дослідницькі аналізи, що значно знижує імовірність появи маніпуляцій з даними, притаманних сумнівним дослідницьким практикам. Коли дослідник проводить аналіз за заздалегідь задекларованим планом, то навіть при отриманні незвичних або небажаних результатів він не може змінити гіпотезу постфактум або подати випадкову закономірність як

підтвердження первинного задуму. Пререєстрація змінює саму структуру психологічного експерименту: дослідник має заздалегідь продумати логіку аналізу, здійснити розрахунок статистичної потужності, чітко визначити очікувані ефекти і сформулювати критерії їхньої оцінки. Це стимулює методологічну дисципліну, якої довгий час бракувало психологічним дослідженням.

Ще одним важливим напрямом подолання кризи є зміна підходів до статистичного аналізу. Психологія протягом багатьох десятиліть значною мірою покладалася на класичну процедуру перевірки нульових гіпотез (NHST) та критерій $p < 0,05$. Унаслідок цього p -value почало сприйматися як універсальний критерій істинності результату, хоча за своєю природою воно не дає інформації про практичну значущість, величину ефекту або ймовірність повторення. Сучасні рекомендації закликають доповнювати p -value іншими статистичними індикаторами, зокрема розмірами ефекту, довірчими інтервалами та попередньо розрахованою статистичною потужністю. Цей підхід дозволяє уникнути надмірно формального тлумачення результатів і забезпечує більш реалістичне уявлення про їхню силу та стабільність. У багатьох журналах вимога надавати розміри ефектів і довірчі інтервали стала обов'язковою, що свідчить про зміни у науковій культурі.

Ще один аспект, який суттєво сприятиме подоланню реплікаційної кризи, полягає у збільшенні статистичної потужності досліджень. Для цього необхідно проводити попередні оцінки розміру вибірки, виходячи з очікуваної величини ефекту, а не з ресурсних обмежень чи традиційних практик лабораторії. Збільшення вибірок зменшує ймовірність випадкових результатів, підвищує точність оцінок і дозволяє отримати ефекти, які можна з високою ймовірністю відтворити незалежно від умов. Проте збільшення вибірок передбачає не лише кількісні, а й якісні зміни: дослідники повинні ретельніше підбирати методи вимірювання, зменшувати рівень шуму в даних, підвищувати стандарти контролю експериментальної ситуації. Особливо важливим є забезпечення репрезентативності вибірок, адже багато психологічних досліджень тривалий час ґрунтувалися на зручних вибірках студентів університетів, що суттєво обмежувало можливість узагальнення.

Одним із найважливіших шляхів подолання кризи є зміна редакційної політики наукових журналів. Традиційно журнали надавали перевагу новим і «привабливим» результатам, тоді як реплікаційні дослідження або роботи з нульовими результатами часто не проходили рецензування. Проте реформи, що почалися у другій половині 2010-х років, спрямовані на те, щоб змінити цю тенденцію. Зокрема, у передових сучасних наукових журналах запроваджено формат пререєстрації, коли рішення про публікацію приймається не після аналізу результатів, а на етапі розгляду методологічного плану. Це дозволяє розірвати залежність між статистичною значущістю та шансом на публікацію. Крім того, більшість сучасних журналів дедалі активніше підтримують публікацію реплікаційних досліджень, що раніше вважалися периферійними.

Такий підхід сприяє формуванню культури критичної перевірки знань і зменшує ризик нагромадження псевдонаукових ефектів.

Розвиток відкритої науки також посилив тенденцію до зберігання й поширення сирих даних, аналітичних скриптів, матеріалів експерименту та протоколів. Це дозволяє не лише перевірити правильність аналізу, але й здійснювати вторинні аналізи, метааналізи, об'єднання даних між дослідницькими групами, що значно підвищує якість наукової аргументації. У свою чергу, це стимулює формування нових підходів до побудови теорій, коли дослідник орієнтується не на одиничні ефекти, а на стійкі закономірності, підтверджені різними методами в різних контекстах.

У межах подолання реплікаційної кризи особливого значення набувають масові міжнародні реплікаційні проєкти, такі як Many Labs, Psychological Science Accelerator, Open Science Collaboration та інші ініціативи, що передбачають координацію десятків або навіть сотень лабораторій. Вони дозволяють перевіряти стабільність ефектів у різних соціальних, культурних, демографічних та методологічних контекстах. Масові реплікації створюють нову модель наукового знання, у якій істинність не визначається успіхом одного експерименту, а перевіряється у широкому спектрі умов. Такі ініціативи значно підвищують якість наукових висновків, адже результати, які успішно відтворюються в різних лабораторіях, мають набагато сильніший статус доказовості.

Подолання реплікаційної кризи передбачає також перегляд теоретичних засад психології. Багато теоретичних моделей, які формувалися на основі нестабільних емпіричних ефектів, мають бути переоцінені. Нові теоретичні підходи повинні бути більш точними, фальсифікованими, перевірюваними й відкритими до систематичної критики. Теорії мають враховувати контекстуальну залежність ефектів, багатовимірність психологічних явищ і необхідність повторної перевірки. Таким чином, реплікаційна криза не лише виявила слабкі місця психології, але й показала, у якому напрямі повинна рухатися дисципліна.

Важливим результатом реакції на кризу стало також посилення уваги до методологічної освіти. Сучасні програми підготовки психологів усе частіше включають розширені курси зі статистики, методів аналізу даних, планування досліджень, критичної інтерпретації результатів. Майбутніх дослідників навчають не лише технічним навичкам, але й науковій етиці, принципам прозорості та відповідальності. Такий підхід дозволяє формувати покоління дослідників, які здатні працювати в умовах високих стандартів доказовості.

У ширшій перспективі подолання реплікаційної кризи передбачає зміну самої філософії наукової діяльності. Наука повинна відмовитися від орієнтації на короткострокову продуктивність і перейти до моделі, де цінуються точність, ретельність і відтворюваність. Це передбачає переосмислення системи мотивації дослідників, коли якість дослідження стає важливішою за кількість публікацій. Криза поставила на порядок денний необхідність створення умов, у яких реплікаційні дослідження, метааналізи та відкриті дані

оцінюються не менш високо, ніж оригінальні відкриття. Така зміна вимагає часу, але вона є необхідною для довгострокового розвитку психології.

Отже, шляхи подолання реплікаційної кризи охоплюють глибоку реформу наукової практики – від запровадження відкритої науки та попередньої реєстрації до зміни редакційної політики, перегляду статистичних стандартів, збільшення потужності вибірок, розвитку міжнародних реплікаційних проєктів та переоцінки теоретичних моделей. Усе це разом формує нову культуру дослідження, яка орієнтується на надійність, прозорість і відтворюваність. Реплікаційна криза стала не лише викликом, але й можливістю для психології – можливістю змінити свої методи, стандарти і цінності, щоб створювати наукове знання, яке дійсно відображає складну природу людської поведінки та психіки.

Завдання на самостійну підготовку

1. У чому полягає основна сутність реплікаційної кризи в психології та чому вона стала одним із найважливіших викликів сучасної науки?
2. Чому багато «класичних» експериментів не змогли пройти перевірку у сучасних реплікаційних проєктах?
3. Як реплікаційна криза виявила розрив між офіційно опублікованими результатами та реальним станом даних у лабораторних практиках?
4. У чому полягає практика р-hacking і чому саме вона є одним із найбільш небезпечних чинників реплікаційної кризи?
5. Що таке HARKing і чому ця практика порушує логіку підтверджувального експерименту?
6. Яким чином вибіркове повідомлення результатів (selective reporting) викривлює наукову літературу?
7. Чому довільне зупинення або продовження збору даних є методологічно некоректним і як це впливає на достовірність результатів?
8. Чому гнучкість статистичних моделей підвищує ризик появи хибнопозитивних ефектів у психологічних дослідженнях?
9. Яким чином інституційний тиск на публікацію стимулює поширення сумнівних дослідницьких практик?
10. Як публікаційний здвиг формує викривлену картину психологічної науки і чому він став одним із ключових чинників реплікаційної кризи?
11. Чому у традиційній редакційній політиці наукові журнали віддавали перевагу «значущим» результатам, а не методологічно строгим дослідженням?
12. Яким чином відсутність публікацій нульових результатів впливає на метааналітичні оцінки психологічних ефектів?
13. Як низька статистична потужність досліджень сприяє появі випадкових, нестабільних та невідтворюваних ефектів?
14. Чому дослідження з малою вибіркою часто призводять до завищення величини ефекту?

15. Які наслідки має відсутність попереднього аналізу потужності під час планування експерименту?

16. У чому полягає основна методологічна помилка у тлумаченні p-value в психологічних дослідженнях?

17. Чому бінарна логіка «значущий/незначущий» результат є непродуктивною для психологічної науки?

18. Як множинні порівняння та неправильне використання NHST підвищують ризик появи хибнопозитивних висновків?

19. Яким чином відсутність відкритих даних, матеріалів та аналітичних скриптів ускладнює реплікацію психологічних досліджень?

20. Які ключові елементи відкритої науки є найбільш ефективними для подолання реплікаційної кризи та підвищення надійності психологічних досліджень?

РОЗДІЛ 3 ПЛАНУВАННЯ НАУКОВОГО ДОСЛІДЖЕННЯ

3.1. Етапи наукового дослідження

Планування наукового дослідження становить одну з базових передумов ефективності пізнавальної діяльності в сучасній психології. Саме від того, наскільки глибоко й методологічно вивірено сплановано дослідницький процес, залежить не лише достовірність і валідність отриманих результатів, а й їхня практична цінність, можливість реплікації, застосування у професійній діяльності. Будь-яке наукове дослідження – це не просто сукупність дій, спрямованих на збирання даних, а цілісний системний процес, у якому кожен елемент пов'язаний із попереднім і наступним логічними ланцюгами. Планування виступає тим організаційним і концептуальним каркасом, що надає дослідженню цілісності, визначає його логіку, структуру, часові та змістовні межі.

Наукове планування у психології має особливу специфіку, адже психічні явища, що виступають предметом дослідження є динамічними, багатовимірними і не завжди безпосередньо спостережуваними. Психологічні процеси не піддаються прямому вимірюванню в тій мірі, як фізичні або хімічні явища, тому дослідник повинен не лише правильно обрати метод, а й адекватно сконструювати дизайн дослідження, передбачити контроль змінних, уникнути суб'єктивних спотворень. Це можливо лише тоді, коли планування є науково обґрунтованим, спирається на логіку гіпотетико-дедуктивного методу, а також на сучасні уявлення про структуру психологічного знання. В цій площині виникає потреба застосування мультиметодового підходу, що поєднує самозвіти, поведінкові спостереження та біологічні показники (наприклад, нейропсихологічні або психофізіологічні дані). Це дозволяє інтегрувати інформацію про психічне явище на різних рівнях аналізу, підвищуючи його конструктивну валідність та зменшуючи ризик інтерпретаційних помилок.

Планування охоплює всі етапи наукового процесу – від вибору теми й формулювання проблеми до оформлення результатів і публікації. Власне, воно є процесом «мислення наперед», у ході якого дослідник моделює майбутній експеримент або емпіричне вивчення, передбачаючи всі можливі труднощі, ризики, варіанти інтерпретації даних. Наука – це не лише накопиченням фактів, а радше системою гіпотез, що перевіряються, і жодна гіпотеза не може бути адекватно перевірена без попереднього логічного плану. Таким чином, планування – це не технічна підготовка до дослідження, а центральний момент наукового мислення.

Наукове мислення в процесі планування – це осмислення природи наукового знання. Коли дослідник вибудовує план, він створює уявну структуру, яка відображає не саму дійсність, а її теоретичну репрезентацію. У

цьому сенсі планування має подвійний характер: воно водночас є пізнавальним процесом і формою контролю над ним.

Методологічна логіка планування передбачає рух від абстрактного до конкретного: спочатку визначається теоретична парадигма (наприклад, когнітивна чи соціально-культурна), потім – категоріальний апарат, далі – емпіричні змінні. Порушення цієї логіки призводить до того, що зібрані дані виявляються непридатними для інтерпретації.

Особливе місце посідає питання відповідності між рівнями теоретичного та емпіричного. У доброму плані кожен термін, що використовується в гіпотезі, має бути операціоналізований, тобто мати конкретний індикатор. Наприклад, якщо говоримо про «мотивацію досягнення», потрібно визначити, як вона буде вимірюється, якими показниками буде визначатися. Без цього логічний зв'язок між поняттям і вимірюванням розривається, і дослідник фактично працює з фікцією.

Планування, таким чином, є проявом наукової рефлексії – здатності мислити про власне мислення. Дослідник оцінює не лише результати, а й способи їх досягнення, постійно запитуючи себе: чи відповідають мої дії принципам наукової раціональності, чи не змішую я рівні аналізу, чи не підмінюю пояснення описом. Ця рефлексивність є фундаментом зрілої наукової культури.

Планування виконує низку функцій. По-перше, воно забезпечує логічну впорядкованість процесу дослідження, допомагаючи визначити послідовність дій. По-друге, воно сприяє методологічній обґрунтованості – кожен вибір методу, інструменту чи способу обробки даних має бути узгоджений із метою й гіпотезою. По-третє, планування дає змогу оптимізувати ресурси – часові, матеріальні, людські та зменшити ризики неефективності чи повторення помилок. Нарешті, воно формує етичну культуру дослідника, адже вже на етапі планування необхідно передбачити процедури забезпечення добровільної згоди, конфіденційності, уникнення шкоди учасникам.

Тобто, планування наукового дослідження – це його каркас, без якого будь-яка ідея ризикує залишитися на рівні спостережень, інтуїції чи випадкових знахідок. Добре сплановане дослідження – це не просто структурований набір дій, а інтелектуальна стратегія, що перетворює хаос фактів на впорядковану систему знань.

Першим і найважливішим рішенням є усвідомлення, яку проблему ми досліджуємо. Проблема – це не просто тема чи цікава ідея, а суперечність між наявним знанням і реальністю, яка потребує пояснення. Формулювання проблеми потребує критичного огляду літератури, аналізу теоретичних підходів і виявлення того, що ще залишається невідомим. У психології, наприклад, проблема може виникнути зі спостереження за поведінкою, яка не вписується в існуючі моделі. Від того, наскільки точно сформульована проблема, залежить успіх усього подальшого дослідження.

Справжня наукова проблема завжди конкретна. Вона не зводиться до загальних формулювань на кшталт «вивчити вплив середовища на

особистість» чи «дослідити особливості мислення студентів». Такі твердження надто широкі, щоб бути операційними. Конкретизація починається з аналізу літератури: дослідник має зрозуміти, що вже зроблено в цій галузі, які питання залишаються відкритими, у чому полягають суперечності між існуючими підходами. Саме ці прогалини формують те, що називають дослідницьким полем. Планування полягає у точному розміщенні власного проєкту в цьому полі – не як дублювання, а як внесок у розвиток науки.

Після усвідомлення проблеми постає питання мети дослідження. Мета – це не просто бажаний результат, а узагальнений образ того, чого прагне досягти дослідник. Вона має бути сформульована чітко, в конкретних термінах, що відображають і зміст, і межі дослідження. Занадто амбітна або розмита мета призводить до методологічної плутанини. Наприклад, мета «виявити психологічні чинники агресивної поведінки» потребує уточнення: у кого саме (підлітків, дорослих, військових), у яких умовах (школа, вулиця, бойова зона), у якому вимірі (мотиваційному, когнітивному, емоційному). Тільки так мета стає операційною і перетворюється на наукову задачу, а не загальну декларацію.

Від мети безпосередньо залежать завдання дослідження. Вони розгортають логіку дослідницького процесу у вигляді конкретних кроків. Зазвичай завдання мають ієрархічну структуру: від загальнотеоретичних – описати явище, уточнити понятійний апарат, розробити концептуальну модель – до емпіричних, які передбачають безпосереднє вимірювання, порівняння, перевірку зв'язків. Грамотно побудований план завжди має внутрішню симетрію: кожне завдання має свою гіпотезу, метод і спосіб перевірки. Якщо завдання сформульовані нечітко або не пов'язані з метою, дослідження розпадається на несумісні частини, і отримані результати не можна узагальнити.

Важливим логічним елементом планування є гіпотеза – наукове припущення, яке пояснює очікувані зв'язки між явищами. Гіпотеза – це серце наукового мислення: без неї дослідження втрачає спрямованість. У природничих науках гіпотеза часто формулюється у вигляді рівняння або моделі; у гуманітарних і психологічних – як вербальне твердження про закономірність чи механізм. Вона повинна бути водночас теоретично значущою і емпірично перевірюваною. Невипадково Карл Поппер називав перевірюваність критерієм науковості: будь-яке твердження, яке не можна спростувати дослідженням, не є науковим.

Планування дослідження передбачає формулювання гіпотези не інтуїтивно, а через аналітичну логіку: від опису фактів до пояснення, від пояснення до передбачення. Гіпотеза виконує одразу кілька функцій: структурну (організує план), прогностичну (передбачає можливий результат) і перевірку (дає критерії істинності).

Жодна гіпотеза не існує поза контекстом об'єкта та предмета дослідження. Об'єкт – це широка область реальності, де проявляється явище,

що нас цікавить. Предмет – конкретна властивість або зв'язок, який дослідник вивчає в межах цього об'єкта. Якщо об'єктом є «психологічна адаптація особистості», то предметом можуть бути «когнітивні механізми адаптації у студентів-першокурсників». Таке розмежування не є формальністю – воно визначає, які змінні ми вимірюватимемо і які методи будемо застосовувати. Без цього дослідження ризикує втратити фокус і стати емпірично розмитим.

Наступним етапом планування є вибір методів і процедур. Кожен метод – це інструмент перевірки гіпотези, а тому його вибір має бути обґрунтований не зручністю, а науковою доцільністю. Загалом психологічна методологія пропонує три основні групи методів: спостереження, експеримент і тестування. Спостереження дозволяє виявити природні прояви поведінки без втручання, експеримент – перевірити причинно-наслідкові зв'язки, а тестування – виміряти психічні властивості кількісно. Вибір залежить від рівня дослідження: на дескриптивному (описовому) етапі переважають спостереження та інтерв'ю; на пояснювальному – експеримент і моделювання; на прогностичному – тести, опитувальники, шкали.

Планування методів включає також визначення дизайну дослідження – загальної логіки побудови дослідження. В загальному вигляді снують експериментальні, квазіекспериментальні, лонгітюдні, поперечні, кореляційні та комбіновані дизайни. Кожен має свої переваги і обмеження. Наприклад, експеримент дозволяє встановити причинно-наслідкові зв'язки, але вимагає контролю змінних і часто має обмежену екологічну валідність. Кореляційне дослідження показує зв'язки, але не дозволяє робити висновки про причинність. На етапі планування необхідно свідомо обрати той дизайн, який найкраще відповідає меті.

Дизайн дослідження – це архітектура наукового експерименту. Його завдання – забезпечити внутрішню логічну зв'язність і контроль зовнішніх змінних. У психології дизайн охоплює всі можливі стратегії емпіричного вивчення поведінки.

Окрему увагу слід приділити вибірці дослідження. Вона є посередником між теорією і реальністю. Саме через вибірку теоретичні припущення перевіряються на емпіричному матеріалі. Основна вимога – репрезентативність, тобто здатність вибірки відображати властивості генеральної сукупності. Це досягається за допомогою випадкового (random) добору або контрольованих квот (стратифікованого добору). План повинен містити детальний опис кількості учасників, вікових, статевих, соціальних характеристик, а також чіткі критерії включення і виключення. Нерідко саме помилки на цьому етапі призводять до спотворення результатів і хибних узагальнень.

Ключовим компонентом плану є опис процедури збору даних. Дослідник має передбачити умови проведення (індивідуальні чи групові), інструкції, тривалість, послідовність завдань. Навіть дрібні деталі – час доби, рівень шуму, інтонація експериментатора – можуть вплинути на результат. Тому всі умови мають бути максимально стандартизовані. У сучасній практиці

часто застосовують протоколи реєстрації (research protocol), які фіксують кожен крок експерименту і дозволяють іншим дослідникам відтворити процедуру.

Планування неможливе без заздалегідь продуманого плану аналізу даних. Багато дослідників припускаються помилки, відкладаючи статистичний аналіз «на потім». Це призводить до ситуації, коли зібрані дані не відповідають вимогам обраних методів. Правильний підхід передбачає вибір статистичних процедур ще на етапі проєктування: чи буде це описова статистика, кореляційний, дисперсійний або регресійний аналіз, факторний або кластерний підхід тощо. Визначення цього заздалегідь дозволяє зрозуміти, які типи шкал потрібні (номінальні, порядкові, інтервальні), який обсяг вибірки забезпечить достатню статистичну потужність і які програмні засоби будуть використані (SPSS, R, Jamovi, Python).

Наукове планування включає також етичний компонент. Будь-яке дослідження, пов'язане з людьми, потребує дотримання принципів добровільності, конфіденційності та недопущення шкоди. Етична експертиза (Institutional Review Board) перевіряє, чи відповідає проєкт цим вимогам. У психології це особливо важливо, адже дослідження може торкатися чутливих тем – травми, насильства, психічного здоров'я. План повинен передбачати механізм отримання інформованої згоди, можливість відмови від участі, захист персональних даних і протокол дій у випадку ризику для респондента.

Планування завершується побудовою календарного графіка та ресурсної структури. Кожен етап має свої часові межі, виконавців, матеріальні й технічні потреби. Грамотно складений графік – це запобіжник від хаотичності. Наукова діяльність вимагає дисципліни не менше, ніж творчості. План – це спосіб упорядкувати інтелектуальний процес, щоб жодна деталь не була випадковою.

Структурна логіка плану наукового дослідження полягає у взаємопов'язаності всіх його компонентів – від вибору проблеми до формулювання висновків. Наукове дослідження можна уявити як систему координат, у якій кожен елемент має своє місце та функцію. Якщо хоча б один компонент буде пропущений або неправильно вибудований, уся система втратить рівновагу. Саме тому досвідчені дослідники говорять: якість плану визначає якість результатів.

У класичному розумінні план має три рівні: методологічний, методичний і процедурний. На методологічному рівні дослідник визначає парадигму – позитивістську, феноменологічну, культурно-історичну, інтегративну тощо. Від цього залежить розуміння того, що є «фактом», які методи допустимі, як інтерпретуються результати. На методичному рівні описується система методів, інструментів і показників, які забезпечують емпіричне втілення теоретичної моделі. На процедурному рівні план деталізується у вигляді послідовності конкретних дій, часу проведення, способів контролю змінних. Таке трирівневе структурування дозволяє уникнути логічної плутанини, коли дослідник плутає «що досліджується» з

«як досліджується». У цьому полягає відмінність між простою програмою дослідження і повноцінним науковим планом.

Після ухвалення цих ключових рішень починається етап практичного втілення – структурування дослідження на послідовні етапи. Хоча різні автори подають різні схеми, більшість погоджується з такою логікою: теоретичний аналіз, формулювання гіпотези, вибір методів, збір даних, аналіз результатів, оцінка й публікація. Проте у реальному житті цей процес не є лінійним – дослідник може повертатися до попередніх етапів, уточнювати гіпотези або модифікувати методiku, якщо виявлено методичні помилки чи нові змінні. Саме така гнучкість відрізняє живу науку від бюрократичної.

Одним із найвідповідальніших і найінтелектуальніших етапів планування є теоретичний аналіз проблеми, який реалізується через огляд наукової літератури. Саме він формує підґрунтя для подальшого розгортання емпіричної частини, дозволяючи зрозуміти, що вже відомо про предмет дослідження, які теорії домінують, які прогалини існують у знаннях і яким чином нова робота може заповнити ці прогалини. Огляд літератури – це не формальний перелік авторів і назв праць, а аналітична, системна, критична діяльність, спрямована на структурування наукового поля.

Етап збору даних вимагає максимальної точності. Будь-яке порушення процедури (зміна інструкції, інтонації, часу проведення тощо) може вплинути на результат. Саме тому план має містити чіткий протокол дій. Контроль умов забезпечує надійність результатів, а повторюваність – їх наукову достовірність.

Наступний крок – аналіз і інтерпретація даних. Це не лише математичні розрахунки, а процес осмислення отриманої інформації. Аналіз дозволяє виявити закономірності, взаємозв'язки, перевірити гіпотези, а інтерпретація – інтегрувати результати в ширший науковий контекст. Іноді аналіз відкриває несподівані результати, що суперечать початковим очікуванням, але саме вони часто ведуть до нових відкриттів і корекції теорії. У цьому полягає цінність емпіричної науки – вона не підтверджує упередження, а випробовує їх.

Завершальним етапом є оцінка результатів і публікація. Публікація – це не просто формальний звіт, а акт комунікації з науковою спільнотою. Вона забезпечує кумулятивний характер знання, тобто можливість перевірки, повторення і подальшого розвитку ідей.

Наукова публікація має бути чесною, прозорою і відкритою для критики, адже лише так наука розвивається. Публікація завершує цикл наукового дослідження, але водночас є початком його життя у науковому просторі. Саме на етапі публікації відбувається перевірка плану на практиці: чи був він достатньо логічним, чи забезпечив відтворюваність, чи дозволив отримати інтерпретовані результати. У якісному плануванні передбачаються навіть формальні аспекти майбутньої публікації: структура статті, спосіб представлення таблиць, тип статистики, формат посилань. Це не дрібниці, а елемент наукової культури. Стандарти APA, Chicago, ДСТУ чи Vancouver

визначають не лише оформлення, а й логіку побудови аргументів – від гіпотези до висновку.

Оформлення результатів – це не риторика, а продовження методології. Як зазначав Л. Виготський: «Наукове мислення завершується у слові». Те, як дослідник структурує текст, відображає його розуміння процесу пізнання. Тому план має включати попередню модель майбутньої публікації – це допомагає мислити системно й не втрачати загальної логіки.

Крім того, планування етапу публікації пов'язане з питанням відкритої науки. У сучасних вимогах усе частіше вимагається публікувати не лише результати, а й самі дані, щоб інші могли їх перевірити. Це сприяє прозорості й підвищує довіру до дослідника. Отже, завершальна фаза планування – це продумана стратегія наукової комунікації, яка забезпечує тривалість і вплив отриманого знання.

Перераховані етапи дослідження не завжди є строго лінійними, часто вони можуть переплітатися, вимагати повернення до попередніх кроків або ітераційного доопрацювання. Однак їх послідовне виконання є запорукою успішного та методологічно обґрунтованого наукового дослідження.

3.2. Роль і значення огляду літератури в сучасній психологічній науці

У сучасній психології якість огляду літератури розглядається як показник наукової зрілості автора. Дослідник, який здатен побачити не лише поверхневі подібності між працями, а й їхню концептуальну глибину, продемонструвати власну позицію у контексті вже існуючих парадигм, фактично створює теоретичний простір для свого дослідження. Теорія в психології має багаторівневу структуру, що включає емпіричний, описово-модельний і метатеоретичний рівні. На емпіричному рівні збираються факти, на описовому – створюються поняття й моделі, а на метатеоретичному рівні формується осмислення самої логіки теоретизування, його методологічних і філософських засад. Відповідно, якісний огляд літератури має охоплювати всі три рівні, адже лише тоді дослідження набуває системності та концептуальної глибини.

Огляд літератури виконує декілька функцій. По-перше, він забезпечує когнітивну орієнтацію дослідника в науковому просторі: дозволяє визначити основні концепції, школи, напрямки, термінологічні відмінності. По-друге, він виконує евристичну функцію, адже саме через аналіз існуючих теорій дослідник виявляє суперечності, недосліджені аспекти, що згодом перетворюються на предмет власного вивчення. По-третє, огляд має методологічну функцію – він показує, які підходи, методи й процедури вже використовувалися, наскільки вони були ефективними, які з них можна адаптувати або вдосконалити.

Під час планування важливо розуміти, що огляд літератури – це не початкова, а постійна складова дослідження. Він починається з моменту вибору теми, триває на етапі формулювання гіпотези, продовжується під час

аналізу даних і навіть на етапі написання висновків. Кожна нова знахідка чи уточнення у процесі дослідження змінює контекст огляду, змушуючи дослідника переглянути певні позиції. Тому огляд – це жива, динамічна частина наукової роботи, яка еволюціонує разом із самим дослідженням.

Однією з головних вимог до якісного огляду є системність. Це означає, що джерела повинні бути відібрані не випадково, а за певним принципом – хронологічним, тематичним або концептуальним. При цьому важливо зберігати баланс між класичними працями, що визначили основи теорії, і сучасними дослідженнями, які відображають актуальний стан науки. Наприклад, досліджуючи стресостійкість, не можна обмежитись лише працями Г. Сельє чи Р. Лазаруса, необхідно також проаналізувати сучасні моделі. Лише поєднання історичної глибини й сучасної емпіричної бази створює цілісну картину.

Другою важливою вимогою є критичність. Дослідник повинен не лише відтворювати позиції авторів, а й оцінювати їхню аргументованість, внутрішню логіку, методологічну послідовність. Критичний аналіз не означає заперечення – він полягає у вмінні бачити межі застосування певної теорії, порівнювати різні підходи, виявляти точки перетину й суперечності. Саме критичність перетворює огляд на аналітичний інструмент, а не на бібліографічний перелік. Наука розвивається не через накопичення фактів, а через переосмислення парадигм. Тому здатність дослідника виявити, де саме в науковій літературі відбувається зміна парадигм, є ознакою його наукової компетентності.

Не менш значущою є методологічна точність під час відбору джерел. У добу цифрових технологій наукова інформація зростає експоненційно, але не кожне джерело має однакову вагу. Плануючи огляд, потрібно керуватися критеріями наукової надійності: рецензовані журнали, індексація у міжнародних базах (Scopus, Web of Science), наявність емпіричних підтверджень, репутація авторів. Використання неперевіраних інтернет-джерел або популярних публікацій без наукового рецензування знижує достовірність роботи. У сучасній академічній культурі вважається, що джерело повинно бути не старше 5–7 років, якщо не йдеться про класичні праці, що мають фундаментальне значення.

Структурно огляд літератури має логічно підводити читача до формулювання гіпотези. Він починається з опису загального контексту проблеми, потім поступово звужується до конкретних аспектів, які дослідник збирається перевірити. Важливо, щоб перехід від загального до приватного був обґрунтований, а не механічний. Наприклад, аналізуючи феномен стресу, спочатку розглядають загальні концепції фізіологічного реагування, далі – когнітивні й емоційні моделі, і лише потім переходять до специфіки професійного стресу в рятувальників або військових. Така структура забезпечує логічність і послідовність викладу.

Ще одним аспектом є інтегративність огляду, тобто вміння поєднати різні підходи. Сучасна психологія характеризується методологічним

плюралізмом: поряд існують когнітивна, гуманістична, біхевіористична, культурно-історична та інші парадигми. Дослідник, який здатен побачити, як вони доповнюють одна одну, а не протиставляються, формує більш глибоке розуміння предмета. Наприклад, вивчення мотивації може одночасно спиратися на біологічні основи (нейрохімічна регуляція), когнітивні механізми (очікування – цінність), і соціальні фактори (культуральні норми успіху). Такий міждисциплінарний підхід особливо важливий для сучасної експериментальної психології.

Важливо також дотримуватись етичних норм цитування під час огляду літератури. Кожна запозичена ідея, навіть якщо вона перефразована, повинна мати посилання на джерело. Це не лише вияв поваги до авторства, а й спосіб продемонструвати глибину власної обізнаності. Коректне цитування допомагає відтворити ланцюг наукових ідей і полегшує перевірку достовірності тверджень. Огляд без джерел або з надмірним використанням невідомих публікацій не може вважатися науковим.

У процесі планування важливо передбачити й технічні аспекти організації огляду. Корисно створювати аналітичні таблиці, де фіксуються автори, рік, методологія, вибірка, результати та обмеження дослідження. Такий підхід дозволяє систематизувати великий масив інформації й швидко орієнтуватися в ньому. Сьогодні для цього активно використовуються цифрові інструменти – Zotero, Mendeley, EndNote, які автоматизують процес цитування та дозволяють створювати бази даних джерел.

Теоретичний етап полягає в глибокому аналізі попередніх досліджень і теорій. Науковець систематизує відомі дані, виявляє суперечності й формує теоретичну рамку, в якій надалі інтерпретуватимуться результати. Це не просто «огляд літератури», а інтелектуальне осмислення стану проблеми, що дозволяє перейти від поверхового спостереження до гіпотетичного пояснення. На цьому етапі створюється понятійний апарат, визначаються змінні, які можна виміряти емпірично.

Глибина теоретичного етапу визначається не лише обсягом опрацьованої літератури, а й умінням дослідника побудувати на її основі цілісну концептуальну модель. Після систематизації попередніх підходів необхідно чітко сформулювати, які саме прогалини залишаються у сучасному науковому знанні, які питання є суперечливими або недостатньо дослідженими. Саме ці прогалини і стають основою для формулювання проблеми та гіпотези дослідження. Важливо, щоб гіпотеза не була довільним припущенням: вона повинна логічно випливати з аналізу літератури, бути теоретично обґрунтованою й водночас емпірично перевірюваною.

У сучасній психології огляд літератури виступає також як інструмент інтеграції різнорідних даних. Завдяки еволюції дисципліни психологічні знання стали багаторівневими: поряд із класичними експериментальними даними активно використовуються якісні дослідження, нейропсихологічні підходи, моделювання когнітивних процесів, великі статистичні вибірки та навіть дані з поведінкової економіки. Здатність науковця об'єднати ці різні

джерела у цілісний огляд демонструє його компетентність і дозволяє створити більш глибоку й багатовимірну картину досліджуваного явища. Такий огляд виходить за рамки традиційного аналізу статей і набуває рівня міждисциплінарного синтезу.

Нарешті, огляд літератури є важливим елементом наукової комунікації. Він дозволяє читачеві зрозуміти логіку дослідника, оцінити рівень його професійної підготовки, визначити, наскільки він занурений у предметну область. У цьому сенсі огляд виконує репутаційну функцію: він відображає інтелектуальну культуру автора, його методологічну грамотність та здатність працювати з великими обсягами інформації. Якісний огляд – це не просто аналіз джерел, а прояв наукової майстерності.

Таким чином, огляд літератури у сучасній психології є комплексним, динамічним і багатофункціональним елементом наукового дослідження. Він забезпечує теоретичну основу, методологічне обґрунтування, аналітичну глибину й інтерпретаційну чіткість. Його якість визначає не лише рівень окремої роботи, а й ступінь наукової зрілості автора, адже саме через огляд літератури вчені долучаються до ширшого наукового дискурсу, формуючи нові підходи, уточнюючи старі теорії та сприяючи розвитку психологічної науки як системи знання.

3.3. Етика наукового дослідження

Проблема етики у психологічних дослідженнях є не менш важливою, ніж методологічна коректність. Наука, яка має справу з людською психікою, не може залишатися морально нейтральною: будь-яке втручання у внутрішній світ людини потенційно впливає на її благополуччя, почуття безпеки, гідність і права. Саме тому етичне планування дослідження є невіддільною складовою його наукової цінності. У сучасному розумінні наукове дослідження вважається повноцінним лише тоді, коли воно не лише методологічно обґрунтоване, а й проведене відповідно до етичних стандартів.

Психологічна етика досліджень бере свій початок у гуманістичній традиції, яка вважає людину не засобом, а метою будь-якої діяльності. Відповідно, ще на етапі планування дослідник зобов'язаний передбачити всі можливі ризики для учасників, як фізичні, так і психічні. Етичність починається не з підпису під інформованою згодою, а з того моменту, коли дослідник лише формує тему і визначає мету свого дослідження. Якщо дослідницьке питання потенційно може поставити учасників у стресову або принизливу ситуацію, воно потребує переосмислення або відмови.

Одним із фундаментальних етичних принципів, що визначає зміст планування, є принцип інформованої згоди. Він закріплений у Кодексі етики Американської психологічної асоціації, а також у Гельсінській декларації Всесвітньої медичної асоціації. Інформована згода означає, що кожен учасник дослідження має отримати зрозуміле, повне й правдиве пояснення суті експерименту, його мети, тривалості, можливих ризиків і переваг, а також має

право відмовитися від участі без будь-яких наслідків. Згода повинна бути добровільною, а не формальною. Це не просто документ, а прояв поваги до автономії людини. Особливої уваги потребують дослідження, де учасниками є діти, люди похилого віку, особи у стресових станах – у цих випадках необхідно враховувати обмежену здатність суб'єкта повною мірою усвідомити умови участі.

Другим ключовим принципом є принцип «не нашкодь». Він передбачає мінімізацію будь-яких негативних наслідків участі у дослідженні. У психологічній практиці відомі численні приклади, коли порушення цього принципу призводило до травматичних ефектів. Класичним антиприкладом є сумнозвісний експеримент Стенлі Мілгрема, де учасники піддавались моральній травматизації, вважаючи, що завдають іншим фізичного болю. Хоча цей експеримент суттєво розширив розуміння феномена підпорядкування авторитету, він поставив під сумнів допустимі межі психологічного експерименту. Відтоді у світовій науці сформувався чіткий консенсус: наукова цінність не може виправдовувати шкоду людині.

Третій важливий елемент етичного планування – принцип конфіденційності та анонімності. Психологічні дані, на відміну від більшості інших типів інформації, часто містять глибоко особистісні відомості: переживання, оцінки, ставлення, біографічні деталі. Порушення конфіденційності може не лише дискредитувати дослідження, а й завдати реальної психологічної шкоди. Тому під час планування необхідно продумати способи зберігання, кодування та знеособлення інформації. Наприклад, замість імен учасників використовуються коди, анкети зберігаються у зашифрованому форматі, а доступ до них мають лише члени дослідницької групи. Будь-яка публікація результатів має бути підготовлена таким чином, щоб жоден учасник не міг бути ідентифікований.

Особливу увагу варто приділяти етичним аспектам маніпуляцій і обману у дослідженнях. Психологічна методологія іноді допускає застосування прихованої мети або «дефінітивного обману», коли учасникам не повідомляють істинну суть експерименту, щоб не спотворити їхню поведінку. Однак такий підхід допустимий лише тоді, коли неможливо досягти наукової мети без нього і коли потенційна користь переважає ризик шкоди. Згідно з етичним кодексом АРА, після завершення експерименту дослідник зобов'язаний провести процедуру дебрифінгу – пояснити справжню мету дослідження, відповісти на всі питання, переконатися, що учасники не зазнали негативного впливу. У разі найменших ознак дистресу дослідник має надати психологічну підтримку або направити до фахівця.

Ще один аспект етичного планування – повага до культурного та соціального контексту. Психологічні дослідження не відбуваються у вакуумі: норми, цінності та звички учасників суттєво впливають на їхню реакцію. Те, що в одній культурі сприймається нейтрально, в іншій може бути травматичним або принизливим. Тому плануючи дослідження, особливо міжкультурне, необхідно враховувати локальні норми, мову, символіку,

особливості комунікації. Дослідник повинен уникати нав'язування ціннісних суджень, стереотипів чи культурної упередженості. Етична чутливість у цьому контексті є невід'ємною складовою наукової доброчесності.

Не менш важливою складовою етичного планування є уникнення конфлікту інтересів. Конфлікт виникає тоді, коли особисті, фінансові чи професійні інтереси дослідника можуть вплинути на об'єктивність інтерпретації результатів. Це може стосуватись фінансування дослідження, співпраці з комерційними структурами, тиску адміністрації або наукових керівників. Щоб уникнути цього, дослідник має відкрито декларувати всі можливі конфлікти й забезпечити прозорість фінансових або організаційних аспектів проекту.

Етичне планування передбачає також дотримання академічної доброчесності. Це поняття охоплює не лише уникнення плагіату, але й чесність у представленні результатів, відсутність фальсифікацій і маніпуляцій даними. Фабрикація результатів (створення неіснуючих даних) або вибіркове подання лише «зручних» фактів є грубим порушенням наукової етики. Плануючи дослідження, дослідник має передбачити процедури контролю за якістю даних: збереження первинних протоколів, ведення електронних журналів спостережень, можливість перевірки з боку рецензентів. У багатьох університетах світу створені етичні комісії, які розглядають плани досліджень і дають дозвіл лише після переконання, що вони відповідають етичним вимогам.

Важливо розуміти, що етика – це не сукупність формальних правил, а внутрішня культура науковця. Справжній дослідник не потребує зовнішнього контролю, щоб дотримуватись чесності – вона є частиною його професійної ідентичності. Етичне мислення в науці передбачає усвідомлення не лише відповідальності перед учасниками, а й перед суспільством, адже результати психологічних досліджень можуть впливати на освітню політику, систему безпеки, медичну практику. Кожне рішення дослідника, починаючи з вибору теми, має потенційні наслідки – і тому планування стає актом морального вибору.

У цьому сенсі етика і методологія є взаємозалежними: добре сплановане дослідження зазвичай є етичним, тому що передбачає чіткість, прозорість, мінімізацію ризиків і повагу до прав людини. Етичні порушення найчастіше виникають там, де планування поверхневе – коли дослідник не передбачає наслідків своїх дій, не аналізує можливі ризики або не узгоджує методiku з етичними нормами. Тому формування етичної компетентності дослідника є обов'язковою складовою його професійної підготовки.

3.4. Академічна доброчесність у плануванні наукового дослідження

Поняття академічної доброчесності є наріжним каменем сучасної науки. У найширшому розумінні воно охоплює сукупність моральних і правових норм, які регулюють чесну, відповідальну та прозору діяльність науковців,

викладачів і здобувачів вищої освіти. Академічна доброчесність виступає не лише етичним, а й методологічним принципом: без неї неможливо забезпечити достовірність знань, репутацію наукової спільноти й ефективність освіти. Вона включає чесність у збиранні та представленні даних, коректність цитування, дотримання авторських прав, об'єктивність у рецензуванні, відмову від фальсифікацій, маніпуляцій та будь-яких форм обману.

У межах планування наукового дослідження академічна доброчесність проявляється вже на початкових етапах: у формулюванні теми, постановці мети, розробці гіпотези, виборі методів. Недоброочесність може проявитись навіть у самому задумі, коли дослідник свідомо обирає маніпулятивний підхід, спрямований на підтвердження бажаного результату, а не на пошук істини. Саме тому у провідних університетах світу існує практика подання на розгляд етичних комісій детального плану дослідження – це запобігає можливим порушенням ще до початку емпіричної роботи.

Одним із найпоширеніших проявів недоброочесності є плагіат – привласнення чужих ідей, текстів, результатів або методик без належного посилання на джерело. Плагіат підриває довіру до науки, дискредитує автора й університет, спотворює реальний стан наукового поступу. У психології це особливо небезпечно, адже повторне використання чужих даних може призвести до хибних висновків і вплинути на практичну роботу психологів.

Планування дослідження передбачає системні дії для запобігання плагіату. По-перше, дослідник повинен формувати власну культуру наукового письма, що передбачає свідоме ставлення до цитування. Усі використані джерела мають бути належно оформлені відповідно до обраного стандарту – APA, MLA, ДСТУ тощо. Це стосується не лише прямих цитат, але й ідей, які були перефразовані. Часто помилкою молодих науковців є уявлення, що зміна кількох слів у реченні звільняє від необхідності посилання. Насправді будь-яка запозичена думка незалежно від способу викладу потребує коректного посилання.

По-друге, під час планування необхідно передбачити інструменти самоперевірки. У більшості сучасних університетів для цього використовують програми Unicheck, StrikePlagiarism, Plag, Turnitin, які дозволяють виявити збіги з опублікованими текстами. Проте автоматична перевірка – лише технічна допомога; головне – сформоване розуміння меж власного та чужого інтелектуального внеску. Академічна доброчесність – це не страх перед санкціями, а свідоме переконання у цінності власної роботи та повазі до чужої.

Третім аспектом доброчесності є уникнення фальсифікацій та фабрикацій даних. Фальсифікація – це навмисне спотворення отриманих результатів (наприклад, відбір лише «зручних» даних), фабрикація – вигадкування даних, яких ніколи не існувало. Обидва ці порушення є грубими злочинами проти науки. Відомі скандали, коли науковці втрачали репутацію через виявлення підроблених експериментів – від психологічних досліджень з ефектом соціального конформізму до медичних випробувань. Щоб запобігти цьому, планування має включати чіткий протокол дослідження, у якому

зафіксовано всі процедури збору даних, а також систему контролю достовірності – наприклад, подвійний запис результатів або незалежну верифікацію статистичних обчислень.

Четвертим важливим принципом є заборона самоплагіату. На перший погляд здається, що автор має право використовувати власні попередні тексти, адже це його інтелектуальна власність. Однак у науковому контексті повторне оприлюднення тих самих результатів без вказівки на джерело створює ілюзію новизни й порушує принцип прозорості. Тому, якщо в новій роботі використано частину попередніх публікацій, необхідно чітко зазначити, що ці дані вже були представлені раніше. Це правило забезпечує чесність наукового обігу та захищає репутацію автора.

Академічна доброчесність також передбачає об'єктивність і неупередженість у процесі аналізу та рецензування. Науковець зобов'язаний представляти результати такими, якими вони є, навіть якщо вони не підтверджують його гіпотези. У цьому полягає справжня наукова чесність. Як наголошував Карл Поппер, жодна гіпотеза не є остаточною – кожна повинна витримувати можливість спростування. Замовчування «незручних» даних або їх виключення з аналізу – це форма маніпуляції, що суперечить суті наукового пізнання.

Суть перевірки гіпотези полягає у демонстрації не того, що вона «правильна», а того, що вона витримує випробування фактами. Це ключова відмінність наукового мислення від буденного: дослідник шукає не підтвердження, а потенційні контраргументи. Лише гіпотеза, що витримала спробу спростування, має наукову цінність.

У психології особливо важливо враховувати ризик ефекту підтвердження – когнітивного викривлення, коли дослідник несвідомо відбирає дані, які підтримують його очікування. Тому план дослідження має включати механізми контролю цього ефекту: подвійне сліпе дослідження, незалежну перевірку даних, повторні вимірювання.

Перевірка гіпотези також передбачає використання надійних критеріїв статистичної значущості. Однак значущість сама по собі не доводить істинності явища – вона лише свідчить про малу ймовірність випадковості. Тому на сучасному етапі дедалі більшої ваги набувають показники розміру ефекту та довірчих інтервалів, які відображають реальну силу зв'язку між змінними.

У контексті планування доброчесність вимагає передбачити процедури прозорості дослідницького процесу. Це означає ведення детального журналу дослідження, де фіксуються всі кроки – від добору учасників до способів обробки даних. Такі записи не лише підвищують якість роботи, а й дають можливість іншим науковцям відтворити дослідження, що є критерієм наукової надійності.

Ще одним виявом академічної чесності є правильне визначення авторства. У колективних проектах часто виникають конфлікти щодо того, хто є автором, а хто лише сприяв роботі. Міжнародна практика передбачає, що

співавтором може бути лише той, хто зробив суттєвий інтелектуальний внесок у дизайн, збір або інтерпретацію даних. Усі інші форми допомоги (надання ресурсів, консультації, редагування) мають зазначатись у подяках, але не в авторстві. Плануючи дослідження, важливо заздалегідь узгодити ролі членів команди, щоб уникнути непорозумінь і порушень етичних норм.

Доброчесність охоплює також питання використання штучного інтелекту та цифрових технологій у науковій роботі. Сьогодні, коли значна частина текстів створюється або редагується за допомогою алгоритмів, дослідники повинні чітко зазначати межі використання таких інструментів. Використання автоматизованих перекладачів, генераторів текстів або систем обробки даних допустиме лише як допоміжний засіб, але не як джерело нових знань. При цьому дослідник несе повну відповідальність за точність і достовірність результатів, навіть якщо частину тексту підготовлено з допомогою технологічних систем.

Отже, академічна доброчесність у плануванні дослідження – це не просто етична вимога, а структуроутворювальний принцип, який визначає якість і наукову вартість роботи. Без неї будь-яке дослідження втрачає сенс, адже наука базується на довірі. Планування, що включає механізми забезпечення чесності, не лише гарантує чистоту результатів, але й формує професійну культуру дослідника, його репутаційну надійність і здатність брати відповідальність за власний інтелектуальний продукт.

Завдання на самостійну підготовку

1. У чому полягає специфіка планування наукового дослідження в психології порівняно з природничими науками?
2. Чому планування вважається не технічним етапом, а центральним моментом наукового мислення?
3. Яке значення має відповідність між теоретичним і емпіричним рівнями дослідження під час планування?
4. У чому полягає роль операціоналізації понять для формулювання гіпотези?
5. Чому мультиметодовий підхід підвищує конструктивну валідність психологічного дослідження?
6. Яким чином конкретизація наукової проблеми визначає якість подальших етапів дослідження?
7. Чим відрізняються об'єкт і предмет дослідження та як це розмежування впливає на вибір методів?
8. Які вимоги висуваються до формулювання мети й завдань дослідження, щоб уникнути методологічної нечіткості?
9. Які основні функції виконує гіпотеза в логіці наукового планування?
10. Які критерії визначають вибір дизайну дослідження під час планування?
11. Чому репрезентативність вибірки є критичною умовою наукової достовірності?

12. Які ключові елементи має включати детальний протокол збору даних?
13. Чому планування статистичного аналізу на початковому етапі є необхідною умовою валідності дослідження?
14. Як у плануванні враховуються етичні вимоги – добровільність, конфіденційність і захист учасників?
15. Які основні функції виконує огляд літератури на різних етапах наукового дослідження?
16. Чому огляд літератури є динамічним процесом, що супроводжує дослідження від вибору теми до публікації?
17. У чому полягають принципи системності та критичності під час формування огляду літератури?
18. Чому принцип інформованої згоди є фундаментальним у плануванні психологічного дослідження?
19. Які етичні вимоги визначають правила зберігання, кодування та публікації персональних психологічних даних?
20. Які форми недоброчесності можуть виникати ще на етапі планування дослідження та якими механізмами їх можна попередити?

РОЗДІЛ 4 ОСОБЛИВОСТІ ВИМІРЮВАННЯ В ПСИХОЛОГІЇ

4.1. Проблема вимірювання у психології

Проблема вимірювання індивідуально-психологічних особливостей є однією з ключових у сучасній психології, оскільки саме від неї залежить можливість перетворення описів психічних явищ на наукове знання. У своїй роботі психолог постійно стикається з необхідністю кількісно оцінювати риси особистості, когнітивні процеси, емоційні стани, мотиваційні тенденції. Для цього в психодіагностиці розробляються спеціальні вимірні процедури, сенс яких полягає в тому, що у процесі дослідження психологічні параметри набувають кількісного вираження. Кількісні дані, отримані внаслідок ретельно спланованого дослідження за допомогою стандартизованих методик, згодом використовуються для математико-статистичної обробки, побудови моделей, перевірки гіпотез. При цьому вимірювання в психології стикається з унікальною складністю: психіка як предмет дослідження не має безпосередньо спостережуваних фізичних характеристик. Якщо у фізиці вимірюють довжину чи масу, у біології – концентрацію речовин чи частоту імпульсів, то в психології йдеться про індивідуальні відмінності у поведінці, мотивації, емоціях та пізнавальних процесах, що не існують у просторово-часовій формі й не піддаються прямому вимірюванню. Психологічні конструкти (тривожність, інтелект, емпатія, самооцінка тощо) не є об'єктами безпосереднього спостереження, а постають як теоретичні узагальнення, що стоять за певними проявами поведінки. Тому вимірювання у психології завжди є опосередкованим і здійснюється через зовнішні індикатори, що породжує принципову проблему валідності: наскільки результати тестів дійсно відображають досліджуване явище, а не лише артефакти методики чи контексту. Це питання має не лише методологічний, а й філософський вимір, адже воно стосується самої сутності психології як науки: якщо психічне можливо виміряти, тоді психологія постає емпіричною дисципліною, спорідненою з природничими науками; якщо ж ні – вона зберігає переважно гуманітарний характер, спираючись на тлумачення, феноменологічний опис і розуміння смислів. Історія психології значною мірою є історією напруги між цими двома підходами – психологією як природничою наукою та психологією як наукою про смисли.

Перші спроби кількісно описати психічні явища з'явилися у XIX столітті в межах психофізики. Її засновники Ернст Вебер і Густав Фехнер прагнули встановити математичну залежність між силою фізичного подразника та інтенсивністю суб'єктивного відчуття. Формула Г. Фехнера $S = k \log R$ стала одним із перших рівнянь, що кількісно описує психічне й демонструє, що суб'єктивний досвід не є хаотичним, а підкоряється певним закономірностям. Е. Вебер ввів поняття «порога розрізнення» як найменшої різниці між двома стимулами, яку може зафіксувати людина, тим самим започаткувавши

традицію, у якій психічні процеси описуються через поведінкові індикатори та реакції суб'єкта. Наприкінці XIX століття вимірювання стає центральною темою диференційної психології завдяки працям Френсіса Гальтона. Він вважав, що психологічні властивості (інтелект, здібності, швидкість реакції тощо) мають спадкову природу й піддаються кількісному порівнянню. Ф. Гальтон пробує за допомогою спеціальних пристроїв вимірювати різні психічні функції. У такий спосіб варіативність перестає вважатися «похибкою» і перетворюється на головний об'єкт психологічної науки. Ідеї Ф. Гальтона надихають Альфреда Біне і Теодора Сімона, які розробляють перший тест інтелекту, а згодом Льюїса Термена, що адаптує його й вводить статистичне нормування, надавши інтелекту числового виміру у вигляді коефіцієнта IQ. Цей крок став переломним моментом, оскільки вперше психічна властивість отримала формальний числовий статус, придатний до порівняння між індивідами.

На початку XX століття формується психометрія – наука про вимірювання психологічних характеристик. Класична психометрія, представлена роботами Ч. Спірмена, Л. Терстоуна, Л. Гуттмана та інших дослідників, виходить із припущення, що будь-який спостережуваний результат (X) є сумою істинного значення (T) і похибки вимірювання (E): $X = T + E$. Такий підхід дозволив не лише враховувати похибку, а й створити основу для стандартизації тестів, розробки норм, оцінки надійності й валідності. На основі ідей Спірмена розвивається факторна теорія інтелекту, у якій результати різних когнітивних завдань інтерпретуються як прояви латентних змінних, зокрема загального фактора розумової енергії g. Подальші дослідження Л. Терстоуна, Р. Кеттелла, Дж. Гілфорда показали, що інтелект є багатовимірною системою, що складається з низки часткових здібностей (просторових, вербальних, аналітичних тощо), а отже, вимірювання має враховувати профіль індивідуальних відмінностей, а не зводитися до єдиного індексу. Паралельно розробляються статистичні методи, зокрема дисперсійний аналіз, які дозволяють відокремлювати істинну варіацію від випадкової, перевіряти гіпотези й кількісно оцінювати вплив факторів. Саме завдяки цьому вимірювання поступово стає ядром емпіричної революції в гуманітарних науках.

У 1930–1950-х роках відбувається теоретичне осмислення вимірювання в психології. Представники операціоналізму наголошують, що психологічні поняття мають науковий сенс лише тоді, коли визначені через сукупність операцій спостереження та вимірювання. У цьому контексті «інтелект», «тривожність» чи «мотивація» розглядаються не як абстрактні сутності, а як результати виконання конкретних завдань, відповідей на пункти опитувальника, швидкості й точності дій. Операціоналізм тим самим перетворює вимірювання на систему процедур, що забезпечують відтворюваність результатів, але водночас звужує розуміння психіки до зовнішньо спостережуваних індикаторів, залишаючи поза увагою феноменологічний вимір досвіду. Важливим кроком стає репрезентативна

теорія вимірювання Стенлі Стівенса (1946), яка визначає вимірювання як присвоєння чисел об'єктам або подіям відповідно до певних правил. Число в такому разі розглядається не як «властивість» самого явища, а як умовний знак, що відображає відношення між явищами. Це призводить до розробки системи шкал (номінальних, порядкових, інтервальних, відношень) та формальних критеріїв для допустимих математичних операцій із кожним із їх типів. Пізніше з'являються гібридні моделі – теорія латентних ознак, моделі структурних рівнянь, які поєднують якісні конструкти з кількісними індикаторами, намагаючись зберегти баланс між філософською глибиною і статистичною строгістю.

На цьому тлі вимірювання дедалі більше перетворюється на центральну проблему психологічної науки. Розвивається тестування інтелекту (шкали Біне–Сімона, адаптації Л. Термена, шкали Д. Векслера), особистісні опитувальники (ММРІ, 16PF Р. Кеттелла, опитувальник Айзенка, шкали депресії А. Бека), інструменти для оцінки тривожності, мотивації, соціальних установок. Для всіх цих методик формуються спільні стандарти: чітко визначена теоретична модель, процедура стандартизації, емпірична перевірка надійності та валідності, побудова тестових норм. Так постає психометрична парадигма конструювання тестів – сукупність підходів, процедур і критеріїв, що визначають, які вимірювальні інструменти можна вважати науково обґрунтованими. Вона розглядає психічні явища як такі, що, попри свою складність і суб'єктивність, можуть бути виміряні за правилами, аналогічними до тих, що застосовуються у природничих науках. Психометрична революція перетворила психологію з дисципліни, що базується на вербальних описах, на науку, здатну оперувати стандартизованими, відтворюваними та кількісно порівнюваними даними. Водночас вона викликала і критику з боку представників гуманітарної, феноменологічної та персонологічної традицій (В. Дільтей, Е. Гусерль, Г. Олпорт та ін.), які застерігали від редукції багатовимірної сутності людини до набору чисел і статистичних моделей. Це протистояння стало підґрунтям розмежування на «кількісну» і «якісну» психологію, однак навіть її критики визнавали: без психометрії психологія не змогла б довести свою науковість і зайняти місце серед емпіричних наук.

Сучасний етап розвитку психометричної парадигми супроводжується не лише вдосконаленням моделей вимірювання, а й зростанням етичної відповідальності психолога. Важливу роль відіграє усвідомлення меж застосування тестових методик, ризику хибних інтерпретацій, можливих соціальних наслідків використання результатів тестування. Ідея вимірюваності психічного має не лише наукові, а й етичні наслідки: вона ставить питання про те, чи об'єктивно ми вимірюємо не фізичні об'єкти, а властивості свідомості, та які гарантії того, що ці вимірювання не будуть використані для дискримінаційних або маніпулятивних цілей. Саме тому в межах психометричної парадигми формуються принципи, які регламентують якість тесту (надійність, валідність, репрезентативність норм), сферу його застосування, вимоги до професійної компетентності користувача. Від

психолога вимагається не лише технічне володіння методиками, а й здатність критично мислити, оцінювати психометричні параметри тестів, розуміти обмеження кожного інструмента.

На цьому тлі особливо загрозливим явищем стає поширення в цифровому та науковому середовищі так званих плацебо-опитувальників – псевдотестів, які імітують справжні психодіагностичні методики, але не відповідають базовим критеріям наукової валідації. Такі інструменти не мають емпіричної підтримки, не проходили стандартизації, перевірки надійності, часто не мають чітко окресленого теоретичного підґрунтя. Їх застосування формує хибне уявлення у психологів-початківців, клієнтів і організацій, що користуються їх результатами, і врешті призводить до дискредитації наукової психології як галузі. Окрім того, масове поширення псевдотестів у медіапросторі сприяє зниженню психометричної культури та посиленню когнітивних спотворень. Користувачі схильні вірити результатам навіть невірогідних тестів через ефект Барнума/Форера, коли загальні, нечіткі описи здаються індивідуально точними. Це підсилює самоідентифікацію з псевдонауковими твердженнями, нормалізує недостовірні уявлення про власну особистість і психічний стан. У результаті підміняється сама ідея наукового вимірювання: замість інструменту об'єктивізації ми отримуємо засіб символічного самопідкріплення та розваги.

Виходом із цієї ситуації є системний розвиток психометричної грамотності у психологів і суміжних фахівців. Йдеться не лише про формальне знання термінів «надійність» і «валідність», а про глибоке розуміння того, як конструюються тести, які вимоги висуваються до вибірки, які статистичні методи використовуються для аналізу структури інструмента, як інтерпретуються шкальні бали. Лише науково обґрунтовані методики здатні забезпечити адекватну оцінку психічних властивостей особистості, підтримати практичні рішення у сфері освіти, клініки, організаційної психології, профвідбору. Поширення доказових підходів до психодіагностики, орієнтація на стандартизовані, валідизовані тести, критичний аналіз їх якості є ключовими чинниками збереження довіри до психологічної науки. Водночас це вимагає від дослідника високого рівня рефлексії щодо власної ролі у процесі вимірювання: будь-який психометричний результат є не «чистим відображенням» психіки, а підсумком складної взаємодії між теорією, методикою, контекстом дослідження та інтерпретацією фахівця. Саме тому вимірювання у психології перестало бути лише математичною процедурою й дедалі більше усвідомлюється як акт інтерпретації, у якому дослідник не лише «рахоує», а й «розуміє», беручи на себе як наукову, так і етичну відповідальність за результати.

4.2. Методи та види вимірювань у психології

У психометричній практиці зазвичай виділяють три основні типи методів вимірювання: суб'єктивні методи – це найбільш поширена група

методів, до якої належать опитувальники, шкали, інтерв'ю, де респондент прямо відповідає на запитання, що стосуються його думок, емоцій, поведінкових намірів або оціночних суджень; об'єктивні методи, що базуються на аналізі продуктів діяльності: результатів виконання стандартних когнітивних або моторних завдань; проєктивні методи, що використовуються для виявлення глибинних, неусвідомлюваних аспектів особистості.

Суб'єктивні методи ґрунтуються на самозвіті – здатності людини усвідомлювати, описувати і оцінювати власні стани, переживання та риси. У практиці вони реалізуються через опитувальники, анкети, інтерв'ю, щоденники самоспостереження.

Основою суб'єктивного підходу є переконання, що людина – найкраще джерело інформації про саму себе, а її внутрішній досвід має наукову цінність, якщо оформлений у стандартизованій формі. Саме на цьому принципі побудовані більшість особистісних тестів: MMPI, 16PF Кеттелла, опитувальник Айзенка, шкали самооцінки Розенберга.

Сильна сторона суб'єктивних методів – висока екологічна валідність, тобто відповідність реальному життєвому досвіду людини. Вони дозволяють фіксувати унікальні переживання, мотиваційні тенденції, установки, що важко виміряти зовні.

Проте слабе місце цих методів – вразливість до спотворень: соціальна бажаність, страх оцінки, низький рівень саморефлексії або бажання дати «правильну» відповідь. Саме тому сучасні дослідники вважають суб'єктивні методи ефективними лише за умов контролю додаткових факторів – наприклад, включення шкал щирості чи комбінування з поведінковими показниками.

Об'єктивні методи на відміну від суб'єктивних не залежать від самооцінки людини. До них належать експериментальні завдання, реакційні тести, вимірювання часу реакції, частоти рухів, кількості помилок тощо. Такі показники є більш надійними, оскільки фіксують поведінкові прояви, що безпосередньо корелюють із функціонуванням нервової системи.

Класичним прикладом є тести на увагу та швидкість реакції (наприклад, тест Бурдона або коректурні проби). Сучасні нейропсихологічні дослідження використовують також комп'ютерні системи, що фіксують мікрорухи очей, зміни електропровідності шкіри, пульс, активність мозку. Такі вимірювання наближають психологію до природничих наук, забезпечуючи високу точність і відтворюваність.

Проективні методи займають особливе місце, оскільки поєднують суб'єктивний і об'єктивний вимір. Вони ґрунтуються на психоаналітичному припущенні, що людина несвідомо проєкує свої внутрішні конфлікти, бажання і установки на неоднозначні стимули. Класичні приклади – тест Роршаха, тематичний аперцептивний тест (ТАТ), малюнкові методики (Дім–Дерево–Людина, Неіснуюча тварина).

Проективний метод не вимірює безпосередньо якусь «величину», як шкала або тест. Його мета – виявлення структури особистості, механізмів

захисту, прихованих мотивів. Інтерпретація результатів тут завжди включає елемент клінічного судження.

З позиції психометрії, проєктивні методи часто критикують за низьку надійність і валідність, адже вони залежать від досвіду та компетентності інтерпретатора. Проте їхня сила – у глибинному розкритті особистісного змісту, який не завжди доступний у тестах чи анкетах.

Сучасна практика намагається подолати цю суперечність, комбінуючи проєктивні методики з кількісними аналізами. Наприклад, у тесті Роршаха зараз застосовують систему Екснера, що дозволяє кодувати відповіді за стандартизованими критеріями (форма, зміст, локалізація), підвищуючи надійність оцінки.

Крім того, у ХХІ столітті з'явилися цифрові проєктивні методики: аналіз малюнків на планшетах, асоціативні тести в онлайн-форматі, навіть текстові проєктиви, де алгоритм штучного інтелекту аналізує семантику відповідей. Це розширює застосування класичних принципів у новому технологічному контексті.

Усі ці типи вимірювань взаємодоповнюють одне одного. Сучасна психологія прагне інтегрувати різні підходи, створюючи мультимодальні системи оцінювання, де результати самооцінки поєднуються з об'єктивними й поведінковими індикаторами. Такий підхід дозволяє компенсувати обмеження кожного методу й підвищити валідність отриманих результатів.

В цілому маємо відмітити, що вимірювання можна визначити як приписування чисел за певними правилами об'єктам або подіям.

Будь-який різновид вимірювання передбачає наявність одиниць вимірювання. Одиниця вимірювання – це та «вимірювальна паличка», як говорив Стенлі Стівенс, яка є умовним еталоном для здійснення тих або інших вимірювальних процедур. У природничих науках і техніці є стандартні одиниці вимірювання, наприклад, градус, метр, ампер тощо.

Психологічні змінні за одиничними винятками не мають власних вимірювальних одиниць. Тому здебільшого значення психологічної ознаки визначається за допомогою спеціальних вимірювальних шкал. Розрізняють декілька типів шкал. Операції, а саме способи вимірювання об'єктів задають тип шкали. Шкала, у свою чергу, характеризується видом перетворень, які можуть бути віднесені до результатів вимірювання. Якщо не дотримуватися цього правила, то структура шкали порушиться, а дані вимірювання не можна буде осмислено інтерпретувати.

Тип шкали визначає сукупність статистичних методів, які можуть бути використані для обробки даних вимірювання. Шкала (латин. *scala*) – інструмент для вимірювання властивостей об'єкта, що являє собою числову систему, у якій відносини між різними властивостями об'єктів виражені властивостями числового ряду. У психології різні шкали використовуються для вивчення різних характеристик соціально-психологічних явищ. Виділяється чотири типи числових систем, що визначають відповідно чотири рівні, або шкали вимірювання: 1) шкала номінальна (номінативна,

найменувань, класифікаційна); 2) шкала порядку (рангова, ординарна); 3) шкала інтервалів (рівних інтервалів); 4) шкала відношень (пропорційна).

Почнемо з найпростішої і найменш інформативної – номінальної шкали. Ця шкала використовується виключно для класифікації об'єктів або подій за категоріями, без будь-якого внутрішнього порядку, ранжування чи кількісного значення. Числа, які можуть бути присвоєні категоріям у номінальній шкалі, слугують лише для їх ідентифікації та розрізнення, вони є просто мітками або кодами. Наприклад, якщо ми класифікуємо людей за кольором очей, ми можемо присвоїти «карим» – 1, «блакитним» – 2, «зеленим» – 3. Проте, ці числа не мають жодного математичного сенсу: ми не можемо стверджувати, що «блакитні очі» (2) є «більшими» за «карі очі» (1) або що різниця між «карими» і «блакитними» очима є такою ж, як між «блакитними» і «зеленими».

Наступним кроком у ієрархії вимірювальних шкал є порядкова шкала. Вона зберігає всі властивості номінальної шкали, але додає до них властивість порядку або ранжування. Це означає, що категорії можуть бути розташовані в певній логічній послідовності – від «меншого» до «більшого», від «нижчого» до «вищого» за певним критерієм. Однак, ключовою відмінністю від більш складних шкал є те, що інтервали між категоріями не є рівними або значущими. Ми знаємо, що «дуже добре» краще за «добре», а «добре» краще за «посередньо», але ми не можемо кількісно оцінити, наскільки саме «дуже добре» краще за «добре». Різниця між «першим місцем» і «другим місцем» на змаганнях не обов'язково є такою ж, як різниця між «другим місцем» і «третьим місцем». У цьому випадку ми знаємо порядок, але не величину різниці. Порядкові шкали не мають справжнього нуля, тобто точки відліку, що означає повну відсутність вимірюваної властивості.

Інтервальна шкала є значним кроком вперед у порівнянні з попередніми, оскільки вона додає до властивостей порядкової шкали рівні інтервали між сусідніми значеннями. Це означає, що різниця між 20°C і 30°C така ж, як різниця між 40°C і 50°C (10°C). Тепер ми можемо не тільки ранжувати дані, але й кількісно оцінювати відстань між ними.

Найвищий рівень вимірювань представлений шкалою відношень. Ця шкала має всі властивості інтервальної шкали – класифікацію, впорядкування та рівні інтервали – але, що найважливіше, вона також володіє істинним, абсолютним нулем. Справжній нуль на шкалі відношень означає повну відсутність вимірюваної властивості. Наприклад, нульова вага означає повну відсутність маси, нульова довжина означає повну відсутність протяжності, а нульовий час означає повну відсутність тривалості.

Номінальний рівень вимірювання використовується для категоріальних змінних і передбачає присвоєння балів, які є мітками категорій. Номінальні шкали таким чином уособлюють найнижчий рівень вимірювання.

У психометричних вимірюваннях дуже важливо розуміти, які математичні операції можна виконувати з числами, що використовуються для представлення психологічних величин, щоб не спотворити їхній сенс. Ці

операції залежать від типу шкали вимірювання. Всі дозволені операції мають не змінювати відносини між об'єктами, тобто порядок, інтервали або пропорції повинні зберігатися.

Тобто, вимірювання – це фундаментальний процес, що дозволяє нам надавати кількісні значення явищам навколишнього світу, перетворюючи якісні характеристики на числові дані, які потім можна аналізувати, порівнювати та використовувати для прийняття рішень. Проте, не всі вимірювання є однаковими за своєю суттю, і розуміння типів вимірювальних шкал є абсолютно критичним для коректної інтерпретації даних, застосування відповідних статистичних методів та уникнення хибних висновків. Саме тип шкали визначає, які математичні операції ми можемо виконувати з отриманими даними та наскільки глибокі висновки ми можемо зробити.

4.3. Види та характеристики шкал відповідей на запитання в опитувальниках

Шкали відповідей – це способи, за допомогою яких респондент виражає свою думку, оцінку або емоційний стан у процесі психологічного вимірювання. Залежно від різних характеристик, шкали можна класифікувати за кількома критеріями: за характером відношень, числовою відповідністю, наявністю та/або сенсу полюсів, а також матеріалом.

Методи бальних оцінок – це психометричні інструменти, які використовуються для оцінки суб'єктивних станів, ставлень, переживань або рівня задоволення респондентів. Основна особливість таких методів – це представлення шкали, що дозволяє респондентові виразити ступінь інтенсивності певного переживання чи ставлення.

У бальних оцінках градації відповіді можуть задаватися різними способами, що впливає як на зручність заповнення, так і на подальшу інтерпретацію даних. Розглянемо три основні підходи, а власне текстові, числові та графічні.

Текстові шкали передбачають, що респонденту пропонуються словесні варіанти відповіді, які відображають різні рівні інтенсивності або якості переживання.

У числових шкалах респондентам пропонується оцінити певний аспект на лінійній шкалі з визначеним діапазоном чисел, наприклад, від 0 до 10 або від 1 до 5.

Графічні шкали – це візуальні засоби, які дозволяють респондентові інтуїтивно обрати варіант, що найбільше відповідає його стану або ставленню. Найпоширеніший приклад – «шкала смайлів», де зображені обличчя з різними емоціями.

За кількістю альтернатив вимірювання виділяють дискретні, в яких респондент обирає одну відповідь із заздалегідь визначеного, обмеженого

набору варіантів. Це можуть бути як числові, так і вербальні опції (наприклад, від 1 до 5, або «згоден / не зовсім згоден / не згоден»).

Неперервні шкали (континуальні) дозволяють респондентові поставити «позначку» у будь-якій точці шкали, тобто число не фіксується у межах конкретного набору, а може бути абсолютно довільним у зазначеному діапазоні (наприклад, від 0 до 100).

Одним із критичних параметрів у конструюванні шкал є кількість градацій, тобто кількість відповідей, серед яких респондент має обрати найближчу до свого внутрішнього стану. Кількість і форма градацій безпосередньо впливають на точність, достовірність, інтерпретацію й навіть емоційний комфорт респондента.

Дихотомічна шкала є найпростішим варіантом, коли респондент має обрати одну з двох альтернатив, наприклад: «так / ні», «згоден / не згоден», «правда / неправда».

В тризначній шкалі до «так» і «ні» додається нейтральний варіант, як-от «так / не впевнений / ні».

У сучасній практиці найчастіше використовуються шкали з 4–7 пунктами, які поєднують достатню чутливість із прийнятною простотою.

Шкали з більшим числом градацій можуть складно диференціюватися, як от 0–10, або 0–100, де респонденти мають обрати будь-яке значення в заданому діапазоні, іноді – навіть у вигляді відсотків.

Розглянемо три класичні моделі, які стали базовими для сучасного тестування: шкали Терстоуна, Гуттмана та Лікерта.

Луїс Терстоун у 1928 році запропонував підхід, який став першою системною спробою виміряти ставлення людини до соціальних або політичних об'єктів. Його метод базується на ідеї психічного континууму.

Шкала Терстоуна передбачає, що експерти оцінюють кожен пункт за рівнем вираженості досліджуваного конструкта. Потім середня експертна оцінка кожного пункту використовується як його вага.

Респонденту пропонують набір тверджень, він обирає ті, з якими згоден, а загальний бал виводиться з підсумку ваг обраних пунктів. Теоретично цей метод дає точне вимірювання, бо пункти розміщуються на числовій осі (наприклад, від 1 до 11) згідно з експертною оцінкою.

Перевага методу – його логічна строгість. Недолік – складність побудови: потрібно залучати велику групу експертів і забезпечувати однорідність суджень.

Попри трудомісткість, шкала Терстоуна відіграла ключову роль у становленні емпіричної соціальної психології. Вона вперше дозволила говорити про «вимірюваний атитюд» як стабільну характеристику, що має числове вираження.

Сьогодні метод Терстоуна застосовується рідше, але принцип рівних інтервалів використовується у побудові сучасних опитувальників і лінгвістичних шкал оцінювання (наприклад, при створенні шкал політичної поляризації або шкал толерантності).

Луї Гуттман у 1944 році створив кумулятивну шкалу, яка передбачає, що згода з певним твердженням означає автоматичну згоду з усіма попередніми за «силою». Це дозволяє не просто вимірювати ставлення, а й досліджувати логічну структуру переконань.

Наприклад, у шкалі патріотизму можуть бути такі твердження:

1. «Мені подобається національна символіка».
2. «Я пишаюся своєю країною».
3. «Я готовий працювати задля її розвитку».
4. «Я готовий пожертвувати собою заради Батьківщини».

Якщо респондент згоден із твердженням №4, то логічно, що він згоден і з усіма попередніми. Таким чином, відповіді можна представити у вигляді ієрархічного шаблону.

Гуттман запровадив поняття коефіцієнта репродуктивності – показника того, наскільки відповіді респондентів відповідають ідеальній кумулятивній моделі. Якщо цей коефіцієнт перевищує 0,90, шкала вважається надійною.

Перевага шкали Гуттмана – її логічна послідовність і одновимірність: усі твердження вимірюють один конструкт, а не кілька різних аспектів. Це робить її цінною у дослідженнях соціальних установок, ступеня залученості, рівня екстремізму чи радикалізації.

Однак така модель передбачає жорстку лінійність розвитку ставлень, що не завжди відповідає реальності. Людина може, наприклад, бути готовою допомагати країні, але не погоджуватись із певними символами чи політикою. Тому у 1970–1990-х роках Гуттманівський підхід еволюціонував у моделі часткової кумуляції та багатовимірного шкалування (MDS), де взаємини між пунктами описуються геометрично, а не строго ієрархічно.

Найвідомішою і найпоширенішою є шкала Лікерта, розроблена Ренсісом Лікертом у 1932 році. Це найпопулярніший і найпрактичніший варіант побудови складених шкал. Шкала Лікерта складається з набору пунктів, кожен з яких оцінюється респондентом за градацією згоди, як правило за п'яти- або семибальною шкалою (від «повністю не згоден» до «повністю згоден»). Кожне твердження формулюється у вигляді суб'єктивного судження: «Я почуваюся спокійно, навіть коли стикаюся з труднощами».

Усі відповіді підсумовуються або усереднюються, що й дає загальний показник рівня певного конструкту.

Перевага цього методу – простота і висока надійність. Хоча окремі пункти не мають рівних інтервалів, у сукупності вони формують інтервальну шкалу. Саме шкала Лікерта стала стандартом для більшості сучасних опитувальників і соціологічних досліджень.

Сьогодні шкали Лікерта застосовують навіть у цифровому форматі – онлайн-опитуваннях, мобільних додатках, нейропсихологічних тестах. З'явилися адаптовані варіанти з візуальними аналоговими шкалами (VAS), що підвищують зручність для респондента і точність результатів.

Загалом, у процесі вимірювання психологічних властивостей перевагу надають багатопунктовим шкалам, бо вони дають більш точні, стабільні та

надійні результати. Історично виникло кілька підходів до побудови таких шкал, але саме модель Лікерта стала домінуючою, завдяки оптимальному балансу між простотою реалізації та психометричною ефективністю.

Попри свої успіхи, психометричний підхід стикається з низкою проблем. По-перше, він передбачає лінійність і стабільність вимірюваних характеристик, тоді як психіка є динамічною і контекстуальною системою. Поведінка людини може змінюватися залежно від ситуації, мотивації, соціального тиску.

По-друге, багато тестів створені в західному культурному контексті і не завжди валідні для інших культур. Крос-культурні дослідження виявили, що поняття «екстраверсії», «лідерства», «емоційного інтелекту» тощо можуть мати різні смисли у різних суспільствах.

По-третє, психометричні методи не враховують індивідуальної унікальності: вони усереднюють результати, підганяючи їх під статистичну норму. Ця стандартизація важлива для науки, але може бути обмеженням у практичній роботі з особистістю.

Після поширення шкал Лікерта в 1960-х роках почалася справжня «психометрична революція». Соціальні науки отримали інструмент, здатний масово фіксувати ставлення, думки, емоції. Проте вже тоді постало питання – чи не надто ми спрощуємо складну внутрішню реальність людини до кількох балів?

Деякі дослідники вказували, що відповіді в анкетах відображають не лише справжні переконання, а й соціальну бажаність – тенденцію демонструвати себе у позитивному світлі. Це явище стало окремою проблемою вимірювання – ефектом соціальної бажаності (social desirability bias).

Для його компенсації розроблено спеціальні шкали-коректори – наприклад, шкалу обману Айзенка або шкалу Марлоу–Крауна, що оцінюють схильність до спотворення відповідей.

У 2000-х роках на основі шкал Лікерта виникла теорія латентних класів, яка дозволила виявляти приховані групи респондентів із подібними профілями відповідей. Це стало проривом у маркетинговій і політичній психології, де вимірювання використовується для прогнозування поведінки виборців чи споживачів.

Саме тому сучасна психологія дедалі більше рухається до інтегративних моделей, що поєднують кількісні й якісні підходи – від експериментів і спостережень до аналізу наративів, біографічних методів і поведінкової аналітики.

Жодне вимірювання не має сенсу без гарантій його точності та відповідності реальності. У психології ці критерії визначаються трьома базовими поняттями: надійність, валідність і стандартизація. Вони утворюють трикутник наукової обґрунтованості, на якому тримається вся психометрія.

4.4. Психометрична надійність тесту

Психометрична надійність тесту є однією з ключових характеристик будь-якого психодіагностичного інструменту і слугує базовим критерієм його наукової якості, практичної придатності та коректності подальших висновків, що ґрунтуються на результатах тестування. У психологічній діагностиці саме надійність визначає, наскільки вимірювання є точним, стабільним і відтворюваним у різних ситуаціях, за умови, що вимірювана властивість залишається сталою. Під надійністю розуміють не абсолютну точність, якої в психології досягти неможливо через багатовимірність і мінливість людської поведінки, а ступінь вільності результатів тесту від випадкових похибок. Це означає, що надійний тест повинен давати подібні оцінки при повторному вимірюванні, незалежно від того, чи йдеться про інше тестування в часі, інший варіант тесту, інший дослідницький контекст або нову вибірку респондентів.

У класичній тестовій теорії, закладеній у працях Лорда й Новіка в 1920–1930-х роках, надійність розглядається через співвідношення між істинним значенням вимірюваної характеристики та випадковою похибкою. Формула $X = T + E$, де X є отриманим (спостережуваним) результатом тестування, T – істинним балом, а E – похибкою вимірювання, відображає основну ідею: жоден тест не може дати абсолютно точного значення, оскільки будь-який результат містить елемент випадковості. Завдання психометрії полягає у зменшенні цієї випадковості, мінімізації її впливу та наблизенні X до T . Саме тому надійність визначає ступінь, у якому результат тестування є вільним від впливу випадкових, неконтрольованих чинників, таких як коливання уваги досліджуваного, зміни умов тестування, специфічні особливості формулювання запитань або інтерпретаційні відмінності тестувальника. Чим меншою є похибка E , тим ближчим до істинного значення T є результат X , а отже, тим вищою є надійність тесту.

Надійність тесту проявляється у тому, наскільки послідовними є результати тестування при повторному вимірюванні. Якщо індивід, проходячи тест повторно, отримує подібні оцінки, якщо зміна формату, інструкції або порядку завдань не призводить до суттєвих коливань у балах, це є свідченням високої стабільності тесту. Дослідження із застосуванням підходу test–retest покликані встановити, наскільки результати залишаються стабільними у часі. Цей метод передбачає повторне проведення тесту на тій самій групі учасників через певний інтервал, після чого обчислюється кореляція між двома наборами оцінок. Висока кореляція означає, що тест фіксує стабільну психологічну властивість, а не випадкові коливання поведінки. Водночас тривалість інтервалу між тестуваннями є критично важливою: якщо проміжок надто малий, може проявитися ефект запам'ятовування; якщо надто великий – реальні зміни в психологічному стані можуть знизити показники надійності, не залежно від якості тесту. Це означає, що test–retest надійність ніколи не є універсальним показником і завжди потребує урахування контексту та особливостей вимірюваної ознаки.

Крім стабільності у часі, надійність також проявляється у внутрішній узгодженості тестових пунктів. Якщо тест складається з кількох пунктів, що мають вимірювати одну й ту саму характеристику, рівень їхньої кореляції між собою є показником того, наскільки послідовно тест оцінює психологічну властивість. Найпоширенішим індикатором внутрішньої узгодженості є коефіцієнт Кронбаха (α), значення якого набуває від 0 до 1. Традиційно прийнятним вважається значення α не нижче 0,70, що свідчить про достатню узгодженість пунктів. Однак надмірно високе значення, близьке до 1, може свідчити про дублювання змісту пунктів, тобто про те, що різні завдання вимірюють занадто подібний аспект властивості, що може знижувати інформативність тесту й обмежувати його діагностичний потенціал. Завдання психометриста полягає в тому, щоб досягти оптимального балансу: тестові пункти мають бути достатньо узгодженими, щоб вимірювати одну конструкцію, але не настільки однотипними, щоб вільно відтворювати одне й те саме.

Ще одним способом оцінки надійності є використання паралельних форм тесту, які мають однакову структуру, рівень складності, змістове наповнення та форму, але являють собою різні варіанти завдань. Якщо результати, отримані за допомогою двох паралельних форм, високо корелюють між собою, це свідчить про те, що тест залишає спроможним вимірювати одну й ту саму характеристику незалежно від варіації пунктів. Цей підхід є особливо актуальним для тестів, що використовуються у ситуаціях, де можливий ефект тренування або де тестування здійснюється багаторазово, як-от у навчальних програмах чи при оцінюванні динаміки професійних компетенцій. Однак створення паралельних форм є надзвичайно трудомістким процесом, який вимагає ретельного добору завдань і математичного вирівнювання складності.

Окремим підходом є оцінка надійності методом розщеплення тесту на частини (split-half reliability), коли тест умовно поділяється на дві рівнозначні частини – наприклад, на основі парних та непарних пунктів. Кореляція між результатами цих частин дозволяє оцінити рівень внутрішньої узгодженості тесту. Оскільки половина тесту містить менше пунктів, отримане значення кореляції коригується за допомогою формули Спірмана–Брауна, щоб екстраполювати його до повної довжини тесту. Цей метод є зручним і ефективним у ситуаціях, коли тест містить достатню кількість пунктів та є однорідним за змістом. Проте він менш придатний для багатовимірних тестів, де різні шкали або підшкали мають вимірювати відмінні аспекти психологічної характеристики. У таких випадках кожна шкала потребує окремої оцінки надійності, оскільки інтегровані показники не відображатимуть реальної якості вимірювання.

У практиці психодіагностики важливо розуміти, що рівень надійності відрізняється залежно від цілей і контексту використання тесту. Інструменти, які застосовуються у високоризикових сферах – наприклад, у відборі персоналу на відповідальні посади, клінічній діагностиці, медико-

психологічних експертизах, прогнозуванні поведінкових тенденцій – повинні демонструвати надзвичайно високі коефіцієнти надійності. Значення α вище 0,90 в таких ситуаціях вважається бажаним, адже від точності результатів залежить ухвалення критично важливих рішень щодо професійної діяльності, лікування або правового статусу особи. Водночас у дослідницькій роботі може бути допустимим використання інструментів із помірним рівнем надійності, якщо метою є виявлення загальних тенденцій, а не ухвалення індивідуальних рішень. У будь-якому разі надійність ніколи не повинна розглядатися як універсальна або постійна властивість тесту: вона залежить від вибірки, від особливостей контексту, від інструкції, від способу застосування тесту і навіть від рівня мотивації досліджуваних. Тому кожен раз, коли тест використовується у новій групі або в нових умовах, його надійність має бути перевірена емпірично.

Питання підвищення надійності тесту передбачає використання кількох стратегій, які спрямовані на мінімізацію випадкових похибок. Одним із найпоширеніших методів є збільшення кількості пунктів, що дозволяє згладити вплив випадкових коливань у відповідях. Проте цей підхід має сенс лише тоді, коли пункти є змістовно якісними та репрезентують різні аспекти вимірюваної конструкції. Уніфікація умов проведення тестування – стандартизація інструкцій, часу виконання, форми подання стимулів – є ще одним важливим чинником підвищення надійності. Нестандартизоване тестування створює додаткові джерела похибок, які можуть суттєво знизити точність результатів. Важливого значення набуває і якісна підготовка адміністраторів тесту, адже будь-які варіації у способі подання завдань, у взаємодії з досліджуваним або в реакції на його питання можуть впливати на результати.

Надійність тесту слід розглядати не лише як властивість інструменту, але і як інтегральний показник якості всієї системи вимірювання. Вона відображає здатність психодіагностичного процесу забезпечити відтворювані результати, незалежно від індивідуальних відмінностей між адміністратором, особливостями ситуації або ситуативним станом досліджуваного. У цьому сенсі надійність є запорукою довіри до психологічних методів оцінювання. Якщо тест не є надійним, то його результати можуть вводити в оману, спричиняючи хибні висновки, некоректні діагнози, необґрунтовані прогнози або невиправдані управлінські рішення. Саме тому психометричні стандарти передбачають ретельну оцінку надійності ще до того, як тест потрапляє до практики, і зобов'язують авторів документувати всі процедури, методи аналізу, статистичні моделі й емпіричні дані, на основі яких було отримано показники надійності.

У сучасних умовах особливої актуальності набуває питання прозорості психометричних досліджень. Відсутність відкритих даних, обмежений опис методик та нестача аналітичних скриптів значно ускладнюють перевірку надійності та відтворюваність психодіагностичних інструментів. Якщо автор не надає інформацію про процедуру формування вибірки, критерії

виключення учасників, методи очищення даних або способи обчислення коефіцієнтів надійності, незалежні дослідники не здатні оцінити, наскільки отримані показники є достовірними. Така ситуація створює ризики аналогічні тим, що лежать в основі реплікаційної кризи: відсутність прозорості унеможливорює перевірку якості інструменту та відтворення результатів у нових умовах. Психометрична надійність тесту, таким чином, стає не лише технічним показником, але й показником наукової відповідальності та відкритості.

Отже, психометрична надійність є фундаментальною умовою наукової цінності будь-якого психодіагностичного інструменту. Вона визначає, наскільки точними, стабільними та звільненими від випадкових похибок є результати тестування. Надійність не є постійною і не може розглядатися як властивість, що раз і назавжди притаманна тесту. Вона завжди залежить від вибірки, контексту, методів вимірювання та умов застосування. Тому кожен тест, який претендує на практичне використання, повинен демонструвати стабільні показники надійності, підтверджені відкритими, прозорими та відтворюваними дослідженнями. У цьому сенсі надійність є не лише психометричним критерієм, але й важливим чинником побудови довіри до психологічної науки загалом.

Зрештою, слід пам'ятати, що надійність – це лише один із елементів психометричної якості тесту. Можна створити тест, який стабільно вимірює певну характеристику, але не ту, яку він мав би вимірювати. Висока надійність без належної валідності не гарантує корисності тесту. Отже, процес створення й оцінки психодіагностичних методик вимагає комплексного підходу, де надійність є обов'язковим, але не єдиним критерієм ефективності.

4.5. Валідність як психометрична властивість тесту

Валідність як психометрична властивість тесту займає важливе місце в сучасній психологічній діагностиці, оскільки визначає не просто технічну якість інструменту, а змістову обґрунтованість і правомірність інтерпретацій, які здійснюються на основі тестових балів. Якщо надійність дає відповідь на запитання, наскільки стабільними та відтворюваними є результати вимірювання, то валідність стосується зовсім іншого виміру – того, чи дійсно тест вимірює те, що він має вимірювати, і чи відображають зроблені на основі його результатів висновки реальні психологічні властивості, процеси або стани. Іншими словами, валідність є характеристикою не стільки самого тесту як набору завдань, скільки інтерпретацій цих завдань у контексті певної теоретичної моделі та практичних рішень. Валідний тест – це інструмент, щодо якого зібрано переконливі й багатовимірні докази того, що його результати можуть бути осмислено використані для опису, пояснення чи прогнозу поведінки, і що ці інтерпретації не є довільними чи випадковими. У класичній психометричній традиції валідність часто описували через три основні її типи – змістову, критеріальну та конструктивну.

Змістова валідність фокусувалася на тому, наскільки зміст завдань тесту охоплює суттєві аспекти теоретично описуваної ознаки: для шкали тривожності важливо, щоб її пункти відображали як емоційні, так і когнітивні, соматичні й поведінкові компоненти тривожності, а не обмежувалися, наприклад, лише суб'єктивним переживанням напруги.

Критеріальна валідність передбачала зіставлення результатів тесту з незалежними зовнішніми критеріями: академічною успішністю, професійними досягненнями, клінічними діагнозами, поведінковими показниками, експертними оцінками. Якщо кореляція між тестовими балами та критерієм є високою й стабільною, можна стверджувати, що тест має високу валідність.

Конструктивна валідність вважалася найглибшим рівнем оцінки: вона передбачала перевірку того, чи відповідає структура тесту теоретичному конструкту, що лежить в його основі, чи узгоджуються між собою емпірично виділені фактори та очікувана модель, чи пов'язаний тест із іншими методиками так, як це впливає з теорії (тобто демонструє конвергентну й дискримінантну валідність).

Однак із часом стало зрозуміло, що такий поділ носить радше методичний, ніж сутнісний характер, і що всі ці аспекти валідності є різними джерелами доказів на користь однієї й тієї самої тези – коректності інтерпретацій тестових результатів. Сучасна психометрія, спираючись на концепцію, розроблену Семюелем Мессіком, розглядає валідність як єдину інтегральну характеристику, яка включає не лише структурні, змістові й критеріальні аспекти, а й ширший контекст застосування тесту – соціальний, етичний, культурний. У цьому підході валідність уже не мислиться як «властивість тесту як такого», вона трактується як властивість висновків, зроблених на підставі тестових балів у певних умовах застосування. Це означає, що не можна один раз «довести валідність» і вважати її сталою; її потрібно постійно підтверджувати, збираючи все нові й нові емпіричні, теоретичні та соціально-етичні докази в різних вибірках, культурах, контекстах використання. Важливим наслідком цього підходу стало усвідомлення, що тест може бути валідним для однієї мети (наприклад, скринінгу загальної тривожності) і недостатньо валідним для іншої (наприклад, диференціальної діагностики тривожних розладів), а отже, валідність завжди має контекстуальний характер.

Процес емпіричного обґрунтування валідності є складним, тривалим і методологічно насиченим. Для оцінки конструктивної валідності широко застосовується факторний аналіз, який дає змогу перевірити, чи відповідає емпірична структура даних теоретично очікуваній. Якщо, наприклад, тест особистості будується на моделі «Великої п'ятірки», то факторний аналіз має продемонструвати виділення п'яти відносно незалежних факторів, які інтерпретуються як екстраверсія, доброзичливість, сумлінність (свідомість), нейротизм і відкритість досвіду. Сучасні підходи здебільшого використовують конфірматорний факторний аналіз (CFA), що дозволяє не

просто «шукати» структуру, а перевіряти відповідність даних наперед заданій моделі, а також моделювання структурними рівняннями (SEM), яке дає змогу розглядати тести не як прості суми пунктів, а як елементи складніших латентних структур. У парадигмі теорії відповіді на тестові завдання (Item Response Theory – IRT) валідність пов'язують з адекватністю моделі (наприклад, моделі Раша або двопараметричної логістичної моделі) до емпіричних даних: якщо параметри завдань та осіб узгоджуються з припущеннями моделі, можна вважати, що інтерпретація латентної риси має психометричні підстави. Усі ці підходи створюють те, що С. Мессік називав «мережею доказів валідності»: жоден окремий аналіз сам по собі не гарантує валідності, але сукупність різних джерел підтверджень – теоретичних обґрунтувань, емпіричних кореляцій, факторних структур, результатів міжгрупових порівнянь, кроскультурних досліджень – формує відносно надійну основу для ствердження, що тестові бали мають той зміст, який їм приписують.

Особливо складним виміром проблеми валідності є кроскультурна валідність, тобто питання про те, чи зберігає тест свою інтерпретаційну адекватність при перенесенні в інші культурні, мовні та соціальні контексти. Тест, створений у певній країні, відображає не лише абстрактну психологічну конструкцію, а й специфіку мови, цінностей, норм вираження емоцій, соціальних очікувань. Наприклад, шкала депресії Бека чи опитувальник для вимірювання професійного вигорання, створені в англomовному середовищі, не можуть автоматично вважатися валідними для використання в українській вибірці – навіть за умови якісного перекладу. Для забезпечення кроскультурної валідності потрібен багаторівневий процес адаптації: прямий переклад, реверсний переклад, експертна оцінка еквівалентності змісту, пілотажне дослідження, статистичний аналіз диференціального функціонування пунктів (DIF), порівняння факторної структури у різних вибірках. Лише тоді можна говорити про те, що інструмент вимірює той самий конструкт у різних культурах, а отже, висновки на основі тестових результатів не спотворюють реальність через лінгвістичні або культурні упередження. У цьому контексті валідність перестає бути суто технічним параметром і набуває виразного етичного змісту: невалідний у певній культурі тест може призвести до стигматизації, неправильної діагностики, несправедливого відбору або дискримінації окремих груп.

Валідність тісно пов'язана з надійністю, але не зводиться до неї. Надійність є необхідною умовою валідності: якщо результати вимірювання є надто варіативними, шумовими, нестабільними, жодні інтерпретації не можуть бути достатньо обґрунтованими. Проте висока надійність сама по собі не гарантує валідності. Цілком можливо створити інструмент, який стабільно й дуже послідовно вимірює феномен, що не має жодного стосунку до теоретично заявленого конструкта. Наприклад, тест, що нібито вимірює креативність, але насправді значною мірою відображає рівень загального інтелекту або швидкість читання, може бути дуже надійним (усі його пункти

узгоджені, результати стабільні), але валідність його висновків як показників саме креативності є сумнівною. Відтак психометрична оцінка тесту має бути комплексною: на першому етапі перевіряють надійність (внутрішню узгодженість, стабільність, еквівалентність форм), але на цьому не зупиняються – далі необхідно збирати докази на користь валідності, аналізуючи структуру тесту, його зв'язки з іншими змінними, адекватність інтерпретацій у різних групах.

Особливого значення питання валідності набуває у контексті сучасної реплікаційної кризи та дискусій про відкритість наукових даних. Недостатня прозорість досліджень, відсутність доступу до сирих даних, аналітичних скриптів та повних описів процедур істотно обмежують можливості незалежної перевірки валідності тестів. Валідність, як уже наголошувалося, є не раз і назавжди встановленою властивістю, а постійним процесом підтвердження інтерпретацій. Цей процес неможливий без того, щоб інші дослідники мали змогу перевіряти розрахунки, повторювати аналізи, застосовувати інструмент у нових вибірках і культурах, тестувати конкуруючі моделі. Коли вихідні дані закриті, опис факторної структури подається лише у вигляді кінцевих коефіцієнтів, а моделі не супроводжуються відкритими скриптами, будь-які твердження про конструктивну або критеріальну валідність залишаються, по суті, неконтрольованими. У такій ситуації валідність перетворюється на декларацію авторів, а не на результат колективної наукової перевірки. Саме рух за відкриту науку, який формувався як відповідь на реплікаційну кризу, наголосив, що валідність не може існувати без прозорості: щоб вважати тест валідним, наукова спільнота повинна мати доступ до повного корпусу доказів, а не лише до вибірових, часто надто «відредагованих» презентацій результатів.

У цьому сенсі відсутність відкритих даних і закритість психометричних досліджень виступають одним із прихованих механізмів підтримання псевдовалідності. У психодіагностичній практиці нерідко трапляється ситуація, коли тест активно використовується десятиліттями, посилаючись на «класичні» дослідження валідації, проте ці дослідження ніколи не були повністю оприлюднені, не перевірялися незалежними групами, не відтворювалися у нових вибірках, не доповнювалися сучасними аналізами. Унаслідок цього валідність тесту фактично спирається на авторитет джерела, а не на сукупність актуальних емпіричних підтверджень. Реплікаційна криза у психології продемонструвала, що значна частина «перевіраних» ефектів виявляється нестійкою при повторних дослідженнях. Ця логіка повною мірою стосується й психодіагностичних тестів: якщо факторна структура, кореляції з критеріями, кроскультурна еквівалентність та інші аспекти валідності не підтверджуються у незалежних дослідженнях, то впевненість у валідності інструменту є ілюзорною. Умовою виправлення цієї ситуації є радикальне підвищення прозорості: відкритий доступ до даних, детальний опис процесу побудови шкал, публікація всіх, а не лише «успішних» результатів валідації.

У практиці тестування питання валідності виходить далеко за межі суто статистичних процедур і торкається сфери етичної та соціальної відповідальності психолога. Коли тест використовується для відбору персоналу, розподілу освітніх можливостей, клінічної діагностики, ухвалення рішень щодо доступу до ресурсів, невірно валідизований або невірно інтерпретований тест може призвести до серйозних негативних наслідків для окремих людей і груп. Саме тому міжнародні стандарти, такі як «Standards for Educational and Psychological Testing» (AERA, APA, NCME), вимагають від розробників тестів та практиків-психологів не лише наводити статистичні показники надійності, а й обґрунтовувати валідність у вигляді системи доказів: теоретичних (опис конструкта і логіка добору пунктів), емпіричних (результати факторного аналізу, кореляцій з критеріями), контекстуальних (дані про групи, для яких тест є репрезентативним) та соціальних (аналіз можливих наслідків використання тесту, наявності або відсутності упереджень, ризик дискримінації). Валідність повинна включати оцінку того, як результати тестування впливають на життя людей, чи не посилюють вони соціальну нерівність, чи не закріплюють стереотипів, чи не використовуються для легітимації несправедливих практик. Таким чином, валідність набуває виміру не лише «істинності інтерпретацій», а й «справедливості їх застосування».

Паралельно з цим змінюється й загальна парадигма психометрії. Якщо раніше основне завдання полягало в тому, щоб створити максимально точний, «твердий» інструмент вимірювання, який, мовби фізичний прилад, однозначно відображає певну латентну властивість, то нині дедалі частіше визнається, що будь-яке психодіагностичне вимірювання є контекстуально обумовленим, інтерпретаційно насиченим і соціально вмонтованим. Перехід від класичної до постнекласичної наукової раціональності означає відмову від ілюзії повної об'єктивності спостерігача, визнання впливу дослідницьких установок, мовних засобів, культурних норм на сам процес вимірювання. У цьому контексті валідність розглядається як постійний діалог між теорією, даними та практикою: жоден тест не може претендувати на статус «абсолютно валідного», але будь-який тест може стати більш або менш валідним у міру накопичення й критичного переосмислення доказів. Недостатня прозорість досліджень, відсутність відкритих даних, закритість аналітичних процедур роблять цей діалог одностороннім, перетворюючи валідність на догмат. Натомість рух до відкритої науки з пререєстрацією досліджень, обов'язковою публікацією даних, кодів, матеріалів, підтримкою реплікаційних проєктів відкриває можливість для справжнього, а не декларованого, підтвердження валідності тестів.

Отже, валідність як психометрична властивість тесту є багатовимірною, динамічною та глибоко контекстуальною характеристикою, що відображає не лише коректність статистичних моделей, а й змістову адекватність, теоретичну вмотивованість, культурну чутливість і соціальну відповідальність тестового вимірювання. Вона не може бути зведена до окремого коефіцієнта

чи одного типу аналізу, а повинна розглядатися як результат тривалого процесу збирання й критичної оцінки доказів, який неможливий без високого рівня прозорості досліджень і відкритості даних. Умовою подолання як локальних проблем валідності окремих тестів, так і ширшої реплікаційної кризи в психології є формування наукової культури, у якій будь-яке твердження про валідність тесту є не кінцевим вердиктом, а запрошенням до подальшої перевірки, співпраці й відповідального використання результатів тестування.

4.6. Тестові норми в психометричному вимірюванні

Тестові норми становлять фундаментальну частину будь-якої психометричної системи, оскільки жоден тестовий бал сам по собі не має змісту, доки не буде співвіднесений із певною референтною групою. Результат, отриманий у тесті, є лише числовим відображенням рівня прояву певної психологічної характеристики, проте без належного інтерпретаційного контексту ця цифра позбавлена психометричної та практичної цінності. Саме система норм забезпечує можливість перетворення так званих сирих балів у стандартизовані показники, які дозволяють визначити, наскільки типовим або нетиповим є рівень розвитку певної властивості в конкретної особи порівняно з вибіркою, що використовується як норма. Норми створюють еталонний контекст для оцінювання індивідуальних даних, дозволяють інтерпретувати отримані результати не як ізольований феномен, а як частину ширшої картини розподілу певної ознаки у популяції. Завдяки нормам тест набуває соціального, культурного та психологічного змісту: замість абстрактних «15 балів» ми можемо стверджувати, що ця особа має рівень показника вищий, рівний або нижчий від певного відсотка людей у репрезентативній вибірці.

Формування поняття норми в психометрії має глибокі історичні корені. Ще Френсіс Гальтон, якого часто вважають одним із засновників психометрії, виходив з припущення, що багато психологічних характеристик, зокрема інтелект, мають нормальний розподіл. На його переконання, властивості людської психіки розташовуються вздовж певної континуальної шкали, де більшість людей демонструє середні значення, а крайні показники зустрічаються рідше. Цю ідею розвинув Джеймс Маккін Кеттел, який сформулював підхід до інтерпретації тестових результатів через співвідношення індивідуального балу зі статистичними параметрами вибірки – середнім арифметичним і стандартним відхиленням. На основі цих принципів виникли z-бали як стандартизований спосіб вираження відхилення індивідуального результату від середнього значення вибірки. Наприклад, якщо інтелектуальний тест має середнє значення 100 і стандартне відхилення 15, то результат 115 означає відхилення на одну стандартну одиницю вгору, що відповідає приблизно верхнім 16% популяції, тоді як результат 130 належить приблизно до верхніх 2%. Таким чином, норми забезпечили можливість математичного порівняння індивідуальної оцінки з емпіричною

вибіркою і заклали підґрунтя для подальшого розвитку стандартизованих шкал, що стали невід'ємною частиною психометричної традиції.

Однак із часом стало зрозуміло, що тестова норма – не лише статистичне поняття, а й глибоко соціально-культурний конструкт. Вона відображає не просто середній рівень у певній вибірці, а й специфічні цінності, норми, поведінкові патерни та очікування, які характерні для певної культури в певний історичний період. Наприклад, у тестах, розроблених у культурі з високим рівнем індивідуалізму, такі риси, як змагальність, самостійність і лідерська активність, можуть оцінюватися як позитивні й характерні для високого рівня функціонування, тоді як у колективістських культурах пріоритетними можуть бути інші соціальні якості – кооперативність, гармонійні міжособистісні стосунки або уникання конфліктів. Це означає, що одні й ті самі завдання, застосовані в різних культурах, можуть мати відмінне значення і формувати різні інтерпретаційні патерни. Унаслідок цього норми повинні бути не лише емпірично обґрунтованими, але й культурно адаптованими, щоб уникнути дискримінаційних або упереджених інтерпретацій. Саме тому сучасна психометрія розрізняє емпіричні норми, які ґрунтуються на реальному розподілі показників у вибірці, та соціальні норми, які відображають соціальні цінності та етичні аспекти застосування тесту.

Процес створення тестових норм є складною і багаторівневою процедурою, яка включає стандартизацію тестування, формування репрезентативної вибірки, перевірку психометричних характеристик інструменту та побудову таблиць стандартизації. Для того щоб тестові норми були валідними й мали реальну інтерпретаційну силу, вибірка для їхньої побудови повинна відображати структуру популяції, для якої призначений тест. Репрезентативність вибірки передбачає залучення осіб різного віку, статі, освіти, соціального статусу, регіону проживання, рідної мови, культурних особливостей. В умовах адаптації тестів до національного контексту важливо враховувати специфіку мовних конструкцій, соціальних норм, культурних моделей поведінки й контекстуальних факторів, які можуть істотно впливати на відповіді респондентів. Якщо тестові норми створюються відповідно до високих наукових стандартів, вони забезпечують можливість об'єктивного вимірювання, тоді як неякісні або неповні норми можуть суттєво викривити результати й сприяти появі інтерпретаційних помилок.

На цьому тлі виникає проблема непрозорості процесу нормування. У багатьох психодіагностичних інструментах норми подаються як усталені величини, проте детальна інформація про вибірку, хід стандартизації, розподіл результатів, методи очищення даних, обробку пропусків, аналіз диференціального функціонування пунктів, доступність вихідних таблиць і емпіричних масивів залишається недоступною. Така ситуація створює серйозні методологічні ризики, оскільки дослідники й практики не можуть перевірити, наскільки отримані норми відповідають сучасній популяції, чи були вони сформовані на репрезентативній вибірці, чи були враховані культурні та соціальні чинники, чи не містить тест упереджених пунктів, що

дискредитують певні групи. Відсутність відкритих даних означає, що норми стають «чорним ящиком»: вони пропонуються користувачам як завершений продукт, але їх наукова обґрунтованість залишається недоступною для критичної оцінки.

Це особливо небезпечно в умовах швидких соціальних змін, оскільки тестові норми мають властивість застарівати. Феномен, відомий як «ефект Флінна», демонструє, що середні результати в інтелектуальних тестах у багатьох країнах мають тенденцію до зростання з покоління в покоління. Це означає, що норми, створені 20 чи навіть 10 років тому, вже можуть не відповідати реальній ситуації й давати хибні інтерпретації. Соціальні трансформації, урбанізація, масова освіта, цифровізація, зміни у способах комунікації та когнітивних навичках роблять популяцію нестатичною. З огляду на це, тестові норми повинні регулярно оновлюватися, однак у багатьох випадках цього не відбувається, або відбувається без належної прозорості і методологічного обґрунтування. Часто розробники тестів оновлюють норми з посиланням на «нові вибірки», проте без публікації сирих даних, опису їх структури, соціально-демографічних характеристик або методів статистичної обробки. У таких умовах наукова спільнота не може незалежно перевірити ці норми, а практичні психологи вимушені покладатися на довіру до авторів тесту, а не на емпірично підтверджену інформацію.

Подібні проблеми поширюються також на різні типи норм. Процентильні норми, хоч і зручні для практичної інтерпретації, можуть бути чутливими до наявності у вибірці екстремальних значень; стенові норми є простішими, але можуть втрачати нюанси реального розподілу; z-бали дають точну математичну характеристику, проте вимагають високої статистичної грамотності для коректного використання. Кожен із цих форматів має свої обмеження, і тому прозорість у їхньому створенні є критично важливою. Наприклад, нерідко трапляються випадки, коли процентильні норми будуються на основі невеликої або нерепрезентативної вибірки, що призводить до викривлення інтерпретацій і появи штучних меж між «високими» й «низькими» показниками. Без можливості перевірити розподіл вибірки, застосовані алгоритми очищення, роботу з пропущеними даними або критерії виключення, незалежні дослідники не можуть оцінити, чи є такі норми науково обґрунтованими.

Особливого значення набуває питання кроскультурної валідності норм, оскільки адаптація тестів, створених у чужому культурному контексті, завжди пов'язана з ризиком втрати інтерпретаційної точності. Тест може бути валідним у своїй початковій культурі, проте його норми можуть бути абсолютно непридатними у новій. Рівень тривожності, депресії або агресивності може мати різні прояви й різне соціальне значення залежно від культурного середовища. Нерідко відмінності між групами у тестових балах пояснюють «реальними» психологічними розбіжностями, тоді як насправді вони можуть бути наслідком невдало адаптованих норм або змістово нееквівалентних пунктів. Саме тому у процесі адаптації тестів сучасні

стандарти вимагають проведення аналізу диференціального функціонування пунктів, порівняння факторної структури різних груп, виявлення культурних упереджень, а також створення локальних норм, які дозволяють точніше інтерпретувати результати у специфічних групах. Локальні норми мають особливу важливість у клінічній діагностиці, роботі з дітьми, оцінюванні представників етнічних меншин або людей з особливими потребами. Однак створення локальних норм також вимагає прозорості – детального опису вибірки, умов тестування, процедур добору учасників і всіх статистичних операцій.

Питання прозорості процесу нормування стає особливо гострим у сучасному українському контексті, де соціальні, культурні й психологічні умови зазнають постійних трансформацій під впливом війни, міграції, зміни демографічної структури, економічної нестабільності та зростання соціальної нерівності. Тестові норми, розроблені у стабільніших соціальних умовах, можуть бути непридатними для сучасного українського суспільства. Доступ до освіти, психосоціальний стрес, травматичний досвід, регіональні відмінності, зміни у когнітивних і поведінкових моделях – усе це може суттєво впливати на результати тестування. Тому створення актуальних, соціально чутливих норм є не лише психометричним завданням, а й елементом соціальної відповідальності психологів. У таких умовах тестові норми повинні бути гнучкими, адаптованими й належно документованими, а їх створення – прозорим і відкритим для наукової спільноти.

Таким чином, тестові норми є не просто технічним інструментом для стандартизації результатів, а складною, динамічною та глибоко контекстуальною системою, що визначає межі й потенціал психологічної діагностики. Вони забезпечують можливість коректної інтерпретації результатів, але водночас стають потенційним джерелом упереджень, помилок або навіть дискримінації, якщо створені без належної репрезентативності й прозорості. Умовою формування якісних норм є відкритість досліджень, доступність вихідних даних, можливість незалежної перевірки та регулярне оновлення нормативної бази відповідно до змін у суспільстві. Без цих умов тестові норми ризикують перетворитися на застарілі або хибні індикатори, що підривають наукову цінність тестування та довіру до психологічної діагностики. У цьому сенсі тестові норми стають одним із ключових полів, на яких вирішується доля сучасної психометрії – чи буде вона прозорою, науково обґрунтованою та відповідальною, чи залишиться у полоні старих підходів, що не витримують перевірки часом і реплікацією.

4.7. Етапи розробки вимірювального інструменту

Етапи розробки вимірювального інструменту в психометрії становлять складний, багаторівневий процес, у якому технічні, теоретичні, методологічні та етичні аспекти переплітаються між собою, утворюючи єдину систему, що визначає наукову якість і практичну придатність тесту. Визначення норм,

перевірка надійності та валідності, хоч і є критично важливими складовими цього процесу, не вичерпують його змісту. Повноцінне конструювання будь-якого вимірювального інструменту починається не з математичних процедур, а з чіткої постановки мети: що саме ми хочемо виміряти, яким чином це поняття визначається у рамках певної теоретичної моделі і для чого саме результати будуть використані. Психологічний тест ніколи не є нейтральним набором завдань; він завжди спирається на певну концепцію особистості, мотивації, інтелекту, емоцій чи поведінки. Якщо, наприклад, розробляється тест мотиваційної сфери, принципово важливо вирішити, якій концепції буде віддано перевагу: ієрархії потреб А. Маслоу, теорії набутих потреб Д. Мак-Клелланда, самодетермінаційній моделі Е. Дісі та Р. Раяна чи іншій теоретичній рамці. Від цього залежить добір індикаторів, формулювання пунктів, логіка структурування шкал і, зрештою, спосіб інтерпретації результатів. За відсутності чіткої теоретичної основи тест перетворюється на випадковий набір запитань, у якому не можна відрізнити суттєве від другорядного, а валідність інтерпретацій втрачається вже на початковому рівні. Саме тому перший етап розробки вимірювального інструменту є концептуальним: він вимагає аналізу психолого-педагогічної літератури, критичного огляду існуючих моделей, операціоналізації ключових понять і формулювання гіпотез щодо структури вимірюваного конструкту.

Після концептуального етапу розпочинається побудова первинного пулу пунктів – це ядро майбутнього тесту, тобто масив завдань, тверджень, запитань або ситуацій, за допомогою яких дослідник прагне «захопити» різні аспекти конструкту. На цьому етапі важливо не стільки досягти фінальної форми тесту, скільки забезпечити надмірне, багате і різноманітне представлення змісту. Формулювання пунктів повинні бути зрозумілими, однозначними, вільними від оцінних суджень і культурно упереджених висловів, які можуть створювати додаткові джерела варіативності, не пов'язані із цільовою властивістю. Створюється великий пул пунктів, який за обсягом, як правило, перевищує майбутню довжину тесту, оскільки подальші статистичні та змістові процедури неминуче приведуть до його скорочення. Водночас на цьому етапі надзвичайно важливою є прозорість: автор тесту повинен документувати джерела ідей (теоретичні моделі, клінічні спостереження, емпіричні дослідження), процес відбору формулювань, критерії виключення та включення пунктів. Відсутність такого опису у технічних звітах і публікаціях ускладнює подальшу оцінку змістової валідності та не дозволяє іншим дослідникам відтворити або переоцінити побудову шкал.

Наступним кроком є експертна оцінка змістової валідності, коли фахівці в галузі відповідного психологічного напрямку аналізують кожен пункт з точки зору його відповідності теоретичній моделі, репрезентативності щодо конструкту, відсутності двозначностей, надмірної складності або культурних викривлень. На цьому етапі часто проводяться процедури контент-аналізу, обчислюються індекси узгодженості експертів, аналізуються розбіжності у

їхніх судженнях. Експертний етап є принципово важливим, однак у традиційній психометричній практиці нерідко описується поверхово: вказується лише кількість експертів, без деталізації їхнього досвіду, критеріїв добору, точних процедур оцінювання. У контексті сучасної вимоги до відкритої науки це є суттєвим недоліком, адже без доступу до протоколів експертної оцінки, критеріїв прийняття рішень щодо редагування або вилучення пунктів неможливо відтворити початкову логіку конструювання інструменту. Таким чином, уже на ранніх етапах розробки тесту відсутність прозорості створює бар'єри для реплікації та зовнішньої перевірки.

Після експертної оцінки здійснюється первинна емпірична апробація тесту на невеликій вибірці. Цей пілотажний етап має на меті виявити технічні, змістові та психометричні недоліки пунктів. Аналізуються середні значення, дисперсії, асиметрія й ексцес розподілу відповідей, кореляції кожного пункту з загальним балом, індекси складності й дискримінативності, а також, за потреби, моделі теорії відповіді на завдання (IRT) для оцінки параметрів пунктів. Пункти, які не диференціюють респондентів, мають низьку кореляцію з сумарним балом, виявляють максимум чи мінімум в різних підгрупах (наприклад, за статтю чи віком), підлягають редагуванню або вилученню. Однак для того, щоб ці рішення були науково обґрунтованими автори тестів повинні надавати доступ до статистичних показників пунктів, описувати критерії їх відбору й відхилення, демонструвати діаграми розподілів, кореляційні матриці, результати факторного аналізу. У реальній практиці часто публікується лише кінцевий варіант тесту з узагальненими показниками надійності, тоді як проміжні етапи очищення пунктів залишаються поза увагою.

Після пілотажу та редагування пунктів настає один із ключових етапів – перевірка тесту на великій репрезентативній вибірці. Саме на цьому рівні закладаються основи для подальшого нормування, перевірки валідності та надійності, створення нормативних таблиць і шкал. За необхідності – редагування, видалення непридатних пунктів. Репрезентативність вибірки на цьому етапі має вирішальне значення: вона повинна відображати структуру тієї популяції, для якої призначений тест (за віком, статтю, освітою, соціально-економічним статусом, регіоном проживання тощо). Проводиться аналіз внутрішньої структури тесту – експлораторний та конфірмаційний факторний аналіз, моделювання латентних змінних, обчислення коефіцієнтів внутрішньої узгодженості, перевірка стабільності факторної структури у підгрупах. Оцінюється також зовнішня валідність – через кореляції з іншими підтвердженими методиками, поведінковими індикаторами, експертними оцінками, критеріями успішності чи клінічними показниками. Якщо результати виявляються задовільними, формується остаточна версія тесту, встановлюються норми, розробляється детальна інструкція для користувачів.

Результатом усіх попередніх етапів має стати підготовка технічного звіту – документа, який можна розглядати як «паспорт» вимірювального інструменту. Технічний звіт містить опис теоретичної моделі, процесу

конструювання пунктів, результатів експертної оцінки, характеристик пілотажної та нормативної вибірок, детальний аналіз надійності (внутрішня узгодженість, тест–retest, паралельні форми), результати факторних і структурних моделей, показники різних видів валідності, відомості про норми й характер їх оновлення. На перший погляд, такий звіт може сприйматися як формальна вимога або додаток до тесту, проте фактично саме він визначає наукову репутацію інструменту і можливість його справжньої наукової перевірки. У культурі відкритої науки технічний звіт не повинен бути лише внутрішнім документом видавництва або компанії-розробника; він має бути доступним для дослідників, рецензентів, фахових спільнот. Лише за цієї умови тест можна вважати повноцінним науковим продуктом, а не просто комерційним інструментом.

Сучасна тенденція у психометрії – перехід від статичних, раз і назавжди фіксованих технічних звітів до так званих «живих» звітів, які регулярно оновлюються в міру накопичення нових даних. Це особливо актуально для конструктів, чутливих до соціальних і культурних змін, таких як установки, цінності, соціальні ролі, когнітивні стилі, ставлення до праці, політики, освіти. Регулярне оновлення технічного звіту дозволяє відстежувати зміни в розподілах результатів, перевіряти стабільність факторної структури, виявляти нові джерела диференційного функціонування пунктів, своєчасно коригувати норми. У цьому сенсі технічний звіт перестає бути архівним документом і стає інструментом постійного моніторингу психометричної якості тесту. У США за протоколами АРА кожні 10 років тест повинен проходити ревізію. В Україні така культура лише формується, але дедалі більше лабораторій психологічного тестування впроваджують стандарти відкритої валідації.

Етичний аспект сьогодні є не менш важливим, ніж технічний. Психометричні інструменти мають величезний вплив на життя людей – вони визначають доступ до освіти, працевлаштування, медичні рішення. Тому тест не може бути лише «нейтральним» інструментом; він є соціальним актом, що вимагає відповідальності. Міжнародна асоціація прикладної психології (IAAP) і Європейська федерація психологічних асоціацій (EFPA) розробили етичні кодекси, які регулюють використання тестів. Вони передбачають інформовану згоду учасника, конфіденційність результатів, обмеження у використанні тестів особами без професійної підготовки.

Особливої уваги потребує етап культурної адаптації інструментів, який у багатьох країнах, зокрема в Україні, історично тривалий час здійснювався формально – шляхом перекладу зарубіжних тестів без повного психометричного відтворення всіх етапів конструювання. Сучасні стандарти вимагають не лише перекладу, а й повторного проходження усіх основних етапів – від експертної оцінки змісту до факторного аналізу на локальних вибірках, перевірки надійності й валідності, створення національних норм, аналізу диференціального функціонування пунктів. Таким чином, адаптація

стає фактично повноцінним процесом розробки тесту в новому культурному контексті.

Отже, етапи розробки вимірювального інструменту – від теоретичного обґрунтування й побудови пулу пунктів до стандартизації, нормування, підготовки технічного звіту утворюють не просто послідовність технічних кроків, а цілісну систему, в якій наукова якість нерозривно пов'язана з прозорістю, відкритістю й етичністю.

4.8. Психосемантична парадигма конструювання психодіагностичних методик

Упродовж більшої частини XX століття провідною методологічною платформою створення психодіагностичних методик залишалася психометрична парадигма, яка виходить із об'єктного уявлення про особистість: досліджуваний розглядається як точка у просторі параметрів, визначених тестом, а достовірність вимірювання забезпечується статистичними нормами, репрезентативністю вибірки та математичною обґрунтованістю шкал. Проте поступово стало очевидно, що психометричні моделі не завжди дозволяють зафіксувати індивідуальну унікальність, смислову своєрідність досвіду та суб'єктивні значення, які відіграють фундаментальну роль у формуванні поведінки та самосвідомості. Саме ці обмеження традиційної психометрії стали передумовою для виникнення психосемантичної парадигми, яка зосереджується не на зовнішніх показниках особистісних властивостей, а на внутрішній структурі індивідуального смислового простору людини.

Психосемантичний підхід постає як суб'єктна парадигма дослідження, що принципово зміщує фокус уваги з порівняння людини з іншими людьми на виявлення унікальної організації її власної системи значень. Він виходить із припущення, що кожна людина володіє специфічною когнітивно-смисловою конструкцією, за допомогою якої інтерпретує світ, будує образи соціальної реальності, ідентифікує себе та інших. Це означає, що діагностика має бути спрямована не стільки на вимірювання локальних характеристик, скільки на реконструкцію структури цих індивідуальних значень, які визначають спосіб переживання, оцінювання та поведінки. У межах психосемантики діагностичний інструмент не задає наперед універсального конструкту, а дозволяє суб'єкту самому розкрити власну систему смислових координат, які в традиційній психометрії залишаються прихованими. Таким чином, якщо психометрична модель розглядає особистість як носія певних параметрів, що підлягають стандартизованому вимірюванню, то психосемантична парадигма визначає її як простір значень, у якому індивід організовує власний життєвий досвід.

Історично психосемантика виникла в 1970-х роках, коли дедалі більше дослідників усвідомлювали, що традиційна психологічна діагностика потребує нових концептуальних рішень. Імпульсом для розвитку

психосемантики стали дві ключові тенденції. По-перше, критика об'єктної парадигми, яка зводила складну і динамічну природу психіки до статичних змінних. По-друге, розвиток когнітивної психології, яка наголошувала на тому, що поведінка визначається не лише властивостями особистості, а й ментальними репрезентаціями, інтерпретаційними схемами та символічними структурами. З одного боку, когнітивна психологія підготувала ґрунт для аналізу того, як людина структурує інформацію, а з іншого – семантичний диференціал Ч. Осгуда дав інструмент для вимірювання індивідуальних значень. Саме на перетині цих підходів і була сформована психосемантична парадигма, яка стала альтернативою традиційній тестології.

Центральним поняттям психосемантики є семантичний простір – багатовимірна координатна модель індивідуальної системи уявлень, значень і ціннісних переваг. Семантичний простір є не просто метафорою когнітивної організації, а математично формалізованою структурою, яка дозволяє аналізувати внутрішні смислові взаємозв'язки між об'єктами, явищами, людьми або власними «Я-образами» досліджуваного. На відміну від тестових норм, що дозволяють порівнювати показники між респондентами, семантичний простір відображає унікальну конфігурацію індивідуального досвіду конкретної людини. У цій моделі точками виступають не люди, а об'єкти оцінювання – поняття, події, ролі, зразки поведінки або навіть емоційні стани. Вектори простору відображають суб'єктивні відмінності між об'єктами, а його осі – узагальнені смислові підстави, які виділяє сама свідомість досліджуваного. Це означає, що семантичний простір виявляє приховану логіку індивідуального світосприйняття, даючи змогу виявити специфічну структуру значень, що визначають поведінку та емоційні реакції людини.

У класичних дослідженнях Ч. Осгуда було встановлено, що найчастіше достатньо тривимірної структури семантичного простору, оскільки більшість емотивно-смислових оцінок зводиться до факторів «Оцінка», «Сила» та «Активність». Проте сучасні психосемантичні дослідження виходять далеко за межі цієї тріади: багатовимірні моделі можуть містити десятки осей, що відображають унікальні індивідуальні конструкти. Завдяки цьому психосемантика дає змогу працювати з тією частиною психічної реальності, яка у традиційних тестах залишається невидимою – із суб'єктивною логікою смислів, цінностей, очікувань і самоозначень.

Процес побудови семантичного простору передбачає три етапи, кожен з яких має як методичний, так і теоретичний зміст. На першому етапі формується масив експериментально-психологічних даних, який відображає семантичні зв'язки між об'єктами. Цей етап є надзвичайно важливим, оскільки саме зміст і спосіб отримання базового матеріалу визначають якість подальшого моделювання. Психосемантика використовує широкий спектр технологічних прийомів: асоціативний експеримент, суб'єктивне шкалювання, ранжування, метод класифікації (сортування), семантичний диференціал, репертуарні решітки та інші інструменти. Кожна з цих процедур

має власну логіку впливу на свідомість досліджуваного, власний спосіб виявлення індивідуальних конструктів і власні можливості для аналітичної інтерпретації. Так, асоціативний експеримент дозволяє реконструювати глибинний семантичний склад слова, тоді як ранжування виявляє структуру переваг. Суб'єктивне шкалювання виявляє ступінь близькості об'єктів за певними параметрами, тоді як сортування дозволяє визначити принципи категоризації.

Другий етап – виявлення структури – є математичною фазою психосемантичного аналізу. На цьому етапі початкові дані піддаються обробці методами багатовимірного статистичного аналізу: факторного аналізу, кластеризації, багатовимірного шкалювання тощо. Метою цього етапу є не створення нових даних, а виявлення латентних структур, що об'єднують окремі елементи у значущі цілісності. Психосемантична матриця схожості, яка формується на першому етапі, перетворюється на багатовимірну модель, що відображає внутрішні координати індивідуального смислового поля. Важливо підкреслити, що на відміну від класичної психометрії, де фактори інтерпретуються як стандартизовані психологічні властивості, у психосемантиці фактори є результатом вибору самого суб'єкта: вони фіксують ті смислові підстави, якими він насправді користується під час інтерпретації реальності. Таким чином, структура не задається ззовні, а витягується зі свідомості досліджуваного, що робить психосемантичний аналіз найбільш суб'єктно-чутливим інструментом сучасної психодіагностики.

Третій етап – інтерпретація – є найбільш складним. Він полягає у визначенні змісту виділених факторів, кластерів або координат індивідуального семантичного простору. Цей процес не може бути механічним, оскільки смислова структура завжди є феноменологічно насиченою. Інтерпретація вимагає високої кваліфікації дослідника, розуміння контексту, індивідуальної історії та особливостей особистості. Часто використовують груповий метод незалежних суддів або метод еталонних об'єктів, які дають змогу підвищити валідність інтерпретації. У разі психосемантичних досліджень сам процес інтерпретації нерідко стає одночасно і діагностичним, і терапевтичним, оскільки допомагає людині усвідомити структуру власної ціннісно-смислової системи.

Психосемантична парадигма має низку суттєвих переваг. По-перше, вона дозволяє уникнути редукціонізму психометричного тестування, розглядаючи особистість не як носія певних властивостей, а як активного творця індивідуального світу значень. По-друге, психосемантика працює з суб'єктивними смислами, які є значно ближчими до реального досвіду людини, ніж абстрактні психометричні показники. По-третє, психосемантичні методи є високочутливими до унікальних особливостей особистості, що робить їх незамінними у клінічній психології, консультуванні та психотерапії. По-четверте, психосемантика дозволяє моделювати індивідуальні концептуальні системи, що є ключовим для розуміння ідентичності, мотивації,

міжособистісних стосунків, професійного самовизначення чи ціннісних конфліктів.

Водночас психосемантична парадигма має й певні обмеження. Суб'єктність даних ускладнює стандартизацію, а інтерпретація результатів потребує високої компетентності. Крім того, психосемантичні методики зазвичай мають низьку групову порівнювальність, що обмежує їхнє застосування у великих вибірках та стандартизованих масових дослідженнях. Проте ці обмеження є природною платою за можливість працювати з індивідуальною структурою значень, унікальною для кожної людини.

Серед технологічних прийомів психосемантики чільне місце посідає суб'єктивне шкалювання, яке дозволяє оцінювати схожість об'єктів та виявляти приховані смислові координати. Ця техніка використовується у багатьох відомих методиках, таких як шкали Дембо-Рубінштейн або методи суб'єктивного семантичного профілю. Ранжування є іншим способом виявлення внутрішньої структури переваг, а парне порівняння дозволяє реконструювати смислові суперечності. Метод сортування, у свою чергу, показує логіку категоризації, тобто спосіб, у який людина структурує різноманіття навколишнього світу. Асоціативний експеримент дозволяє моделювати значення через аналіз асоціативних мереж, а семантичний диференціал – є виміром особистісних смислів, емоційних реакцій і ціннісних ставлень. Репертуарні решітки, розроблені Дж. Келлі, дозволяють відтворювати систему особистісних конструктів, через які людина інтерпретує себе та інших.

Кожен із цих методів має свою діагностичну специфіку, але всі вони об'єднані ключовим принципом: психологічні явища вивчаються не як об'єкти зовнішнього спостереження, а як індивідуальні смислові феномени. Завдяки цьому психосемантика дозволяє виявляти те, що неможливо виміряти стандартним тестом: спосіб, у який людина конструює світ.

Сучасні тенденції розвитку психосемантики свідчать про її активну інтеграцію з нейропсихологічними, когнітивними та соціокультурними науками. Сьогодні психосемантичні моделі використовуються для вивчення політичних орієнтацій, етнічної ідентичності, професійної мотивації, соціальних установок, образу «Я», ціннісних конфліктів, емоційних переживань та ін. Психосемантика активно застосовується в консультуванні: вона дозволяє реконструювати внутрішню картину світу клієнта, виявити смислові суперечності, визначити джерела внутрішньої напруги та психологічної дезадаптації. У клінічній психології психосемантичні методи допомагають зрозуміти індивідуальні механізми переживання хвороби, страхів, депресивних схем або травматичного досвіду. У дослідницькій практиці психосемантика дозволяє моделювати зміни ціннісних систем, формування групових норм або динаміку соціальних уявлень у кризових ситуаціях.

Отже, психосемантична парадигма конструювання психодіагностичних методик є однією з найперспективніших у сучасній психології. Вона дає

інструменти для роботи з суб'єктивними смислами, які визначають внутрішню організацію особистості, дозволяє виявити індивідуальні структури значень і реконструювати спосіб, у який людина «бачить» світ. На відміну від традиційних психометричних методів, психосемантика не прагне до стандартизації людського досвіду, а навпаки – прагне зрозуміти його унікальність. Це робить її не лише науковою методологією, а й гуманістичним шляхом пізнання особистості, у центрі якого знаходиться не абстрактний показник, а жива, мисляча і відчуюча людина, що творить власну систему смислів.

4.9. Сучасні тенденції та виклики в психометрії

Сучасна психометрія поступово переходить від тестів-зрізів до динамічного вимірювання станів. Методи моментальної екологічної оцінки (Ecological Momentary Assessment – ЕМА) дозволяють фіксувати настрій, рівень тривоги чи стресу кілька разів на день через мобільні додатки. Такі дані створюють часові ряди, з яких будуються індивідуальні моделі емоційної регуляції. Це відкриває можливість точнішої діагностики й інтервенцій у режимі реального часу. Наприклад, у проектах з дослідження посттравматичного стресу використовують смартфони, що фіксують фізіологічні сигнали (пульс, частоту дихання) і співвідносять їх із самооцінками стану. Так психометрія переходить у фазу «психофізіологічного моніторингу», де межа між психологічним і біологічним вимірюванням стирається.

Одним із сучасних викликів психометрії є вимірювання складних, динамічних властивостей – емоційного інтелекту, резильєнтності, креативності тощо. Ці конструкти не мають стабільних поведінкових індикаторів, а тому важко піддаються стандартизації. Нові методи, як-от динамічне тестування (dynamic assessment) або гібридні підходи, що поєднують кількісні та якісні дані, дозволяють краще фіксувати зміну стану в часі. Такі підходи використовують у клінічній практиці, де важливо не лише оцінити рівень тривоги чи депресії, а й відстежити процес одужання.

Психометрична парадигма сьогодні трансформується у більш гнучку і контекстуальну систему. Вона визнає, що вимірювання не може бути абсолютно об'єктивним, оскільки психологічні явища залежать від культури, мови, часу. Проте це не заперечує науковість психології, а навпаки, робить її глибшою.

Усвідомлення неможливості повної об'єктивності призвело до переосмислення поняття валідності. Якщо у класичній психометрії валідність ототожнювали із статистичною точністю прогнозу, то сучасна наука розглядає її як єдність змістової, структурної та соціальної узгодженості. С. Мессік у своїй інтегративній моделі довів, що жоден тип валідності не може існувати ізольовано, оскільки будь-який тест є водночас і науковим інструментом, і соціальним артефактом. Наприклад, тест на професійну придатність може бути статистично надійним, але невалідним етично, якщо він систематично

занижує шанси певних груп. Таким чином, валідність має не лише емпіричний, а й моральний вимір.

Змістова валідність передбачає логічну відповідність пунктів тесту сутності вимірюваного конструкта. Для її перевірки залучають експертів, які оцінюють, наскільки кожен пункт відображає теоретичний зміст. У цьому сенсі психометрія не є лише статистикою, а способом формалізації змісту психологічної теорії. Критеріальна валідність визначається шляхом співвідношення з зовнішнім критерієм – наприклад, оцінками в навчанні, результатами роботи або клінічними діагнозами. Конструктна валідність є більш складною: вона підтверджується за допомогою факторного аналізу, перевірки гіпотез про взаємозв'язки змінних і диференціальної чутливості тесту до окремих характеристик. Саме цей тип валідності визначає, чи є тест справжнім виміром певного психологічного явища, а не поверхневим набором запитань.

У ХХІ столітті до цієї тріади валідності додають ще один вимір – екологічну валідність, тобто відповідність ситуації тестування реальним умовам життя. Зростання інтересу до поведінкової екології та нейроетології показало, що результати, отримані в лабораторії, не завжди відтворюються у повсякденному середовищі. Тому сучасні дослідження активно розробляють польові дизайни та мобільні платформи, що дозволяють фіксувати дані у природному контексті. Наприклад, застосунки типу Experience Sampling Method (ESM) або Ecological Momentary Assessment (EMA) дають змогу відстежувати емоційні реакції в реальному часі. Це наближає психометрію до того, що можна назвати «динамічною валідністю» – здатністю інструмента залишатися релевантним у змінних умовах.

Об'єктивність у сучасному сенсі означає не усунення суб'єктивного, а контроль і прозорість процедур. Саме тому у психометрії дедалі частіше використовують відкриті бази даних, публічні звіти та алгоритми відтворення результатів.

Розвиток відкритої науки змінив спосіб перевірки психометричних моделей. Традиційно тести розглядалися як «чорні скриньки» – користувач бачив лише інструкцію та ключ, тоді як алгоритм оцінювання залишався комерційною таємницею. Сьогодні дедалі більше вчених виступають за відкритість усіх параметрів – від формули розрахунку балів до вихідних даних. Це не лише підвищує довіру, а й створює можливість реплікації результатів. У сучасних базах, як-от OpenPsychometrics, публікуються не тільки тексти тестів, а й сирі відповіді тисяч респондентів, що дозволяє іншим дослідникам перевіряти факторні структури, проводити метааналізи, оновлювати норми.

У психометрії з'явилися навіть окремі проєкти, наприклад, Open Test Initiative, які просувають принцип «прозорих тестів». Його суть полягає в тому, що інструмент, яким вимірюють людину, не може бути недоступним для самої людини. Відкритість тестових моделей сприяє формуванню критичного мислення у користувачів і зменшує ризик маніпуляцій. Адже, як показує історія, найгірші наслідки тестування виникали тоді, коли його процедури

були прихованими: расові тести початку XX століття чи псевдонаукові оцінки інтелекту, що використовувалися для дискримінації.

Сучасні методи статистичного моделювання розширили арсенал психометрії далеко за межі класичної теорії тестів. Наприклад, теорія IRT дозволяє оцінювати не лише рівень здібностей респондента, а й властивості самого пункту – його складність, дискримінативність, ймовірність випадкової відповіді. Моделі MIRT (Multidimensional IRT) розглядають одразу кілька латентних факторів, що дозволяє створювати багатовимірні профілі особистості. Генералізована теорія надійності (G-Theory), запропонована Кронбахом, вийшла за межі простої оцінки похибки, розкривши її структуру: похибка може походити від особливостей спостерігача, контексту завдання, взаємодії між оцінювачами. Це дало можливість значно точніше визначати джерела варіативності результатів.

Нові методи не скасовують класичну психометрію, а розширюють її можливості. Завдяки поєднанню структурного моделювання (SEM) та факторного аналізу з'явилася можливість оцінювати, наскільки добре дані відповідають теоретичній моделі. Таким чином, валідність тесту тепер перевіряється не лише через кореляції, а через відповідність емпіричної структури передбачень. Це підвищує вимоги до якості тестових даних, але водночас робить психологію більш точною наукою.

Окремий напрям розвитку – застосування машинного навчання у психометрії. Алгоритми здатні виявляти складні закономірності у великих масивах даних, що перевищують можливості традиційних методів. З їх допомогою можна прогнозувати поведінкові патерни, класифікувати респондентів, автоматично оцінювати відповіді на відкриті питання. Проте тут знову постає етична дилема: алгоритми навчаються на даних, що можуть містити соціальні упередження, і відтворюють їх у прогнозах. Тому нова ера психометрії вимагає не лише технічної досконалості, а й моральної відповідальності.

Сучасна парадигма психометрії поєднує кілька принципів: теоретичну обґрунтованість, статистичну перевірку, культурну чутливість і етичну прозорість. Вона відмовляється від ілюзії універсальності, натомість прагне гнучкості та відкритості. У цьому сенсі психометрія перестає бути лише «технологією тестів» і перетворюється на науку про якість психологічного знання.

Водночас дедалі більше дослідників звертаються до філософських основ вимірювання. Постнекласична наука розглядає будь-який вимір як інтерпретацію, а не як просте відображення реальності. З цього погляду тест не є дзеркалом психіки, а радше моделлю, яка створює свій власний простір смислів. Психометричні показники стають мовою, у якій описується людина, і тому вони формують те, як суспільство її бачить. Саме тому психометрію називають не лише технічною, а й дискурсивною практикою.

Деякі теоретики, як-от Пол Клайн і Джон Равен, наголошували, що тести формують культурні очікування. Наприклад, поняття «інтелекту» у тестах

Векслера відображає західну традицію раціональності, тоді як у східних культурах більш важливою є мудрість, соціальна гармонія, інтуїція. Це означає, що тест не просто вимірює, а й задає нормативний стандарт. Тому новітні підходи вимагають, щоб конструювання тестів супроводжувалося міжкультурним аналізом – наскільки конструкт, який ми вимірюємо, має універсальний зміст.

Філософський зсув у психометрії полягає також у переході від лінійного мислення до системного. Сьогодні психологічні властивості розглядають як мережеві структури, де симптоми або риси пов'язані між собою не через єдиний латентний фактор, а через взаємні впливи. Наприклад, у мережевих моделях депресії не існує «центрального» симптому – безсоння, апатія чи відчуття провини взаємно підтримують одне одного, утворюючи систему. Цей підхід змінює саму логіку тестування: замість пошуку одного показника дослідники аналізують топологію зв'язків між пунктами, що дозволяє бачити психічну структуру як живу динамічну систему.

Мережевий підхід також створює нові виклики для статистики. Якщо класична психометрія працює з латентними змінними, то мережеві моделі розглядають симптоми або риси як самостійні елементи, що взаємно впливають. Для цього використовують складні методи графового аналізу – від оцінки центральності вузлів до моделювання зважених зв'язків. Це відкриває шлях до персоналізованої психодіагностики: у кожної людини може бути власна структура взаємодії між елементами особистості.

Цей підхід має спільні риси з теорією хаосу та синергетикою – галузями, що вивчають самоорганізацію у нелінійних середовищах. Людська психіка, з цього погляду, не є сукупністю незалежних рис, а радше системою із взаємозалежних елементів, що постійно змінюють один одного. Тому будь-яке вимірювання є лише «знімком» певного стану системи в конкретний момент часу.

Таке розуміння призводить до появи динамічної психометрії – напрямку, який намагається поєднати кількісні показники із часовими характеристиками змін. Наприклад, дослідження емоційних коливань за допомогою щоденникових або мобільних методів дозволяє не просто виміряти рівень емоцій, а й побачити закономірності їхньої еволюції протягом дня чи тижня. Це відкриває нові горизонти для клінічної практики, оскільки дозволяє виявляти не лише симптоми, а й тригери – події чи ситуації, що запускають певні психологічні реакції.

Динамічний підхід також пов'язаний із розвитком часових моделей у психометрії. Традиційно тести розглядалися як позачасові: передбачалося, що властивість особистості є стабільною. Проте сучасні дані показують, що навіть такі риси, як екстраверсія чи нейротизм, можуть змінюватися залежно від життєвого контексту, соціальних ролей і стану здоров'я. Це означає, що тест має враховувати часову мінливість, а не ігнорувати її.

У цьому напрямі розвиваються моделі латентного зростання (Latent Growth Models), що дозволяють відстежувати зміни психологічних показників

у часі та порівнювати динаміку між групами. Такі моделі дедалі частіше використовують у дослідженнях розвитку особистості, емоційного вигорання та ефективності психотерапії.

Нарешті, сучасна психометрія виходить за межі індивідуального вимірювання. Вона дедалі частіше використовується у великих соціальних дослідженнях – від оцінювання освітніх систем (PISA, TIMSS) до аналізу колективних настроїв у соціальних мережах. Алгоритми обробки мовлення, емоційного тону текстів, патернів взаємодії стають новими формами тестів – неявних, але потужних. Це породжує питання: чи залишається людина суб'єктом вимірювання, якщо її поведінку оцінює машина? І хто відповідає за інтерпретацію таких результатів?

У цьому питанні концентрується головна дилема цифрової епохи – взаємовідносини між автономією людини та алгоритмічним контролем. Автоматизовані системи тестування здатні аналізувати величезні масиви даних і видавати прогнози з вражаючою точністю, але вони не мають морального судження. Алгоритм може ідеально відтворити кореляції, але не розуміє, що таке справедливість. Тому психометрія XXI століття потребує не лише фахівців з даних, а й фахівців з етики.

Стає очевидним, що тестування майбутнього повинно бути «людиноцентричним» – орієнтованим не лише на ефективність, а й на збереження гідності особистості. У цьому контексті розвивається концепція «етичної психометрії», яка пропонує оцінювати не лише достовірність тесту, а й наслідки його застосування. Зокрема, дослідники пропонують вводити індекс соціальної відповідальності тесту – показник, що відображає його вплив на різні групи населення, прозорість використання та можливість апеляції результатів. На відміну від минулих десятиліть, коли тести вважалися нейтральними інструментами, сьогодні дедалі більше уваги приділяють тому, як результати впливають на суспільство. У науковій літературі з'являється поняття «testing impact» – оцінювання наслідків застосування тесту для індивіда, групи й культури. Наприклад, широке використання тестів здібностей у рекрутингу може створювати невидимі бар'єри для кандидатів з низьким соціально-економічним статусом, а стандартизовані освітні оцінювання – сприяти відтворенню нерівності між школами.

Сучасна етична психометрія відповідає на ці виклики через розробку соціально-чутливих тестів. Такі методики створюються у співпраці з представниками різних спільнот, включають пункти, що враховують культурний контекст, і проходять незалежний аудит на справедливість. У деяких країнах (Канада, Нідерланди) вже існують національні етичні комісії, які затверджують нові психометричні інструменти перед їх масовим використанням.

Психометрична парадигма конструювання тестів пройшла тривалий шлях від механічного вимірювання до складної системи наукових і етичних принципів. Вона стала одним із головних інструментів, які забезпечують об'єктивність психологічної науки, проте сама ця об'єктивність із часом

змінити зміст. Сьогодні психометрія розглядає людину не як носія стабільних властивостей, а як динамічну систему, у якій поєднуються біологічні, когнітивні й соціальні фактори. Її мета – не лише фіксувати відмінності, а й розуміти механізми, що стоять за ними.

Сучасні тенденції демонструють перехід від кількісної до інтегративної логіки вимірювання, де статистика поєднується зі змістовним аналізом, а точність – із гуманістичною чутливістю. Психометрія дедалі більше визнає контекст, у якому створюються тести, їхній вплив на суспільство та відповідальність дослідника. Вона стає не просто технічною дисципліною, а культурною практикою, що формує уявлення про людину, її здібності й межі.

Тест, створений за принципами психометричної парадигми, – це не лише інструмент вимірювання, а й дзеркало наукового мислення своєї епохи. І від того, наскільки чесно й усвідомлено ми користуємося цим дзеркалом, залежить не лише точність наших даних, а й гуманність психології як науки про людину.

Майбутнє психометрії визначатиметься балансом між технологічним прогресом і гуманістичними цінностями. Ймовірно, вже найближчими десятиліттями тести стануть частиною інтегрованих цифрових систем, де штучний інтелект автоматично аналізуватиме поведінкові й мовні патерни людини у повсякденному житті. Такі системи потенційно можуть допомогти у ранній діагностиці психічних розладів, оцінці ризику вигорання чи навіть у профілактиці самогубств. Але водночас вони створюють ризики тотального контролю, втрати приватності, комерціалізації особистісних даних.

Тому головне завдання психології – не лише вдосконалювати методи вимірювання, а й формувати етичні рамки їхнього використання. У цьому контексті особливої ваги набуває поняття психометричної грамотності – здатності критично розуміти тести, їхні обмеження, контекст створення та наслідки інтерпретації. Психометрична освіта має стати частиною професійної культури кожного психолога, адже саме вона забезпечує довіру до науки у суспільстві.

Формування психометричної культури передбачає не лише знання формул і моделей, а й розуміння історичного контексту науки. Психометрія народилася у добу модерну, коли людство вперше повірило, що психіку можна виміряти так само точно, як довжину чи температуру. Сьогодні ж ми живемо у світі, де межа між людиною і даними стає розмитою, а отже, психологові необхідно не лише володіти інструментом, а й розуміти його силу.

Університетські навчальні програми мають включати курси з критичного аналізу тестів, де здобувачі освіти вчаться не просто користуватися готовими шкалами, а перевіряти їхню валідність, розуміти джерела похибок, етичні ризики. Це дозволить створити покоління фахівців, які зможуть не лише оцінювати людину, а й захищати її від помилок вимірювання. У кінцевому підсумку саме це визначить, чи залишиться психологія гуманною наукою у добу алгоритмів.

Завдання на самостійну підготовку

1. У чому полягає принципова складність вимірювання психічних явищ порівняно з природничими науками?
2. Що таке психологічні конструкти та чому вони не можуть бути виміряні безпосередньо?
3. У чому полягає значення репрезентативної теорії вимірювання Стенлі Стівенса для розвитку шкал вимірювання?
4. Які переваги та недоліки мають текстові, числові та графічні шкали відповідей?
5. У чому полягають принципові відмінності між шкалою Терстоуна, Гуттмана та Лікерта?
6. У чому полягає сутність психометричної надійності тесту та чому вона є критичною для практичної психодіагностики?
7. Яку роль відіграє коефіцієнт Кронбаха (α) у визначенні внутрішньої узгодженості тесту, і чому надто високі значення α можуть бути проблематичними?
8. У чому полягає логіка оцінювання надійності за допомогою паралельних форм тесту?
9. Які стратегії використовуються для підвищення психометричної надійності тесту?
10. Як змістова, критеріальна та конструктивна валідність взаємодоповнюють одна одну у класичній психометрії?
11. Чому сучасний підхід розглядає валідність не як властивість тесту, а як властивість інтерпретацій тестових балів?
12. У чому полягає специфіка кроскультурної валідності та чому її забезпечення є етично важливим?
13. Чому сирий тестовий бал не має значення без системи норм і яку функцію виконують стандартизовані шкали?
14. Чим відрізняються емпіричні та соціальні норми, і чому обидва типи є важливими для інтерпретації результатів?
15. Які методологічні ризики виникають при відсутності прозорості інформації про вибірку та процедури нормування?
16. Чому тестові норми повинні регулярно оновлюватися та як ефект Флінна впливає на їхню актуальність?
17. У чому полягають переваги та обмеження різних типів норм (z-бали, проценти, стени)?
18. Як культурні та соціальні чинники можуть спотворювати інтерпретацію тестових балів без належної адаптації норм?
19. Які ключові етапи включає процес розробки психологічного тесту – від теоретичного обґрунтування до створення технічного звіту?
20. У чому полягає принципова відмінність психосемантичної парадигми від традиційної психометричної, і як суб'єктний підхід змінює логіку психологічної діагностики?

РОЗДІЛ 5

НЕЕКСПЕРИМЕНТАЛЬНІ МЕТОДИ ДОСЛІДЖЕННЯ ПСИХІКИ

5.1. Основні методи збору даних у психологічних дослідженнях

Психологія як наука про психічну реальність має особливий статус серед інших дисциплін: її об'єкт не піддається безпосередньому вимірюванню, а проявляється через поведінку, мовлення, емоційні реакції та продукти діяльності. Саме тому методи дослідження в психології не можуть бути зведені лише до експерименту. Неекспериментальні методи становлять рівноправну, а в деяких галузях навіть провідну частину методологічного арсеналу науки. Вони дозволяють аналізувати психічні явища в їхньому природному контексті, зберігаючи автентичність досвіду людини.

Поняття «неекспериментальний метод» не означає «ненауковий» або «менш достовірний». Навпаки – це підхід, який спирається на інші принципи науковості. Якщо експеримент прагне до контролю змінних і причинно-наслідкових зв'язків, то неекспериментальне дослідження прагне до повноти опису, екологічної валідності та інтерпретаційної глибини. У термінах сучасної методології йдеться не лише про спостереження фактів, а про реконструкцію смислів, які стоять за ними.

Історично неекспериментальні методи формувалися одночасно з експериментальними. На початку ХХ століття спостереження, інтерв'ю, анкетування стали базовими інструментами емпіричної психології. Саме на їх основі будувалися перші соціометричні методики, опитувальники особистості (Айзенк, Кеттелл), педагогічні й клінічні спостереження. Згодом ці методи отримали систематичне обґрунтування у працях К. Левіна, Г. Олпорта, А. Валлона, де неекспериментальне дослідження розглядалося як спосіб фіксації реальної поведінки у її природних зв'язках.

Неекспериментальні методи ґрунтуються на принципі спостереження психічного через зовнішні прояви поведінки, мови, емоційної експресії, продуктів діяльності або текстів. Вони передбачають систематичний, але не маніпулятивний підхід – дослідник фіксує вже наявні феномени, не змінюючи середовище. Це наближає неекспериментальні методи до гуманітарних наук, зокрема антропології, соціології, культурології, де важливо зрозуміти суб'єктивний досвід, а не лише його кількісні параметри.

Як зазначає С. Д. Максименко, «психіка не може бути вивчена лише як об'єкт – вона є одночасно суб'єктом власного прояву». Тому будь-яке спостереження або опитування у психології має не лише фіксувати зовнішні ознаки, але й урахувати внутрішній світ особистості – її мотивацію, установки, культурні коди, досвід.

У системі наукового пізнання неекспериментальні методи посідають місце між описовими і пояснювальними підходами. Їхнє завдання полягає не лише в реєстрації фактів, а у виявленні закономірностей психічного

функціонування в реальних, природних умовах. На відміну від експерименту, що спирається на контроль змінних, неекспериментальний підхід прагне до екологічної валідності – тобто відповідності результатів реальній поведінці людини у життєвих ситуаціях.

Основою неекспериментальних методів є ідея, що психічне не можна «виміряти» безпосередньо, а лише опосередковано через його прояви. Цей підхід спирається на феноменологічну і герменевтичну традиції. Феноменологічна установка передбачає опис досвіду «таким, як він є», без редукції до змінних, а герменевтична – тлумачення цього досвіду в контексті культури, мови, міжособистісних взаємодій.

У сучасній науковій методології неекспериментальні методи входять до складу якісних і квазікількісних дослідницьких стратегій. Вони забезпечують доступ до суб'єктивного змісту досвіду – мотивацій, цінностей, уявлень, що не можуть бути зведені до числових показників. Їх мета – не статистичне узагальнення, а аналітична репрезентативність, тобто розкриття типових механізмів у межах конкретного випадку.

Важливою рисою неекспериментальних методів є те, що вони нерідко комбінуються між собою або інтегруються в більш складні програми змішаного типу. Наприклад, дослідник може поєднувати аналіз текстових матеріалів із напівструктурованим інтерв'ю, або використовувати анкетування разом із довготривалим спостереженням. У такому разі йдеться вже не про ізольований метод, а про неекспериментальний дизайн дослідження, в межах якого кожен з елементів підпорядкований загальній логіці перевірки гіпотези.

Слід підкреслити, що успішність неекспериментального підходу залежить від точності інструментарію, етичної чутливості дослідника, його спостережливості й уміння інтерпретувати складні, контекстуально насичені дані. Саме тому неекспериментальні методи часто пов'язуються з високими вимогами до професійної культури дослідника: недостатньо лише зафіксувати поведінку чи відповідь – потрібно зрозуміти, що вона означає, у якому контексті вона з'явилась, які змістові навантаження несе.

Неекспериментальні методи – це не просто альтернативні техніки збору даних, а повноцінна стратегія дослідницької роботи, що забезпечує глибоке, природне й контекстуально точне вивчення психіки людини. Її ефективність залежить від чіткого методологічного бачення, коректного використання інструментів і здатності дослідника мислити в межах якісної та комплексної моделі реальності.

Неекспериментальні методи охоплюють широкий спектр технік збору даних, серед яких основними є:

- спостереження,
- анкетування,
- інтерв'ю (бесіда),
- архівний метод,
- контент-аналіз як допоміжна аналітична процедура.

Кожен із них має власні цілі, процедури, етичні вимоги та можливості для узагальнення результатів. Неекспериментальні методи є різноманітними за своєю природою, функціями, джерелами інформації, типами змінних і контекстом використання. Відтак, виникає необхідність у систематизації цього різноманітного масиву – тобто в класифікації, що дозволить зрозуміти, за якою логікою і з якими цілями ці методи застосовуються в психології.

Ключовим критерієм такої класифікації виступає форма взаємодії з джерелом психологічної інформації. За цим принципом неекспериментальні методи можна умовно поділити на три великі групи:

- 1) методи прямого спостереження;
- 2) методи вербального опитування;
- 3) методи аналізу вторинних даних.

До першої групи належать усі форми фіксації поведінки без її безпосереднього стимулювання: спостереження в різних формах – від вільного, наративного до строго структурованого, з використанням технічних засобів. Цей метод є базовим для психології, адже саме через спостереження дослідник отримує первинне уявлення про прояви психіки в дії. Особливо важливо те, що спостереження дозволяє фіксувати не лише усвідомлену, а й автоматичну поведінку – жести, міміку, рухову активність, інтонаційні особливості, емоційну динаміку. Саме тому цей метод широко застосовується в дитячій, клінічній, соціальній, педагогічній психології. Залежно від контексту та рівня формалізації спостереження може бути включеним (тобто таким, де дослідник є учасником подій) або невключеним (тобто зовнішнім щодо досліджуваної ситуації), відкритим або прихованим, структурованим або неструктурованим, тривалим або епізодичним. Відповідно до цього змінюються і способи фіксації даних: від вільного опису до використання чек-листів, таблиць, аудіо– чи відеозапису. Спостереження особливо цінне у тих випадках, коли поведінка не може бути відтворена за вимогою або ж може змінитися під впливом експериментатора.

До другої групи належать методи вербального опитування – тобто такі, у яких дослідник отримує психологічну інформацію безпосередньо від суб'єкта через запитання й відповіді. Сюди належать анкетування, тестування, інтерв'ю. Кожен із цих методів має власні можливості і обмеження. Анкетування дозволяє отримати великі масиви даних за короткий час, особливо у соціально-психологічних та демографічних дослідженнях. Воно є зручним для аналізу установок, оцінок, переконань і типових способів реагування. Проте анкетування надзвичайно чутливе до форми запитань, послідовності їх подачі, формулювань, варіантів відповідей. Брак контролю за ситуацією відповіді, неможливість уточнення або перевірки щирості роблять цей метод вразливим щодо викривлення результатів. Інтерв'ю, у свою чергу, дозволяє глибше зануритися у внутрішній світ особистості. У стандартизованих варіантах воно забезпечує порівнюваність даних, а у вільній або клінічній формі – створює простір для виявлення унікальних інтерпретацій, глибинних переживань, особистісних сенсів. Якість інтерв'ю

значною мірою залежить від рівня підготовки дослідника, його комунікативної компетентності, вміння слухати, підтримувати діалог і не нав'язувати очікувану відповідь. Усі методи вербального опитування мають спільну рису – вони ґрунтуються на самозвіті, а тому завжди містять певний суб'єктивний компонент: у сприйнятті запитання, у виборі формулювання, у представленні себе перед дослідником.

Третю групу становлять методи аналізу вже існуючих джерел інформації, які не були створені для наукового дослідження, але можуть містити цінні психологічні дані. Ідеться про так звані архівні або документальні методи, що охоплюють аналіз текстів, записів, щоденників, листування, медичних карт, судових протоколів, соціальних мереж, художніх творів тощо. Такі джерела часто є єдино можливими у вивченні минулих подій, біографій, соціокультурних змін. Їх аналіз дозволяє реконструювати психологічні характеристики особистості або групи на основі змісту, стилю, структури тексту, мовленнєвих маркерів, частотності тем. Цінність цього методу полягає у його неінвазивності – дослідник не втручається в життя суб'єкта, не впливає на його відповіді. Проте складність методу полягає в неоднозначності інтерпретацій, складності перевірки гіпотез і труднощах з переведенням якісних даних у кількісну форму.

Класифікація неекспериментальних методів у психології дозволяє впорядкувати різномірний інструментарій пізнання психіки, виходячи з типу джерела, способу доступу до інформації та ролі дослідника у ситуації. Ці методи не є заміною експерименту, але вони забезпечують доступ до психологічної реальності у її природному, нефальсифікованому вигляді. Уміння правильно обрати метод, усвідомлювати його обмеження і переваги, коректно інтерпретувати результати – ось що визначає рівень професійної зрілості дослідника у неекспериментальній парадигмі.

Головною перевагою неекспериментальних методів є їхня екологічна валідність – здатність відображати реальну поведінку людей без спотворення контексту. Такі дослідження часто проводяться у природних умовах: у школах, на виробництві, у медичних закладах, у соціальних групах. Це дозволяє врахувати комплекс соціальних, культурних і ситуаційних чинників, що формують психіку.

Крім того, неекспериментальні методи є етично безпечнішими, оскільки не передбачають маніпуляцій, що можуть вплинути на стан учасників. Вони особливо корисні для дослідження вразливих груп – дітей, осіб із ПТСР, пацієнтів із психічними розладами, де експеримент може бути неприйнятним.

Ще однією перевагою є можливість ретроспективного аналізу. Наприклад, архівні методи дозволяють досліджувати події минулого, реконструювати психологічні риси особистості на основі її текстів, щоденників, висловлювань. Такі підходи мають особливе значення для історичної психології та психології культури.

Разом із перевагами неекспериментальні методи мають певні обмеження. Найважливішим із них є суб'єктивність інтерпретації. Оскільки

дослідник не контролює змінні, велика частина аналітичної роботи полягає у тлумаченні. Це створює ризик упередженості, особливо при якісному аналізі текстів або поведінки.

Другою проблемою є складність верифікації результатів. Повторити неекспериментальне дослідження з тими самими умовами майже неможливо, адже соціальний контекст постійно змінюється. Тому вимога до детального опису процедури (методологічна прозорість) є обов'язковою умовою наукової надійності.

Третя проблема – репрезентативність вибірки. У спостереженнях або інтерв'ю часто беруть участь обмежені групи, що не завжди відображають ширшу популяцію. Відтак узагальнення результатів повинне бути обережним і чітко обґрунтованим.

5.2. Метод спостереження

Метод спостереження посідає одне з найважливіших місць у психологічній науці, адже саме через нього здійснюється безпосереднє пізнання поведінки, емоційних реакцій і соціальної взаємодії людини в природному середовищі. На відміну від експерименту, який передбачає активне втручання дослідника у хід подій, спостереження є більш «делікатним» інструментом, що націлений на вивчення психіки у її природному вияві. У науковому сенсі воно визначається як систематичне, цілеспрямоване, організоване сприймання дослідником дій, висловлювань, невербальних сигналів і ситуацій взаємодії з метою отримання достовірних даних про внутрішні психічні процеси людини. Таке розуміння спирається на класичні підходи Б.Г. Ананьєва, який підкреслював, що спостереження – це форма непрямого пізнання психічного, де зовнішні прояви поведінки є лише засобом доступу до внутрішнього світу особистості.

У найширшому значенні спостереження в психології – це метод цілеспрямованого, систематичного і контрольованого сприймання та фіксації поведінкових актів, які є зовнішніми проявами психічних процесів. Важливо розуміти, що об'єктом спостереження виступає не психіка безпосередньо, а її зовнішні вираження – вербальна, моторна, мімічна, пантомімічна, емоційна та ситуаційна поведінка, що має психічну основу. Через такі прояви дослідник реконструює уявлення про внутрішні стани, мотиви, установки, процеси сприймання, уваги, мислення тощо.

Історично метод спостереження був першим науковим інструментом, яким користувалася психологія. Уже в античності філософи, зокрема Арістотель і Платон, описували емоції, мислення, волю на підставі емпіричних спостережень за людьми. В епоху Відродження ці спроби набули системності: спостереження стало застосовуватись у педагогічних і медичних практиках, де лікарі та вихователі робили нотатки про поведінку, темперамент і реакції людей. Однак справжній методологічний статус воно отримало у XIX столітті, коли Вільгельм Вундт запровадив у Лейпцизькому університеті правила

фіксації спостережень. Саме з цього моменту психологія перестала покладатися лише на інтуїцію чи описові враження й почала вимагати планомірності, повторюваності й об'єктивності.

У XX столітті спостереження розвивалося в різних напрямках. Представники біхевіористичної школи намагалися надати йому максимальної точності. Вони наполягали, що лише те, що можна спостерігати і вимірювати, може вважатися науковим фактом. Цей підхід призвів до розробки стандартизованих форм реєстрації поведінки, коли кожна дія, слово чи мімічна реакція фіксувались за визначеною схемою. Натомість гуманістична та феноменологічна психологія наголошувала, що спостереження не може бути цінним без розуміння внутрішнього смислу поведінки. Вони пропонували інтегрувати до спостереження емпатійний компонент, коли дослідник намагається не лише бачити, а й «відчувати» психологічний стан іншої людини. Такий підхід заклав основу сучасного якісного аналізу поведінки, що поєднує фактичну фіксацію подій із глибоким смисловим осмисленням.

У вітчизняній психології спостереження набуло виняткового значення завдяки культурно-історичному напрямку Льва Виготського та його послідовників. Вони вважали, що поведінку людини не можна зрозуміти поза її соціальним контекстом. Тому спостереження повинно не лише фіксувати дії, а й аналізувати умови, у яких вони відбуваються: мову, правила, традиції, цілі діяльності. Саме звідси походить поняття «включеного спостереження», коли дослідник не стоїть осторонь, а сам стає учасником середовища, що вивчається. Наприклад, спостерігаючи за поведінкою дітей у шкільному класі, психолог може тимчасово виконувати роль учителя або асистента, щоб природніше інтегруватися в ситуацію та не викликати напруження. У цьому полягає одна з головних переваг неекспериментального підходу: він дає змогу побачити людину не у штучній лабораторній ситуації, а в реальному потоці її життя.

Щоб спостереження мало наукову цінність, воно повинно бути цілеспрямованим, плановим і систематичним. Цілеспрямованість означає, що спостереження здійснюється не випадково, а виходячи з чітко визначеної дослідницької мети. У процесі планування вказується, за якими саме проявами поведінки ведеться спостереження, у яких умовах, протягом якого часу. Плановість передбачає створення програми, де визначено послідовність дій дослідника, критерії реєстрації та форми запису. Систематичність означає, що кожен випадок спостереження повинен бути частиною ширшої серії, щоб забезпечити можливість статистичного узагальнення. Якщо дослідник фіксує, скажімо, частоту проявів агресії у підлітків, він має дотримуватись одних і тих самих критеріїв оцінки для всіх учасників і всіх днів, коли ведеться спостереження.

Особливу увагу приділяють об'єктивності. Спостереження завжди супроводжується ризиком суб'єктивних спотворень: дослідник може несвідомо інтерпретувати поведінку через власні уявлення або очікування. Щоб мінімізувати такі ефекти, застосовують кількох спостерігачів,

стандартизовані категорії оцінки або технічні засоби – відеозаписи, цифрові сенсори, автоматизовані системи аналізу рухів. Сучасна психологія дедалі частіше залучає методи комп'ютерного бачення для кількісного аналізу невербальної поведінки, що підвищує точність і зменшує вплив людського фактору.

Розглядаючи різновиди спостереження, слід пам'ятати, що вони не є жорстко відокремленими. Дослідник може комбінувати їх залежно від завдань. Включене спостереження дає змогу краще зрозуміти смислову сторону поведінки, але водночас підвищує ризик упередженості, адже спостерігач стає частиною групи. Невключене – забезпечує об'єктивність, але часто позбавлене емоційного контексту. Польове спостереження дозволяє бачити справжню динаміку поведінки у природних умовах, проте важко контролювати зовнішні змінні, що можуть впливати на результати. Лабораторне – дає змогу відтворити ситуацію багаторазово, але іноді втрачає природність. Тому у сучасних дослідженнях переважають гібридні форми, коли природне спостереження поєднується з елементами контролю, наприклад, використання відеозаписів у реальних умовах із подальшою кількісною обробкою.

Надзвичайно важливою частиною є фіксація даних. У класичній психології застосовувались так звані щоденники або карти спостережень – структуровані таблиці, куди заносили час, подію, реакцію, коментар. Згодом ці форми були стандартизовані: з'явилися чек-листи, кодувальні таблиці, а пізніше – комп'ютерні протоколи. Завдання дослідника полягає не лише в тому, щоб записати подію, а й у тому, щоб визначити її категорію. Наприклад, у дослідженні поведінки дітей фраза «не хочу» може бути віднесена до категорії «опір», «висловлення емоцій» або «комунікативна ініціатива» – залежно від контексту. Тому важливо, щоб категоріальна система була узгоджена між усіма спостерігачами й відповідала меті дослідження.

Після фіксації даних починається інтерпретація. Саме тут проявляється теоретична компетентність психолога. Зібрані факти не мають цінності без розуміння причин, мотивів і контексту. Інтерпретація може здійснюватися як у якісній формі (словесний опис, смисловий аналіз), так і в кількісній (статистична обробка частот, середніх значень, кореляцій). У складних дослідженнях ці підходи поєднуються, утворюючи змішану методологію. Особливо цінною є перевірка достовірності інтерпретації: результати одного спостерігача зіставляються з іншими або з даними, отриманими за допомогою тестів, анкет, експериментів.

Процес наукового спостереження у психології є не просто механічним фіксуванням поведінкових проявів, а складною системою пізнавальних дій, що вимагають від дослідника високої підготовленості, уважності та здатності до аналітичного мислення. Після завершення етапу підготовки дослідник переходить до безпосереднього збору емпіричних даних. На цьому етапі важливо забезпечити постійну об'єктивність та сталість критеріїв фіксації спостережуваних явищ. Досвід показує, що навіть незначні зміни у способі

реєстрації поведінки можуть суттєво вплинути на результати. Тому психологи часто застосовують декілька взаємодоповнюючих засобів – ручне протоколювання, аудіо- та відеозапис, цифрові програми хронометражу. Це дозволяє отримати точніший матеріал для подальшого аналізу.

Однак саме фіксація поведінки є лише частиною загального процесу. Не менш значущим етапом виступає інтерпретація, тобто перехід від опису зовнішніх дій до розуміння внутрішніх психічних процесів. Тут психолог виступає не як пасивний реєстратор, а як аналітик, що співвідносить конкретні факти з теоретичними моделями поведінки. Наприклад, якщо під час спостереження за дитиною у груповій грі виявляється тенденція до уникнення контактів і швидкого відходу від конфліктних ситуацій, дослідник може інтерпретувати це як прояв підвищеної тривожності або низького рівня соціальної впевненості. Проте така інтерпретація має спиратися не лише на інтуїцію, а на систему попередньо визначених критеріїв, валідизованих у психологічних дослідженнях.

Питання валідності спостереження, тобто відповідності між тим, що спостерігається, і тим, що дійсно відображає досліджуване явище, є ключовим. Валідність забезпечується через точність формулювання дослідницьких завдань, адекватність інструментів реєстрації, стабільність поведінкових показників у часі та контроль за впливом зовнішніх факторів. Надійність же означає повторюваність результатів у разі проведення аналогічних спостережень за подібних умов. Якщо два незалежні дослідники, використовуючи однакову методику, доходять подібних висновків, це свідчить про високу надійність спостереження. Саме завдяки таким критеріям спостереження набуває статусу наукового методу, а не просто описового прийому.

Однією з основних проблем спостереження є вплив особистості дослідника на результати. Навіть при високому рівні професіоналізму дослідник схильний несвідомо інтерпретувати поведінку інших через призму власного досвіду, очікувань чи емоційного стану. Цей феномен, відомий як ефект упередження спостерігача, може суттєво спотворювати об'єктивність висновків. Для його мінімізації застосовують «сліпі» спостереження, коли дослідник не знає певних характеристик учасників (наприклад, групової приналежності), або подвійно сліпий метод, коли й дослідник, і учасники не знають про ключову гіпотезу. Крім того, залучення кількох спостерігачів і порівняння їхніх записів дозволяє виявити системні розбіжності й усунути їх у процесі узгодження даних.

Ще одним аспектом, який визначає якість спостереження, є наявність чітких критеріїв категоризації поведінкових проявів. У науковій практиці широко використовуються так звані карти спостереження – стандартизовані схеми, де кожен тип поведінки кодується за певним номером або категорією. Наприклад, у педагогічній психології існують схеми, що дозволяють фіксувати рівень активності учня, характер його взаємодії з учителем, частоту звернень до однолітків, емоційний стан під час уроку. У соціальній психології

подібні карти допомагають оцінювати лідерські ролі, ініціативність або схильність до конформізму. Такі інструменти забезпечують порівнюваність даних і спрощують подальший статистичний аналіз.

Важливу роль у розвитку методу спостереження відіграли роботи Альберта Бандури, який використовував спостереження для дослідження процесів соціального наочіння. У своєму відомому експерименті з «лялькою Бобо» він показав, як діти імітують агресивну поведінку дорослих. Цей експеримент продемонстрував не лише значення спостереження як способу збору даних, а й те, що поведінкові моделі можуть формуватись через пасивне спостереження без безпосереднього підкріплення. У подальших дослідженнях Бандура удосконалив техніку відеоспостереження, яка сьогодні є стандартом у багатьох лабораторіях світу.

Психологічне спостереження має також велику цінність у клінічній практиці. Під час діагностики поведінкових і емоційних розладів лікар спостерігає за пацієнтом у природному спілкуванні, фіксуючи особливості міміки, рухів, інтонації, темпу мовлення, структури пауз. Ці прояви дають змогу зробити висновки про внутрішній стан, навіть якщо вербальна інформація обмежена.

Важливою тенденцією сучасної психології є поєднання спостереження з іншими неекспериментальними методами – інтерв'ю, контент-аналізом, анкетуванням, аналізом продуктів діяльності. Це підвищує достовірність результатів і забезпечує багатовимірне бачення психічних процесів. Наприклад, у дослідженнях дитячої агресивності спостереження може бути доповнене опитувальником для батьків і аналізом малюнків дітей, що дозволяє простежити як зовнішню, так і внутрішню динаміку поведінки.

Метод спостереження залишається одним із базових у психологічних дослідженнях саме тому, що він дає змогу наблизитись до природного функціонування психіки, не спотворюючи його лабораторними умовами. Його сила – у здатності вловлювати нюанси людських дій, які не завжди піддаються кількісному виміру, але мають величезне значення для розуміння внутрішніх механізмів поведінки. У сучасних умовах, коли дедалі більшого значення набувають міждисциплінарні підходи, спостереження поступово поєднується з технологічними інструментами – цифровими системами аналізу рухів, автоматичним розпізнаванням емоцій, штучним інтелектом, який допомагає виявляти закономірності у великих масивах відеоданих. Таким чином, старий класичний емпіричний метод отримує нове життя, залишаючись основою емпіричної достовірності психологічної науки.

5.3. Анкетування як метод дослідження психіки

Метод анкетування є одним із найпоширеніших і водночас найгнучкіших неекспериментальних методів дослідження психіки. Його сутність полягає у збиранні емпіричної інформації про внутрішні стани, переживання, установки, мотивацію, цінності чи поведінкові наміри

респондентів за допомогою спеціально розроблених анкет – стандартизованих опитувальних форм, що містять систему запитань і варіантів відповідей. Анкетування дозволяє досліднику звернутися до великої кількості людей, отримати узагальнені дані та статистично опрацювати результати, що робить цей метод особливо придатним для масових соціально-психологічних досліджень.

Анкетування вирізняється високим ступенем формалізації, що дозволяє збирати дані в масштабі, недоступному для більшості інших методів. Саме тому воно широко використовується у соціальній, політичній, організаційній, педагогічній та медичній психології. Цей метод придатний як для первинного виявлення дослідницької проблематики (розвідувальні анкети), так і для перевірки гіпотез, виявлення тенденцій, порівняння груп, а також для оцінки динаміки змін у повторних замірах.

Своєю появою анкетування завдячує розвитку соціології та експериментальної психології XIX–XX століття. Уже на початку XX століття Чарльз Дарвін, Вільям Джеймс і Альфред Біне використовували опитувальні форми для збору даних про індивідуальні відмінності у сприйнятті, пам'яті та емоційних реакціях. У радянській і українській психології метод набув популярності завдяки працям Б. Ананьєва, О. Леонтьєва, К. Платонова, які підкреслювали його цінність як способу виявлення внутрішніх психологічних установок, не доступних прямому спостереженню. Згодом анкетування стало невід'ємною складовою більшості емпіричних досліджень у соціальній, педагогічній, медичній, військовій та організаційній психології.

Головна перевага анкетування полягає в можливості отримати інформацію про суб'єктивні стани людини – її переконання, емоційні оцінки, самоусвідомлення, тобто ті феномени, які неможливо спостерігати зовні. Анкета фактично виступає інструментом вербалізації внутрішнього світу респондента. При цьому ефективність методу залежить від якості побудови самої анкети, логіки запитань, точності формулювань і довіри між дослідником і опитуваним. Будь-яка двозначність або навідний характер запитання може призвести до викривлення результатів, тому створення анкети – це складний процес, що вимагає методологічної культури й розуміння психології комунікації.

Побудова анкети традиційно починається з визначення мети дослідження і формулювання робочих гіпотез. Далі дослідник розробляє логічну структуру опитувальника, яка має відповідати принципу поступового розгортання від загального до конкретного. Початкові запитання, як правило, мають бути нейтральними, простими для розуміння, аби створити сприятливий психологічний фон і знизити рівень напруженості. Лише після цього переходять до запитань, які торкаються суті досліджуваної проблеми. Така побудова анкети допомагає мінімізувати ефект соціальної бажаності (тенденцію відповідати так, як «правильно», а не як є насправді), ефект стомлення (особливо в довгих анкетах), некоректну вибірку (невідповідність характеристик респондентів меті дослідження), випадковість відповідей або

імітацію співпраці. Для виявлення таких ефектів застосовуються контрольні шкали, спеціальні індикатори нещирості, статистичні перевірки внутрішньої узгодженості відповідей.

У структурі анкети зазвичай виокремлюють три основні частини: вступ, основний зміст і соціально-демографічний блок. У вступі респондентові коротко пояснюється мета опитування, наголошується на добровільності участі та конфіденційності результатів. Основна частина містить змістові запитання, які можуть бути відкритими, закритими або змішаними. Відкриті запитання дозволяють отримати розгорнуту відповідь у довільній формі, що дає можливість фіксувати не лише зміст, а й стиль мислення, рівень рефлексії, словниковий запас респондента. Закриті запитання обмежують вибір варіантами відповідей (так/ні, шкала від 1 до 5, множинний вибір тощо) – вони зручні для кількісної обробки й статистичного порівняння. Змішані запитання поєднують обидва підходи, даючи респонденту змогу доповнити готові варіанти власним коментарем.

Наукова коректність анкети передбачає використання різних типів шкал залежно від мети аналізу. Наприклад, для вимірювання рівня тривожності або задоволеності роботою часто застосовують шкалу Лікерта, що пропонує оцінити ступінь згоди з певними твердженнями. Для визначення переваг чи інтенсивності емоційного ставлення використовують семантичний диференціал Ч. Осгуда, де респондент оцінює об'єкт за біполярними характеристиками («приємний – неприємний», «активний – пасивний», «сильний – слабкий»). Такі методи дозволяють перевести якісні судження у кількісні показники, що робить можливим подальший статистичний аналіз.

Важливим етапом є апробація анкети – її попереднє тестування на невеликій вибірці, що дозволяє виявити нечіткі або некоректні формулювання, визначити тривалість заповнення та оцінити розуміння запитань. Результати апробації використовують для удосконалення структури анкети, після чого вона проходить процедуру стандартизації. Стандартизація передбачає створення єдиних умов проведення опитування, інструкцій, способів кодування та обробки результатів, що забезпечує можливість порівняння даних різних груп респондентів і повторення дослідження іншими науковцями.

Анкетування може проводитись у письмовій, електронній або усній формі. Класичний письмовий варіант передбачає роздачу анкет у паперовому вигляді, тоді як сучасні цифрові технології дозволяють реалізовувати онлайн-опитування за допомогою спеціальних платформ (Google Forms, Qualtrics, LimeSurvey). Електронні анкети значно спрощують обробку даних, автоматично формуючи бази відповідей і дозволяючи одразу проводити статистичний аналіз. Проте дослідник повинен зважати на технічні та етичні аспекти: інтернет-опитування часто стикаються з проблемою вибіркової (доступ мають лише користувачі певних платформ) та анонімності (учасники можуть заповнювати анкету кілька разів або надавати неправдиву інформацію).

Особливу роль у психологічних дослідженнях відіграє форма формулювання запитань. Вони повинні бути короткими, однозначними, граматично коректними, без спеціальних термінів, що можуть бути незрозумілими респонденту. Забороняється поєднувати в одному запитанні дві теми («Чи вважаєте ви, що навчання в університеті корисне і водночас занадто складне?»), адже така структура не дозволяє визначити, на яку частину питання респондент відповідає. Крім того, важливо уникати запитань-оцінок, що передбачають моральне судження, – вони часто активують механізми самозахисту і викликають соціально бажані відповіді.

Достовірність результатів анкетування залежить також від мотивації респондента. Якщо людина не зацікавлена у дослідженні або сумнівається в конфіденційності, вона може давати випадкові або поверхові відповіді. Тому у вступній частині анкети необхідно створити атмосферу довіри, пояснити, що всі відповіді є анонімними, а результати використовуються лише у наукових цілях. У практиці прикладних досліджень часто застосовують невеликі винагороди чи бонуси за участь – не як матеріальний стимул, а як жест поваги до часу респондента.

Важливо наголосити, що анкетування, як і будь-який інший метод, не може розглядатись у відриві від контексту дослідження. Воно не замінює спостереження чи експеримент, але чудово їх доповнює, дозволяючи отримати суб'єктивні дані, які неможливо зібрати іншим шляхом. Найвищу наукову цінність мають дослідження, у яких анкетування поєднане з поведінковими або фізіологічними вимірюваннями – наприклад, коли самооцінка рівня тривожності зіставляється з показниками серцевого ритму чи реакцією шкіри на стрес.

Загалом метод анкетування в сучасній психології є важливим засобом емпіричного пізнання. Його ефективність залежить не лише від технічної досконалості анкети, а передусім від теоретичного мислення дослідника, здатності формулювати запитання, що дійсно відображають суть досліджуваного явища. Правильно спроектована анкета перетворюється на інструмент не просто збору даних, а на засіб проникнення у внутрішній світ людини, виявлення закономірностей її психічної діяльності й побудови узагальнених наукових висновків.

5.4. Метод інтерв'ю (бесіди)

Метод інтерв'ю (або бесіди) є одним із найглибших способів дослідження психіки, який базується на безпосередній вербальній взаємодії між дослідником і респондентом. Його сутність полягає в отриманні достовірної, змістовної інформації через діалог, у якому психолог виступає не лише як запитувач, а як уважний слухач, спостерігач і аналітик. На відміну від анкетування, інтерв'ю дозволяє глибше проникнути у внутрішній світ людини, дослідити не лише зміст її відповідей, а й спосіб їх подання, інтонацію, емоційну реакцію, невербальні прояви. Саме тому цей метод вважають одним

із найбільш гнучких і психологічно інформативних інструментів у неекспериментальних дослідженнях.

Психологічне інтерв'ю має глибокі історичні корені. Уже в працях Зигмунда Фрейда, Альфреда Адлера, Карла Юнга та Карла Роджерса інтерв'ю використовувалося як основний метод пізнання внутрішнього світу особистості. У класичному психоаналізі воно мало форму вільної асоціативної бесіди, де клієнт висловлював будь-які думки, а аналітик інтерпретував їхній прихований зміст. У гуманістичній психології інтерв'ю набуло нового змісту – як рівноправна комунікація, у якій створюється атмосфера прийняття, емпатії та безумовної поваги до особистості. У наукових дослідженнях інтерв'ю еволюціонувало у стандартизований метод збору даних, що поєднує елементи контролю та гнучкості, дозволяючи водночас отримувати як якісну, так і кількісну інформацію.

Основна цінність методу інтерв'ю полягає у його здатності виявляти смисли, мотиви, установки та індивідуальні інтерпретації подій. Людина не просто повідомляє факти, вона конструює свій досвід у процесі розмови. Саме через цю конструктивну функцію інтерв'ю дає змогу психологу зрозуміти, як респондент осмислює власну поведінку, як вибудовує причинно-наслідкові зв'язки у власній біографії, як формує ставлення до інших і до себе. Таким чином, інтерв'ю – це не лише метод збору інформації, а й інструмент реконструкції психологічної реальності.

У науковій психології прийнято розрізняти кілька видів інтерв'ю за ступенем структурованості. Стандартизоване інтерв'ю – це форма, у якій усі запитання сформульовані заздалегідь і подаються респонденту в чітко визначеній послідовності. Відповіді зазвичай кодуються для подальшого статистичного аналізу. Такий підхід забезпечує порівнянність даних між різними респондентами та знижує вплив суб'єктивного чинника дослідника. Проте його недолік – обмежена глибина: респондент може не мати можливості вільно висловитися, а важливі деталі залишаються поза увагою.

Напівструктуроване інтерв'ю поєднує стандартизовані запитання з відкритими. Дослідник має певний план, проте може змінювати послідовність, уточнювати або розширювати запитання залежно від відповідей респондента. Такий формат забезпечує баланс між науковою системністю і живою, гнучкою бесідою. Він найчастіше використовується у психологічних дослідженнях, оскільки дозволяє виявити як кількісні закономірності, так і якісні особливості досвіду.

Неструктуроване інтерв'ю (або клінічна бесіда) максимально наближене до природного спілкування. У ньому немає фіксованих запитань; дослідник лише задає загальний напрям розмови, а респондент вільно висловлює свої думки. Така форма особливо ефективна у психотерапевтичних, феноменологічних і біографічних дослідженнях, де важливо розкрити внутрішній світ людини. Проте вона потребує високої кваліфікації дослідника, вміння слухати, не нав'язуючи власних інтерпретацій, і водночас непомітно спрямовувати розмову у потрібне русло.

Метод інтерв'ю має складну психологічну структуру. Його ефективність визначається не лише запитаннями, а й усім контекстом спілкування: атмосферою, тоном голосу, невербальними сигналами, взаємною довірою. Психолог повинен створити умови, за яких респондент почувається прийнятим і безпечним, інакше відповіді будуть поверховими або соціально бажаними. Для цього інтерв'юер використовує техніки активного слухання – перефразування, уточнення, вербалізацію емоцій («Ви маєте на увазі, що це викликало у вас тривогу?»). Такі прийоми дозволяють не лише глибше зрозуміти зміст висловлювання, а й стимулюють респондента до подальших роздумів.

Під час планування інтерв'ю необхідно чітко визначити його мету, тему, вибірку та процедуру проведення. Вибір респондентів повинен бути репрезентативним щодо досліджуваної проблеми – наприклад, якщо йдеться про вивчення професійного вигорання, доцільно опитувати фахівців, які працюють у стресових умовах. Важливо заздалегідь продумати місце проведення інтерв'ю: воно має бути спокійним, без сторонніх шумів і відволікань. Оптимальна тривалість бесіди – від 30 до 60 хвилин; коротші інтерв'ю не дозволяють досягти глибини, а надто довгі можуть викликати втому та втрату уваги.

Розробка запитань – найвідповідальніший етап підготовки інтерв'ю. Вони повинні бути відкритими, стимулювати респондента до роздумів і водночас не нав'язувати певну позицію. Недопустимими є запитання, що містять оцінку («Ви згодні, що так чинити неправильно?»), або передбачають лише одну правильну відповідь. Ефективні запитання починаються словами «як», «чому», «що ви відчували, коли...», «як ви ставитеся до...». Саме вони відкривають простір для осмислення.

У науковій практиці під час інтерв'ю важливим є процес фіксації відповідей. Це може бути стенографування, аудіозапис, відеозйомка або одночасне занотування ключових тез. Використання технічних засобів підвищує точність, однак потребує обов'язкової згоди респондента. Після інтерв'ю проводиться транскрипція – детальне відтворення тексту розмови, що зберігає не лише слова, а й паузи, емоційні реакції, інтонаційні особливості. Такий матеріал є основою для подальшого якісного аналізу – контент-аналізу, тематичного аналізу, дискурсивного чи наративного підходів.

Контент-аналіз дозволяє виділити повторювані теми, категорії та смисли в текстах інтерв'ю. Тематичний аналіз спрямований на виявлення глибинних структур досвіду – наприклад, мотивів поведінки, переживання втрати, етапів прийняття складної ситуації. Дискурсивний аналіз досліджує, як через мову респондент конструює свою ідентичність, а наративний – як вибудовує власну життєву історію. Таким чином, інтерв'ю є універсальним джерелом даних для багатьох підходів сучасної психології.

Варто підкреслити, що інтерв'ю має не лише пізнавальну, а й терапевтичну функцію. У процесі розмови людина часто усвідомлює власні

переживання, структурує досвід, знаходить нові сенси. Це явище отримало назву «ефект інсайту» – раптового усвідомлення проблеми під час її вербалізації. Саме тому інтерв'ю вимагає високого рівня етичної відповідальності: дослідник повинен бути готовим підтримати респондента, якщо розмова торкається болісних тем.

Етичні стандарти проведення інтерв'ю передбачають добровільну участь, анонімність, право на відмову від відповіді й захист персональних даних. Респондент має бути попереджений про мету дослідження, тривалість бесіди, можливість запису. Особливої уваги потребують інтерв'ю з дітьми, людьми похилого віку або представниками вразливих груп – у цих випадках дослідник повинен отримати дозвіл опікунів або дотримуватися спеціальних протоколів етичного супроводу.

Серед основних переваг інтерв'ю – гнучкість, глибина та контекстуальність отриманих даних. Жоден стандартизований тест не може дати такої кількості нюансів, як жива розмова. Проте цей метод має і свої обмеження: він є трудомістким, суб'єктивним, вимагає високої підготовки дослідника і складний для статистичного узагальнення. Крім того, існує ризик інтерпретаційних упереджень – коли дослідник несвідомо спотворює відповіді відповідно до власних очікувань. Щоб мінімізувати ці ризики, застосовують процедури міжсуб'єктної перевірки (*inter-rater reliability*), коли кілька незалежних аналітиків аналізують ті самі інтерв'ю і порівнюють результати.

У сучасній психології метод інтерв'ю активно використовується у вивченні особистісних криз, професійного вигорання, сімейних стосунків, адаптації до стресу, процесів саморефлексії, ідентичності та соціальної взаємодії. Його застосовують у клінічній практиці, організаційній психології, педагогіці, геронтопсихології, нейропсихологічній діагностиці. Наприклад, клінічне інтерв'ю допомагає зібрати анамнез, діагностувати психічні розлади, визначити ресурси особистості. У соціальній психології воно використовується для аналізу міжособистісної взаємодії та соціальних ролей, а в педагогічній – для вивчення ставлення учнів до навчання, учителів, однолітків.

Загалом метод інтерв'ю (бесіди) у психологічному дослідженні – це не просто спосіб збору інформації, а складна форма міжособистісного пізнання, де головним інструментом виступає сам дослідник. Його професійність визначається не лише здатністю ставити запитання, а й умінням слухати, помічати невимовлене, будувати довіру та підтримувати етичний баланс між науковою цікавістю і повагою до людської гідності. Правильно проведене інтерв'ю перетворюється на дзеркало людської свідомості – у ньому відбиваються не лише думки, а й смисли, емоції, переживання.

5.5. Архівний метод в психології

Архівний метод у психології – це специфічний неекспериментальний спосіб дослідження, який ґрунтується на аналізі вже наявних документальних

матеріалів, створених у процесі життєдіяльності людини або інституцій, без прямого втручання дослідника у поведінку досліджуваних осіб. Іншими словами, це метод вторинного аналізу даних, де джерелом інформації є тексти, записи, звіти, щоденники, листування, стенограми, протоколи, електронні бази чи навіть соціальні мережі. Його головна мета – виявити психологічні закономірності, що відображаються у цих матеріалах, і реконструювати психологічні аспекти людської діяльності в історичному, культурному або соціальному контексті.

Архівний метод має довгу історію в гуманітарних науках. Ще у XIX столітті історики культури, філологи та соціологи використовували письмові джерела для вивчення менталітету епохи, світогляду та емоційних установок людей минулого. У психології цей підхід розвинувся завдяки працям В. Вундта, Г. Еббінгауза, З. Фрейда та Л. Виготського, які підкреслювали цінність культурно-історичних артефактів для розуміння психічного розвитку. Фрейд, наприклад, аналізував особисті листи та автобіографічні матеріали, щоб реконструювати несвідомі мотиви поведінки. Виготський розглядав архівні тексти як прояви вищих психічних функцій, що відображають суспільно-історичну природу мислення. Таким чином, архівний метод має не лише емпіричну, а й теоретико-методологічну вагу – він показує, що психіка людини залишає сліди не лише у поведінці, а й у культурних продуктах.

З погляду сучасної методології, архівний метод – це систематичний процес виявлення, відбору, аналізу та інтерпретації документів, що містять психологічно значущу інформацію. Він може застосовуватися як самостійний спосіб дослідження (наприклад, у психології творчості, історії психології, політичній чи соціальній психології), так і як допоміжний – у поєднанні зі спостереженням, контент-аналізом, інтерв'юванням чи експериментом.

Основна перевага архівного методу полягає у його ненав'язливості: дослідник не впливає на поведінку суб'єкта, а аналізує природно створені матеріали. Це дає змогу отримати достовірні дані про психічні процеси у реальних умовах життя, без спотворень, що часто виникають під час експерименту чи опитування. Наприклад, аналіз листів фронтовиків часів Другої світової війни дозволяє досліднику відтворити емоційну динаміку страху, надії, адаптації та згуртованості без ретроспективних спотворень, властивих пізнішим спогадам.

Джерела архівних даних у психології надзвичайно різноманітні. До них належать:

- офіційні документи (протоколи засідань, звіти, накази, судові справи);
- особисті записи (щоденники, листи, мемуари, автобіографії);
- медичні чи психіатричні карти;
- навчальні журнали, характеристики, шкільні твори;
- записи психологічних експериментів минулого;
- друковані або електронні публікації в засобах масової інформації;
- соціальні медіа – пости, коментарі, блоги, цифрові щоденники.

Сучасна психологія активно розширює поняття «архіву», включаючи цифрові простори. Так, інтернет-платформи, форуми чи соціальні мережі розглядаються як новітні «цифрові архіви», що відображають колективну психологію суспільства. Аналіз таких даних дає змогу вивчати емоційні тенденції, мовні патерни, соціальні ідентичності, колективні страхи чи надії. Наприклад, під час пандемії COVID-19 архіви публікацій у Twitter та Facebook використовували для виявлення рівня тривожності, панічних настроїв та соціальної підтримки у різних країнах.

Однак використання архівних даних вимагає особливої критичності. На відміну від експерименту, де змінні контрольовані, у документальних джерелах інформація є фрагментарною, контекстуально обмеженою і часто створювалася не для наукових цілей. Тому дослідник повинен оцінювати достовірність джерела, його авторство, мету створення, стиль і потенційні упередження. Наприклад, лист, написаний під цензурою або в умовах страху, може не відображати справжніх емоцій. У цьому сенсі архівний метод вимагає високого рівня герменевтичної (тлумачної) культури – уміння «читати між рядків», враховувати соціальний контекст і мовні особливості епохи.

Аналіз архівних даних може здійснюватися кількома способами. Найпоширеніші – контент-аналіз і дискурсивний аналіз. Перший полягає у кількісному підрахунку частоти появи певних слів, тем або категорій, що дозволяє виявити домінантні мотиви, емоційні тенденції чи установки. Другий орієнтований на якісне вивчення смислів, які створюються у тексті – наприклад, як людина через мову конструює власну ідентичність, як формуються дискурси влади, гендеру, моралі. У поєднанні ці підходи дають змогу відтворити не лише зміст повідомлень, а й глибинну структуру мислення авторів.

Важливим етапом роботи з архівами є систематизація та кодування матеріалів. Дослідник створює класифікаційні схеми, що дозволяють об'єднувати фрагменти за темами, періодами, психологічними показниками (емоції, мотиви, цінності, типи поведінки). Наприклад, у дослідженні дитячих щоденників воєнного часу можна виокремити коди «страх», «надія», «сім'я», «втрати», «звичне життя». Це дозволяє простежити динаміку психічних станів дітей залежно від обставин.

Архівний метод особливо цінний у психології розвитку та історії науки. Завдяки йому можна простежити еволюцію понять, теорій, напрямів – наприклад, як змінювалося розуміння «особистості» від психоаналітичної моделі до когнітивної. Аналіз архівних праць видатних психологів дозволяє реконструювати логіку їхнього мислення, виявити вплив культурних і соціальних факторів на наукові ідеї. У цьому сенсі архівний метод виступає інструментом метанаукового пізнання – вивчення самої науки як психологічного процесу.

Сучасна наука активно застосовує архівний метод у прикладних сферах. У кримінальній психології він допомагає аналізувати судові справи, свідчення,

поліцейські протоколи, щоб виявити психологічні механізми злочинної поведінки. У медичній психології – використовується для вивчення історій хвороб, зокрема з метою відстеження впливу травматичних подій на психічний стан пацієнтів. У педагогічній психології архіви шкільних творів або звітів про навчання дозволяють відтворити історію розвитку освітніх цінностей і дитячих уявлень про навчання.

Використання архівного методу пов'язане також із низкою етичних питань. Аналіз особистих документів (листів, щоденників) потребує згоди власника або його спадкоємців, а також збереження анонімності. Навіть якщо документи є публічними, психолог повинен дотримуватись принципу поваги до приватності, не спотворювати смисли і не робити висновків, які можуть принизити честь чи гідність особи. У цифрову добу ця проблема стає особливо актуальною, адже великі масиви даних у соцмережах є формально відкритими, проте їхнє використання без згоди користувачів може порушувати етичні стандарти.

Як і будь-який метод, архівний має свої обмеження. Він не дозволяє досліднику безпосередньо перевіряти гіпотези експериментальним шляхом, обмежує можливість контролю змінних і часто залежить від випадковості збережених документів. Крім того, дані можуть бути неповними, фрагментарними або упередженими. Тому важливо поєднувати архівний метод з іншими підходами – спостереженням, біографічним аналізом, контент-аналізом, анкетуванням. У такому поєднанні він набуває найбільшої наукової сили, оскільки дозволяє підтвердити архівні висновки сучасними емпіричними даними.

Попри ці обмеження, архівний метод залишається незамінним для дослідження феноменів, які неможливо відтворити експериментально – наприклад, історичних подій, масових соціальних рухів, психології лідерства чи колективної поведінки. Він допомагає побачити, як психічне життя індивіда вплітається у культурну тканину часу, як індивідуальні переживання стають частиною колективного досвіду.

5.6. Метод контент-аналізу в психологічних дослідженнях

Контент-аналіз посідає у сучасній психології особливе місце, оскільки дозволяє дослідникові працювати не лише з поведінковими реакціями чи кількісними показниками, а й безпосередньо з продуктами діяльності людини – текстами, висловлюваннями, комунікативними повідомленнями, соціальними наративами, образами, візуальними матеріалами. Його унікальність полягає в тому, що він надає можливість працювати з продуктами діяльності – текстами, висловлюваннями, розповідями, щоденниковими записами, повідомленнями у соціальних мережах, реакціями досліджуваних у проєктивних методиках, стенограмами інтерв'ю чи навіть невербальними фрагментами комунікації. На відміну від традиційних експериментальних методів, що фіксують поведінкові реакції у контрольованих умовах, контент-

аналіз дозволяє проникнути у глибинні смисли – не лише те, що людина робить або говорить, а й те, як вона структурує свій внутрішній світ через вербальні та семантичні патерни.

Витоки контент-аналізу пов'язані з розвитком соціальних наук ХХ століття, коли постала потреба у систематичному вивченні масових комунікацій. Проте у психології цей метод набув значного розвитку після робіт Л. Виготського, Дж. Келлі, Ч. Осгуда, Дж. Лассуелла та представників наративної психології. Поступово він перетворився на інструмент дослідження як індивідуальної, так і колективної свідомості. На когнітивному рівні контент-аналіз дозволяє реконструювати смислові категорії, які суб'єкт використовує для розуміння світу; на емоційному рівні – визначати настрій, емоційний тон, афективні домінанти; на соціально-психологічному – виявляти установки, цінності, стереотипи, структуру соціальних уявлень. У цьому сенсі контент-аналіз часто розглядають як один із найточніших методів доступу до латентних змістів свідомості, оскільки людина не завжди може або хоче прямо висловити свої внутрішні стани, але вони проявляються в мовленнєвій побудові тексту.

Теоретичною засадою контент-аналізу є припущення, що будь-який текст – це не випадковий потік слів, а смислова система, у якій окремі елементи взаємопов'язані. Психологічний зміст таких елементів проявляється через їх частоту, контекст, сполучуваність, позиційні особливості, емоційну забарвленість або стилістичні характеристики. Тому аналіз може включати як просте підрахування одиниць (наприклад, частоту слів, категорій, «смислових маркерів»), так і складні процедури, спрямовані на виявлення глибинних структур – когнітивних схем, наративних стратегій, культурних кодів, особистісних смислів.

На практичному рівні контент-аналіз реалізується через низку послідовних етапів. Першим кроком є формування дослідницької проблеми й операціоналізація категорій. Психолог визначає, які аспекти особистості чи поведінки він хоче вивчити, і перетворює ці абстрактні поняття на конкретні одиниці аналізу – слова, фрази, тематичні блоки, смислові фрагменти. Часто на цьому етапі проводиться пілотне читання матеріалу для з'ясування, які категорії природно виникають у тексті. Другим етапом є формування системи кодування – створення інструкцій або «кодбуку», що описує правила віднесення текстових елементів до тих чи інших категорій. Надійність у цьому етапі є критично важливою, оскільки нечіткі правила призводять до суб'єктивних спотворень, особливо коли кодувальників декілька.

Третій етап передбачає безпосереднє кодування матеріалу. Тексти розчленовуються на одиниці – окремі слова, речення або більш комплексні смислові структури. Залежно від завдання, кодування може бути номінальним (віднесення до категорії), порядковим (ранжування смислів), кількісним (підрахунок частот), або якісним (інтерпретація контексту). На четвертому етапі здійснюється підрахунок частот, співвіднесення категорій, побудова матриць, визначення співвідношень між різними смисловими параметрами

тексту. Це може включати кластерний аналіз, факторний аналіз, кореляції, мережевий аналіз, контекстуальну комп'ютерну обробку (наприклад, аналіз тональності). Завершальним етапом є інтерпретація отриманих результатів – перехід від формальних частот або структурних схем до психологічного змісту, формулювання висновків щодо особистісних особливостей автора тексту або соціально-психологічних характеристик групи.

Однією з ключових переваг контент-аналізу є можливість поєднання об'єктивності кількісних методів із глибиною якісної інтерпретації. Частотні показники зменшують ризик суб'єктивного трактування даних: якщо певне смислове слово або емоційний маркер зустрічається у тексті вдвічі частіше, ніж інше, це об'єктивний факт. Однак справжня психологічна інтерпретація стає можливою лише тоді, коли кількісні дані поєднуються з аналізом контексту – функції, ролі та позиції смислових елементів у загальній структурі тексту. Такий підхід дозволяє розкривати глибинні ставлення, що не завжди усвідомлюються досліджуваними. Наприклад, у дослідженнях тривожності виявлено, що люди з високим рівнем тривоги частіше використовують слова, пов'язані з контролем і безпекою, а також більшою кількістю заперечень («не хочу», «не знаю», «не впевнений»). У соціально-психологічних дослідженнях ворожості ефективним показником є перевага слів зі знеособлюючою семантикою: «вони», «інші», «чужі». У клінічних наративах пацієнтів із депресією часто домінують теми втрати, провини та безнадії, які можуть не проявлятися у прямих відповідях опитувальників, але чітко проступають у структурі розповіді.

Застосування контент-аналізу в психології охоплює широке коло напрямів. У клінічній практиці він використовується для оцінки емоційного стану, визначення когнітивних викривлень, аналізу динаміки терапії за зміною мовленнєвих патернів. У дослідженнях розвитку – для вивчення формування Я-концепції, моральних суджень або уявлень дитини про світ. У соціальній психології – для аналізу стереотипів, масових уявлень, соціальних наративів, дискурсів влади, маніпулятивних технологій у медіа. У психолінгвістиці – для дослідження образної структури мови, мовленнєвих стратегій, стилістичних особливостей індивідуального письма. З розвитком цифрових технологій метод отримав новий імпульс: автоматизовані системи контент-аналізу дозволяють працювати з величезними масивами текстів, включно з повідомленнями в соціальних мережах, онлайн-форумах і базах даних клінічних нотаток.

Одним із важливих методологічних питань контент-аналізу є вибір одиниць аналізу. Традиційно використовують слова, словосполучення або теми, але іноді ефективніше працювати з категоріями, які не є прямо вираженими у тексті, а реконструюються під час аналізу. Наприклад, у наративній психології одиницями аналізу є «наративні сегменти» – цілі фрагменти історії, що несуть певний психологічний сенс. У дослідженнях мотивації – «сміслові фрейми», які організують інтерпретацію подій. У

вивченні ідентичності – «позиціонування», тобто те, як автор описує себе щодо інших.

Ще однією проблемою є небезпека надмірної формалізації результатів. Некритичне підрахування слів може створити ілюзію об'єктивності, але ігнорувати змістовий аспект. Наприклад, одна й та сама емоційно забарвлена лексема може мати різний психологічний сенс залежно від контексту. Тому сучасний контент-аналіз прагне до інтеграції обох підходів: кількісного – для забезпечення об'єктивності, і якісного – для змістової інтерпретації. У цьому значенні контент-аналіз є не просто технікою, а методологічною структурою, що дозволяє поєднувати статистичні й інтерпретаційні компоненти.

Особливу цінність у психології має здатність контент-аналізу виявляти латентні структури свідомості, які не завжди проявляються у прямій поведінці. У дослідженнях особистісних конструктів, наприклад, було показано, що люди, які будують наративи з високою часткою абстрактних лексем, частіше демонструють гнучкі когнітивні стилі, тоді як ті, хто використовує багато конкретних описів дій, мають більш поведінково-орієнтований стиль. У терапевтичній практиці зміна мовлення пацієнта – поява слів, пов'язаних із контролем, перспективою, надією – часто є об'єктивним маркером прогресу. У дослідженнях міжгрупових конфліктів аналіз текстів дозволяє ідентифікувати ступінь поляризації через збільшення кількості морально насичених слів («зрада», «борг», «справедливість»), що часто передують загостренню соціальної напруги.

Критично важливим аспектом застосування контент-аналізу є забезпечення надійності та валідності. Надійність підвищують через стандартизацію кодувальних процедур, перевірку узгодженості роботи різних кодувальників (коефіцієнт Каппа Коена), повторне кодування матеріалів або автоматизований аналіз. Валідність забезпечують через експертні оцінки категорій, логічну відповідність операціоналізації дослідницькій меті, співвіднесення з іншими методами (наприклад, психосемантичними або проєктивними). У цьому сенсі контент-аналіз вписується у широку тенденцію сучасної психології до мульти-методологічності та взаємного підтвердження даних з різних джерел.

Таким чином, контент-аналіз у психології є не просто технікою опрацювання текстової інформації, але складною методологічною системою, що дозволяє досліджувати психологічні явища через їх смислове втілення у мовленні. Він поєднує формальну об'єктивність кількісних методів із гнучкістю якісного аналізу, дозволяючи досягти глибинного розуміння індивідуальної та групової свідомості. Контент-аналіз відкриває можливість реконструювати суб'єктивні структури уявлень, виявляти емоційні патерни, діагностувати когнітивні особливості й аналізувати соціально-психологічні процеси, що формують людську поведінку. У цьому сенсі він стає одним із найбільш потужних інструментів у сучасній психології – інструментів, що поєднують наукову точність із можливістю проникнення у смислову глибину людської психіки.

Завдання на самостійну підготовку

1. Чому неекспериментальні методи забезпечують високу екологічну валідність і як це впливає на достовірність результатів?
2. Як феноменологічна та герменевтична традиції вплинули на методологію неекспериментальних досліджень?
3. Які переваги та труднощі виникають під час дослідження психіки у природних умовах?
4. У чому полягає логіка класифікації неекспериментальних методів за формою взаємодії з джерелом психологічної інформації?
5. Чому точність інтерпретації є ключовою вимогою до дослідника, що працює з неекспериментальними даними?
6. Які основні історичні етапи розвитку методу спостереження у психології та як вони вплинули на його сучасну форму?
7. У чому полягає різниця між включеним і невключеним спостереженням та які переваги й ризики притаманні кожному з них?
8. Як стандартизовані форми фіксації поведінки (карти спостереження, чек-листи) підвищують наукову надійність цього методу?
9. Чому об'єктивність є критично важливою для методу спостереження і як її можна забезпечити?
10. У чому полягає ефект упередження спостерігача та якими способами його можна мінімізувати?
11. Як поєднання спостереження з іншими методами (інтерв'ю, аналіз продуктів діяльності) підвищує достовірність результатів?
12. У чому полягає наукова цінність анкетування для вивчення суб'єктивних станів, мотивацій та установок респондентів?
13. Які ризики виникають через ефекти соціальної бажаності, стомлення або невідповідної вибірки під час анкетування?
14. Чому апробація анкети на пілотній вибірці є необхідною умовою її подальшого використання?
15. Які особливості інтерв'ю дозволяють досліднику отримати глибшу інформацію порівняно з анкетуванням?
16. У чому полягають відмінності між стандартизованим, напівструктурованим і неструктурованим інтерв'ю?
17. Які техніки активного слухання забезпечують якість і глибину психологічного інтерв'ю?
18. Які вимоги ставляться до етичної організації інтерв'ю, особливо в роботі з вразливими групами?
19. У чому полягає сутність архівного методу та які переваги він має порівняно з іншими неекспериментальними методами?
20. Які основні труднощі пов'язані з інтерпретацією архівних даних і як дослідник може забезпечити їхню валідність?

РОЗДІЛ 6

ХАРАКТЕРИСТИКА МЕТОДУ ПСИХОЛОГІЧНОГО ЕКСПЕРИМЕНТУ

6.1. Проблема вивчення причинно-наслідкових зв'язків у психології

Проблема причинно-наслідкових зв'язків (каузальних зв'язків) є центральною не лише для психології, а й для всієї наукової традиції. Пізнати причинність означає зрозуміти, чому явище відбувається, а не просто як воно проявляється. Власне, перехід від описового рівня знання до пояснювального завжди знаменує становлення науки у повному сенсі цього слова. Якщо філософія обмежується спробою осмислити сенс подій, а спостереження фіксує їхню зовнішню послідовність, то експеримент як метод доказу наявності каузальних зв'язків прагне проникнути у сам механізм взаємодії – визначити, який саме фактор викликає зміни в іншому.

Для психології цей перехід був особливо складним. Психічні явища не мають безпосередньої тілесної форми, їх неможливо побачити або виміряти безпосередньо. Протягом століть уявлення про душу чи свідомість залишалися у сфері філософських роздумів або теологічних концепцій. Спроба пояснити поведінку людини в термінах причин і наслідків вимагала радикального методологічного зсуву – від споглядального аналізу до емпіричного експерименту. Саме завдяки цьому переходу психологія здобула статус самостійної науки. Якщо спостереження чи анкетування лише описують уже наявні явища, то експеримент дозволяє перевірити, чому вони відбуваються. Його головна мета – не фіксація фактів, а встановлення причинно-наслідкових зв'язків. Такий підхід дозволяє зробити висновок про причинність, що є недосяжним у межах спостереження чи опитування.

У традиції наукового пізнання причинність завжди була критерієм зрілості науки. Коли ми говоримо, що певна подія викликає іншу, ми припускаємо наявність закономірності, що може бути відтворена. Саме це відтворення і є серцевиною експериментального підходу. Психологічний експеримент дозволяє перевірити, чи зберігається результат при повторенні процедури, чи залежить він від контексту, чи є стабільним у різних вибірках. Таким чином, експеримент не лише виявляє причинно-наслідкові зв'язки, а й перевіряє їхню стійкість і універсальність.

Ідея причинності у психології виникла під впливом природничих наук XIX століття. Ісаак Ньютон сформулював класичний детермінізм – уявлення про те, що кожна подія має конкретну причину, і якщо її знайти, можна передбачити наслідок. Ця логіка спонукала дослідників шукати аналогічні закони в психіці. Вільгельм Вундт вважав, що психічні процеси можна досліджувати так само точно, як фізичні явища, якщо створити контрольовані умови для їх вивчення. Відкриття його лабораторії стало символічним актом відокремлення психології від філософії – переходом до емпіричного пошуку причин.

Однак перенесення механістичного розуміння причинності у психологію виявилось не таким простим. Людська поведінка не завжди підпорядковується лінійним законам: між стимулом і реакцією існують посередники – свідомість, досвід, соціальний контекст, емоційний стан. Саме тому психологічний експеримент ніколи не був просто копією фізичного. Його головне завдання полягає у виявленні не зовнішніх, а внутрішніх детермінант поведінки – тих механізмів, через які зовнішні події впливають на суб'єктивний світ людини.

На відміну від природничих дисциплін, у психології об'єкт дослідження – людина, суб'єкт із власною свідомістю, емоціями, інтенціями. Це накладає суттєві методологічні обмеження: неможливо повністю стандартизувати умови або усунути вплив очікувань і мотивацій. Тому логіка організації психологічного експерименту включає не лише технічні, а й етичні, комунікативні та когнітивні аспекти. Психіка людини не піддається безпосередньому спостереженню; вона виявляється лише через поведінку, реакції, вербальні повідомлення, фізіологічні показники. Тому будь-який експеримент вимагає складної операціоналізації – тобто перетворення абстрактних понять («тривожність», «мотивація», «самооцінка») у вимірювані показники.

Проблема причинності у психології має кілька рівнів. На емпіричному рівні це питання про те, як одна змінна (наприклад, рівень стресу) впливає на іншу (увагу, пам'ять, працездатність). На теоретичному рівні – це пошук пояснювальних моделей, що описують механізм цього впливу (наприклад, теорія когнітивного навантаження чи модель емоційного збудження). А на метатеоретичному рівні – це роздуми про саму природу причинності у психіці: чи є вона механічною, системною, самоорганізованою, нелінійною? Таким чином, експеримент у психології виступає не лише інструментом перевірки гіпотез, а й засобом осмислення природи психічного детермінізму.

Експеримент у психології дозволяє перейти від кореляційного мислення до причинного. У кореляційних дослідженнях ми можемо лише спостерігати, що дві змінні змінюються разом (наприклад, чим вищий рівень тривожності, тим гірші результати навчання), але не можемо сказати, чому це відбувається. Експеримент же дає змогу створити ситуацію, у якій ми контролюємо всі сторонні фактори й навмисно змінюємо одну змінну, щоб побачити, чи дійсно вона викликає зміни в іншій. Саме завдяки цьому експеримент є єдиним методом, що дозволяє встановити причинно-наслідковий зв'язок із науковою точністю.

Водночас причинність у психології має не лише лінійний, а й системний характер. Поведінка людини визначається не однією причиною, а цілою мережею взаємодіючих факторів – біологічних, соціальних, когнітивних, культурних. Тому сучасні експериментальні дослідження дедалі частіше використовують багатофакторні дизайни, у яких одночасно варіюються кілька змінних. Наприклад, вивчаючи вплив стресу на когнітивну ефективність, дослідник може враховувати не лише інтенсивність стресора, а й тип

особистості, рівень мотивації, соціальну підтримку. Це дозволяє моделювати складні реальні ситуації і виявляти не просто прямі, а модеруючі та медіаторні зв'язки між змінними.

Питання про те, що саме є причиною у психологічному процесі, залишається складним і дискусійним. Якщо у фізиці причинність зазвичай розуміють як послідовність подій у часі (подія А викликає подію Б), то в психології причинність часто має смисловий характер: мотив викликає дію, переконання впливає на емоцію, очікування змінює сприйняття. Тобто причиною виступає не об'єктивний факт, а його інтерпретація свідомістю. Саме тому експериментальна психологія повинна враховувати як об'єктивні умови (стимули, ситуацію), так і суб'єктивні значення, які надає їм людина.

Тому одним з найважливіших методологічних викликів, що постають перед психологом-дослідником, є відмежування справжніх причинно-наслідкових зв'язків від випадкових або уявних співвідношень.

Для того щоб зробити обґрунтоване твердження про наявність причинного зв'язку між двома змінними, повинні бути дотримані три базові умови, що стали класичними у філософії науки та психології.

По-перше, змінна, яка розглядається як причина, має передувати у часі змінній, що визначається як наслідок. Це правило хронологічного пріоритету гарантує напрямок впливу та виключає зворотну причинність.

По-друге, між змінними повинен бути наявний стабільний статистичний зв'язок, який підтверджується достовірними методами аналізу даних. Іншими словами, спостерігається регулярна, а не випадкова взаємозалежність показників

По-третє, цей зв'язок не повинен бути зумовлений якимись іншими, сторонніми чинниками. Тобто, дослідник має виключити або принаймні врахувати всі альтернативні пояснення, які можуть зумовити зміну залежної змінної.

В експериментальній психології ці три умови реалізуються шляхом ретельного проєктування дослідження: незалежна змінна вводиться в ситуацію свідомо й керовано, залежна змінна вимірюється з урахуванням надійності та валідності, а потенційні побічні змінні – або строго контролюються, або ізолюються. Саме експериментальний контроль дозволяє об'єктивно інтерпретувати отримані результати та стверджувати, що виявлений ефект є наслідком саме експериментального впливу, а не результатом сторонніх або випадкових факторів.

6.2. Характеристика та логіка психологічного експерименту

Поява експериментального методу в психології стала однією з найважливіших подій у розвитку цієї науки. До другої половини XIX століття знання про психіку формувалися переважно у межах філософських систем і описових підходів, що ґрунтувалися на інтерпретаціях та внутрішньому спостереженні. І хоча такі методи дозволяли накопичувати феноменологічні

уявлення про свідомість, вони не могли забезпечити наукової точності або перевірюваності. Саме тому роботи Г. Гельмгольца, Е. Вебера і Г. Фехнера, які запровадили перші кількісні закономірності психофізики, стали тим інтелектуальним підґрунтям, що підштовхнуло Вільгельма Вундта до створення першої експериментальної лабораторії психології в 1879 році. Відтоді експериментальний метод почав виконувати роль основного інструмента наукового відкриття причинних механізмів поведінки та психічних процесів.

Психологічний експеримент – це така форма наукового дослідження, у якій дослідник активно змінює певні умови, вводячи експериментальний вплив, та вимірює наслідки цієї зміни за допомогою стандартизованих методів. У центрі експерименту лежить ідея керованої маніпуляції незалежною змінною та реєстрація змін залежної змінної. Якщо зміни у залежній змінній стабільно з'являються після введення впливу, і якщо всі інші фактори контролюються або утримуються сталими, дослідник отримує можливість зробити висновок про причинний характер взаємозв'язку.

Логіка експерименту ґрунтується на тому, що дослідник ізолює одну змінну й оцінює ту зміну, яка виникає після її варіювання. Такий підхід дає змогу відтворювати явище багаторазово, змінювати рівні впливу, створювати контрольні групи та перевіряти альтернативні пояснення.

Одним із найважливіших аспектів експериментальної логіки є операціоналізація. Психолог працює з такими поняттями, як «увага», «мотивація», «стрес», «самоконтроль», які не мають безпосередньої фізичної форми. Тому він повинен розробити індикатори, за допомогою яких абстрактні процеси можна зробити доступними для вимірювання: час реакції, кількість правильних відповідей, фізіологічні параметри, поведінкові показники, інтенсивність вербальних відповідей тощо. Якість експерименту значною мірою залежить від того, наскільки коректно побудовано цю операціоналізацію.

Ще однією ключовою складовою логіки експерименту є варіювання незалежної змінної. Дослідник не просто спостерігає її прояви, а створює її рівні або умови. Наприклад, змінює інтенсивність звукового стимулу, тривалість експозиції візуального матеріалу, рівень стресогенності ситуації або складність когнітивного завдання. Порівнюючи результати в умовах різної інтенсивності впливу, він може визначити не лише факт впливу, а й форму залежності – лінійну, криволінійну, оптимальну (наприклад, інвертовану U-подібну, як у законі Єркса–Додсона).

Центральним елементом експериментальної логіки є принцип контролю. Він полягає в тому, що всі чинники, окрім незалежної змінної, мають бути усунуті або утримані сталими. Наприклад, якщо експеримент вивчає ефект шуму на продуктивність, то рівень освітлення, час доби, температура приміщення, інструкція, тривалість завдання та інші умови мають бути однаковими для всіх учасників. Саме контроль можливих побічних

впливів робить експеримент досить «чистим» для того, щоб робити причинні висновки.

Порівняння є центральним механізмом, що дає змогу виявити різницю, яка виникла після маніпуляції. Дослідник порівнює поведінку двох або більше груп, які відрізняються лише рівнем впливу. Таким чином, класифікація експериментальних умов – це не формальна процедура, а спосіб реалізувати логіку зіставлення.

Усі експериментальні плани, попри різноманіття, базуються на однаковій структурі: введення впливу, вимірювання результату, аналіз отриманої різниці. Однак реалізувати цю структуру можна по-різному. У найширшому розумінні види психологічного експерименту поділяють за умовами проведення, за метою, за способом контролю незалежної змінної та за способом реалізації.

Якщо розглядати експеримент за умовами проведення, то виділяють лабораторний, природний і польовий варіанти. Лабораторний експеримент проводиться у спеціально обладнаному просторі, де всі змінні можуть бути стабілізовані та жорстко контрольовані. Він дає змогу досягти високої внутрішньої валідності, оскільки дослідник точно знає, які фактори впливають на результат. Проте лабораторні умови часто створюють ризик низької зовнішньої валідності: поведінка в штучному середовищі може відрізнитися від поведінки у реальному житті. Природний експеримент, навпаки, дозволяє вивчати психічні процеси у звичних для людини умовах, що підвищує екологічну валідність, але часто обмежує можливості контролю побічних змінних. Польовий експеримент займає проміжну позицію: він проводиться в реальному середовищі, але за умов навмисно організованої маніпуляції. Наприклад, у школах, організаціях чи на вулиці дослідник може змінювати певні параметри ситуації, проте контроль усіх сторонніх чинників залишається неможливим.

За метою проведення експеримент може бути пошуковим, підтверджуючим або критичним. Пошуковий експеримент спрямований на виявлення невідомих закономірностей або перевірку загальних можливих механізмів. Він є гнучким, допускає коригування методики в процесі роботи. Підтверджуючий експеримент, навпаки, спрямований на перевірку строго сформульованої гіпотези, яка ґрунтується на теоретичних передумовах. Він вимагає жорсткого контролю та стандартизації. Критичний експеримент застосовується в ситуаціях теоретичної конкуренції, коли необхідно емпірично відрізнити дві або більше теорії, які прогнозують різні результати. Такий формат дослідження має високу пояснювальну силу, оскільки він дозволяє виявити, яка система припущень точніше відображає реальні механізми психіки.

За способом реалізації експерименти поділяють на реальні, уявні та ідеальні. Реальний експеримент здійснюється у фактичному середовищі за участю досліджуваних. Ідеальний експеримент – це теоретична модель, яка демонструє логіку причинно-наслідкових зв'язків, але не може бути проведена

буквально (наприклад, ізолювання всіх сторонніх чинників у чистому вигляді). Уявний експеримент виконується як інтелектуальна модель, коли дослідник створює уявний набір умов і прогнозує можливі результати (метод широко використовувався у когнітивній психології). Хоча він не є емпіричною процедурою, уявний експеримент часто є інструментом побудови теоретичних гіпотез, які згодом перевіряються реальним експериментом.

Інший важливий критерій – ступінь контролю незалежної змінної. У спровокованому експерименті незалежна змінна задається дослідником штучно: створюється маніпуляція, умова або стимул. У експерименті, на який посилаються (*ex post facto*), дослідник не може втручатися в незалежну змінну, оскільки вона вже існує природно. Наприклад, він порівнює людей із різним рівнем тривожності або типом темпераменту. Хоча такі дослідження називають експериментами умовно, вони не забезпечують повного експериментального контролю, оскільки незалежна змінна не варіюється довільно. Це значно обмежує можливість робити причинні висновки, тому такі дослідження часто класифікують як квазіекспериментальні.

У сучасній психології експеримент відіграє не лише дослідницьку, а й практичну функцію. Саме він лежить в основі доказових протоколів у клінічній психології, психотерапії, нейропсихології, навчанні та організаційних інтервенціях. Наприклад, ефективність когнітивно-поведінкової терапії була неодноразово підтверджена у рандомізованих контрольованих випробуваннях – найсуворішому типі експериментальних планів, який дозволяє виключити альтернативні тлумачення. У галузі освітньої психології експеримент дозволив виявити вплив способів подання інформації, методів навчання, форм зворотного зв'язку, особливостей мотивації на успішність учнів. У психології праці експериментальні методи застосовуються для оцінювання умов робочого середовища, вивчення впливу шуму, освітлення чи ритму праці на продуктивність. У соціальній психології експеримент дав змогу виявити закономірності групового тиску, конформності, міжособистісного впливу, агресії та альтруїзму.

Логіка експерименту, таким чином, включає кілька взаємопов'язаних етапів: формування гіпотези, визначення незалежної та залежної змінних, конструювання умов, вибір плану експерименту, добір вибірки, реалізація маніпуляції, реєстрація реакцій, аналіз даних та інтерпретація результатів. Усі ці етапи взаємопов'язані та підпорядковані єдиній меті: максимально точно і мінімально викривлено виявити причинний механізм.

Окрему увагу приділяють питанням внутрішньої валідності. Вона означає, що встановлений у дослідженні причинний зв'язок є справжнім, а не наслідком впливу сторонніх факторів. Для забезпечення внутрішньої валідності використовуються рандомізація, стандартизація умов, контроль конфаундів, використання контрольних груп, подвійний сліпий метод, автоматизація процедур. Зовнішня валідність – це можливість узагальнити результати експерименту на ширшу сукупність ситуацій, груп чи контекстів. Між внутрішньою та зовнішньою валідністю часто існує напруження: чим

сильніше контроль, тим штучніші умови і тим важче перенести результати на реальне життя. Тому різні типи експериментів намагаються збалансувати ці дві вимоги.

Наприкінці важливо підкреслити, що психологічний експеримент не є самоціллю. Він – це метод мислення, спрямований на відкриття закономірностей психічного життя. Його головна сила полягає у поєднанні теоретичної строгості та емпіричної перевірюваності. Завдяки експерименту психологія рухається від інтуїтивних уявлень до опису реальних механізмів людської поведінки, від вражень – до доказів, від припущень – до закономірностей. У цьому полягає її наукова природа і провідна роль експериментального методу у структурі психологічного знання.

6.3. Види змінних у психологічному експерименті

У будь-якому науковому експерименті поняття змінної є центральним, оскільки саме через її контроль, маніпуляцію та вимірювання відбувається перевірка гіпотези. У психологічному експерименті змінні мають особливе значення, адже вони дозволяють зв'язати абстрактні теоретичні поняття, такі як «мотивація», «стрес» чи «увага» тощо, з конкретними емпіричними показниками, які можна спостерігати та кількісно оцінити.

Змінна – це будь-яка властивість, характеристика або фактор, що може набувати різних значень у різних умовах або в різних осіб. Логіка експерименту ґрунтується на взаємозв'язку між змінними: зміна однієї веде до зміни іншої. Таким чином, змінні виступають не просто об'єктами вимірювання, а структурними елементами причинно-наслідкового ланцюга. Будь-яке експериментальне дослідження відбувається у просторі взаємодії змінних, де дослідник прагне ізолювати та виміряти ефекти певного впливу.

У цій системі виділяються три основні категорії змінних, які необхідні для перевірки експериментального впливу:

- незалежна змінна;
- залежна змінна;
- побічна змінна.

Незалежна змінна – це змінна, яку дослідник навмисно вводить або модифікує в експериментальній ситуації з метою виявлення її впливу на інші психологічні характеристики чи процеси. Вона виконує функцію причинного чинника в експериментальній моделі. Незалежна змінна задається як така, що викликає або ініціює зміни в інших змінних, і саме її ефект є предметом аналізу. Вона може бути бінарною (присутність або відсутність впливу), багаторівневою (різні типи стимулів чи умов), модифікованою за інтенсивністю або якістю.

У психології незалежна змінна може мати найрізноманітнішу природу. Це не лише фізичні стимули (світло, звук, час реакції), а й соціальні, когнітивні чи емоційні фактори. Дослідник може змінювати тип інструкції, характер підкріплення, рівень складності завдання, вид стимульного матеріалу, умови

взаємодії тощо. Важливо, щоб незалежна змінна була чітко операціоналізована – тобто щоб було зрозуміло, у якій формі вона подається учасникам і як саме варіюється.

У структурі незалежних змінних прийнято виділяти: структуровані (маніпульовані) та неструктуровані (суб'єктні) змінні.

Структуровані або маніпульовані змінні – ті, які повністю контролюються експериментатором.

Неструктуровані або суб'єктні змінні – ті, які не змінюються в межах експерименту, але фіксуються як індивідуальні характеристики досліджуваних (наприклад, стать, рівень освіти, особистісні риси, когнітивні стилі). Хоча вони не є повноцінними незалежними змінними, вони часто використовуються для стратифікації або групування даних та подальшого аналізу їх впливу.

Ключова вимога до незалежної змінної – чистота її маніпуляції. Вона має змінюватися лише в одному параметрі, без супутніх ефектів. Якщо дослідник одночасно змінює кілька характеристик стимулу (наприклад, зміст і тон висловлювання), то стає неможливо визначити, яка саме з них спричинила зміну поведінки. Таке явище називають змішуванням змінних (confounding). Саме тому психологічні експерименти ретельно проектується, часто вдаючись до попереднього тестування стимулів або пілотних досліджень.

З методологічного погляду незалежна змінна – це інструмент перевірки теоретичних гіпотез. Вона є формою переведення теоретичного поняття у вимірювану площину. Наприклад, теорія когнітивного навантаження передбачає, що обмеження робочої пам'яті знижує ефективність засвоєння матеріалу. Незалежною змінною тут виступає рівень когнітивного навантаження, який можна змінювати, ускладнюючи завдання або додаючи відволікаючі стимули. Таким чином, незалежна змінна стає механізмом перевірки теоретичного припущення.

Ще однією вимогою є внутрішня логічна узгодженість незалежної змінної з рештою елементів експерименту. Вона має відповідати гіпотезі, об'єкту та предмету дослідження, не виходячи за межі концептуальної моделі. Якщо дослідник, наприклад, вивчає вплив емоційного стану на пам'ять, але як незалежну змінну вводить фізичну активність, то логічний зв'язок руйнується.

Залежна змінна – це показник, який демонструє ефект, що виникає під впливом незалежної змінної, яку дослідник вимірює після здійснення експериментального впливу. Вона виступає як індикатор змін, викликаних незалежною змінною, і відображає ті параметри психіки або поведінки, які є критичними для підтвердження або спростування дослідницької гіпотези. У психології залежна змінна може мати як кількісне, так і якісне вираження: це можуть бути числові показники (час реакції, кількість правильних відповідей, рівень фізіологічної активації), вербальні характеристики (оцінка по шкалах, висловлювання), поведінкові прояви (частота дій, типові реакції), а також зміни в емоційній або мотиваційній сфері.

Ключове завдання дослідника – забезпечити валідність і надійність залежної змінної. Валідність означає, що змінна справді відображає те явище, яке передбачається виміряти, а не побічний ефект. Надійність – це стабільність результатів при повторному вимірюванні. Якщо тест або показник дає випадкові результати, то жодні висновки про вплив незалежної змінної не будуть науково обґрунтованими.

Залежна змінна також має бути чутливою до змін. Саме тому у психологічних експериментах часто комбінують кілька типів вимірювань: наприклад, фіксують не лише кількість правильних відповідей, а й швидкість їх подачі, самооцінку учасника та фізіологічну реакцію. Такий підхід дозволяє побачити більш повну картину процесу, що досліджується.

Ще одна важлива властивість залежної змінної – відповідність рівня вимірювання типу гіпотези. Якщо гіпотеза передбачає лінійну залежність, необхідно мати метричну шкалу, яка дозволяє застосовувати кореляційний або дисперсійний аналіз. Якщо ж змінна номінальна (наприклад, тип відповіді: «так» чи «ні»), то доцільно використовувати непараметричні методи. Невідповідність між рівнем вимірювання і статистичною обробкою є частою методологічною помилкою, яка призводить до хибних висновків.

У широкому сенсі залежна змінна – це не просто дані, а модель поведінки системи. Вона відображає, як психіка реагує на вплив середовища, як внутрішні процеси перетворюють стимул у відповідь. У цьому розумінні правильне визначення залежної змінної – це, по суті, визначення того, що саме психолог вважає суттю досліджуваного процесу.

Побічні змінні – це всі ті змінні, які можуть вплинути на результати експерименту, але не є предметом дослідження. Вони створюють методологічні ризики, оскільки можуть спотворити залежність між незалежною і залежною змінними, породжуючи хибні висновки. Побічні змінні можуть бути зовнішніми (наприклад, умови проведення експерименту, фізичне середовище, час доби) або внутрішніми (індивідуальні особливості учасників, особливості взаємодії з експериментатором). Їхній вплив часто прихований, проте саме вони становлять головну загрозу внутрішній валідності. Якщо побічні змінні не контролюються, то жодна впевненість у тому, що спостережуваний ефект зумовлений саме незалежною змінною, неможлива.

У психології, через високу варіативність людської поведінки та суб'єктивного досвіду, побічні змінні набувають особливої ваги. Їхні джерела надзвичайно різноманітні: когнітивні установки учасників, соціальні очікування, мотиваційні стани, рівень тривожності, попередній досвід, емоційне забарвлення ситуації тощо. Особливу загрозу становлять побічні змінні, які залишаються неврахованими або неусвідомленими самим дослідником.

Загалом, на етапі планування експерименту дослідник має не лише чітко визначити основні змінні, але й скласти повну схему потенційних джерел варіації та впливу. Це включає побудову концептуальної моделі змінних,

визначення способів їх вимірювання, засобів контролю та системи аналізу отриманих даних. Такий підхід забезпечує логічну цілісність дослідження, його валідність, надійність і наукову значущість.

6.4. Психологічні та ситуативні джерела побічних впливів в експерименті

Організація психологічного експерименту передбачає створення умов, у яких дослідник може контролювати незалежну змінну, вимірювати зміни залежної та мінімізувати вплив побічних чинників. Однак у реальному процесі взаємодії з людьми уникнути додаткових впливів майже неможливо. Людська поведінка чутлива до контексту, соціальних сигналів, інтерпретацій, очікувань і навіть до тонких невербальних реакцій експериментатора. Усе це створює широкий спектр побічних змінних, які можуть викривити результати дослідження, зменшити внутрішню валідність та поставити під питання достовірність причинного висновку.

Побічні змінні – це будь-які чинники, окрім незалежної змінної, які можуть систематично впливати на залежну змінну. На відміну від випадкових помилок, побічні впливи мають структурну природу: вони діють неоднаково в різних експериментальних умовах або по-різному для різних учасників. Саме тому їх не можна усунути простим збільшенням вибірки або повторним проведенням експерименту. Вони вписані в саму логіку взаємодії між дослідником, процедурою та учасником, і тому мають бути ретельно усвідомлені та контрольовані.

Сучасна експериментальна психологія розрізняє три великі групи джерел побічних впливів: чинники, що походять від експериментатора; чинники, пов'язані з особистістю або поведінкою досліджуваних; та феномени, що виникають унаслідок взаємодії між ними. Кожен із цих блоків має власну специфіку, пов'язану з психологічними процесами інтерпретації, соціальної перцепції, мотивації та комунікативної взаємодії.

Першу групу становлять впливи, пов'язані з фігурою експериментатора. Дослідник, навіть не усвідомлюючи цього, може виступати активним джерелом систематичних помилок. Найпоширенішим проявом є ефект очікування, відомий також як ефект Пігмаліону. Суть його полягає в тому, що експериментатор, переконаний у правильності своєї гіпотези або у здатностях певних учасників, може ненавмисно передавати їм свої очікування через інтонацію, міміку, стиль спілкування або акценти у зворотному зв'язку. Цей вплив часто діє на рівні несвідомих сигналів: легка посмішка, більш доброзичливий тон, підбадьорливий жест – усе це може підвищувати мотивацію учасника та впливати на результат. Класичні дослідження Розенталя показали, що очікування експериментатора здатне збільшувати або зменшувати продуктивність учасників навіть у тих ситуаціях, де об'єктивних підстав для цього немає. Ефект Пігмаліону демонструє, наскільки

експериментатор може стати невидимою змінною, що визначає поведінку досліджуваних.

Іншим джерелом систематичних впливів є особистісні особливості експериментатора, які формують так званий «ефект експериментатора». Темперамент, рівень домінантності, стиль комунікації, соціальна чуйність, рівень впевненості або тривожності – усі ці характеристики можуть впливати на атмосферу експерименту. Наприклад, дослідження показали, що тривожні експериментатори частіше передають учасникам невпевнені сигнали, що може впливати на виконання складних когнітивних завдань. Домінантні експериментатори, навпаки, інколи стимулюють підвищену напругу, що може зменшувати точність виконання завдань, пов'язаних зі швидкістю реакції. Суттєвим чинником є також стать експериментатора. Відомо, що чоловіки-експериментатори можуть збільшувати ризик соціально бажаних відповідей у жінок-досліджуваних, тоді як жінки-експериментатори інколи викликають підвищені змагальні або репрезентативні установки у чоловіків. Ці ефекти можуть не лише викривляти результати, а й породжувати хибні інтерпретації, якщо експериментатор не враховує власний внесок у перебіг дослідження.

Також суттєвою загрозою є інтерпретація результатів з боку дослідника. Ефект підтвердження – схильність шукати, помічати і надавати більшої ваги тим даним, які підтримують попередні очікування – є серйозною небезпекою для об'єктивності. Навіть при дотриманні процедурних стандартів експериментатор може ненавмисно впливати на вибір способу статистичного аналізу, критерії відбракування даних або інтерпретацію маргінально значущих результатів. Саме тому сучасні дослідження вимагають передреєстрації гіпотез, використання осліплених процедур аналізу та суворих критеріїв обробки даних, щоб мінімізувати вплив експериментатора на результат.

Друга велика група побічних змінних пов'язана з поведінкою, очікуваннями та інтерпретаціями самих досліджуваних. Людина в експерименті рідко є пасивним об'єктом: вона активно формує уявлення про ситуацію, ставить гіпотези щодо мети дослідження, оцінює очікування дослідника і модифікує свою поведінку відповідно до соціальних норм чи бажання справити враження. Найвідомішим є Хоторнський ефект – підвищення продуктивності учасників просто через факт участі у дослідженні та відчуття уваги з боку спостерігачів. Ефект отримав свою назву від досліджень на фабриці Western Electric у Хоторні, де будь-які зміни умов праці, в тому числі навіть нейтральні, спричиняли підвищення продуктивності працівників лише тому, що вони знали про експериментальне спостереження. У психологічних експериментах цей ефект проявляється у прагненні учасників демонструвати більш старанну, зосереджену або соціально бажану поведінку.

Іншим впливом з боку учасників є ефект плацебо. Він виникає тоді, коли очікування учасника щодо впливу стають самостійним джерелом змін у поведінці чи переживаннях. У клінічній психології та нейропсихології плацебо-ефект може суттєво спотворювати результати експериментів, якщо не

створено контрольної групи з нейтральною процедурою. Проте аналогічні механізми діють і в соціальній, і в когнітивній психології. Наприклад, учасник, переконаний, що певна стимуляція покращує увагу, демонструватиме кращі результати навіть за відсутності реальної маніпуляції, а той, хто вважає, що завдання є складним, може погіршувати результати через підвищену тривожність.

Важливим джерелом систематичних впливів є мотивація досліджуваних. Деякі учасники прагнуть максимально точно виконати інструкції, інші – мінімізувати зусилля, а треті – навмисно змінювати поведінку, щоб заплутати експериментатора або прискорити завершення процедури. Суттєвим є також рівень інтерпретації ситуації: досліджувані нерідко намагаються зрозуміти «справжню» мету експерименту, і цей процес може радикально змінити їхню поведінку. Наприклад, якщо учасник вважає, що дослідник очікує високих результатів, він може несвідомо підвищувати старанність, тоді як у ситуаціях, де учасник відчуває, що експеримент спрямований на виявлення його слабких сторін, може виникати захисна поведінка, уникнення або демонстрація соціально бажаних відповідей.

Чинником, який часто недооцінюється, є соціальні порівняння між умовами експерименту. Учасники можуть співвідносити свою поведінку з поведінкою інших, яких вони бачать перед собою або про яких чують у процесі експерименту. Наприклад, якщо учасник дізнається, що інші виконують завдання швидше, він може несвідомо прискорювати темп, навіть якщо це не відповідає природному стилю роботи. Такі порівняння можуть відбуватися і на рівні очікувань: учасники, які потрапляють до умов, що «виглядають складнішими», можуть демонструвати вищу напруженість або тривожність.

Нарешті, третя група охоплює феномени, що виникають унаслідок взаємодії експериментатора і досліджуваного. Ці процеси є соціально-психологічними за своєю природою і часто мають двосторонній характер. Найвідомішим є явище соціальної фасилітації – тенденція до підвищення продуктивності в присутності спостерігача при виконанні простих або добре засвоєних завдань. Однак цей ефект має дзеркальну протилежність – соціальну інгібіцію, коли складні, незвичні або нові завдання виконуються гірше в присутності інших. Присутність експериментатора як соціального спостерігача може стимулювати підвищення швидкості реакції, але зменшувати точність або збільшувати кількість помилок. Така взаємодія є надзвичайно складною, оскільки залежить від індивідуальних особливостей учасників (рівня екстраверсії, тривожності, потреби у схваленні) та характеристик експериментатора (домінантності, професійності, соціальної привабливості).

Взаємодія може набувати і складніших форм. Наприклад, учасник може прагнути сподобатися експериментатору, демонструючи версію поведінки, яка, на його думку, відповідає очікуванням дослідника. Це називають ефектом соціальної бажаності. Інші учасники можуть сприймати експеримент як

змагання, намагаючись перевершити уявні стандарти. Іноді дослідники спостерігають так званий ефект компенсації – коли контрольна група, яка не отримує експериментального впливу, намагається докласти більше зусиль, щоб «не відстати», створюючи додаткові конфаунди.

Таким чином, побічні впливи в експерименті є комплексними та багаторівневими, оскільки пов'язані з трьома джерелами: експериментатором, досліджуванним і взаємодією між ними. Вони можуть суттєво викривлювати результати, знижувати внутрішню валідність і ставити під питання висновки про причинність. Саме тому сучасні методологічні стандарти вимагають структурованих засобів контролю: подвійних сліпих процедур, автоматизації інструкцій, стандартизованої невербальної поведінки експериментатора, використання контрольних груп плацебо, ретельної рандомізації, попередньої реєстрації гіпотез, оцінювання мотивації та інтерпретацій досліджуваних, а також моніторингу соціальної динаміки під час експерименту.

6.5. Методи контролю побічних змінних в експерименті

Організація психологічного експерименту потребує не лише точного конструювання незалежної змінної та ретельного вимірювання залежної, але й системного контролю низки побічних впливів, які можуть спотворювати причинно-наслідкові висновки. Будь-яка зміна в поведінці досліджуваних може бути спричинена не експериментальним фактором, а побічними чинниками – очікуваннями, інтерпретаціями, мотивацією, особистісними характеристиками, соціальними реакціями або властивостями експериментатора. Саме тому методи контролю побічних змінних є центральним елементом логіки експерименту: вони дозволяють мінімізувати конфаунди, забезпечити внутрішню валідність та отримати результати, які можна інтерпретувати як наслідок впливу незалежної змінної. Контроль побічних змінних має на меті створення максимальної еквівалентності умов, за яких спостерігаються зміни у залежній змінній, та усунення впливів, які можуть маскувати або імітувати експериментальний ефект. Цей контроль може мати різну природу – процедурну, психологічну, технічну, організаційну чи статистичну, – проте його цілі завжди залишаються однаковими: забезпечення чистоти причинного висновку та збереження логіки експериментального методу.

Одним із найефективніших засобів контролю є подвійний сліпий метод, який передбачає, що ні досліджувані, ні експериментатори, які взаємодіють з ними, не знають, у яку групу призначений кожен учасник – експериментальну чи контрольну. Цей підхід усуває дві групи критично важливих побічних впливів: очікування експериментатора та очікування досліджуваних. Якщо експериментатор не знає, хто отримує реальний вплив, він не може, навіть ненавмисно, передати учасникам підказки щодо бажаної поведінки. Так само учасники, не знаючи, чи отримують вони активний вплив, не можуть скоригувати свої реакції відповідно до очікуваних ефектів. Наприклад, у

дослідженні впливу тренінгу уважності на рівень стресу, експериментатор, який знає, що певні учасники отримали «справжній» тренінг, може неусвідомлено демонструвати цим учасникам більш підтримувальну поведінку, сприяючи покращенню їхніх результатів. Подвійний сліпий метод усуває ймовірність таких сигналів, забезпечуючи реальну ізоляцію впливу незалежної змінної від соціальних очікувань. Цей метод є стандартом у клінічних дослідженнях, але дедалі частіше застосовується і в когнітивних та соціально-психологічних, де експериментальна ситуація може стати джерелом навіювання або упереджень.

Іншою важливою стратегією є автоматизація експерименту, яка передбачає використання комп'ютеризованих процедур подання стимулів, фіксації реакцій і збору даних. Автоматизовані експерименти усувають вплив експериментатора на перебіг дослідження, оскільки людина не бере участі в безпосередньому пред'явленні стимулів, поясненні завдань або реєстрації відповідей. Комп'ютер забезпечує однакові інструкції всім учасникам, стабільний час подання стимулів, точні параметри експозиції та відсутність соціального тиску. Наприклад, у дослідженні когнітивної інгібіції завдання Струпа, подане через комп'ютер, гарантує, що кожен стимул з'являється з однаковим інтервалом та в однакових умовах освітлення й контрастності, а реакції реєструються з точністю до мілісекунд. Це усуває можливість того, що експериментатор своїм невербальним стилем або манерою спілкування вплине на швидкість і точність відповідей. У клінічних дослідженнях автоматизація також мінімізує соціальну бажаність відповідей, оскільки учасник взаємодіє не з людиною, а з неперсоніфікованою системою, що знижує тривожність та зменшує вплив мотиваційних чинників.

У деяких випадках забезпечення чистоти експериментальної ситуації потребує застосування методу обману, який полягає у наданні учасникам неправдивої або частково спотвореної інформації про справжню мету дослідження. Метод обману використовується тоді, коли знання учасника про реальну гіпотезу може радикально змінити його поведінку і зробити результати недостовірними. Наприклад, у класичних дослідженнях конформності Ашера учасникам повідомляли, що вони беруть участь у тесті на зорове сприйняття, хоча справжньою метою було вивчення впливу групового тиску на індивідуальні рішення. Без обману учасники могли б навмисно протидіяти тиску групи або адаптувати свою поведінку, знаючи про інтерес дослідників до конформності. Метод обману дозволяє виявити природну поведінку у соціально значущих ситуаціях, проте потребує обов'язкового постекспериментального інтерв'ю, у межах якого учасники пояснюють справжню мету дослідження, відповідають на його запитання і знімають можливі негативні переживання, пов'язані з тимчасовим введенням в оману. Етичний кодекс АРА (American Psychological Association) наголошує на необхідності мінімізації психологічного дискомфорту та забезпечення добровільної згоди учасників після пояснення сутності експерименту.

Близьким, але не тотожним до методу обману є метод прихованого експерименту, який використовується тоді, коли сам факт участі в експерименті може істотно змінити поведінку учасників. У таких випадках досліджувані не знають, що перебувають у складі експериментальної ситуації, і поведуться природно. Наприклад, у дослідженні соціальної фасилітації можна встановити приховані камери у коридорі університету, щоб аналізувати, як студенти проходять повз незнайомих людей або як зупиняються перед інформаційними стендами. Або у дослідженні допоміжної поведінки можна створити ситуацію випадкового падіння предметів у громадському місці і спостерігати, чи допоможуть перехожі. Такі методи дозволяють уникнути штучності лабораторних умов, але застосовуються лише тоді, коли не порушують етичних норм, не створюють небезпеки і не вторгаються в приватне життя учасників. Після завершення експерименту дослідники повинні провести інформування та отримати постфактум згоду на використання зібраних даних.

Ще одним важливим засобом контролю є оцінка того, як учасники сприйняли експериментальну ситуацію. Дослідник не може бути впевнений, що учасники зрозуміли інструкцію, не здогадалися про гіпотезу або не сформували альтернативні очікування щодо мети дослідження. Тому після завершення експерименту часто проводиться постекспериментальне інтерв'ю, у межах якого учасників запитують, що вони думали про мету дослідження, чи відчували вони вплив експериментальної маніпуляції, чи здогадалися про структуру завдань. Це дозволяє виявити учасників, які зрозуміли гіпотезу або поводитися не відповідно до інструкцій, а на основі власних припущень. Таких учасників можуть виключати із фінального аналізу, якщо їхня поведінка потенційно спотворює результати. Наприклад, у дослідженні впливу підкріплення на мотивацію учасник, який здогадався, що дослідник прагне вивчити ефект похвали, може навмисно змінювати свою поведінку, щоб відповідати уявленню про бажану реакцію. Виключення таких випадків із вибірки підвищує чистоту експериментального ефекту.

Контроль побічних змінних може здійснюватися і через методологічні прийоми рандомізації та балансування. Рандомізація дозволяє випадково розподілити учасників по експериментальних умовах так, щоб індивідуальні відмінності не спричиняли систематичних перекосів. Якщо учасники вибірки випадковим чином потрапляють у різні групи, ризик того, що певні характеристики (наприклад, інтелект, мотивація, рівень тривожності) будуть зосереджені в одній групі, мінімізується. Рандомізація застосовується також у внутрішньогрупових планах для визначення порядку подання умов, щоб уникнути ефектів послідовності, таких як втома чи практика. Балансування доповнює рандомізацію, забезпечуючи рівну представленість певних факторів у різних умовах. Наприклад, якщо експеримент має дві умови, порядок яких може вплинути на результати, дослідник формує дві підгрупи: одна виконує спочатку умову А, потім Б, а інша – навпаки. Це дозволяє нейтралізувати ефекти порядку та стабілізувати варіативність даних.

Існують і більш складні методи контролю, наприклад стандартизація інструкцій, яка передбачає, що всі учасники отримують однакові пояснення, зачитані дослівно або записані на аудіо чи відео. Такий підхід усуває можливість індивідуальних варіацій у вербальному стилі експериментатора. Стандартизація стосується також місця, освітлення, часу проведення, температури та інших фізичних умов, які можуть впливати на результати. Використання подвійного сліпого кодування даних, коли особа, яка вводить дані або виконує статистичний аналіз, не знає, до якої групи належать учасники, додатково знижує ризик упередженості при інтерпретації. У деяких випадках застосовують контрольні групи плацебо, які отримують імітацію впливу, що дозволяє розмежувати реальний ефект маніпуляції і ефект очікування. Також існує метод контрбалансування соціальної взаємодії, коли різні експериментатори рівною мірою представлені в усіх умовах, що дозволяє виключити вплив особистості експериментатора на результати.

Особливе місце у контролі побічних змінних займають статистичні методи, такі як коваріаційний аналіз (ANCOVA), який дозволяє коригувати результати дослідження з урахуванням міжіндивідуальних відмінностей за певними параметрами, що могли вплинути на поведінку. Наприклад, якщо в експерименті з навчання увазі учасники суттєво відрізняються за базовим рівнем концентрації, цей параметр може бути включений у модель як співзмінна, що дозволяє оцінити реальний вплив тренінгу незалежно від початкових відмінностей.

Таким чином, контроль побічних змінних є складною, багатокомпонентною системою, що включає технічні, процедурні, психологічні та аналітичні механізми. Експериментальна логіка передбачає не лише створення впливу, але й ретельну організацію умов, у яких усі альтернативні пояснення змін у поведінці виключені або зведені до мінімуму. Успішність експериментального дослідження визначається тим, наскільки уважно дослідник передбачає, ізолює та контролює всі можливі побічні впливи. Лише за таких умов результати експерименту можуть бути інтерпретовані як надійний доказ причинно-наслідкового зв'язку між змінними, що робить експеримент фундаментальним інструментом наукової психології.

6.6. Валідність психологічного експерименту як критерій наукової достовірності

У науковій психології поняття валідності є одним із ключових методологічних принципів, що визначає, наскільки експеримент здатний давати істинні, обґрунтовані та узагальнювані висновки. Валідність концептуальний фундамент наукового дослідження, визначаючи, чи відповідає отриманий результат тому явищу, яке дослідник прагнув вивчити. Власне, валідність є критерієм того, чи можемо ми інтерпретувати дані як такі, що відображають причинно-наслідкові залежності, психологічні конструкти

або закономірності поведінки. Без достатнього рівня валідності експеримент втрачає епістемологічний статус наукового доказу й перетворюється лише на послідовність контрольованих спостережень, які не несуть пояснювальної сили.

Традиційно валідність поділяють на кілька видів: внутрішню, зовнішню, конструктну та статистичну валідність. Кожен вид відіграє власну роль в обґрунтуванні того, що результати відображають саме те, що мав на меті дослідник.

Внутрішня валідність визначає, чи можна вважати виявлені зміни в залежній змінній наслідком маніпуляції незалежною змінною, а не результатом дії сторонніх або неконтрольованих факторів. Іншими словами, вона відповідає на питання: «Чи справді X є причиною Y?» Тільки за наявності високої внутрішньої валідності можна робити обґрунтовані висновки про каузальні залежності. У контексті психологічних експериментів внутрішня валідність означає методологічну чистоту дослідження, відсутність альтернативних пояснень та контроль побічних змінних. Оскільки психічні явища є надзвичайно чутливими до індивідуальних, ситуативних та контекстуальних впливів, питання внутрішньої валідності набуває центрального значення. У разі низької внутрішньої валідності сам факт отримання статистично значущого результату не має наукової цінності, адже не можна сказати, чим саме цей результат зумовлений.

Зовнішня валідність, зі свого боку, визначає, наскільки результати експерименту можуть бути узагальнені на інші вибірки, умови, ситуації та контексти, відмінні від тих, що були використані у дослідженні. Якщо внутрішня валідність дає відповідь на питання «Чи є ефект реальним?», то зовнішня відповідає на питання «Де і для кого цей ефект буде проявлятися?». Зовнішня валідність охоплює різні рівні узагальнення: від можливості повторення ефекту в іншій лабораторії до застосування результатів у реальному житті, освітній, медичній, організаційній чи соціальній практиці. У психології, де контекст поведінки надзвичайно важливий, зовнішня валідність стає не просто додатковим критерієм, а необхідною умовою наукового та практичного значення результатів. Лабораторія, хоч і створює оптимальні умови для контролю змінних, не завжди відтворює природні умови, у яких функціонує поведінка людини.

Ці два види валідності перебувають у складному й часто напруженому взаємозв'язку. У класичних експериментальних умовах підвищення внутрішньої валідності досягається шляхом максимального контролю змінних, стандартизації умов і зменшення варіативності контексту. Проте саме ці дії можуть знизити зовнішню валідність, оскільки створюють ситуації, які мало нагадують реальні умови поведінки. Наприклад, дослідження впливу шуму на концентрацію в умовах ізолюваної лабораторної кімнати з 70 дБ білого шуму має високу внутрішню валідність, але зовнішня валідність може бути низькою, адже в повсякденному житті шум має інший характер, іншу структуру джерел і змінюється непередбачувано. З іншого боку, дослідження

в реальних умовах, наприклад у школі чи на робочому місці, може демонструвати високу екологічну валідність, але водночас мати низьку внутрішню валідність через неможливість контролю побічних змінних. Таким чином, внутрішня та зовнішня валідність часто перебувають у компромісних відносинах: максимізація однієї може призводити до зниження іншої.

Конструктна валідність стосується того, наскільки експериментальні маніпуляції й вимірювання адекватно відображають теоретичні конструкти, які лежать в основі дослідження. Це питання не про правильність причинного висновку чи можливість узагальнення, а про відповідність між теоретичною моделлю і процедурою експерименту. У психології конструктна валідність має фундаментальне значення, оскільки конструкти – такі як тривожність, мотивація, самоконтроль, емпатія – не мають прямої фізичної форми й завжди опосередковуються поведінкою, висловлюваннями або психофізіологічними показниками. Якщо операціоналізація конструкта слабка, навіть експеримент із високою внутрішньою та зовнішньою валідністю може бути беззмисловим, оскільки він вимірюватиме не те, що передбачала теорія. Проблеми конструктної валідності часто виникають у дослідженнях афективних і когнітивних процесів, де експериментальні задачі можуть бути надто штучними, а індикатори – недостатньо чутливими або надмірно спрощеними.

Статистична заключна валідність стосується правильності застосування статистичних методів і точності інтерпретації статистичних результатів. Вона включає такі питання, як достатність вибірки, чутливість тесту, ризик помилок першого й другого роду, адекватність статистичних моделей та відповідність обраного критерію розподілу даних. Статистична валідність є критичною для того, щоб оцінити, чи знайдений ефект є реальним, чи він є наслідком випадковості або неправильного аналізу. У психології проблема статистичної валідності особливо актуальна у зв'язку з поширеністю практик «p-hacking», неадекватного використання t-тестів, неправильного застосування непараметричних методів або ігнорування множинних порівнянь.

Окремий вимір валідності становить екологічна валідність, що охоплює відповідність експериментальних умов природним ситуаціям, у яких зазвичай відбувається поведінка. Екологічна валідність не є тотожним поняттям зовнішньої валідності, але є її важливою складовою. Висока екологічна валідність означає, що експериментальна ситуація близька до реальних соціальних, когнітивних або емоційних умов, у яких функціонує людина. Це особливо важливо в дослідженнях соціальної поведінки, міжособистісних взаємодій, групових процесів, навчання та мотивації. Наприклад, дослідження групового мислення в умовах лабораторії може недостатньо відтворювати реальні динаміки, притаманні робочим командам, студентським групам або військовим підрозділам.

У структурі психологічної науки всі види валідності утворюють складну систему, де кожен елемент доповнює інші. Внутрішня валідність дає можливість стверджувати причинно-наслідкові взаємозв'язки. Зовнішня валідність дозволяє поширювати ці взаємозв'язки на ширші контексти.

Конструктна валідність забезпечує відповідність між теорією та емпіричними процедурами. Статистична валідність гарантує, що зроблені висновки є результатом коректної інтерпретації даних, а не статистичного шуму. Відсутність хоча б одного виду валідності істотно знижує наукову цінність експерименту. Наприклад, експеримент з високою статистичною валідністю, але низькою конструктною валідністю може давати точні, але беззмістовні висновки; експеримент із високою конструктною валідністю, але низькою зовнішньою валідністю може бути цікавим теоретично, але не мати жодного практичного значення; експеримент, що демонструє високу зовнішню валідність, але низьку внутрішню, не може стверджувати причинних залежностей.

У практичному аспекті психологічних досліджень питання валідності набуває не лише наукового, а й етичного значення. Висновки, що ґрунтуються на невалідних експериментах, можуть вводити в оману фахівців, впливати на рішення щодо психотерапевтичних підходів, освітніх програм чи організаційних інтервенцій. Тому забезпечення високої валідності є не лише науковим обов'язком, а й елементом професійної відповідальності психолога.

Таким чином, валідність психологічного експерименту – це складна, багатовимірна характеристика, що визначає наукову силу причинних висновків, концептуальну адекватність операціоналізації, точність статистичної аргументації та застосовність отриманих результатів. У поєднанні внутрішня, зовнішня, конструктна та статистична валідності формують епістемологічний каркас експериментального методу. І хоча дослідник часто вимушений балансувати між ними, оптимально вибудована експериментальна логіка дозволяє досягти достатньої рівноваги, яка забезпечує наукову достовірність і практичну цінність психологічного знання.

6.7. Загрози достовірності висновків психологічного експерименту

Сама по собі експериментальна процедура не гарантує істинності висновків. Будь-яке дослідження здійснюється в умовах, де на поведінку учасників, хід маніпуляцій, якість вимірювань та інтерпретацію результатів впливає велика кількість факторів. Ці фактори, які не є частиною дослідницького задуму, але здатні спотворити логіку причинного висновку або обмежити можливість узагальнення результатів, отримали назву загроз валідності. У психології, де поведінка завжди є функцією складного взаємозв'язку індивідуальних, ситуативних та соціальних змінних, контроль таких загроз стає ключовим критерієм наукової строгості.

Традиційно більшість уваги приділяється загрозам внутрішньої та зовнішньої валідності, описаним Д. Кемпбеллом. Внутрішня валідність визначає, чи може дослідник достовірно стверджувати, що зміни в залежній змінній є наслідком маніпуляції незалежною змінною. Якщо цей критерій не дотримано, будь-які ефекти можуть бути результатом сторонніх процесів, а не причинного впливу, що його хотів перевірити дослідник. Зовнішня валідність,

зі свого боку, вказує на межі узагальнення експериментальних результатів. Навіть якщо виявлений ефект є причинним, він може бути характерним лише для специфічної категорії учасників, окремої експериментальної ситуації або особливого порядку впливів, що різко обмежує наукову та практичну цінність експерименту. Розглянемо ці загрози детальніше, з опорою на класичну типологію Кемпбелла.

Однією з ключових загроз внутрішній валідності є ефект історії, або ефект фону. Він полягає в тому, що між моментом експериментального впливу та моментом вимірювання можуть відбутися зовнішні події, які незалежно від експерименту змінюють поведінку учасників. Наприклад, дослідник вивчає вплив тренінгу емоційної регуляції на рівень тривожності студентів, але в період між тренінгом і наступним тестуванням трапляється значуща соціальна подія, яка підвищує тривожність у всіх членів групи. У такій ситуації зміни поведінки не можна однозначно приписати експериментальному впливу. Особливо небезпечним цей фактор стає у тривалих лонгітюдних експериментах, де проміжних подій набагато більше.

Тісно пов'язаний із попереднім фактором ефект природного розвитку. Він полягає в тому, що незалежно від експерименту учасники можуть змінюватися внаслідок вікових, когнітивних, емоційних або фізіологічних процесів. Наприклад, у дослідженні ефективності програми розвитку пам'яті у дітей молодшого шкільного віку природний когнітивний розвиток сам собою може покращити результати повторного тестування. У такому випадку збільшення показників пам'яті не обов'язково відображає дію експериментального фактора. Ефект розвитку є особливо чутливим у дослідженнях когнітивної психології, нейропсихології та вікової психології.

Ще однією фундаментальною загрозою внутрішньої валідності є ефект тестування, або реактивність попередніх вимірювань. Цей ефект визначає ситуації, коли саме перше вимірювання змінює поведінку учасників, впливаючи на результати подальших тестів. Наприклад, якщо людина виконує тест на увагу вдруге, вона може демонструвати кращий результат не через експериментальний вплив, а завдяки ефекту тренування. Окремою варіацією цього феномену є так звана реакція на тестування – ситуація, коли експериментальні результати можуть узагальнюватися лише на людей, які вже проходили тест. Це створює обмеження як для внутрішньої валідності (бо тест впливає на поведінку), так і для зовнішньої (бо результат не можна переносити на реальних людей, які не були попередньо протестовані).

Іншою суттєвою загрозою є регресія до середнього, яка виникає в експериментах, де групи формуються на основі екстремальних показників. Учасники з дуже високими або дуже низькими результатами мають статистичну тенденцію повертатися до більш середніх значень при повторному вимірюванні. Наприклад, якщо експериментатор відібрав до групи лише студентів з найнижчими результатами тесту, їхній наступний тестовий показник може підвищитися незалежно від експериментального впливу просто внаслідок статистичних закономірностей. Ця загроза є

надзвичайно поширеною в дослідженнях освітніх інтервенцій та клінічній психології.

Серйозною загрозою внутрішній валідності є також нееквівалентність груп, яка виникає у випадках не випадкового відбору учасників. Якщо люди самі записуються до експериментальної групи або їх розподіл здійснюється за ознаками, пов'язаними з досліджуванним феноменом, групи можуть істотно відрізнятися ще до початку експерименту. Наприклад, у дослідженні ефективності антистресової програми група, яка добровільно обирає участь у тренінгу, може першочергово мати вищу мотивацію до змін або вищий рівень тривожності. У такій ситуації неможливо стверджувати, що виявлені відмінності зумовлені експериментальною маніпуляцією.

Не менш небезпечною є загроза, відома як ефект відсіву або вимирання. Вона полягає в тому, що в процесі експерименту частина учасників може вийти з дослідження, і ці втрати зазвичай не є випадковими. У клінічних дослідженнях із тривалими втручаннями найчастіше вибувають учасники з найнижчою мотивацією або з найменшими позитивними змінами. Якщо ці учасники вибувають переважно з однієї групи, це зумовлює систематичне спотворення результатів. У деяких випадках ефект відсіву може радикально змінити ефект експериментального втручання – навіть перетворити негативний вплив на позитивний або навпаки.

Ще однією загрозою є взаємна інтерференція (взаємовплив) експериментальних впливів, яка виникає тоді, коли один учасник зазнає кількох впливів. Зміна порядку, тривалості або структури втручань може змінювати поведінку учасника так, що стає неможливо визначити, який саме вплив спричинив зміни. Цей ефект особливо важливий у дизайнах з повторними вимірюваннями, багатофакторних експериментах, а також у дослідженнях навчання, де кожне нове завдання може змінити когнітивний стан учасника.

Проблемою, що об'єднує всі попередні, є взаємодія загроз, коли кілька побічних процесів одночасно формують складні й непередбачувані спотворення. Наприклад, у дослідженні без рандомізації одночасно можуть діяти нееквівалентність груп, регресія до середнього та ефект розвитку. У таких випадках відокремити вплив незалежної змінної від побічних факторів майже неможливо.

Окрім внутрішньої валідності, Кемпбелл виділяє чотири головні загрози зовнішній валідності, які визначають межі узагальнення експериментальних результатів на інші ситуації, популяції та контексти. Першою з них є вплив тестування, або реактивний ефект попереднього вимірювання, який обмежує можливість узагальнення результатів на людей, що не проходили це тестування. Якщо експериментальні ефекти виникають лише тому, що тест привернув увагу учасника до проблеми, вони не будуть проявлятися в реальних умовах, де такої підготовчої процедури немає.

Другою загрозою є взаємодія відбору та експериментального впливу. Ефект може бути типовим лише для певної групи людей зі специфічними

соціальними, індивідуальними або мотиваційними характеристиками. Наприклад, програма розвитку лідерства може працювати тільки з високомотивованими студентами, але не давати результатів у вибірці школярів або працівників підприємств.

Третя загроза – це вплив умов організації експерименту, тобто ситуаційна унікальність експериментального середовища. Якщо ефект виникає тільки в лабораторії або тільки у присутності експериментатора, він може не повторитися в реальному контексті. Наприклад, вплив моделювання агресії в експериментальному середовищі не завжди екстраполюється на поведінку в сім'ї чи школі.

Четвертою загрозою зовнішній валідності є взаємна інтерференція експериментальних впливів, коли порядок або структура втручень визначає те, чи проявляється ефект у реальному житті. Якщо експеримент моделює умови, яких у природних контекстах не буває, узагальнення результатів стає сумнівним.

На додачу до зовнішньої та внутрішньої, у психологічних експериментах існують загрози конструктній і статистичній валідності, які можуть спотворювати результати незалежно від ситуаційних факторів. Конструктні загрози пов'язані з тим, що досліджуваний конструкт може бути операціоналізований неправильно або надто вузько. Наприклад, використання одного тесту на тривожність може відображати не саму тривожність, а рівень тестової мотивації чи інтелектуальну гнучкість. Статистичні загрози виникають тоді, коли застосовані методи аналізу не відповідають структурі даних – наприклад, надто мала вибірка, ігнорування асиметрії розподілів, неправильні критерії або відсутність контролю за множинними порівняннями.

Таким чином, загрози валідності формують складну систему факторів, які можуть істотно знизити достовірність експериментальних висновків, навіть якщо процедура проведена технічно коректно. Усвідомлення та аналіз кожної з них є необхідною умовою для побудови науково строгого експериментального дизайну. Власне, контроль таких загроз і становить серцевину експериментальної методології, відмежовуючи справжні причинні залежності від випадкових або хибних пояснень.

6.8. Етапи реалізації психологічного експерименту

Реалізація психологічного експерименту є складним, багатоступеневим процесом, що вимагає ретельного планування, точного дотримання процедурних стандартів і глибокого методологічного розуміння. На відміну від описових або кореляційних досліджень, де збір даних часто зводиться до фіксації природних проявів поведінки чи суб'єктивних оцінок, експеримент передбачає активне конструювання ситуації, у якій незалежна змінна маніпулюється цілеспрямовано, а залежна – вимірюється з науковою точністю. Тому процес проведення експерименту традиційно поділяють на кілька послідовних етапів: планування, пілотажне дослідження, відбір та

інструктування учасників, реалізацію експериментального впливу та вимірювання, постекспериментальні процедури й аналіз результатів. Кожен з цих етапів не лише виконує певну функцію, а й забезпечує логіку причинного висновку, що лежить в основі експериментального методу.

Першим і найбільш критичним етапом є планування та підготовка дослідження. На цьому етапі формулюється гіпотеза, визначається незалежна та залежна змінні, обирається тип експериментального дизайну і встановлюються заходи контролю побічних змінних. Планування включає також розробку експериментальної маніпуляції: дослідник має точно визначити, яким способом буде вводиться незалежна змінна, які рівні вона матиме та як забезпечуватиметься її стабільність. Наприклад, у дослідженні впливу стресу на продуктивність уваги потрібно заздалегідь обрати тип стресора (соціальний, когнітивний, сенсорний), його інтенсивність та тривалість. Наукова проблема полягає в тому, що психологічні ефекти часто є тонкими, тому навіть незначні коливання інструкцій, інтонації експериментатора чи умов середовища можуть «розмити» результат. Важливою частиною планування є розробка інструкцій, які мають бути чіткими, однозначними та однаковими для всіх учасників. Саме на цьому етапі визначаються також способи реєстрації залежної змінної – за допомогою тестів, реакційного часу, шкал самооцінки, поведінкових індикаторів чи фізіологічних показників. Поширена складність полягає в правильній операціоналізації – перетворенні теоретичного поняття (наприклад, «увага») у конкретний і надійний вимірювальний інструмент. Якщо операціоналізація слабка, навіть ідеально реалізований експеримент може не дати змоги отримати валідні висновки.

Наступним етапом є пілотажне дослідження, яке виконує функцію перевірки життєздатності експериментального плану. Пілотаж дозволяє емпірично переконатися, що маніпуляція працює так, як задумано, і що інструкції є зрозумілими. У цьому сенсі пілотаж – це тест стійкості експериментальної процедури до реальних умов. Під час пілотажного дослідження можуть виявитися проблеми, про які дослідник не здогадувався: учасники можуть інтерпретувати інструкції не так, як очікувалось; незалежна змінна може не спричиняти достатньо сильного ефекту; залежна змінна може виявитися недостатньо чутливою; час експозиції стимулів може бути занадто великим або занадто малим. Наприклад, у дослідженні впливу шуму на логічне мислення пілотаж може показати, що рівень шуму недостатній для викликання когнітивного навантаження, або що учасники використовують стратегії компенсації, які зменшують вплив. Без пілотажу такі проблеми можуть проявитися вже після збору даних, що зробить дослідження марним. Пілотаж також дозволяє уточнити структуру експерименту, скоригувати інструкції, змінити формулювання завдань, підібрати оптимальний часовий режим і вдосконалити методи контролю побічних змінних.

Третім етапом є відбір, інструктування та мотивування досліджуваних. На цьому етапі вирішальне значення має еквівалентність груп, якщо

дослідження є міжгруповим. Важливо забезпечити, щоб учасники були подібними за ключовими характеристиками, що можуть впливати на результат: наприклад, рівнем інтелекту, віком, статтю, темпераментом, попереднім досвідом. Типовою методологічною складністю є проблема селекційного упередження – коли у групи потрапляють учасники з певними прихованими характеристиками. Рандомізація значною мірою вирішує цю проблему, проте навіть рандомізація не гарантує абсолютної еквівалентності при малих вибірках. Інструктування досліджуваних – один з найбільш вразливих моментів експериментального процесу. Нечіткі інструкції можуть призвести до того, що учасники виконуватимуть завдання не так, як потрібно, або будуть використовувати неочікувані стратегії. Особливу увагу потрібно приділяти мотивації: надмірно зацікавлені учасники можуть демонструвати підвищену старанність, а незацікавлені – пасивність, що спотворює дані. У деяких експериментах використовується символічна винагорода, інші ж прагнуть мінімізувати зовнішню мотивацію, щоб уникнути небажаного ефекту очікувань.

Після відбору учасників та їх інструктування проводиться основний етап – експериментальний вплив та вимірювання. Цей етап є серцевиною дослідження, оскільки саме тут вводиться незалежна змінна та реєструється залежна. У деяких випадках проводиться попереднє тестування (pretest), яке дозволяє фіксувати початковий рівень залежної змінної. Після впливу проводиться повторне тестування (posttest), що дозволяє оцінити зміни, які можна віднести до дії незалежної змінної. Особливе значення має процедура перевірки маніпуляції – методологічна практика, спрямована на визначення того, чи дійсно учасники пережили той стан або отримали той вплив, який задумував дослідник. Наприклад, у дослідженні індукції емоцій учасників після маніпуляції можуть попросити оцінити свій емоційний стан; у дослідженні когнітивного навантаження – оцінити суб'єктивну складність завдання; у дослідженні впливу соціальної ізоляції – оцінити відчуття відторгнення. Якщо маніпуляція не спрацювала, результати буде неможливо інтерпретувати, навіть якщо статистичні тести покажуть відсутність ефекту. Складність цього етапу полягає також у необхідності контролю сторонніх впливів, які можуть змінювати залежну змінну незалежно від експериментального фактора. Навіть дрібні фактори – шум у коридорі, зміна освітлення, технічні збої, випадкові комунікації між учасниками – можуть спотворити поведінкові або когнітивні реакції.

Після завершення експериментальних процедур проводиться постекспериментальне інтерв'ю. Ця процедура має подвійну функцію: етичну та методологічну. На етичному рівні дослідник зобов'язаний пояснити справжню мету експерименту, особливо якщо використовував метод обману або прихованого впливу. Учасникам слід надати можливість поставити запитання, висловити свої переживання та отримати підтримку у разі психологічного дискомфорту. На методологічному рівні постекспериментальне інтерв'ю дозволяє виявити учасників, які здогадалися

про гіпотезу, зрозуміли структуру експерименту або діяли не відповідно до інструкцій. Такі дані можуть бути виключені або розглянуті окремо, щоб уникнути змішування ефектів. Наприклад, у дослідженні впливу підкріплення на мотивацію учасник може повідомити, що зрозумів: підкріплення було частиною маніпуляції. Це означатиме, що його поведінка була не «чистою», а орієнтованою на уявлення про мету експерименту.

Заключним етапом є аналіз результатів та статистичний контроль. На цьому етапі залежна змінна перетворюється на числові дані, які аналізуються відповідно до експериментального плану. Основним завданням є визначення, чи існують статистично значущі відмінності між умовами, які можна інтерпретувати як наслідок впливу незалежної змінної. Статистичний аналіз може включати t-тести, дисперсійний аналіз (ANOVA), коваріаційний аналіз (ANCOVA), аналіз змішаних моделей, корекції на множинні порівняння та аналіз розміру ефекту. Складність цього етапу полягає у необхідності врахування структури даних, типу шкал, перевірки нормальності розподілу та вибору правильного статистичного критерію. Крім того, потрібно оцінити надійність вимірювальних інструментів, перевірити відсутність пропусків, артефактів та аутлаєрів. Якщо дослідник не контролює побічні змінні, статистичний аналіз може виявити ефекти, що не є причинно обумовленими, або навпаки – приховати реальний вплив. Правильний статистичний контроль дозволяє підсилити внутрішню валідність експерименту й уникнути помилкових висновків.

Таким чином, реалізація психологічного експерименту – це багатофазний процес, у якому кожен етап відіграє принципову роль. Від якості планування, ретельності пілотажу, точності інструктажу, послідовності маніпуляції та валідності вимірювання залежить не лише успішність дослідження, а й можливість робити науково обґрунтовані висновки про причинно-наслідкові зв'язки. Експеримент – це система жорстких методологічних вимог і водночас інструмент найвищої наукової точності, здатний виявляти закономірності поведінки й психічних процесів, недоступні жодному іншому методу психології.

Завдання на самостійну підготовку

1. У чому полягає специфічність проблеми причинно-наслідкових зв'язків у психології порівняно з природничими науками?
2. Чому експеримент вважається єдиним методом, який дозволяє встановити причинно-наслідкові зв'язки?
3. Які три класичні умови необхідні для доведення причинності між змінними?
4. Чим відрізняється кореляційне мислення від причинного у контексті психологічних досліджень?
5. У чому полягає логіка психологічного експерименту та які її ключові елементи?

6. Назвіть основні види експериментів за умовами проведення та поясніть їх особливості.
7. Що таке незалежна змінна та які її основні характеристики в експериментальному дизайні?
8. Які вимоги висуваються до залежної змінної?
9. Що таке побічні змінні та чому вони становлять загрозу для дослідження?
10. Які основні психологічні та ситуативні фактори, що походять від експериментатора, можуть викривляти результати експерименту?
11. У чому полягає Хоторнський ефект та як він впливає на поведінку досліджуваних?
12. Як проявляються ефект плацебо та соціально бажана поведінка в експериментальних ситуаціях?
13. У чому полягає логіка подвійного сліпого методу та які проблеми він допомагає усунути?
14. Як автоматизація процедур сприяє підвищенню внутрішньої валідності експерименту?
15. Які вимоги етичного характеру виникають під час застосування методу обману та прихованого експерименту?
16. Які види валідності (внутрішня, зовнішня, конструктна, статистична) потрібно враховувати під час побудови експерименту?
17. У чому полягає протиріччя між внутрішньою та зовнішньою валідністю?
18. Які основні загрози внутрішній валідності виділяв Д. Кемпбелл та як вони впливають на результати дослідження?
19. Як загрози зовнішній валідності обмежують можливість узагальнення експериментальних результатів?
20. Які основні етапи реалізації психологічного експерименту та чому кожен із них є критично важливим для забезпечення наукової достовірності?

РОЗДІЛ 7

ДИЗАЙНИ ПСИХОЛОГІЧНОГО ЕКСПЕРИМЕНТУ

7.1. Основні вимоги до експериментального дизайну

Експериментальний дизайн у психології є методологічною основою для встановлення причинно-наслідкових залежностей між явищами, що вивчаються. Від того, наскільки грамотно побудований експеримент, залежить достовірність отриманих результатів, можливість їх узагальнення, відтворюваність та інтерпретаційна чіткість. Експеримент базується на активному втручанні дослідника в умови перебігу психологічного процесу, що дозволяє ізольовано оцінити вплив окремих змінних. Проте для того, щоб зробити такі висновки обґрунтованими, сам дизайн має відповідати низці наукових, логічних і статистичних вимог. Ці вимоги не є технічними деталями, а становлять фундамент, який забезпечує внутрішню та зовнішню валідність експерименту, мінімізує імовірність систематичних похибок, контроль сторонніх чинників та точний вимір ефектів.

Першою ключовою вимогою експериментального дизайну є чітке визначення незалежної та залежної змінних. Незалежна змінна – це фактор, який дослідник маніпулює і який передбачається як причина змін поведінки чи психологічного стану. Залежна змінна фіксує реакцію, відгук або результат впливу. Якщо ці змінні визначено нечітко або внутрішній зміст вимірюваних конструкцій розмитий, експеримент втрачає логіку причинності й перетворюється на ситуативне спостереження. У психологічних працях часто трапляються помилки, коли «незалежною» називають змінну, яку насправді неможливо контролювати (наприклад, темперамент), або коли залежна змінна вимірюється інструментом із низькою валідністю. У таких випадках будь-який ефект стає сумнівним, навіть якщо статистично значущим. Тому теоретичне обґрунтування природи змінних є центральною умовою наукової коректності дизайну.

Наступною фундаментальною вимогою є контроль сторонніх чинників, що забезпечує внутрішню валідність експерименту. У психологічних дослідженнях сторонні змінні можуть бути численними: індивідуальні відмінності учасників, мотиваційні стани, контекст проведення, ефекти очікування, експериментаторські впливи, історичні події, коливання уваги, попередній досвід тощо. Якщо ці фактори не враховано, вони можуть стати альтернативним поясненням змін залежної змінної, що позбавляє експеримент можливості продемонструвати причинність. Тому контроль – через рандомізацію, константування, контрбалансування або створення контрольної групи – є одним із наріжних каменів якісного експериментального дизайну. Наприклад, у дослідженнях ефективності когнітивно-поведінкової терапії контрольна група дозволяє відрізнити ефект власне терапії від ефектів часу, очікування або соціальної взаємодії. Без такого контролю будь-яке зростання

психологічного благополуччя могло б бути приписане терапії, навіть якщо воно виникло спонтанно.

Третьою вимогою є адекватна рандомізація – процедура випадкового розподілу учасників між експериментальними умовами. Саме рандомізація забезпечує рівномірний розподіл індивідуальних властивостей між групами, зменшуючи вірогідність систематичних відмінностей. Якщо дослідник нехтує рандомізацією або проводить її поверхово, він ризикує отримати групи, які відрізняються не через експериментальний вплив, а через випадкові особливості складу. Наприклад, якщо в експериментальну групу потраплять більш мотивовані студенти, а в контрольну – менш мотивовані, це створить передумови для хибних висновків щодо ефективності втручання. Рандомізація підвищує внутрішню валідність настільки суттєво, що саме вона лежить в основі «золотого стандарту» – рандомізованого контрольованого експерименту.

Наступною вимогою є стандартизація процедур, що передбачає однорідність умов проведення експерименту для всіх учасників. У психологічних дослідженнях стандартизація охоплює тривалість експериментального впливу, інструкції, матеріали, просторове оточення, поведінку експериментатора, інтервали між вимірюваннями тощо. Якщо умови відрізняються між групами або між різними етапами, виникають артефакти, що ускладнюють інтерпретацію. Навіть інтонація експериментатора може змінити реакцію учасника, якщо інструкція сформульована неоднозначно. Тому експериментальні протоколи повинні бути не лише стандартизованими, а й описаними з максимальною точністю, забезпечуючи можливість відтворення.

Ще однією критичною вимогою є достатній обсяг вибірки, що забезпечує статистичну потужність дослідження. У психології часто трапляється ситуація, коли експерименти проводять із надто малою кількістю учасників, що призводить до нестабільності ефектів, високої ймовірності помилок першого і другого роду та неможливості виявити реальні закономірності. Вимога достатньої вибірки має не лише статистичний характер. Вона визначає відтворюваність дослідження, знижує флуктуації даних і дозволяє оцінити ефект з належною точністю. Потужність стає особливо важливою у дослідженнях, де очікувані ефекти є слабкими або помірними, що характерно для більшості психологічних втручань.

Суттєвим аспектом є також забезпечення зовнішньої валідності, тобто можливості узагальнення результатів експерименту на ширшу популяцію. Ця вимога пов'язана зі способом формування вибірки, характеристиками учасників, екологічною валідністю ситуацій, культурною відповідністю матеріалів. Експеримент, проведений на вибірці студентів у лабораторних умовах, може демонструвати високий рівень внутрішнього контролю, але його результати можуть бути погано узагальнюваними на інші соціальні групи. Вимога зовнішньої валідності передбачає баланс між строгістю дизайну та реалістичністю досліджуваної ситуації. Наприклад, експерименти зі стресом

часто критикують за штучність лабораторних умов, що не відображають складності реального життя.

Одним із центральних елементів якісного експериментального дизайну є запобігання очікувальним ефектам – як з боку учасників, так і з боку дослідника. Ефект Пігмаліона, ефект Розенталя та ефект Хоторна добре відомі в психології як фактори, що здатні суттєво змінити результати експерименту незалежно від маніпуляції. Щоб мінімізувати ці впливи, застосовують сліпі та подвійно сліпі процедури, автоматизовані системи подання стимулів, стандартизовані інструкції. Ці заходи забезпечують нейтралізацію очікувань і дозволяють отримати результати, що дійсно відображають вплив експериментального чинника.

Важливою вимогою є також коректне проєктування контрольних умов, оскільки саме вони дозволяють відокремити ефект незалежної змінної від інших можливих пояснень. Контрольні умови можуть бути нейтральними, активними, плацебо або альтернативними залежно від того, яку саме гіпотезу перевіряє дослідник. Неправильний вибір контрольної умови здатен повністю змінити висновки експерименту. Наприклад, у дослідженнях психологічних тренінгів контрольний варіант «нічого не робити» може виявитися методологічно некоректним, оскільки не враховує ефекти соціальної взаємодії. У таких випадках доцільні активні контрольні групи, що отримують інше втручання з таким самим рівнем соціальної активності.

Окремою вимогою є чіткість гіпотез, що передбачає їх формулювання у вигляді емпірично перевірюваних тверджень. Гіпотези мають бути конкретними, однозначними та відповідати теоретичній моделі, що лежить в основі експерименту. Аморфні або надмірно широкі гіпотези унеможливають коректний вибір статистичного критерію та ускладнюють інтерпретацію результатів. Водночас попередня реєстрація гіпотез і аналітичних планів значно підвищує прозорість дизайну, запобігає пост-хок маніпуляціям та сприяє реплікованості.

Експериментальний дизайн потребує також уважного ставлення до етичних норм. Дослідження не може завдати шкоди учаснику, повинно передбачати інформовану згоду, можливість припинення участі та гарантувати конфіденційність даних. Особливо критичними є етичні вимоги у дослідженнях із вразливими групами, у клінічних експериментах та у випадках використання обману як методичного прийому. Дизайн має забезпечити баланс між науковими цілями та етичними стандартами, і цей баланс є невід'ємною характеристикою професійної роботи психолога-дослідника.

У підсумку експериментальний дизайн – це не просто структурована схема дослідження, а цілісна система логічних, методологічних, статистичних та етичних рішень, що забезпечують можливість отримання валідних, надійних і відтворюваних результатів. Відповідність вимогам дизайну визначає не лише якість окремого експерименту, а й рівень довіри до психологічної науки в цілому. Це складний, але необхідний процес, що

вимагає від дослідника високої компетентності, продуманості та відповідальності.

7.2. Види експериментальних дизайнів

Експериментальний дизайн у психологічному дослідженні виступає фундаментальною основою для встановлення причинно-наслідкових зв'язків між змінними, адже саме він задає внутрішню логіку процедури, визначає рівень контролю над експериментальними умовами і ступінь валідності отриманих висновків. У межах експериментальної методології надзвичайно важливим є питання структури експерименту, адже від того, як організовано вплив незалежної змінної, як сформовано групи досліджуваних і як здійснюється контроль за сторонніми чинниками, залежить можливість дослідника інтерпретувати виявлені ефекти саме як наслідок експериментального втручання, а не як результат дії випадкових або неконтрольованих чинників. У сучасній психології найбільш авторитетною є класифікація експериментальних планів, запропонована Дональдом Кемпбеллом, яка включає три великі групи: доекспериментальні плани, істинні експерименти та квазіекспериментальні плани. Кожна з цих груп має свої методологічні особливості, рівень внутрішньої валідності, типові сфери застосування, переваги та обмеження.

Для розуміння експериментальних планів представлених нижче нам знадобляться три основні позначення, запропоновані Д. Кемпбеллом:

R – рандомізація. Це випадкове призначення учасників до різних груп, що допомагає зробити групи максимально схожими перед початком експерименту.

X – експериментальний вплив. Це наша незалежна змінна, те, що ми змінюємо або вводимо в експерименті.

O – реєстрація вимірювання. Це наша залежна змінна, те, що ми вимірюємо до або після впливу.

Доекспериментальні плани становлять найнижчий рівень експериментального контролю, оскільки в них відсутня ключова характеристика справжнього експерименту – порівняння з контрольними групами, а також бракує рандомізації. Здебільшого такі плани використовуються тоді, коли дослідження перебуває на початкових етапах, коли оцінюється потенційність впливу, апробується методика або коли з етичних причин неможливо впроваджувати складнішу структуру. Найпростішим прикладом є план з одноразовим вимірюванням після впливу, коли дослідник застосовує певний стимул або методику і відразу після цього фіксує результати. На практиці такі схеми використовуються в межах психотерапевтичних сесій, коли психолог спостерігає миттєві зміни у стані клієнта після інтервенції. Проте такий план не дозволяє достовірно інтерпретувати ефект, оскільки будь-які зміни могли бути зумовлені

зовнішніми чинниками – наприклад, коливаннями настрою, впливом соціального контексту або ефектом новизни.

На противагу цьому істинні експерименти являють собою максимальний рівень методологічної строгості, адже вони обов'язково включають рандомізацію, наявність експериментальної та контрольної групи, а також можливість чітко маніпулювати незалежною змінною. Рандомізація – це випадковий розподіл учасників між групами – дозволяє створити еквівалентні групи, вирівнюючи їх за усіма можливими характеристиками, як вимірними, так і невимірними. Завдяки цьому дослідник отримує можливість приписувати зміни залежної змінної саме впливу незалежної змінної, адже альтернативні пояснення були мінімізовані. Типовим прикладом істинного експерименту є план з пре-тестом і посттестом, у якому дві групи, сформовані шляхом рандомізації, проходять однакову процедуру вимірювання, але лише одна група отримує експериментальний вплив. У такому плані співставлення змін у групах дозволяє сформулювати висновок про ефективність втручання.

Квazіекспериментальні плани займають проміжне положення між доекспериментальними та істинними експериментами. Вони використовуються тоді, коли з етичних, соціальних або організаційних причин неможливо здійснити рандомізацію. Наприклад, дослідник не може випадково розподілити школярів між різними формами навчання або організувати експериментальне втручання у лікарні шляхом випадкового призначення пацієнтів на різні протоколи лікування. У таких випадках дослідник працює з уже сформованими групами, що створює ризик їхньої нееквівалентності. Щоб компенсувати цю загрозу, квazіексперименти включають додаткові процедури контролю: вимірювання до та після впливу, статистичний контроль (наприклад, ANCOVA), використання великої кількості вимірювань у часових рядах, підбір контрольних груп за ключовими характеристиками. Прикладом квazіекспериментального підходу є план перерваних часових рядів, у якому дослідник багаторазово вимірює залежну змінну до і після впливу, що дозволяє бачити динаміку змін і визначати, чи були вони спричинені саме експериментальною подією.

Квazіексперименти менш контрольовані, ніж істинні експерименти, але мають високу екологічну валідність, адже відбуваються у реальних умовах. Вони дозволяють вивчати ефекти широкомасштабних соціальних явищ, реформ або природних подій. Наприклад, досліджуванню можна піддати вплив пандемії, нової освітньої реформи або політичної кампанії на психологічні показники населення. У таких випадках незалежна змінна не створюється експериментатором, а виникає як результат реальних подій, і завдання дослідника полягає у виявленні її ефектів у природному контексті.

Окремим виміром експериментальних дизайнів є питання того, як організовано роботу з досліджуваними – у формі міжгрупових або внутрішньогрупових планів. У міжгрупових (крос-індивідуальних) планах різні люди виконують різні умови дослідження. Наприклад, одна група отримує тренінг емоційної регуляції, а інша – не отримує, або одна група

виконує завдання з високим рівнем стресу, а інша – з низьким. Основною проблемою таких планів є індивідуальні відмінності між учасниками, які можуть випадково зробити групи нееквівалентними. Щоб контролювати ці відмінності, дослідник може використовувати кілька стратегій. Рандомізація забезпечує випадковий розподіл людей по групах, вирівнюючи їх за всіма характеристиками. Попарний добір передбачає послідовне формування пар учасників за важливою властивістю (наприклад, рівень тривожності), а потім випадкове призначення одного з кожної пари до експериментальної групи, а іншого – до контрольної. Стратометричний відбір дозволяє контролювати групи за статтю, віком, освітою, соціальним статусом чи іншими соціально-демографічними параметрами. Перевагою міжгрупових планів є відсутність ефектів втоми, навчання або взаємного впливу умов, але вони потребують більших вибірок, що не завжди можливо.

На відміну від міжгрупових, внутрішньогрупові (інтраіндивідуальні) плани передбачають, що всі досліджувані виконують усі експериментальні умови. Наприклад, учасники можуть спочатку виконувати завдання у стані спокою, а потім – у стані стресу, або навпаки. Це дозволяє порівнювати результати тієї самої особи в різних умовах, що суттєво підвищує статистичну потужність і зменшує вплив індивідуальних відмінностей. Проте такі плани вразливі до позиційних ефектів, коли порядок пред'явлення стимулів впливає на результати, і до ефектів послідовності, коли схожі або контрастні стимули можуть викликати праймінг або зміну установки. Щоб уникнути цих загроз, дослідники використовують рандомізацію порядку стимулів для кожного учасника, позиційне вирівнювання, балансування і контрбалансування. Наприклад, у контрбалансуванні група учасників випадково поділяється на підгрупи, кожній із яких пред'являється різна послідовність умов, що дозволяє нейтралізувати вплив порядку.

Таким чином, експериментальні дизайни у психології є різноманітними інструментами, кожен із яких має свої методологічні переваги та обмеження. Доекспериментальні плани служать для попереднього вивчення явищ, істинні експерименти забезпечують найвищий рівень контролю і дають змогу формулювати причинно-наслідкові висновки, а квазіексперименти дозволяють досліджувати природні події або соціальні процеси у реальних умовах. Міжгрупові та внутрішньогрупові плани визначають спосіб організації вибірки і також впливають на внутрішню валідність. Розуміння цих різновидів дозволяє досліднику грамотно будувати логіку дослідження, ефективно контролювати сторонні змінні та відповідально інтерпретувати результати.

7.3. Характеристика доекспериментальних планів

Доекспериментальні плани, які Д. Кемпбелл розглядає як найнижчий за рівнем контролю тип експериментальної організації, займають важливе місце в історії розвитку емпіричних методів у психології, хоча й характеризуються

низькою здатністю забезпечувати внутрішню валідність. Їхня роль полягає не у строгій перевірці причинно-наслідкових гіпотез, а у формуванні попередніх уявлень, зборі початкових даних, описі феноменів та розробці тестових впливів, які згодом можуть бути перевірені у більш строгих експериментальних схемах. Власне доекспериментальні схеми не включають рандомізації, оскільки за визначенням не містять контрольної групи й не забезпечують механізмів урівноваження зовнішніх впливів. Тому вони є вразливими майже до всіх загроз валідності, описаних Д. Кемпбеллом, зокрема фону, природного розвитку, впливу попереднього тестування, інструментальної похибки, регресії до середнього, селекції, а також взаємодій між тестуванням і впливом незалежної змінної.

Найпростішим видом доекспериментального плану є дослідження одиничного випадку, позначене як:

X O,

де X – експериментальний вплив,

O – одиничне післятестове вимірювання.

Це найбільш елементарна форма емпіричного дослідження, у якій відсутнє будь-яке попереднє тестування, не передбачено групового порівняння та не використовуються механізми контролю за змінними. Дослідник вводить певний вплив і після цього фіксує певний показник, припускаючи, що зафіксована зміна може бути результатом цього впливу. У психології така схема часто фігурує у форматі описових клінічних випадків, природних експериментів або ілюстративних прикладів дії психокорекційної методики. Наприклад, психолог може впровадити релаксаційну техніку для зменшення гострої тривожності у пацієнта, який звернувся у стані сильного стресу, а після застосування методики зафіксувати рівень тривожності за відповідною шкалою. На перший погляд може здатися, що зменшення тривожності свідчить про ефективність втручання. Проте така інтерпретація є методологічно вразливою, адже відсутні дані про природне відновлення, взаємодію з іншими чинниками, вплив часу або спонтанні коливання психічного стану.

Недоліки цього плану охоплюють практично повну відсутність контролю внутрішньої валідності. Будь-які зміни у вимірюванні можуть мати інші причини: природне згасання емоційної реакції, зміни зовнішнього середовища, спонтанний регрес до середнього або навіть артефакти вимірювального інструменту. Так само не можна виключати ефект експериментатора, очікувану реакцію досліджуваного, або вплив сторонніх подій, які відбулися паралельно з експериментом. Дослідження одиничного випадку, однак, зберігає значну цінність для прикладної психології, оскільки дозволяє отримати первинне уявлення про феномен, сформулювати робочу гіпотезу, протестувати новий інструмент або методику, а також зібрати матеріал, який згодом стане основою для справжнього експерименту з використанням контрольної групи та рандомізації.

Другим типом доекспериментального плану є схема з попереднім і підсумковим тестуванням однієї групи, що позначається як:

O X O.

Цей план частково компенсує слабкість першої схеми шляхом включення попереднього тестування, що дозволяє порівнювати стан досліджуваних до і після впливу. Психологічний приклад використання такого дизайну можна знайти в оцінці ефективності короткочасної тренінгової програми з підвищення самооцінки: спочатку учасники проходять опитувальник самооцінки (O), потім беруть участь у тренінгу (X), а після цього знову оцінюються за цією ж шкалою (O). Якщо показники зросли, дослідник може припустити, що тренінг був ефективним. Проте, навіть за наявності попереднього тестування, план не дозволяє виключити такі альтернативні пояснення, як ефекти тестування (покращення результатів лише завдяки повторному тестуванню), природний розвиток, вплив зовнішніх факторів або селекція учасників.

Цей вид плану частково підвищує рівень логічної аргументації щодо змін, але залишається вразливим до низки загроз. Наприклад, якщо досліджувані проходили тест двічі, можливий ефект навчання, за якого повторне тестування призводить до покращення результатів без реального впливу тренінгу. Також залишається ризик регресії до середнього, коли особливо низькі або високі показники під час першого тестування згодом наближаються до середніх значень незалежно від впливу. Не менш суттєвим є вплив одноразової події, яка могла відбутися між двома вимірюваннями і вплинути на вибірку. Тому хоча план O X O виглядає переконливішим, ніж одиничне вимірювання, він не забезпечує строгих доказів причинності.

Третій вид доекспериментального плану – це порівняння статичних груп, позначене як:

X O
O

Одна група отримує вплив і проходить вимірювання, а інша – лише вимірювання без впливу. Здавалося б, включення контрольної групи має підвищити валідність дослідження, але оскільки групи не формуються за допомогою рандомізації або інших методів вирівнювання, порівняння стає вкрай вразливим до ефектів селекції. Наприклад, якщо психолог досліджує ефективність програми розвитку когнітивної гнучкості для школярів, обравши одну класну групу як експериментальну та іншу як контрольну, то їхні відмінності після впливу можуть бути зумовлені початковими особливостями класів: різним рівнем мотивації, успішності, соціального клімату або стилю викладання. У такому випадку навіть значущі відмінності не можна інтерпретувати як наслідок застосованої програми.

Цей план є вразливим практично до всіх загроз валідності, але особливо критичною є загроза нееквівалентності груп, яку Д. Кемпбелл вважав однією з найнебезпечніших у природних дослідженнях без рандомізації. Групи можуть різнитися за вихідними параметрами, що й зумовить результати

незалежно від застосованого впливу. Тому будь-які причинні твердження у межах цього плану є слабо обґрунтованими. Проте порівняння статичних груп може бути корисним на початкових етапах вивчення феномену, коли неможливо організувати строгий експеримент або коли йдеться про природні групи, що вже сформовані незалежно від дослідницького задуму.

Загалом доекспериментальні плани практично не контролюють внутрішню валідність, оскільки не усувають впливу фону, природного розвитку, інструментальної похибки, регресії до середнього, селекції або взаємодій між фактором впливу та процесом тестування. Щодо зовнішньої валідності, їхня здатність до узагальнення також обмежена, оскільки відсутність контролю за змінними робить отримані результати специфічними для конкретних умов і вибірки. Ці плани практично не контролюють жоден із ключових джерел невалідності: у першому плані (X O) внутрішня валідність майже повністю відсутня; у другому (O X O) частково компенсується вплив селекції, але зберігається ризик ефектів тестування та природного розвитку; у третьому (X O / O) є видимість контролю завдяки наявності групи порівняння, але через відсутність рандомізації загроза селекції залишається критичною.

Попри методологічні обмеження, доекспериментальні плани мають свою педагогічну та прикладну цінність. По-перше, вони дозволяють здобувачам освіти і молодим дослідникам опанувати базове мислення у термінах X, O та логіки експериментального впливу. По-друге, вони можуть бути незамінними в умовах, коли неможливо застосувати рандомізацію: у клінічних дослідженнях, у ситуаціях з обмеженою вибіркою, в польових умовах або при роботі з групами, де втручання не може бути довільно розподілено. По-третє, вони слугують попереднім етапом перед побудовою квазіекспериментальних або істинних експериментальних планів.

Важливо розуміти, що доекспериментальні плани не є «поганими» чи «неправильними» – вони виконують іншу функцію: описову, гіпотезогенеруючу, попередню. У сучасних дослідженнях їх інколи застосовують для оцінки нових інструментів, прогнозування ефекту втручання або у випадках, коли дослідник має доступ до дуже обмеженої вибірки. Проте вони ніколи не можуть служити остаточною основою для встановлення причинно-наслідкових зв'язків.

7.4. Характеристика істинних експериментальних планів для однієї незалежної змінної

Істинні експериментальні плани становлять центральний інструмент доказової психології, оскільки саме вони забезпечують найвищий рівень контролю за сторонніми змінними та дозволяють здійснювати валідні висновки про каузальні зв'язки між незалежною та залежною змінними. У концептуальній класифікації Дональда Кемпбелла ця група планів займає фундаментальну позицію, оскільки являє собою найсуворіший стандарт експериментальної логіки. Ключовою ознакою істинного експерименту є

використання процедури рандомізації, яка забезпечує рівномірний розподіл індивідуальних відмінностей між групами та нівелює загрози валідності, характерні для доекспериментальних та квазіекспериментальних схем. Незалежно від того, чи включає дизайн попереднє тестування, чи обмежується лише післятестовим вимірюванням, чи має комбіновану багатогрупову структуру Соломона, його внутрішня логіка завжди ґрунтується на припущенні, що рандомізований розподіл є критичним елементом для достовірного встановлення причинно-наслідкових залежностей. У межах істинних експериментів з однією незалежною змінною Д. Кемпбелл виділяє три базові плани: двогруповий дизайн з попереднім і підсумковим тестуванням, план Соломона з чотирма групами та план з контрольною групою і тестуванням лише після впливу. Кожен із них має свою логіку, різний рівень контролю потенційних джерел артефактів та різні сфери доцільного застосування в психологічних дослідженнях.

У найпростішій та найелегантнішій формі істинний експеримент представлений планом двох рандомізованих груп із тестуванням тільки після впливу. Його нотація за Кемпбеллом має вигляд:

$$\begin{array}{ccc} R & X & O \\ R & & O \end{array}$$

Одна група отримує вплив незалежної змінної, а інша виступає контрольною. Відсутність попереднього тестування дозволяє повністю усунути ефекти тестування та виключити взаємодію між вступним вимірюванням і маніпуляцією. Усі потенційні загрози валідності, пов'язані зі складанням груп, природним розвитком, регресією до середнього чи фоновими подіями, контролюються завдяки рандомізації, яка гарантує рівномірний розподіл характеристик досліджуваних між умовами. Це означає, що будь-які посттестові відмінності О можуть бути інтерпретовані як результат впливу Х, а не побічних чинників. Саме тому цей дизайн часто визначають як «золотий стандарт» для ситуацій, у яких попереднє тестування недоцільне або потенційно шкідливе – наприклад, у дослідженнях первинної реакції на стимул, у вивченні одноразових ефектів стресора, у випадках, коли вимірювання саме по собі може змінити поведінку учасників.

Психологічний приклад використання цього плану може стосуватися експерименту з вивчення ефективності короткого мотиваційного відео на зниження прокрастинації у студентів: експериментальна група дивиться відео Х, а контрольна – нейтральний ролик, після чого всі учасники виконують реальне навчальне завдання О. Якщо рандомізація проведена коректно, відмінності у продуктивності можуть бути з великою ймовірністю приписані саме впливу мотиваційного матеріалу. Подібні дизайни широко застосовуються й у клінічній психології: наприклад, при дослідженні впливу короткої психологічної інтервенції на зниження гострої тривоги перед медичною процедурою.

Разом із тим цей дизайн має певні обмеження: хоча він ефективно нейтралізує більшість внутрішніх загроз валідності, він не дозволяє оцінити

початковий рівень залежної змінної, що може бути критичним у певних дослідженнях, де важливо контролювати індивідуальні відмінності, які, попри рандомізацію, все ж можуть виникнути в малих вибірках. Крім того, він недостатньо інформативний щодо стабільності впливу, щодо динаміки змін або щодо специфічних механізмів, через які відбуваються виявлені ефекти.

Для ситуацій, коли важливо оцінити зміну стану чи поведінки в часі, застосовується один із найпоширеніших істинних експериментальних планів – двогруповий дизайн з попереднім і підсумковим тестуванням. Його схема має вигляд:

R O X O
R O O

Цей план дозволяє порівнювати не лише посттестові значення, але й прирости у кожній групі, що робить його особливо привабливим для досліджень у сфері освіти, психології розвитку, клінічної психології та психотерапії. Попередній тест O_1 дає можливість контролювати первинний стан учасників, а також підвищує статистичну потужність аналізу завдяки використанню коваріації між до- та післятестовими значеннями. Водночас саме попереднє тестування створює ризик специфічних загроз валідності – передусім взаємодії тестування з експериментальним впливом: O може зробити учасників більш чутливими, більш мотивованими або змінити їхню поведінку таким чином, що вплив X спрацює інакше, ніж у реальних життєвих умовах. Це становить загрозу зовнішній валідності, оскільки втручання може виявитися ефективним лише за умови попереднього тестування, але не працювати у ситуаціях «чистого» впливу без попередньої підготовки.

У психологічних дослідженнях план O X O часто застосовується для оцінки ефективності психоосвітніх та тренінгових програм, розвитку когнітивних навичок, формування емоційної компетентності чи зміни поведінкових стратегій. Наприклад, при дослідженні впливу програми розвитку емоційної регуляції у підлітків попереднє тестування дозволяє зафіксувати вихідний рівень навичок, а післятестове – оцінити приріст у порівнянні з контрольною групою. У цьому плані можна використовувати аналіз зміни $\delta = O_2 - O_1$ для кожної групи, t-тест для незалежних вибірок або ANCOVA, де O_1 виступає коваріатою.

План із попереднім тестуванням дуже сильний з точки зору внутрішньої валідності: рандомізація нейтралізує індивідуальні відмінності, попереднє тестування забезпечує базову точку порівняння, фонові події та природний розвиток однаково впливають на обидві групи, а регресія до середнього контролюється завдяки R. Проте загроза взаємодії тестування з маніпуляцією залишається принциповою та обмежує можливість генералізації результатів на ситуації, де попереднє тестування відсутнє.

Саме для розв'язання цієї проблеми був розроблений найглибший та найповніший істинний експериментальний дизайн – чотиригруповий план Соломона. Він з'єднує в собі переваги попереднього тестування та переваги чистого післятестового дизайну, дозволяючи оцінити як основний ефект

впливу, так і ефект попереднього тестування, і навіть взаємодію між тестуванням і впливом. Його схема:

R	O	X	O
R	O		O
R		X	O
R			O

У цьому плані перша пара груп реалізує двогруповий план O X O, а друга пара – план X O / O. Це дозволяє оцінювати ефект втручання одразу чотирма способами, що суттєво підсилює надійність висновків. Наприклад, порівняння O₂ і O₄ дає інформацію про ефективність впливу з урахуванням попереднього тестування, тоді як порівняння O₅ і O₆ дозволяє визначити вплив X без ризику ефекту тестування. Порівняння O₂–O₁ та O₅–O₃ дозволяє оцінити, наскільки попереднє тестування могло посилити або послабити ефект втручання. Нарешті, порівняння O₆ з O₁ або O₃ дає інформацію про природний розвиток чи фонові впливи, що дозволяє здійснювати аналіз умовної стабільності поведінки.

Чотиригруповий план часто використовується у високоточних психолого-педагогічних експериментах, дослідженнях у когнітивній психології, а також при оцінці ефективності терапевтичних інтервенцій, де необхідно враховувати можливий вплив тестів на поведінку учасників. Наприклад, у дослідженні ефективності програми когнітивно-поведінкової терапії для зниження соціальної тривоги попереднє тестування може активізувати тривожні реакції, що впливатиме на результати втручання. План Соломона дозволяє відокремити ефект терапії від ефекту тестування, забезпечуючи найвищу валідність при одночасному контролі всіх головних загроз.

Перевагою цього плану є ідеальне управління як внутрішньою, так і зовнішньою валідністю: він контролює всі основні джерела артефактів, включно з ефектом тестування, взаємодією тестування з впливом, регресією, фоновими подіями та природним розвитком. Його головний недолік – це потреба в дуже великій вибірці та складність реалізації на практиці, а також значне навантаження на ресурси дослідника.

Усі істинні експериментальні плани, попри відмінності у структурі, мають принципову спільну рису: вони створені для того, щоб максимально чітко відокремити ефект незалежної змінної від усіх інших можливих факторів. Завдяки цьому істинні експерименти залишаються найнадійнішою основою доказових висновків у психології, освіті, клінічній практиці, нейропсихологічних дослідженнях, соціальній та прикладній психології.

7.5. Декілька незалежних змінних: повні факторні плани

Факторні експериментальні дизайни посідають особливе місце в арсеналі методологічних інструментів експериментальної психології, оскільки саме вони дають можливість дослідникові вийти за межі простого аналізу

ізолюваного впливу однієї незалежної змінної й перейти до виявлення більш складних, дво- або багатовимірних закономірностей поведінки. Коли психологічна реальність формується під впливом багатьох взаємодіючих чинників, дослідники потребують таких планів, які дозволяють одночасно варіювати кілька незалежних змінних і перевіряти не лише їх окремий (головний) ефект, а й ефект їх взаємодії. Такий підхід дозволяє досягти високого ступеня внутрішньої валідності та має суттєві переваги над одновимірними планами, адже дозволяє зрозуміти, як саме один фактор може змінювати вплив іншого, або як поєднання двох умов створює новий, специфічний ефект, який не заперечує окремих впливів, але не може бути пояснений ними поодиноці.

Важливо підкреслити, що використання факторних планів логічно випливає з сутності психологічного експерименту як процесу моделювання складних причинних структур. Оскільки поведінка, когнітивні процеси, емоційні стани та соціальна взаємодія завжди формуються під дією багатьох змінних, ізолювані експериментальні маніпуляції нерідко спрощують реальність до такої міри, що результати дослідження стають складноінтерпретованими або обмеженими з погляду зовнішньої валідності. Тому факторні плани дозволяють одночасно контролювати вплив сторонніх змінних шляхом включення їх у структуру експерименту як нових факторів, а також встановлювати кількісну природу залежності між змінними за наявності кількох їх рівнів. Саме багаторівневість є ключовим елементом, коли дослідник прагне не лише констатувати наявність ефекту, а й оцінити його форму – чи має вона лінійний, квадратичний або більш складний характер.

Повний факторний план означає, що всі рівні однієї незалежної змінної поєднуються з усіма рівнями іншої. Якщо дві незалежні змінні А і В мають по два рівні, план буде позначатися як 2×2 ; якщо одна змінна має три рівні, а інша два – план позначається як 3×2 ; якщо три змінні мають по два рівні – це $2 \times 2 \times 2$. У кожній комірці такого плану формується окрема група досліджуваних, яким пред'являється специфічна комбінація рівнів незалежних змінних. Позначення плану, таким чином, показує кількість груп, але ще важливіше – кількість теоретичних порівнянь, які може здійснити психолог. Кожен фактор створює можливість для аналізу головних ефектів, тоді як комбінації факторів створюють можливість аналізувати взаємодії. У двофакторному плані це взаємодія $A \times B$; у трьохфакторному – $A \times B \times C$, а також вторинні взаємодії $A \times B$, $A \times C$, $B \times C$.

З погляду валідності факторні експерименти мають суттєву перевагу завдяки кращому контролю над сторонніми впливами. Замість того, щоб намагатися утримувати сторонню змінну на постійному рівні або усувати її, дослідник вводить її як фактор у план, тим самим перетворюючи потенційне джерело артефактів на структурний елемент дизайну. Це дозволяє не лише уникнути загроз внутрішній валідності, але й зробити моделі більш узагальненими й застосовними до ширших контекстів.

Одним із найпоширеніших факторних експериментів є двофакторний план 2×2 . Його привабливість полягає у простоті реалізації та надзвичайно високій інформативності. Він дозволяє виявити три типи ефектів: головний ефект першої змінної, головний ефект другої змінної та їх взаємодію. Наприклад, у дослідженні впливу типу навчальної інструкції (інтерактивна vs. традиційна) та рівня тривожності студентів (низька vs. висока) на академічну успішність створюються чотири групи досліджуваних. Якщо ефект взаємодії виявиться значущим, це означатиме, що ефективність інтерактивної інструкції залежить від рівня тривожності. Наприклад, інтерактивні методи можуть бути вкрай ефективними для студентів із високою тривожністю та не давати суттєвого ефекту серед студентів із низькою тривожністю. Ця взаємодія була б неможливою для інтерпретації в межах будь-якого одновимірнього експерименту.

Трифакторні плани $2 \times 2 \times 2$ значно складніші, але вони дозволяють охопити три незалежні змінні, що є корисним при дослідженні складних когнітивних або соціально-психологічних процесів. Наприклад, у дослідженні впливу типу роботи пам'яті (робоча, довготривала), типу стимульного матеріалу (словесний, невербальний) та рівня мотивації учасника (висока, низька) створюється вісім груп, що дозволяє проаналізувати індивідуальні ефекти кожного фактора та численні взаємодії. Такий план дозволяє описати не лише те, чи змінюється пам'ять під співдією кількох чинників, але й те, як саме ця зміна відбувається на різних комбінаціях умов.

Плани 3×2 або 3×3 дають можливість встановити форму залежності між незалежною та залежною змінними, оскільки три рівні змінної вже дозволяють припустити тип функціональної залежності. Наприклад, у вивченні впливу рівня стресу (низький, середній, високий) на когнітивну продуктивність можна виявити не лише факт впливу стресу, але й форму цього впливу – наприклад, класичну інвертовану U-подібну криву, відповідно до якої низький та високий рівні стресу знижують продуктивність, а середній оптимізує її.

Особливо цінними у психологічних дослідженнях є плани, що дозволяють контролювати міжособистісні відмінності шляхом рандомізації або введення змінної як фактора. Наприклад, стать, вік, рівень освіти чи вихідний рівень навичок нерідко стають контрольними факторами, що дозволяє зрозуміти, як ефект незалежної змінної варіює у різних категоріях учасників. Такі плани значною мірою підвищують зовнішню валідність, адже результати стають застосовними до ширших популяцій.

Разом із розширенням кількості незалежних змінних зростають і вимоги до забезпечення внутрішньої валідності. Чим більше факторів включено в експеримент, тим більший ризик того, що ефекти сторонніх змінних будуть неоднаково проявлятися у різних комбінаціях умов. Проте саме факторний підхід значною мірою вирішує цю проблему, оскільки стороння змінна може бути не усунена, а інтегрована в дизайн як повноцінний фактор. Наприклад, якщо дослідник підозрює, що рівень мотивації може впливати на ефективність когнітивного тренування, він не намагається «утримати мотивацію сталою», а

перетворює її на незалежну змінну з кількома рівнями, тим самим перетворюючи потенційне джерело варіацій на конструктивний елемент плану. Такий підхід збільшує як внутрішню, так і зовнішню валідність експерименту.

Переваги факторних планів також полягають у можливості аналізувати взаємодії між змінними, адже психологічні процеси рідко формуються під впливом одного чинника. Взаємодія є одним із найцінніших результатів експериментального дослідження, оскільки вона демонструє, що дія одного фактора не є сталою, а змінюється залежно від рівнів іншого. У соціальній психології, наприклад, можна дослідити, як стиль керівництва (авторитарний vs. демократичний) взаємодіє з груповою згуртованістю (висока vs. низька), створюючи різні моделі продуктивності групи. У когнітивній психології часто вивчають взаємодію типу завдання (вербальне vs. просторове) та часу на його виконання (короткий, середній, тривалий), що дозволяє виявити, для яких типів завдань часові обмеження є найбільш критичними. У клінічній психології факторні плани дозволяють досліджувати взаємодію типу терапії та рівня вихідної симптоматики, демонструючи, що деякі методи ефективні лише для клієнтів із певними характеристиками.

Застосування багатофакторних планів суттєво підвищує вимоги до організації вибірки. Оскільки кількість комірок зростає пропорційно добутку кількості рівнів усіх факторів, дослідникам потрібні більші вибірки для збереження статистичної потужності. Наприклад, план $2 \times 2 \times 2$ містить вісім груп, і навіть якщо у кожній має бути мінімум 20 учасників для забезпечення адекватної потужності, загальна кількість досліджуваних становитиме не менше 160 осіб. Це одна з основних практичних труднощів, яка часом спонукає дослідників до компромісів, таких як зменшення кількості рівнів або застосування внутрішньогрупових планів, що дають змогу задіяти менше учасників.

У внутрішньогрупових факторних планах кожен учасник проходить усі комбінації умов. Це дає змогу різко зменшити загальну кількість учасників, але водночас створює ризики появи послідовних ефектів – втоми, практики, переносу й антиципації. У таких випадках необхідно застосовувати рандомізацію порядку пред'явлення умов, латинські квадрати або повне контрбалансування. Наприклад, якщо дослідник вивчає вплив типу стимулу (емоційний vs. нейтральний) та рівня когнітивного навантаження (низьке vs. високе) на швидкість реакції, він повинен подбати про те, щоб порядок комбінацій умов був різним для різних учасників. Це запобігає тому, щоб емоційний стимул завжди пред'являвся першим і не створював стабільної різниці, пов'язаної лише з ефектом порядку.

Загрози валідності у міжгрупових та внутрішньогрупових планах мають різну природу. У міжгрупових планах головною загрозою є нееквівалентність груп, яка може виникнути навіть за наявності рандомізації, якщо вибірка надто мала або випадково сформована група відрізняється від іншої за важливими змінними. Тому інколи додатково застосовують метод попарного добору

(matching), де учасників підбирають парами за ключовими ознаками, або стратифікований відбір, що забезпечує пропорційне представлення соціальних чи психологічних характеристик. У внутрішньогрупових планах навпаки, еквівалентність гарантується тим, що кожен учасник є власним контролем, але загрози пов'язані з послідовністю умов, ефектами перенесення та збільшенням обсягу роботи для кожного учасника.

Багатофакторні плани, зокрема трифакторні й чотирьохфакторні, дають змогу моделювати реальні психологічні ситуації з високою точністю, але їх практичне застосування обмежується ресурсами дослідника. У більшості експериментів використовуються плани 2×2 або 2×3 , оскільки вони забезпечують оптимальне співвідношення між складністю, інформативністю та зручністю інтерпретації. При цьому навіть найпростіший план 2×2 може виявитися надзвичайно продуктивним, якщо дослідник правильно обирає незалежні змінні та передбачає можливі взаємодії.

Повнофакторні плани також дозволяють значною мірою підвищити зовнішню валідність експерименту, оскільки вони створюють модель поведінки, ближчу до реальних умов, де результати людської діяльності визначаються не одним фактором, а сукупністю взаємодіючих умов. Наприклад, у вивченні професійного вигорання можна одночасно досліджувати вплив типу завдань, соціальної підтримки та індивідуальних характеристик особистості. Такий дизайн створює можливість описати, у яких саме комбінаціях факторів ризик вигорання є максимальним, а в яких – зменшується.

Не менш важливим є питання статистичного аналізу результатів факторних планів. Найчастіше застосовується багатофакторний дисперсійний аналіз (ANOVA), який дозволяє одночасно оцінити головні ефекти та їхні взаємодії. У більш складних випадках використовуються узагальнені лінійні моделі, моделі змішаних ефектів, MANOVA або навіть структурне моделювання, якщо залежна змінна має складну структуру. Обов'язковим компонентом аналізу є оцінка розміру ефекту – η^2 , ω^2 або інші індекси, які дозволяють зрозуміти реальний, а не лише статистично значущий вплив факторів.

7.6. Неповні, змішані та інші експериментальні плани

Неповні, змішані та інші експериментальні плани становлять важливу категорію сучасної експериментальної методології, оскільки дають змогу дослідникам працювати з реальними умовами, де суворі вимоги класичного істинного експерименту або надто дорогі, або надто складні, або взагалі неможливі до реалізації. Розвиток психології як науки поступово привів до усвідомлення того, що далеко не всі експериментальні ситуації допускають повне перетворення дослідження на контрольовану лабораторну процедуру, у якій кожен учасник може пройти всі стимули в ідеально збалансованому порядку, або де всі фактори можна рандомізувати чи вирівняти. Через це

виникають неповні системи варіацій, змішані структури факторів і багато типів гнучких експериментальних конфігурацій, які на перший погляд здаються менш строгими, але при правильному застосуванні здатні давати не менш переконливі й валідні висновки. Саме такі плани – неповні, змішані та рідкісні різновиди експериментальних дизайнів – становлять окрему методологічну сферу, що поєднує вимоги експериментального контролю з обмеженнями реальної практики дослідження поведінки.

Неповні експериментальні плани виникають там, де застосування повного факторного плану є неможливим або недоцільним. У класичному вигляді повний факторний план передбачає, що всі рівні всіх незалежних змінних поєднуються один з одним, утворюючи матрицю $N \times M$ або складніші комбінації на кшталт $L \times M \times N$, коли факторів декілька. Але в реальних дослідженнях психолог часто стикається з тим, що кількість рівнів фактора або кількість факторів є настільки великою, що повний план перетворюється на практично нездійсненну конструкцію, вимагаючи десятків або навіть сотень експериментальних груп. Неповні плани дозволяють частково відмовитися від повної комбінації рівнів, зберігаючи при цьому можливість оцінити вплив факторів, взаємодії або структуру різниць між умовами. Принцип побудови таких планів полягає у системному скороченні кількості необхідних комбінацій на основі математично обґрунтованих схем.

Одним із найвідоміших неповних планів є латинський квадрат. Його логіка ґрунтується на тому, що два набори рівнів незалежних змінних розташовуються у квадраті так, щоб кожен рівень однієї змінної з'являвся в кожному рядку та кожному стовпці лише один раз. Хоча такий план складається з меншої кількості умов, ніж повний факторний, він забезпечує контроль за двома послідовними ефектами – порядковим і блоковим. У психології латинський квадрат використовується для контролю порядку подання стимулів або рівнів експериментального впливу. Наприклад, якщо дослідник вивчає продуктивність запам'ятовування текстової інформації залежно від трьох типів навчальних інструкцій і трьох складностей матеріалу, він може зіткнутися з тим, що повний план потребуватиме дев'яти окремих груп або дев'яти повторних тестувань. Натомість латинський квадрат дозволяє мінімізувати кількість умов і водночас контролювати вплив порядку.

Греко-латинський квадрат є розширенням латинського і дозволяє контролювати вже три незалежні джерела варіації. У такому плані до латинських символів додається третя система маркування (грецькі символи), що дає змогу врівноважити три фактори таким чином, що кожна комбінація зустрічається лише один раз. Греко-латинські структури рідко використовуються у повномасштабних психологічних експериментах через складність побудови, але вони мають важливе методологічне значення саме тому, що демонструють можливість оптимального вирівнювання трьох порядкових або системно варіативних параметрів. У когнітивній психології їх іноді застосовують при вивченні індивідуальних відмінностей у декількох

когнітивних процесах (наприклад, тип стимула, його позиція в послідовності та спосіб інструкцій), коли неможливо дозволити собі повний факторний план.

Ще одним видом неповних планів є збалансовані неповні блокові плани, де кожен учасник проходить лише частину умов, але ця частина підібрана так, щоб загалом усі комбінації факторів були представлені пропорційно. У психології такі плани застосовують при тестуванні великої кількості стимулів – наприклад, у лінгвістичних дослідженнях, де учасник не може прослухати сотні аудіостимулів без втрати продуктивності; або у дослідженнях емоційного сприйняття, де надмірна кількість стимулів може викликати ефект звикання чи емоційне виснаження. Збалансовані неповні плани дозволяють уникнути цього, розподіляючи набори стимулів між учасниками таким чином, щоб жоден не був перевантажений, але всі умови були достатньо репрезентовані.

Неповні плани не забезпечують такого рівня контролю загроз внутрішній валідності, як істинні експериментальні дизайни, але вони дозволяють значною мірою контролювати ефекти порядку, історії, очікування та позиційні фактори. Їхня основна слабкість – обмежена можливість виявлення взаємодій між факторами, адже неповна структура не завжди дозволяє математично коректно оцінити ефект кожної комбінації факторів. Проте в багатьох реальних психологічних дослідженнях жорсткий контроль усіх можливих комбінацій є менш важливим, ніж можливість втримати експеримент у реалістичних межах часу, зусиль і когнітивної витривалості учасників.

Змішані плани становлять інший великий пласт сучасної експериментальної методології. Їхня суть полягає у поєднанні міжгрупових і внутрішньогрупових факторів у межах одного дослідження. Найважливіша перевага змішаних планів – можливість водночас контролювати індивідуальні відмінності (завдяки внутрішньогруповим елементам) і виявляти ефекти, стабільні для різних підгруп або умов (завдяки міжгруповим факторам). У психологічній практиці змішані плани використовують дуже часто, тому що багато процесів мають як індивідуальну, так і міжіндивідуальну динаміку. Наприклад, план, у якому одна група отримує когнітивний тренінг, а інша – контрольне завдання, але обидві групи вимірюються у кілька моментів часу, є класичним змішаним дизайном.

Загрози валідності у змішаних планах частково відмінні від неповних або простих факторних систем. Вони краще контролюють ефекти індивідуальних відмінностей, але гірше – ефекти порядку, якщо незалежна змінна має повторювані рівні. Тому змішані плани часто поєднують із методами контрбалансування або випадкової зміни порядку стимулів. У когнітивній психології змішані плани застосовують, коли дослідник хоче вивчити вплив того самого фактора на різні групи учасників у різні моменти часу. У дослідженні ефективності мнемонічних технік однією з груп учасників може бути студентська молодь, іншою – літні люди; обидві групи проходять кілька сесій тестування, що дозволяє оцінити різну динаміку навчання.

Змішаний план дає змогу аналізувати ефект факторів окремо і в їхній взаємодії.

Інші нестандартні плани включають перехресні (crossover) дизайни, де учасники по черзі отримують різні рівні незалежної змінної; плани зі складною стратифікованою рандомізацією; багатоетапні плани, де ефект попередньої маніпуляції стає новим фоном для наступної; адаптивні експериментальні плани, у яких параметри впливу можуть змінюватися залежно від індивідуальної реакції учасника.

Перехресні плани особливо важливі в психологічних дослідженнях, де ефект лікування або втручання швидко зникає. Наприклад, у дослідженнях впливу короточасних практик медитації на рівень тривожності той самий учасник може у різні дні пробувати різні види вправ, а дослідник порівнює ефекти всередині однієї особи. Однак такі плани погано підходять для втручань із довготривалим ефектом, наприклад, із психотерапією.

Неповні, змішані та інші плани відіграють винятково важливу роль там, де метою є зберегти високу внутрішню валідність, але при цьому максимально врахувати поведінкову реальність. Психологічні дослідження рідко бувають ідеальними лабораторними моделями, і часто найбільш валідним є саме той план, який найкраще узгоджується з природою феномену, що вивчається. Це означає, що вибір між повним, неповним і змішаним планом не є питанням математичної чистоти; це питання відповідності дослідницькій логіці, умовам реального експерименту та поведінковим властивостям учасників.

7.7. Характеристика квазіекспериментальних планів

Квазіекспериментальні плани посідають важливе місце в арсеналі методології психологічних досліджень, оскільки вони дозволяють здійснювати аналіз причинно-наслідкових зв'язків у ситуаціях, коли повноцінна рандомізація груп є неможливою або етично недоцільною. На відміну від істинних експериментів, де рандомізація виступає головним принципом еквівалентності груп, квазіекспериментальні плани компенсують її відсутність завдяки багаторазовим вимірюванням, включенню контрольних груп, поєднанню часових рядів, використанню природних подій як джерел варіювання або застосуванню статистичних методів контролю. Вони стали особливо затребуваними в освітній, клінічній, соціальній, організаційній і когнітивній психології, де суворий контроль за умовами дослідження часто неможливий. Ключовою ознакою квазіекспериментів є те, що незалежна змінна має природне або адміністративне походження, а дослідник не може керувати випадковим розподілом досліджуваних у групі, що створює потенційні загрози внутрішній валідності та потребує складніших компенсаторних механізмів.

Окреме місце у квазіекспериментальній методології посідають перервані часові ряди, які дозволяють аналізувати ефекти впливу на основі

багаторазових вимірювань до і після введення незалежної змінної. Їх класична структура має форму:

$O_1 O_2 O_3 O_4 X O_5 O_6 O_7 O_8$.

Такий підхід дає змогу спостерігати тенденції зміни залежної змінної, оцінювати стабільність початкового тренду і визначати, чи відбулася різка або поступова зміна після введення впливу. У психології подібні плани застосовуються, наприклад, для оцінювання впливу впровадження нової системи профілактики вигорання у великій організації або для дослідження ефектів запуску освітньої реформи на загальну успішність студентів. Сильна сторона плану – у контролі фону: якщо до втручання спостерігається стабільна тенденція, а після – різка зміна, її важко пояснити лише історією чи природним розвитком. Проте значна небезпека пов'язана з тим, що зовнішня подія, яка збіглася у часі з впливом, може бути помилково інтерпретована як його ефект.

Ще більш надійним варіантом є перервані часові ряди з нерівноеквівалентною контрольною групою, які мають структуру:

$O_1 O_2 O_3 O_4 X O_5 O_6 O_7 O_8$

$O_9 O_{10} O_{11} O_{12} O_{13} O_{14} O_{15} O_{16}$.

Застосування контрольної групи дозволяє значно підвищити внутрішню валідність, оскільки зміни в експериментальній групі порівнюються з динамікою контрольної. Якщо в експериментальній групі після введення впливу спостерігається різкий злам тренду, а в контрольній – продовження початкової тенденції, причинні висновки стають значно більш обґрунтованими. У соціальній психології цей план може бути використаний для перевірки ефективності масштабної антикризової програми у школах, де експериментальні школи отримують нову систему роботи з булінгом, а контрольні – ні. Такий план, попри високу валідність, вимагає тривалого періоду збору даних, що не завжди є можливим.

Важливою категорією квазіекспериментів є регресійні дискретизовані плани, що виникають у випадках, коли групи формуються на основі певної порогової змінної, яка впливає на включення в групу впливу або контроль. Їх базова структура може мати вигляд $O_1 X O_2$ для учасників, що перевищили поріг, і $O_3 O_4$ для учасників нижче порогу. Типовим прикладом є ситуація, коли студентів із низькими балами за діагностичною шкалою тривожності відправляють на додатковий курс тренінгів, а тих, хто має вищі показники, не включають у програму. Аналіз таких даних потребує застосування регресійних та коваріаційних моделей, які дозволяють компенсувати різні початкові рівні залежної змінної. Проте такі плани не забезпечують контролю регресії до середнього і мають високий ризик артефактів, пов'язаних із селекцією.

Досить поширеним у практичних психологічних дослідженнях є план ротації або чергування впливів, який може мати структуру:

$O_1 X_A O_2 X^b O_3 X_A O_4$ або подібні варіанти.

Він дозволяє оцінювати не лише вплив різних рівнів незалежної змінної, але й те, як зміна послідовності впливає на результати. У нейропсихології

подібні плани використовуються, коли потрібно порівняти ефективність двох методик тренування робочої пам'яті: одні учасники спочатку працюють у режимі А, потім – В, а інші – у зворотному порядку. Такий підхід частково контролює ефекти порядку, але не ліквідує загрози, пов'язані з природним розвитком та історією.

У деяких випадках квазіекспериментальні плани ґрунтуються на природних подіях, які виконують роль незалежної змінної, а дослідник лише фіксує їхній вплив. Структура таких планів наближається до перерваних часових рядів і може виглядати як $O_1 O_2 O_3 O_4 X O_5 O_6 O_7 O_8$, але X є природною подією – військовою кризою, зміною законодавства, масовою міграцією, стихійним лихом. У кризовій психології подібний підхід дозволяє, наприклад, аналізувати динаміку ПТСР серед населення до і після трагічної події. Хоча такі плани мають високу екологічну валідність, вони абсолютно не контролюють фактори складу груп, а будь-які зміни можуть бути результатом сукупності багатьох неконтрольованих впливів.

Квазіексперименти також можуть застосовувати складні змішані конструкції, які поєднують елементи міжгрупових і внутрішньогрупових планів. Структура подібних планів може мати вигляд $O_1 X O_2 / O_3 O_4$ з додатковими повторними вимірюваннями або варіюванням рівнів впливу. У дослідженнях когнітивної науки та психології розвитку такі плани використовуються для оцінювання навчальних програм, де одна група отримує інтенсивний курс навчання, а інша – стандартний, але дослідження включає кілька точок вимірювання, що дає змогу оцінити темпи прогресу. Подібні плани частково контролюють історію, природний розвиток і селекцію, але не можуть забезпечити повного контролю всіх загроз валідності.

Аналіз валідності квазіекспериментів вимагає наголосу на тому, що вони ніколи не досягають рівня методологічної чистоти істинних експериментів. Проте вони можуть забезпечити достатньо високу внутрішню валідність, якщо містять відображення структури часових змін, наявність контрольних груп, статистичне коригування, використання порогових або дискретизованих процедур. Для контролю історії та природного розвитку перевагу мають перервані часові ряди; для контролю складу груп – плани з попереднім тестуванням; для контролю селекції – регресійні дискретизовані конструкції; для підвищення зовнішньої валідності – природні експерименти.

У психологічних дослідженнях квазіексперименти дозволяють вивчати ефективність шкільних програм розвитку соціально-емоційних компетентностей, терапевтичних втручань у клінічних умовах, профілактичних заходів у сфері ментального здоров'я, вплив змін у політиці безпеки, ефективність організаційних реформ, вплив медійних кампаній та інші складні явища, де рандомізація є неможливою. Їхній розвиток свідчить про розширення методологічної свідомості психологів та прагнення працювати з реальними контекстами, де контроль за умовами часто обмежений, а дослідницькі ситуації визначаються природними обставинами.

7.8. Дизайн досліджень для маленьких вибірок

Дизайни досліджень для маленьких вибірок посідають особливе місце в методології психологічних експериментів, оскільки значна частина реальних дослідницьких ситуацій не дозволяє набрати достатньо великий обсяг учасників. Це особливо характерно для клінічної психології, нейропсихології, досліджень рідкісних патологій, спеціалізованих груп, а також прикладних тестувань нових психотерапевтичних або реабілітаційних методів. На відміну від класичних експериментальних планів, що покладаються на статистичну силу великих вибірок, дослідження з малою кількістю учасників зосереджені на детальному аналізі кожного окремого випадку або на порівнянні невеликих груп за умов максимально жорсткого контролю внутрішньої валідності. Такі дизайни мають власну логіку побудови, що передбачає багаторазові вимірювання, повторення експериментальних впливів, ротацію умов або глибокий аналіз змін у межах однієї особи. Саме тому вони дозволяють формувати висновки навіть за умов, коли статистичні критерії, що потребують великих вибірок, стають непридатними.

Одним із найвідоміших підходів є дизайн одиничного випадку (Single-Case Experimental Design), який широко поширений у клінічній практиці та когнітивній нейропсихології. Він базується на детальному багатократному вимірюванні поведінки одного індивіда до, під час та після втручання. Його класична структура передбачає чергування фаз базового рівня та фаз впливу, що дозволяє відстежувати зміну показника у динаміці. Найпростішим його варіантом є план АВА: спочатку здійснюється вимірювання базового стану O_1 O_2 $O_3 \dots$, потім вводиться вплив X і проводяться повторні вимірювання O_4 O_5 $O_6 \dots$, після чого вплив припиняється, і знову фіксується повернення до базового рівня O_7 O_8 O_9 . Ідея полягає в тому, що якщо поведінка змінюється після введення впливу і повертається до вихідного рівня після його припинення, можна зробити обґрунтований висновок про причинність. Такий план добре контролює загрози історії та природного розвитку, оскільки зміни, що не пов'язані з впливом, не мають чіткої точки зламу. Він також частково контролює індивідуальні відмінності, адже в межах одного індивіда немає проблеми нееквівалентності груп. Проте він погано контролює регресію до середнього, а ефекти тестування можуть впливати на результати, оскільки багаторазове вимірювання здатне призвести до навчання або втоми.

Існує також розширена версія такого дизайну – план АВAB, що передбачає дворазове введення та дворазове скасування впливу. Його структура виглядає як O_1 O_2 $O_3 - X - O_4$ O_5 $O_6 - \text{відміна } X - O_7$ O_8 $O_9 - \text{повторне введення } X - O_{10}$ O_{11} O_{12} . Повторний ефект впливу дозволяє зробити причинні висновки більш переконливими, оскільки ймовірність того, що два ідентичні злами тренду пояснюються випадковими подіями, є значно меншою. У поведінковій психотерапії цей план часто застосовується для оцінювання ефективності індивідуальних методів: наприклад, дослідник вивчає, чи допомагає техніка релаксації знизити частоту панічних атак у конкретного

клієнта. Якщо під час фаз X спостерігається зменшення симптомів, а під час фаз без X – їхнє повернення, це свідчить про причинний ефект.

Особливої уваги заслуговують багатобазові дизайни, які дають змогу уникнути етичної проблеми, пов'язаної зі скасуванням терапевтичного впливу. У таких планах дослідник створює кілька базових ліній по різних учасниках, різних поведінкових реакціях однієї особи або різних умовах середовища. Класична структура плану може мати вигляд $O_1 O_2 O_3 O_4 X O_5 O_6 O_7$ для першого учасника, $O_1 O_2 O_3 O_4 O_5 X O_6 O_7 O_8$ – для другого та більш довгу базову лінію для третього. Вплив вводиться у різний час, що дозволяє виключити вплив випадкових подій: якщо зміна настає тільки після X, але не під час тривалого попереднього спостереження, це є доказом причинності. Перевагою цього дизайну є висока внутрішня валідність за рахунок часової диференціації впливу, а також відсутність необхідності скасовувати лікування. Його активно використовують у роботі з дітьми з розладами спектра аутизму, коли оцінюють ефективність втручання щодо різних поведінкових патернів – наприклад, розвиток навичок комунікації, зменшення стереотипій або підвищення соціальної взаємодії.

Для ситуацій, у яких необхідно дослідити вплив різної інтенсивності або різних доз втручання, застосовуються плани чергування умов. Їхня структура полягає у випадковій зміні рівнів впливу X_A , X_B , X_C у коротких блоках, наприклад: $O_1 X_A O_2 X_B O_3 X_C O_4 X_A O_5 X_B O_6$. Такі дизайни дозволяють порівняти ефективність кількох форм впливу у межах одного учасника. У когнітивній психології подібні плани використовують для аналізу того, як різні режими тренування робочої пам'яті впливають на продуктивність: учасник може виконувати завдання з різною швидкістю, різним рівнем навантаження або різним форматом стимулів. Основна загроза валідності пов'язана з ефектами порядку та можливим перенесенням ефектів між умовами, проте ротаційний принцип, якщо він рандомізований, дає змогу частково компенсувати ці недоліки.

У випадках, коли вибірка складається з дуже малої кількості учасників, але кожен з них проходить кілька умов, застосовуються латинські квадрати та збалансовані послідовні плани. Вони використовують математичні принципи балансування, що дозволяють кожному впливу з'явитися в кожній позиції порядку однакову кількість разів. Це дозволяє контролювати позиційні та послідовні ефекти, що особливо важливо у дослідженнях когнітивних процесів, де виконання одного типу завдання може впливати на виконання наступного. Наприклад, у дослідженні впливу видів праймінгу на швидкість розпізнавання емоційних стимулів дослідник може збалансувати порядок появи позитивних, негативних і нейтральних праймів так, щоб кожен тип мав однакові умови подання.

Незважаючи на різноманіття схем, усі дизайни для маленьких вибірок мають характерні обмеження. Найсуттєвішою загрозою для валідності є історія: зовнішні події, що відбуваються під час тривалих спостережень, можуть змінювати поведінку незалежно від впливу. Також природний

розвиток може провокувати зміни у поведінці без участі незалежної змінної. Ефекти тестування стають особливо помітними, оскільки багаторазове вимірювання стимулює ефекти звикання, практики або втоми. Проблемою є і регресія до середнього, особливо в клінічних зразках, коли екстремальні показники з часом повертаються до типових значень. Усі ці загрози вимагають багаторазових вимірювань, ретельного документування трендів та використання структур, які дозволяють відокремити власне ефект впливу від спонтанних коливань.

Проте дизайни для малих вибірок мають і сильні методологічні переваги. Вони забезпечують глибокий аналіз динаміки окремого випадку, що є цінним у клінічній практиці. Вони дозволяють виявляти ефекти, які неможливо помітити у групових середніх, особливо в ситуаціях високої індивідуальної варіативності. Вони є незамінними у випадках, де неможливо контролювано маніпулювати групами, але можна аналізувати траєкторію змін. Крім того, вони забезпечують високу екологічну валідність, оскільки часто проводяться у природних умовах і відображають індивідуальну поведінку в реальних контекстах.

Загалом, дизайни для маленьких вибірок відіграють ключову роль у сучасній психологічній методології, забезпечуючи можливість наукового інструментального аналізу там, де класичні групові експерименти не можуть бути реалізовані. Незважаючи на обмеження, вони дозволяють формувати причинні висновки, якщо структура плану передбачає достатній контроль за внутрішньою валідністю, а кількість вимірювань дає змогу відокремити закономірні зміни від випадкових. Їхнє грамотне використання забезпечує досліднику можливість працювати з унікальними групами, рідкісними випадками та індивідуальними поведінковими профілями, що робить їх незамінним інструментом у тих галузях психології, де індивідуальність та контекст відіграють центральну роль.

7.9. Експериментальні плани для одного досліджуваного

У психології такі плани часто застосовуються у клінічній практиці, у роботі з дітьми, у дослідженнях рідкісних когнітивних порушень, а також у ситуаціях, коли дослідник прагне простежити індивідуальну динаміку реагування на втручання. На відміну від традиційних групових планів, вони не спираються на статистичну силу великих вибірок; їхня логіка побудована на послідовному зіставленні фаз спостереження, що дозволяє відокремити ефект незалежної змінної від природних флуктуацій поведінки. Саме тому вони використовують багаторазові вимірювання, поетапні зміни експериментальних умов, контроль часових закономірностей та докладний аналіз форм поведінкових кривих.

У своїй класичній формі план для одного досліджуваного починається з фази базового рівня, який реєструється багаторазово та відображає природний стан залежної змінної до введення експериментального впливу. Такі

вимірювання, позначені як $O_1 O_2 O_3 O_4 \dots$, дозволяють зафіксувати тенденції, спричинені природним розвитком, спонтанними коливаннями чи ситуативними чинниками. За відсутності стабільного або передбачуваного базового рівня експериментальний вплив неможливо інтерпретувати, оскільки будь-які зміни після введення незалежної змінної можуть бути спричинені внутрішніми коливаннями залежної змінної, а не впливом X .

Однією з найпростіших форм планів для одного досліджуваного є дизайн типу АВ, де А позначає фазу базового рівня ($O_1 O_2 O_3 \dots$), а В – фазу впливу ($X O_4 O_5 O_6 \dots$). Це найменш сильний серед експериментальних планів такого типу, адже він демонструє зміну поведінки лише після введення X , але не містить механізму для перевірки того, чи є ця зміна наслідком саме втручання. Загрози валідності – історія, природний розвиток, регресія до середнього – залишаються неконтрольованими. Наприклад, у дослідженні ефективності короткої техніки дихальної релаксації при епізодах тривоги у підлітка спостерігач може побачити зменшення симптомів під час фази В, але без додаткової інформації неможливо встановити причинність, адже симптоми могли зменшитися самі собою.

Набагато сильнішою конструкцією є план типу АВА, у якому після введення впливу дослідник припиняє його та знову спостерігає зміну поведінки. Структура виглядає як $O_1 O_2 O_3 - X - O_4 O_5 O_6 - \text{скасування } X - O_7 O_8 O_9$. Якщо поведінка повертається до початкового рівня на останній фазі, це надає вагомих аргументів на користь причинної ролі впливу. Наприклад, якщо у клієнта з логофобією спостерігається зменшення страху мовлення під час застосування прийому систематичної десенсибілізації, а після припинення впливу страх знову збільшується, це значно підсилює інтерпретацію причинності. План контролює загрози історії та природного розвитку, оскільки випадкові події не можуть пояснити чіткі переломні точки між фазами, але ефекти тестування та регресія до середнього лишаються потенційними конфаундерами.

Ще більш переконливим є класичний план АВ АВ, або план повторного введення та скасування впливу: $O_1 O_2 O_3 - X - O_4 O_5 O_6 - \text{без } X - O_7 O_8 O_9 - \text{повторне } X - O_{10} O_{11} O_{12}$. Таке повторення дозволяє довести, що вплив має систематичний ефект, адже незалежне повернення поведінки до базового рівня після кожного скасування впливу є вкрай мало ймовірним результатом випадкових подій. У психології цей підхід активно використовують для оцінки ефективності індивідуальних терапевтичних втручань – наприклад, для аналізу зменшення тикових реакцій у дітей із синдромом Туретта після застосування техніки зворотного формування звички. Якщо під час фаз впливу кількість тиков стабільно знижується, а під час фаз без впливу – повертається, це створює надзвичайно сильне підґрунтя для причинних висновків.

Окрему групу становлять плани з множинною базовою лінією, що застосовуються тоді, коли з етичних або практичних причин не можна скасовувати терапевтичний вплив. У таких планах дослідник спостерігає кілька базових ліній, які відрізняються за тривалістю: для різних осіб, різних

поведінкових реакцій або різних умов. Типова структура виглядає як $O_1 O_2 O_3 O_4 X O_5 O_6 O_7$ для одного учасника, $O_1 O_2 O_3 O_4 O_5 X O_6 O_7 O_8$ для другого та ще довший базовий відрізок для третього. Важливо, що зміни відбуваються тільки після введення X , тоді як під час розширеної фази базового рівня ніяких тенденцій не спостерігається. Наприклад, у дослідженнях когнітивно-поведінкових інтервенцій при аутизмі можуть бути використані різні поведінкові цілі (зоровий контакт, використання жестів, імітація). Якщо кожна з них демонструє зміни лише після впливу та лише після того моменту, коли X вводиться саме для цієї цілі, це дозволяє сформулювати причинний висновок навіть за наявності одного учасника. Такий план забезпечує високу внутрішню валідність за рахунок часової ізоляції впливів, але він не може повністю контролювати історію або природний розвиток, якщо зовнішні події впливають на всі поведінкові реакції одночасно.

Дизайни чергування умов становлять іншу важливу групу підходів, що дозволяють дослідити порівняльну ефективність різних типів впливу у коротких інтервалах часу. Вони можуть мати структуру $X_A O_1 - X_B O_2 - X^c O_3 - X_A O_4 - X_B O_5 - X^c O_6$. Такі плани використовуються в експериментальній аналізі поведінки, у нейропсихологічних тестових батареях, у порівнянні різних навчальних програм. Наприклад, якщо дослідник хоче визначити, яка з двох когнітивних технік (словесна репетиція або візуалізація) ефективніша для запам'ятовування слів у пацієнта з амнестичними порушеннями, він може чергувати ці техніки у випадковому порядку. Якщо один з видів впливу послідовно передбачає кращий результат незалежно від позиції у послідовності, це дозволяє робити локальні причинні висновки. Основною загрозою валідності є ефекти порядку, які частково компенсуються рандомізацією та збалансованими схемами.

Своєрідним різновидом планів для одного учасника є зигзагоподібні або градуйовані плани, у яких дослідник маніпулює інтенсивністю впливу послідовно: $X_1 O_1 O_2 - X_2 O_3 O_4 - X_3 O_5 O_6$. У цьому випадку метою є виявлення дозозалежної реакції. У нейрокогнітивних дослідженнях подібні підходи використовують для оцінки ефектів різної складності завдань на швидкість обробки інформації у пацієнтів після черепно-мозкової травми. Якщо збільшення складності викликає систематичне погіршення продуктивності, це дозволяє досліднику зробити висновки про збережені та порушені когнітивні механізми.

Усі ці плани мають спільну сильну сторону – вони забезпечують глибоку чутливість до індивідуальних змін, що часто недоступно груповим методам. Вони контролюють індивідуальні відмінності завдяки тому, що одна й та сама особа виступає і експериментальною, і контрольною умовою. Водночас вони мають свої слабкі місця. Найбільш критичними загрозами валідності є історія, природний розвиток та ефекти тестування, особливо за тривалих спостережень. Регресія до середнього також може суттєво вплинути на висновки, коли початкові значення є екстремальними. Зовнішня валідність у таких планах є обмеженою, адже отриманий ефект може бути унікальним для

одного конкретного учасника і не узагальнюватися на інші групи. Проте саме глибина аналізу одиничного випадку дає можливість формувати теорії у ситуаціях, коли групові експерименти недоступні.

У підсумку плани експериментів для одного досліджуваного є важливою складовою методологічного арсеналу психології. Вони дозволяють отримати причинні висновки там, де групові дослідження неможливо провести, забезпечують тонкий аналіз поведінкових змін у часі, створюють основу для клінічних рішень та індивідуалізованих інтервенцій. Попри обмеження, грамотне застосування структур типу АВ, АВА, АВАВ, множинних базових ліній або чергування умов забезпечує високу якість внутрішньої валідності та дозволяє досліднику виявити закономірності, які втрачаються при усередненні у великих вибірках.

Завдання на самостійну підготовку

1. Чому чітке визначення незалежної та залежної змінних є ключовою вимогою експериментального дизайну?
2. Які ризики виникають, якщо залежна змінна вимірюється інструментом із низькою валідністю?
3. Чому рандомізація вважається основним механізмом забезпечення еквівалентності груп?
4. Чому недостатній обсяг вибірки призводить до зниження статистичної потужності експерименту?
5. У чому полягає головна відмінність між доекспериментальними, істинними та квазіекспериментальними планами?
6. Чому доекспериментальні плани забезпечують найнижчий рівень внутрішньої валідності?
7. Які методи вирівнювання груп використовуються у міжгрупових планах?
8. Які загрози характерні для внутрішньогрупових планів і як їх можна мінімізувати?
9. Які головні обмеження має план одиничного вимірювання після впливу (X O)?
10. Чому план O X O залишається вразливим до ефектів тестування та природного розвитку?
11. Яку головну загрозу валідності не контролює план порівняння статичних груп (X O / O)?
12. Які переваги має план з рандомізованими групами та тестуванням лише після впливу?
13. У чому полягає логіка двогрупового плану з попереднім і підсумковим тестуванням?
14. Чому план Соломона є найбільш потужною конструкцією для контролю загроз валідності?
15. Які переваги повних факторних планів порівняно з одновимірними експериментами?

16. У чому полягає сенс аналізу взаємодії незалежних змінних у факторному дизайні?

17. У чому полягає принцип побудови латинського квадрата та які ефекти він дозволяє контролювати?

18. Які особливості мають змішані експериментальні плани та у яких випадках вони є найбільш доцільними?

19. Які методологічні переваги мають перервані часові ряди порівняно з одноразовими вимірюваннями?

20. Чому в дослідженнях із малими вибірками особливо важливими є багаторазові вимірювання та аналіз індивідуальних траєкторій змін?

РОЗДІЛ 8

ПЛАНИ КОРЕЛЯЦІЙНИХ ДОСЛІДЖЕНЬ В ПСИХОЛОГІЇ

8.1. Теоретико-методологічні засади кореляційних досліджень у психології

Кореляційні дослідження посідають одне з центральних місць у сучасній психологічній науці, оскільки спрямовані на виявлення закономірностей між різними психологічними характеристиками та поведінковими проявами, не втручаючись у природний хід подій і не змінюючи експериментально умови їх виникнення. Саме тому кореляційний підхід став одним з найбільш поширених інструментів у тих галузях психології, де пряме маніпулювання змінними є неможливим або етично неприйнятним. Він дозволяє відтворити природну динаміку взаємозв'язків у складних психічних явищах, досліджувати індивідуальні відмінності, структуру особистісних рис, рівень адаптації, когнітивні процеси, емоційні стани, показники психічного здоров'я й багато інших феноменів, що проявляються в реальних соціальних і життєвих контекстах. Проте попри універсальність і високу дослідницьку цінність кореляційного підходу, він ґрунтується на складних теоретико-методологічних засадах, що зумовлюють як його силу, так і його ключові обмеження.

Передусім кореляційний метод виходить з припущення, що психологічні змінні можуть бути представлені у вигляді кількісних показників, які здатні відображати індивідуальні відмінності з достатньою точністю для подальшого математичного аналізу. Це означає, що будь-яке кореляційне дослідження передбачає операціоналізацію теоретичних конструктів – складний процес, який перетворює абстрактні психологічні явища на вимірювальні індикатори. На цьому етапі виникає фундаментальна методологічна проблема: якою мірою емпіричні показники дійсно відображають поняття, яке дослідник прагне вивчати? Відповідь на це питання залежить від валідності та надійності використовуваних методик, від точності перекладу й культурної адаптації тестів, від умов проведення дослідження і навіть від мотивації респондентів. Через це будь-яка кореляція ніколи не може розглядатися у відриві від якості психологічного вимірювання, а помилки вимірювання неминуче зменшують величину кореляції й створюють уявну слабкість зв'язків, яких насправді може не бути.

Водночас кореляційні дослідження базуються на ймовірнісній природі статистичних залежностей: вони не виявляють жорстких детермінованих законів, а відображають тенденції, що проявляються в середньому, на рівні вибірки. Кореляція завжди є узагальненим індикатором, і навіть висока величина зв'язку не означає, що він однаково проявляється у всіх учасників. Це принципово відрізняє психологію від точних наук, у яких причинно-наслідкові залежності мають однозначний характер. Психологічні процеси завжди є багатофакторними, нелінійними, контекстуально залежними, а

кореляція лише фіксує співзміність показників, не розкриваючи природу цієї співзміності й не пояснюючи механізмів, які лежать в її основі.

Одним з найважливіших теоретичних обмежень кореляційних досліджень є неможливість зробити висновок про причинність. Кореляція відображає спільну варіативність змінних, але не встановлює, яка зі змінних є причиною, а яка – наслідком, або чи не є обидві вони наслідком третьої, невимірюваної змінної. Це означає, що навіть висока кореляція між двома змінними не дозволяє стверджувати, що одна з них впливає на іншу. У психології це явище часто називають проблемою «третьої змінної», коли за видимим зв'язком між двома показниками приховано інший фактор, який одночасно визначає їх обидва. Наприклад, кореляція між тривожністю та успішністю може бути спричинена не прямим зв'язком, а загальним рівнем стресу, підтримкою в сім'ї або індивідуальними відмінностями у мотивації. Саме тому кореляційні дослідження потребують особливо ретельного теоретичного обґрунтування: без чіткої концептуальної моделі вони легко перетворюються на набір статистичних фактів, які не мають пояснювальної сили.

Не менш важливим обмеженням є залежність кореляції від форми зв'язку. Традиційна кореляція Пірсона фіксує лише лінійні залежності, тобто такі, за яких зміна однієї змінної супроводжується пропорційною зміною іншої. Проте багато психологічних відносин мають нелінійний характер – наприклад, рівень стресу може позитивно впливати на продуктивність до певної межі, після чого надмірний стрес різко знижує ефективність. У таких випадках коефіцієнт кореляції Пірсона може бути близьким до нуля, хоча між змінними існує чітка функціональна залежність. Це означає, що кореляційний підхід не може автоматично застосовуватися без попереднього аналізу характеру розподілу та форми зв'язку, а дослідник має враховувати можливість криволінійних, порогових або складних інтерактивних відносин, які не піддаються простому лінійному опису.

Кореляційний аналіз також суттєво залежить від статистичних припущень – нормальності розподілу, відсутності значних викидів, гомогенності дисперсій. У реальних психологічних вибірках ці припущення часто порушуються. Психологічні показники нерідко мають асиметричні розподіли, оскільки емоційні, когнітивні або поведінкові стани можуть мати природні межі або специфічні патерни розподілу в популяції. Крім того, у дослідженнях особистості, клінічній психології та нейропсихології часто трапляються викиди – унікальні або крайні значення, що суттєво спотворюють оцінки кореляції. У таких випадках традиційні методи кореляції можуть давати хибні результати, а досліднику необхідно використовувати робастні підходи, непараметричні коефіцієнти або перетворення даних, щоб забезпечити коректність висновків.

Структурна складність психологічних явищ зумовлює ще одне обмеження кореляційного підходу – можливість виявлення спуріозних зв'язків. У великих масивах даних, особливо за умови багатократного

тестування, значущі кореляції можуть виникати випадково. Це є однією з методологічних причин реплікаційної кризи в сучасній психології, коли результати, отримані в окремому дослідженні, не відтворюються при повторних спробах, оскільки відображали випадкові співпадіння, а не реальні психологічні закономірності. Звідси випливає необхідність контролю за множинними порівняннями, застосування корекцій рівня значущості, а також підтвердження будь-яких кореляційних висновків незалежними вибірками або альтернативними методами аналізу.

Ще однією важливою особливістю кореляційних досліджень є їхня чутливість до ширших соціальних і контекстуальних впливів. Психологічні змінні не існують у вакуумі: на їх прояви впливають культура, соціальне середовище, економічні умови, геополітичні події, міжособистісні стосунки та життєвий досвід. Через це кореляції, отримані в одній популяції, не завжди можна переносити на іншу. Емпіричні зв'язки часто залежать від часу, контексту і навіть від способу формулювання запитань у тестах. Це ставить перед дослідником завдання врахування культурних, етнічних і соціальних особливостей вибірки, уважного аналізу методів репрезентативності та доцільності узагальнення висновків.

З іншого боку, кореляційні дослідження є незамінними для виявлення структури психологічних конструктів. Багатофакторні методи, такі як факторний аналіз, структурне моделювання, латентні змінні та інші складні статистичні підходи, ґрунтуються на кореляційних матрицях і потребують саме кореляційного дизайну. У цьому контексті кореляційний підхід стає ключовим інструментом для валідації психологічних тестів, для розуміння архітектури інтелекту, структури емоційних переживань, організації особистісних рис. Кореляції дозволяють побачити приховані патерни, виявити латентні зв'язки, сформулювати нові гіпотези та перевірити теоретичні моделі щодо відповідності емпіричним даним.

Водночас кореляційний підхід має важливе методологічне обмеження, яке полягає в тому, що він не здатний повністю врахувати вплив часу. Більшість кореляційних досліджень є поперечними, тобто виконуються в один момент часу. Такі дослідження фіксують одночасні значення змінних, але не здатні виявити, як вони змінюються у часі, що з'являється раніше, а що пізніше, і який характер мають причинно-наслідкові зв'язки. Саме тому дедалі більшої популярності набувають лонгітюдні кореляційні дизайни, які дозволяють у перспективі досліджувати зміни психологічних характеристик, простежувати розвиток явищ, бачити динаміку, а не лише статичні зв'язки.

Таким чином, кореляційні дослідження у психології мають глибокі теоретико-методологічні засади, що визначають їх можливості та обмеження. Вони дозволяють виявляти закономірності, структурувати психологічні конструкти, досліджувати індивідуальні відмінності та аналізувати взаємозв'язки між численними аспектами психічного життя. Однак їхня сила водночас є й їхньою слабкістю: вони не виявляють причинності, чутливі до помилок вимірювання, залежні від форми розподілу, вразливі до випадкових

співпадінь, контекстуальних впливів і культурних відмінностей. Тому кореляційний підхід повинен застосовуватися з високою методологічною компетентністю, із глибоким розумінням його суті, обмежень і можливостей. Лише за умови такого відповідального застосування кореляційні дослідження здатні не лише виявляти статистичні зв'язки, а й надавати змістовну інформацію про психологічні механізми, структуру психіки та закономірності людської поведінки.

8.2. Коефіцієнт кореляції як показник статистичних взаємозв'язків

Коефіцієнт кореляції виступає кількісним індикатором того, чи існує статистичний взаємозв'язок між феноменами. Його значення формує вихідну точку для прийняття рішень щодо подальшої моделі аналізу, розуміння структури явища та побудови теоретичних моделей.

Класичний коефіцієнт кореляції Пірсона оцінює силу лінійного зв'язку між двома змінними та передбачає, що кожна з них вимірювана в інтервальній шкалі або шкалі відношень, має приблизно нормальний розподіл і не містить серйозних відхилень у вигляді викидів, які здатні суттєво змінити результат. Діапазон значення r коливається від -1 до $+1$: крайні значення свідчать про ідеальні залежності, а значення, близькі до нуля, – про їх відсутність. Проте нульове значення r зовсім не означає відсутності зв'язку як такого: воно свідчить лише про відсутність лінійного зв'язку, водночас нелінійні закономірності можуть залишатися прихованими для цього показника. Так, у психології нерідко зустрічаються U-подібні або інверсні криволінійні залежності, наприклад, між рівнем стимуляції та продуктивністю (класична крива Йеркса–Додсона), що демонструє обмеженість виключно лінійного підходу.

Знак коефіцієнта кореляції відображає напрямок статистичної асоціації між двома змінними, тобто спосіб, у який вони змінюються одна відносно одної. У кореляційному аналізі дослідника цікавить не лише сила зв'язку, яку показує абсолютне значення коефіцієнта, але й те, чи підвищення однієї змінної супроводжується підвищенням іншої, чи, навпаки, їхні значення змінюються у протилежних напрямках. Саме знак дозволяє зрозуміти цю сторону зв'язку.

Позитивний знак (+) у кореляційному коефіцієнті свідчить про те, що зміни змінних відбуваються у погодженому напрямку: коли перша змінна збільшується, друга також має тенденцію зростати; коли одна зменшується – зменшується і інша. Такий тип зв'язку поширений у ситуаціях, пов'язаних із розвитком, навчанням, тренінговими впливами, мотиваційними процесами. Наприклад, у дослідженнях академічної успішності часто фіксують позитивний зв'язок між обсягом навчальної мотивації та рівнем успішності: чим вищою є мотивація, тим, як правило, вищі оцінки. Аналогічно у дослідженнях психічного здоров'я позитивний зв'язок між соціальною підтримкою та суб'єктивним благополуччям означає, що вищий рівень

підтримки пов'язаний із вищим рівнем благополуччя. Таким чином, позитивний знак не просто фіксує співнапрямленість змін, а часто вказує на потенційний механізм взаємопосилюючих психологічних процесів.

Негативний знак (–) у кореляційному коефіцієнті означає протилежний напрямок зв'язку: збільшення однієї змінної супроводжується зменшенням іншої. Такий зв'язок особливо характерний для феноменів, у яких ресурси або енергія розподіляються між різними процесами, або де одна характеристика стримує іншу. Наприклад, у дослідженнях емоційного вигорання зазвичай фіксують негативний зв'язок між рівнем вигорання і задоволеністю роботою: чим вищий рівень вигорання, тим нижча задоволеність. У когнітивній психології негативний зв'язок між рівнем тривожності та концентрацією уваги означає, що за підвищення тривожності здатність до концентрації погіршується. Отже, знак «–» не лише відображає статистичний напрямок змін, а й вказує на психологічну взаємопов'язаність феноменів, де підвищення одного чинника призводить до виснаження або зниження іншого.

Ключовим аспектом будь-якого кореляційного аналізу є оцінка статистичної значущості отриманого коефіцієнта. Значення p -value дозволяє визначити, наскільки ймовірно те, що виявлений зв'язок міг виникнути випадково, за умови відсутності реальної залежності у популяції. Традиційно в соціальних науках використовується рівень $\alpha = 0,05$, однак сліпе дотримання цього порога може призвести до помилкової інтерпретації, особливо за умов множинних порівнянь. Якщо в дослідженні розраховується десятки або сотні кореляцій, ймовірність хибнопозитивних висновків різко зростає, і у таких ситуаціях застосовуються поправки на множинні порівняння, зокрема поправка Бонферроні або менш консервативні методи, такі як поправка Бенджаміні–Хохберга. Це особливо важливо в психологічних дослідженнях особистості, де автори часто аналізують десятки шкал та підшкал, створюючи передумови для статистичного шуму.

Ще одним методологічним аспектом є проблема інтерпретації величини кореляції. Навіть дуже низькі значення можуть досягати статистичної значущості за великого обсягу вибірки, тоді як у малих вибірках сильні ефекти можуть залишатися несприйнятними статистичними тестами через низьку потужність. Тому в сучасній психології дедалі частіше наголошують на важливості оцінки розміру ефекту та довірчих інтервалів. Кредит довірчому інтервалу полягає в тому, що він демонструє, у яких межах може коливатися справжній коефіцієнт у генеральній сукупності. Це є більш інформативним підходом, оскільки дозволяє оцінити точність вимірювання та визначити ступінь невизначеності, що супроводжує будь-які кореляційні висновки.

Коли постає задача усереднення декількох коефіцієнтів, наприклад, у метааналізі або в аналізі внутрішньої структурної узгодженості поведінкових показників, застосовується перетворення Фішера. Це перетворення переводить коефіцієнт Пірсона у z -шкалу, яка має властивості нормального розподілу і дозволяє обчислювати середні значення без порушення статистичних припущень. Після усереднення показник повертають до

вихідної шкали r , що робить його інтерпретацію звичною для психологів. Таке перетворення необхідне через неадитивність коефіцієнта Пірсона: середнє арифметичне кількох r без перетворення може бути математично некоректним і вводити дослідника в оману.

Безперечно важливим напрямом застосування кореляційних коефіцієнтів є аналіз внутрішньокласової кореляції. Цей показник використовується в ситуаціях, коли дані мають групову структуру: учні в класах, пацієнти в лікарнях, працівники в підрозділах або оцінки, виставлені кількома експертами. Внутрішньокласова кореляція дає можливість визначити, наскільки значна частка варіації зумовлена груповою належністю або схожістю відповідей у межах кластера. Особливо важливими такі коефіцієнти є у дослідженнях інтеррейтної надійності: оцінка психологічного стану, соціальних навичок, емоційної регуляції або поведінки дитини вимагає гарантії того, що різні спостерігачі інтерпретують поведінку схожим чином.

У психології кореляційні коефіцієнти відіграють не лише описову роль, а й виконують функцію будівельних блоків для складніших моделей. Регресійний аналіз, факторний аналіз, моделювання структурних рівнянь, аналіз медіації та модерації – усі ці підходи ґрунтуються на концепції асоціації між змінними, яку відображає кореляція. Водночас важливо пам'ятати, що кореляції, особливо високі, можуть бути симптомом не причинного зв'язку, а впливу третьої змінної, спільного джерела варіації або методологічного артефакту. Розуміння природи цих обмежень є ключовим для того, щоб уникнути надмірних узагальнень або хибних інтерпретацій.

Таким чином, коефіцієнти кореляції становлять складну систему індикаторів, яка дозволяє досліднику отримати багатовимірну картину взаємозв'язків між психологічними явищами. Це не лише числа, що демонструють силу зв'язку, а інструменти, які вимагають точності вимірювання, методологічної обережності та глибокого розуміння природи досліджуваних змінних. Інтерпретація кореляцій повинна враховувати шкали вимірювання, обсяг вибірки, форму розподілу, наявність викидів і структурні властивості даних. Відповідальне застосування цих коефіцієнтів робить кореляційні дослідження надійною основою для побудови теорій, планування емпіричних перевірок і розвитку психологічної науки в цілому.

8.3. Основні джерела похибок у кореляційних дослідженнях

Попри свою універсальність і високу дослідницьку цінність, кореляційний підхід залишається значною мірою вразливим до різноманітних методологічних і статистичних похибок, які здатні суттєво спотворювати результати й інтерпретації дослідження. У сучасній психології, яка переживає період активної методологічної переоцінки, проблема некоректних або хибно інтерпретованих кореляцій набула особливої гостроти. Кореляційний зв'язок може виглядати статистично значущим, але фактично бути побудованим на систематичних помилках збору даних, неправильному виборі методів, низькій

надійності інструментів, неоднорідності вибірки або на впливі неконтрольованих факторів. Це створює широке поле для хибних висновків, у яких кореляції приймаються за каузальні зв'язки, а виявлені закономірності не відтворюються в інших вибірках або культурах.

Однією з найбільш поширених причин похибок у кореляційному аналізі є явище множинних порівнянь. У багатьох емпіричних роботах дослідники вимірюють десятки змінних одночасно – показники тривожності, самооцінки, стресостійкості, когнітивної гнучкості, академічної мотивації, емоційної регуляції тощо. За таких умов утворюється велика кореляційна матриця, яка містить сотні або навіть тисячі парних кореляцій. Навіть якщо всі змінні повністю незалежні одна від одної, статистичні критерії все одно дадуть кілька «значущих» кореляцій при $\alpha = .05$ виключно випадковим чином. У психології це особливо небезпечно, адже дослідник має спокусю інтерпретувати ці випадкові відхилення як наукові закономірності. У дослідженні університету Джорджії, де студенти заповнили 56 різних опитувальників емоційних, особистісних і поведінкових показників, отримано понад півтори тисячі кореляцій. Близько 80 із них виявилися статистично значущими на рівні $p < .05$, але після корекції множинних порівнянь за Бонферроні статистично значущими залишилися лише три. Це означає, що приблизно 95% «значущих» зв'язків були хибнопозитивними, тобто такими, що виникли виключно через випадковість. У реальних психологічних дослідженнях це означає, що багато опублікованих «висновків» про нібито важливі зв'язки між особистісними особливостями, поведінковими тенденціями чи когнітивними параметрами можуть бути наслідком не реальних закономірностей, а статистичних артефактів.

Ще одним важливим джерелом помилок є проблеми, пов'язані з розподілом даних та аутлайєрами (викидами). Психологічні змінні майже ніколи не розподілені рівномірно або нормально; багато з них мають асиметрію, скошення, платокуртичність або мультимодальність. Ці особливості спотворюють розрахунок кореляційних коефіцієнтів, що особливо критично для показників Пірсона, який вимагає приблизної нормальності обох змінних. Значний вплив може мати навіть одна-єдина аномальна точка. У дослідженні Ларсона та колег, що аналізувало взаємозв'язок між використанням соціальних мереж і суб'єктивним відчуттям щастя, було виявлено сильну негативну кореляцію $r = -0,42$. Однак подальший аналіз показав, що це спотворення виникло через одного респондента, який проводив у соціальних мережах понад 12 годин на добу й водночас мав надзвичайно низький показник позитивного афекту. Після видалення цього аутлайєра кореляція впала до $-0,17$ і перестала бути значущою. У контексті психологічних досліджень аутлайєри можуть виникати через різні фактори: неуважне заповнення анкети, випадкове внесення помилкових значень, неправдиві відповіді, специфічні індивідуальні стани або навіть технічні похибки електронних систем збору даних. Ігнорування таких точок

призводить до того, що одна-єдина особа здатна створити ілюзію закономірності, якої в реальності не існує.

Важливою передумовою коректності кореляційного аналізу є надійність вимірювальних інструментів. Кореляція між двома змінними не може бути сильною, якщо принаймні одна з них вимірюється з низькою точністю. Тест із низьким показником α Кромбаха або низькою стабільністю у повторному вимірюванні неминуче занижуватиме кореляції, роблячи їх слабшими, ніж вони є насправді. Показовим є дослідження Томпсона та Снайдера, які порівнювали коротку 4-пунктову шкалу позитивного афекту з повною 20-пунктовою версією. Коротка шкала мала $\alpha = 0,48$ і показувала кореляцію з самооцінкою $r = 0,19$. Повна версія мала $\alpha = 0,88$ і давала кореляцію $r = 0,48$. Така різниця означає, що у першому випадку слабкий зв'язок був зумовлений не особливостями вибірки чи психологічним змістом конструктів, а виключно низькою надійністю інструменту. Ненадійні шкали, швидкі онлайн-опитувальники часто породжують хибне враження про слабкість або відсутність зв'язків між важливими психологічними характеристиками.

Наступним джерелом похибок є неоднорідність вибірки. Кореляційні дослідження передбачають, що змінні мають регресувати одна щодо одної послідовно й у межах однієї сукупності. Якщо вибірка складається з різко відмінних груп – наприклад, учнів загальноосвітніх шкіл і підлітків із центрів ресоціалізації – то їх об'єднання може створити штучні кореляції, яких немає всередині кожної групи окремо. Це добре ілюструє дослідження самооцінки та агресивності: у загальній вибірці кореляція становила $r = 0,31$, тоді як у школярів – $r = 0,03$, а у підлітків з девіантною поведінкою – $r = 0,62$. Це означає, що загальний зв'язок був не реальною психологічною закономірністю, а статистичним ефектом змішування двох різних груп. Подібні артефакти можуть виникати у дослідженнях студентів різних спеціальностей, респондентів з різних регіонів, різних рівнів соціально-економічного статусу чи навіть у вибірках, що формувалися у різні періоди часу. Неоднорідність вибірки руйнує припущення про єдину структуру даних і робить кореляції непорівнюваними та теоретично неінтерпретованими.

Однією з найсерйозніших і найчастіше недооцінених проблем є вплив третьої змінної. Кореляція між двома змінними може бути наслідком впливу фактору, про який дослідник не знає або який не контролює. Відомим є приклад, де кореляція між фізичною активністю та самооцінкою становила $r = 0,34$, але після контролю індексу маси тіла зменшилася до $r = 0,05$. Це означає, що індекс маси тіла одночасно впливав на обидві змінні: люди з нижчим індексом маси тіла частіше займалися фізичними вправами й мали вищу самооцінку. Інший класичний приклад – кореляція між рівнем стресу та кількістю випитої кави серед офісних працівників. На перший погляд здається, що кава збільшує стрес або стрес стимулює споживання кофеїну. Однак після контролю робочого навантаження виявилось, що жодного прямого зв'язку між стресом і кавою немає: обидві змінні були лише індикаторами інтенсивності робочих вимог. У цьому сенсі приватна змінна може бути як психологічною

(мотивація, темперамент, стиль мислення), так і соціальною (доходи, освіта), когнітивною (інтелект), демографічною (вік, стать), а іноді навіть технічною (час заповнення анкети).

Окрема група похибок кореляційних досліджень пов'язана з тим, що кореляція дуже чутлива до лінійності. Багато психологічних залежностей мають нелінійний характер, але традиційний коефіцієнт Пірсона не виявляє таких зв'язків, або ж інтерпретує їх як слабкі. Наприклад, зв'язок між рівнем тривожності та ефективністю виконання завдання часто має вигляд інверсної U-подібної кривої: дуже низький і дуже високий рівень тривожності погіршує виконання, а помірний рівень оптимізує його. Кореляція Пірсона тут може бути майже нульовою, тоді як реальний механізм дуже сильний. Аналогічно, у дослідженнях творчості та інтелекту часто спостерігається пороговий ефект: після певної точки приріст інтелекту вже не визначає рівень творчої продуктивності, і кореляція знижується.

Загалом кореляційні дослідження дають цінні інсайти лише тоді, коли враховано всі перелічені джерела похибок. Необхідно уважно перевіряти розподіл даних, очищувати їх від аутлайєрів, застосовувати корекцію множинних порівнянь, контролювати третю змінну, оцінювати надійність інструментів і уникати об'єднання структурно різних груп респондентів. Саме такий комплексний підхід дозволяє підвищити достовірність і теоретичну значущість кореляційних досліджень, зробити висновки справді науковими й такими, що здатні відтворитися в різних вибірках і контекстах. Інакше кореляційний аналіз перетворюється на процедуру, що генерує ілюзії закономірностей, які спокушають дослідника, але не мають емпіричної реальності.

8.4. Типологія та конструктивні варіації планів кореляційних досліджень

Упродовж останніх десятиліть кореляційні моделі суттєво ускладнилися: від простих одночасних оцінок двох змінних до масштабних системних структур, що враховують багатовимірність поведінки, багатоетапність розвитку психічних процесів та ієрархічність соціальних контекстів. Така еволюція зумовлює потребу у чіткій типології кореляційних планів, яка дозволяє досліднику усвідомлено обирати адекватний дизайн залежно від теоретичної моделі, структури вибірки та логіки гіпотез. Найбільш поширеними та методологічно обґрунтованими серед них є зрізові дизайни, лонгітюдні плани та змішані кореляційні стратегії, які разом охоплюють більшість практичних запитів психологічного дослідження.

Зрізові кореляційні дизайни становлять найпростішу й водночас найпоширенішу форму кореляційних досліджень, оскільки передбачають одночасне вимірювання змінних у певній вибірці і дають змогу швидко отримати уявлення про зв'язки між ознаками. У межах такого підходу особливе значення мають дослідження, де всі змінні виміряні в однієї групи

респондентів у один момент часу. Це дозволяє моделювати структурні зв'язки між кількома психологічними конструктами, знаходити як прямі лінійні кореляції, так і складніші асоціативні мережі, включно з потенційними медіаційними й модераційними ефектами. Наприклад, у дослідженні студентської молоді часто одночасно вимірюють самооцінку, рівень соціальної підтримки та показники психологічного благополуччя. Таке дослідження дає змогу встановити, що між цими трьома змінними існують суттєві зв'язки, і в рамках одного зрізу можна побудувати кілька альтернативних моделей: одна передбачає, що соціальна підтримка визначає самооцінку, яка, своєю чергою, впливає на благополуччя; інша – що саме благополуччя є передумовою позитивного сприйняття соціального оточення. Зрізовий дизайн у цьому випадку не дозволяє вирішити питання причинності, але створює основу для подальших аналітичних кроків і розвитку теорії.

Поширеним різновидом зрізових планів є дослідження, у яких беруть участь декілька груп. На відміну від попереднього, такий підхід дозволяє не лише оцінювати загальні кореляції між змінними, але й порівнювати силу та структуру зв'язків у різних підвибірках. Це відкриває можливість аналізувати, чи зберігається асоціація між змінними в групах, що відрізняються за статтю, віком, професійним статусом, рівнем тривожності чи іншими соціально-психологічними характеристиками. Наприклад, можна виявити, що кореляція між емоційним вигоранням і рівнем задоволеності роботою є значно вищою в групі початківців-рятувальників, ніж у досвідчених працівників служби реагування на надзвичайні ситуації. Або ж що зв'язок між академічною мотивацією та тривожністю присутній у жінок, але значно слабший у чоловіків. Такі результати дають змогу дослідникові формувати гіпотези про групові модерації та уточнювати психологічні моделі, пов'язані з віковими, соціальними або професійними відмінностями.

Складнішими, але надзвичайно продуктивними є багаторівневі кореляційні дослідження, у яких дані мають ієрархічну структуру: учні вкладені в класи, класи – у школи, працівники – у підрозділи, підрозділи – у організації. У таких випадках змінні вимірюються на кількох рівнях одночасно – наприклад, рівень психологічного благополуччя на рівні індивідів і рівень підтримки менеджменту на рівні організації. Ігнорування цієї багаторівневості призводить до помилок, оскільки кореляції можуть суттєво відрізнятися на різних рівнях. Наприклад, у дослідженнях освітніх процесів часто виявляється, що кореляція між академічною самооцінкою та успішністю є значною на рівні учнів, але майже відсутньою на рівні класів або шкіл. Це означає, що індивідуальні відмінності учнів істотно перевищують відмінності між навчальними закладами. Водночас у дослідженні робочого клімату може виявитися протилежна ситуація, коли саме групові характеристики (наприклад, стиль керівництва) значною мірою визначають психологічні стани співробітників. Багаторівневий кореляційний підхід дозволяє досліднику виявляти такі закономірності, створювати точніші моделі та

уникати хибних висновків, що виникають при нехтуванні груповими ефектами.

На відміну від зрізових планів, лонгітюдні дослідження мають часовий вимір і дають можливість відстежити тривалі процеси, внутрішні зміни та індивідуальні траєкторії розвитку психологічних конструктів. Такі плани є незамінними у випадках, коли метою є виявлення стійкості характеристик, динаміки емоційних станів або наслідків життєвих подій. Вони дозволяють контролювати попередні значення змінної, оцінювати вплив минулих станів на майбутні, а також застосовувати крос-лагові моделі, що відкривають можливість аналізувати потенційну причинну послідовність без експериментального втручання. Наприклад, у дослідженні студентів першого курсу було проведено три хвили вимірювання академічної тривожності та самоконтролю протягом року. У результаті виявилось, що високий рівень початкової тривожності передбачає зниження самоконтролю через пів року, тоді як зворотний зв'язок не був статистично значущим. У рамках одного зрізу цей висновок був би неможливим, тоді як лонгітюдний підхід дозволив виявити потенційну каузальну асиметрію.

Лонгітюдні плани також дозволяють аналізувати індивідуальні траєкторії змін, виявляти латентні класи осіб із подібною динамікою, а також оцінювати, чи є зміни стабільними або флуктуючими. Наприклад, у дослідженні працівників однієї організації, що тривало півтора року, було показано, що для більшості респондентів рівень емоційного вигорання зростає поступово, але в окремої частини спостерігався різкий стрибок після серії стресових подій. Без лонгітюдного підходу ці патерни залишилися б непоміченими, оскільки вони не проявляються у середніх значеннях у межах одного зрізу.

Змішані кореляційні плани становлять методологічний перехід між зрізовими й лонгітюдними моделями, поєднуючи їхні сильні сторони. Вони передбачають одночасне використання декількох когортно-зрізових груп, вимірювання тих самих змінних у різні моменти часу та порівняння різних вікових або професійних груп. Така стратегія дозволяє розділити ефекти віку та часу, а також аналізувати як стабільність, так і змінність психологічних характеристик у різних популяціях. Наприклад, дослідження розвитку когнітивної гнучкості може включати три вікові групи – дітей, підлітків та дорослих – і передбачати повторне тестування протягом року. Це дозволяє виявити не лише загальну вікову динаміку, а й різні траєкторії змін у кожній групі, включаючи можливі періоди прискороного або уповільненого розвитку. Змішаний підхід також дозволяє компенсувати недоліки тривалих лонгітюдних досліджень, зменшуючи вплив вибуття учасників і прискорюючи отримання результатів.

У психології змішані плани широко застосовуються у дослідженнях професійного становлення, де одночасно досліджують декілька когортно-вікових груп працівників і протягом короткого періоду спостерігають зміни у професійній мотивації, рівні стресу або задоволеності роботою. Такий підхід

дозволяє встановити, чи пов'язані зміни з віком, стажем чи зовнішніми факторами, наприклад, реформами в організації або змінами в умовах праці. Водночас у дослідженнях соціальної поведінки змішані дизайни використовують для аналізу змін у соціальній довірі або політичній участі серед різних поколінь протягом певного періоду. Завдяки цьому дослідник може відстежувати не лише вікові відмінності, а й макросоціальні ефекти, що по-різному впливають на різні вікові групи.

Таким чином, кожен тип кореляційного дизайну виконує специфічну методологічну функцію: зрізові плани дозволяють швидко отримувати інформацію про структуру зв'язків, багатогрупові та багаторівневі зрізи розкривають складні асоціативні ієрархії, а лонгітюдні схеми дають змогу аналізувати часову динаміку та передбачення, тоді як змішані стратегії забезпечують найповніше розуміння розвитку психологічних характеристик у реальному контексті. Кожен із цих планів має власні аналітичні можливості, обмеження та вимоги до інтерпретації, але разом вони формують фундаментальний методологічний інструментарій сучасної психології, що дозволяє дослідникам отримувати глибоке й багатовимірне бачення складних психічних процесів.

8.5. Психогенетичні та крос-культурні дослідження як специфічні види кореляційних досліджень

Психогенетичні та крос-культурні дослідження становлять особливі різновиди кореляційних планів, оскільки вони зосереджені на пошуку закономірностей природних варіацій поведінки, психічних процесів та індивідуальних відмінностей у середовищах, які неможливо відтворити експериментально. Обидва типи досліджень характеризуються складною структурою вибірок, багаторівневістю даних, необхідністю контролю значної кількості зовнішніх чинників та особливою увагою до інтерпретації отриманих асоціацій. Вони мають спільну методологічну рису: виявлені кореляції не дозволяють говорити про причинність, але дають змогу інтерпретувати внесок різних детермінант спадковості, середовища або культурних факторів у формування психологічних характеристик людини.

Психогенетичні дослідження спрямовані на виявлення того, якою мірою індивідуальні відмінності обумовлені генетичними чинниками, а якою – впливами навколишнього середовища. У традиційних експериментальних дослідженнях неможливо контролювати або маніпулювати генетичними особливостями людини, тому психогенетика використовує природну варіативність у ступені спорідненості між людьми. Найбільш відомим підходом є близнюковий метод, у якому порівнюються рівні кореляцій між монозиготними (MZ) і дизиготними (DZ) парами близнюків. Монозиготні близнюки мають практично ідентичний генотип, тоді як дизиготні близнюки, як і звичайні брати й сестри, поділяють у середньому лише половину генетичних варіацій. Якщо певна психологічна ознака демонструє значно

вищу кореляцію в монозиготних парах порівняно з дизиготними, це свідчить про значний спадковий компонент.

Внесок генетичних чинників часто оцінюється за допомогою коефіцієнта спадковості. Наприклад, якщо кореляція рівня тривожності між монозиготними близнюками становить $r = 0,60$, а для дизиготних – $r = 0,30$, спадковість цього показника може сягати 60%. Це означає, що понад половину індивідуальних відмінностей у тривожності можна пояснити генетичними факторами. Проте важливо розуміти, що коефіцієнт спадковості не стосується конкретної людини, а описує варіативність у популяції, і що високі показники спадковості не виключають суттєвого впливу середовища.

Психогенетичні дослідження також враховують роль середовища – спільного та індивідуального. Спільне середовище – це чинники, які однаково впливають на всіх дітей у сім'ї: соціально-економічний статус, стиль виховання, освітні можливості. Індивідуальне середовище включає унікальні досвіди: різні вчителі, групи однолітків, травматичні події, індивідуальні життєві траєкторії. Часто дослідники поєднують близнюкові, сиблінгові та інші споріднені вибірки, що дозволяє побудувати складні кореляційні моделі, які розділяють вплив генетичних та середовищних детермінант. Наприклад, метааналізи близнюкових досліджень показали, що успішність у школі визначається приблизно порівну спадковістю та індивідуальним середовищем, тоді як риси особистості демонструють високий спадковий компонент та відносно слабкий вплив спільного середовища. У психології розвитку психогенетичні дослідження дозволили з'ясувати, що інтелект має тенденцію до «розкриття» генетичного потенціалу з віком: у дитинстві спадковість нижча, ніж у дорослому віці, що свідчить про складну динамічну взаємодію середовища та генів.

Крім близнюкових планів, психогенетика застосовує прийомні дослідження, в яких порівнюють дітей, вихованих генетично нерідними батьками, та дітей, що залишилися у біологічних сім'ях. Наприклад, якщо рівень агресивності у прийомних дітей корелює з агресивністю біологічних, а не прийомних батьків, це свідчить про вплив генетичних чинників. Однак такі висновки потребують обережності, оскільки навіть у прийомних сім'ях можуть виникати систематичні відмінності, пов'язані з відібраністю середовищ. У сучасних дослідженнях активно використовуються діадичні моделі, у яких як одиниця аналізу розглядається пара – наприклад, мама-дитина, двоє партнерів або двоє братів. Такі моделі дозволяють досліджувати, яка частина зв'язків між психологічними змінними зумовлена спільним середовищем, а яка є специфічною для кожного члена діади.

Як і будь-який кореляційний дизайн, психогенетичні дослідження стикаються з обмеженнями, що ускладнюють інтерпретацію результатів. Одним із ключових викликів є припущення про еквівалентність середовищ, у яких зростають монозиготні та дизиготні близнюки. Якщо монозиготні близнюки зазнають більш подібних умов виховання (що часто трапляється через їхню зовнішню подібність), то різницю між кореляціями MZ і DZ може

частково пояснити не генетика, а соціальний контекст. Крім того, моделі спадковості не враховують можливих епігенетичних ефектів або генетично зумовлених реакцій на середовище, які можуть істотно модифікувати фенотип протягом життя.

Крос-культурні дослідження становлять інший великий напрям специфічних кореляційних методів. Їхня головна мета – вивчити, як психологічні конструкти проявляються у представників різних культур, і чи зберігаються зв'язки між змінними у різних культурних контекстах. Такі дослідження необхідні не лише для порівняння середніх рівнів певних характеристик, а й для аналізу структурних відмінностей у взаємозв'язках між змінними, що дозволяє виявити, чи є певні закономірності універсальними або культурно специфічними. У крос-культурних дизайнах кореляційні плани використовуються у кількох типових аналітичних завданнях: порівняння середніх значень і кореляцій між змінними у різних культурах; вивчення зв'язків між макрорівневими культурними індикаторами (наприклад, індексом індивідуалізму) та психологічними показниками; аналіз багаторівневих зв'язків між індивідуальними та культурними змінними.

Одним із ключових викликів у крос-культурних дослідженнях є забезпечення еквівалентності вимірювальних інструментів. Для того щоб порівняння між культурами було коректним, необхідно бути впевненим, що конструкт вимірюється однаково, має аналогічну факторну структуру та не містить перекручень, спричинених мовними або культурними ідіосинкразіями. Тому широко застосовується процедура прямого та зворотного перекладу психологічних методик, однак навіть досконалий переклад не гарантує повної еквівалентності змісту. Для оцінки структурної еквівалентності використовуються методи конфірматорного факторного аналізу, які дозволяють перевірити рівні інваріантності: конфігураційний (подібність структури конструкту), метричний (рівність факторних навантажень), скалярний (рівність пунктових середніх) та строгий (рівність залишкових варіацій). Якщо інваріантність не підтверджена, порівняння середніх значень у вибірках є методологічно некоректним.

Крос-культурні дослідження також потребують ретельного контролю методологічних джерел варіабельності: відмінностей у соціально бажаних стилях відповіді, різної мотивації участі у дослідженні, різного рівня абстрактності конструктів у культурних групах, а також ефектів перекладу. Важливою проблемою є феномен диференціального функціонування пунктів (DIF), коли формально однаковий пункт опитувальника має різні психометричні властивості у різних культурах. Це може призводити до спотворення кореляційних структур, особливо у дослідженнях міжкультурних відмінностей у рисах особистості або ціннісних орієнтаціях.

У крос-культурних дослідженнях застосовуються два основні стратегічні підходи – емічний і етичний. Емічний підхід виходить із припущення, що кожна культура має власні, унікальні конструкти та специфічні прояви поведінки, тому методики повинні бути адаптовані або

створені спеціально для кожної культури. Такий підхід забезпечує глибину розуміння культурних феноменів, але унеможлиблює точні порівняння між культурами. Етичний підхід, навпаки, передбачає використання універсальних інструментів, валідизованих для різних культур, що дозволяє здійснювати кількісне порівняння, але вимагає особливо ретельної перевірки еквівалентності вимірювань.

Приклади крос-культурних досліджень охоплюють широкий спектр психологічних тем: від вивчення універсальності структури особистісних рис (модель Big Five) до дослідження культурних відмінностей у стилях емоційної регуляції або схильності до соціальної тривожності. Наприклад, дослідження показали, що позитивні емоції пов'язані з самооцінкою в індивідуалістичних культурах значно сильніше, ніж у колективістичних, де більшу роль відіграють міжособистісні зв'язки. Інше крос-культурне дослідження виявило, що рівень життєвої задоволеності може мати різні предиктори у різних країнах: у країнах Північної Європи він частіше пов'язаний з автономією, тоді як у країнах Азії – з гармонією соціальних відносин.

Обидва типи специфічних кореляційних досліджень – психогенетичні та крос-культурні – висувають високі вимоги до методологічного контролю та інтерпретації. У психогенетиці це контроль еквівалентності середовища, вибір коректних моделей спадковості, урахування епігенетичних факторів та індивідуальних середовищних впливів. У крос-культурних дослідженнях – це забезпечення інваріантності вимірювань, дотримання правильних процедур перекладу, контроль за соціально бажаними відповідями та врахування культурних стилів мислення.

Важливо підкреслити, що, хоча ці дизайни не дозволяють робити висновків про причинність у класичному розумінні, вони є незамінними в психологічній науці. Психогенетика дає змогу оцінити роль природи та виховання у формуванні психологічних характеристик, тоді як крос-культурні дослідження дозволяють визначити універсальні та культурно специфічні закономірності психічного життя. Поєднання цих підходів формує комплексне розуміння людини як біологічної, так і культурної істоти, чия поведінка одночасно зумовлена генетичною структурою та культурним контекстом. Саме тому психогенетичні й крос-культурні кореляційні дослідження відіграють ключову роль у сучасній психологічній науці, забезпечуючи глибину і широту аналізу, недосяжні в межах лабораторної експериментальної парадигми.

Завдання на самостійну підготовку

1. У чому полягає сутнісна відмінність кореляційного підходу від експериментального в психологічних дослідженнях?
2. Чому помилки вимірювання призводять до заниження величини коефіцієнта кореляції та як це варто враховувати при інтерпретації результатів?
3. У чому полягає проблема «третьої змінної» в кореляційних дослідженнях і якими способами дослідник може намагатися її контролювати?

4. Чому коефіцієнт кореляції Пірсона вважається адекватним показником лише для лінійних зв'язків і які наслідки це має для аналізу психологічних даних?

5. Наведіть приклад психологічної залежності, яка може мати нелінійний характер, і поясніть, чому традиційний кореляційний аналіз може її «не бачити».

6. У чому полягає небезпека множинних порівнянь у кореляційних матрицях і як корекція рівня значущості допомагає зменшити цю проблему?

7. Поясніть різницю між статистичною значущістю кореляції та її практичною (психологічною) значущістю.

8. Чому довірчі інтервали для коефіцієнта кореляції вважаються більш інформативними, ніж лише значення p -value?

9. У яких ситуаціях доцільно застосовувати перетворення Фішера для коефіцієнтів кореляції і чому просте арифметичне усереднення r є некоректним?

10. Як низька надійність психодіагностичних інструментів впливає на інтерпретацію сили кореляційних зв'язків між психологічними змінними?

11. Чому структурна неоднорідність вибірки (наявність різко відмінних груп) може створювати штучні або спуріозні кореляції?

12. У чому полягає сутність зрізового кореляційного дизайну і які його ключові переваги та обмеження для психологічних досліджень?

13. Які додаткові можливості дають багатогрупові кореляційні плани порівняно з простим зрізом однієї вибірки?

14. Чим багаторівневі (ієрархічні) кореляційні дослідження відрізняються від традиційних і які питання вони дозволяють розв'язувати в освітній чи організаційній психології?

15. Які методологічні переваги лонгітюдних кореляційних досліджень у виявленні динаміки психологічних характеристик?

16. У чому полягає специфіка змішаних (когортно-зрізових) кореляційних планів і які задачі розвитку особистості чи професійного становлення вони дозволяють досліджувати?

17. Які ключові методологічні ідеї лежать в основі психогенетичних кореляційних досліджень (зокрема близнюкових) і як за їхньою допомогою оцінюється спадковість психологічних рис?

18. Чим відрізняються впливи спільного та індивідуального середовища в психогенетичних моделях і як це відображається в кореляційних показниках?

19. Які основні методологічні виклики пов'язані з інтерпретацією коефіцієнта спадковості в психогенетичних дослідженнях?

20. Чому забезпечення інваріантності вимірювань (конфігураційної, метричної, скалярної) є критично важливим для крос-культурних кореляційних досліджень?

РОЗДІЛ 9

БАЗОВІ МЕТОДИ СТАТИСТИЧНОГО АНАЛІЗУ РЕЗУЛЬТАТІВ ПСИХОЛОГІЧНИХ ДОСЛІДЖЕНЬ

9.1. Особливості підготовки результатів дослідження до статистичного аналізу

Підготовка емпіричних даних до статистичного аналізу в психологічному дослідженні є не другорядним технічним кроком, а центральною ланкою всього наукового процесу, від якої безпосередньо залежать валідність висновків та рівень довіри до проведеної роботи. Сам по собі масив отриманих відповідей, балів чи показників є лише «сирим матеріалом», у якому неминуче присутні помилки введення, випадкові артефакти, пропуски, а також особливості розподілу змінних, що можуть або відповідати припущенням обраних статистичних методів, або грубо їх порушувати. Тому перш ніж переходити до перевірки гіпотез, порівняння груп чи встановлення зв'язків, дослідник має виконати цілу серію процедур: розрахувати описові статистики, перевірити дані на наявність помилок, визначити тип вимірювальних шкал, оцінити властивості розподілів, ідентифікувати та проаналізувати пропущені значення, оцінити нормальність розподілу, знайти аутлаєри, а за потреби – трансформувати дані або обрати непараметричні (робастні) методи аналізу. Така послідовна робота із «чисткою даних» не лише підвищує якість конкретного дослідження, а й відповідає сучасним стандартам наукової доброчесності, що передбачають прозорість усіх етапів роботи з даними.

Першим кроком підготовки даних є розрахунок описових статистик, які забезпечують базове, але критично важливе уявлення про структуру вибірки. Міри центральної тенденції – середнє арифметичне, медіана та мода – дозволяють зрозуміти, де «зосереджені» значення змінної. Середнє арифметичне фіксує середній рівень показника, медіана ділить вибірку навпіл, показуючи значення, нижче і вище якого знаходиться по 50 % спостережень, а мода вказує на найчастіше значення. У випадку нормального розподілу ці три показники збігаються, і саме ця властивість часто використовується як один з індикаторів близькості розподілу до нормального. Однак обмежитися лише характеристиками центральної тенденції було б фундаментальною помилкою. Для адекватної підготовки даних необхідно оцінити також мінливість: дисперсію, стандартне відхилення, міжквартильний розмах, розмах значень. Дисперсія як середній квадрат відхилень окремих значень від середнього показника дає уявлення про загальний розкид даних, а стандартне відхилення, будучи квадратним коренем з дисперсії, повертає цю оцінку в ті ж одиниці вимірювання, що й сам показник, роблячи її більш інтерпретованою. Міжквартильний розмах дозволяє оцінити розкид «середніх» 50 % значень, а розмах (різниця між мінімумом і максимумом) часто є першим сигналом про

потенційні помилки введення або наявність аномальних спостережень, які потребують додаткової уваги.

На цьому етапі таблиця описових статистик – із зазначеними мінімальними й максимальними значеннями, середнім, медіаною, стандартним відхиленням, можливо, асиметрією й ексцесом – стає своєрідним «діагностичним екраном» для дослідника. Уважний аналіз такої таблиці дозволяє виявити неможливі або вкрай мало ймовірні значення, наприклад бали, що виходять за межі шкали, перевищуючи максимально передбачене завданнями значення, або нульові значення там, де за інструкцією мінімальний бал не може бути меншим за одиницю. Часто подібні аномалії виникають через «злипання» цифр під час введення (наприклад, замість 12 – 112), неправильне кодування пропущених відповідей або технічні помилки при імпорті даних із зовнішніх джерел. Пошук і виправлення таких артефактів – одна з базових, але необхідних дій у процесі чистки даних.

Після первинної перевірки на наявність очевидних помилок постає питання про природу самих змінних і типи вимірювальних шкал, у яких вони представлені. Правильне визначення, чи є змінна номінальною, порядковою, інтервальною чи відношення, є не лише теоретичним завданням, а й ключовою умовою коректного вибору статистичних процедур. Номінальні змінні, що відображають лише належність до категорії без порядку (наприклад, стать, тип професії, форма навчання), вимагають опису у термінах частот і відсотків, а також використання відповідних методів аналізу (χ^2 -критерії, коефіцієнти асоціації). Порядкові змінні, як-от рівень задоволеності або ступінь тривожності, дозволяють впорядкувати респондентів, але не гарантують рівності інтервалів між сусідніми градаціями. Інтервальні змінні, зокрема шкали типу IQ або деякі стандартизовані психологічні тестові бали, вже допускають операції із різницями між значеннями, але мають умовний нуль. Шкала відношень, що включає абсолютний нуль (наприклад, час реакції, кількість правильних відповідей), відкриває можливість інтерпретацій типу «у скільки разів більше/менше». На етапі підготовки даних дослідник має не лише класифікувати змінні за типом шкали, а й усвідомити, які припущення він готовий прийняти, наприклад, трактуючи підсумковий бал за шкалою Лікерта як інтервальний показник. Від цього безпосередньо залежать як вибір статистичного критерію, так і можливість коректної інтерпретації результатів.

Наступним важливим кроком є аналіз пропущених значень. Пропуски в психолого-педагогічних дослідженнях є радше нормою, ніж винятком: респонденти можуть не відповідати на окремі запитання, пропускати блоки завдань, достроково переривати участь або механічно пропускати окремі позиції. Проте принципово важливо не лише зафіксувати факт наявності пропусків, а й зрозуміти їхню структуру. Пропуски можуть бути випадковими, частково залежними від значень інших змінних або систематичними. Якщо вони розподілені рівномірно і не пов'язані зі

специфічними характеристиками респондентів, їх вплив на результати може бути відносно невеликим. Якщо ж пропуски концентруються в певних групах або пов'язані з високим чи низьким рівнем досліджуваної властивості, ігнорування цього факту призводить до систематичних зміщень. На етапі чистки даних необхідно оцінити частку пропусків, порівняти респондентів із повними та неповними даними, а далі свідомо обрати стратегію: видалення випадків з великою кількістю пропусків, виключення змінних з надмірною кількістю відсутніх значень, застосування методів імпутації. Важливо підкреслити, що автоматичне заповнення пропусків середнім значенням без аналізу їх природи може значно зменшити дисперсію, зробити розподіл більш «штучно нормальним» і тим самим спотворити результати, зокрема кореляційний аналіз та оцінку ефектів.

Після того як базову структуру даних упорядковано, а основні помилки виявлено й виправлено, постає питання про властивості розподілу змінних, зокрема про їх відповідність нормальному розподілу. Це важливий етап, оскільки багато класичних статистичних методів – t -критерій Стюдента, однофакторний ANOVA, лінійна регресія, кореляція Пірсона – спираються на припущення про нормальність розподілу або принаймні його близькість до нормального. Оцінка нормальності не обмежується використанням одного формального критерію; натомість сучасний підхід передбачає поєднання графічних і кількісних методів. Візуальна оцінка за допомогою гістограм, на яку накладено криву теоретичного нормального розподілу, дозволяє побачити загальний характер розподілу: симетричний він чи скошений, має один чи декілька піків, є «плоским» чи надто «загостреним». Графіки Q–Q (квантиль-квантиль) і P–P (ймовірність-ймовірність) дають змогу оцінити, наскільки емпіричні квантілі збігаються з теоретичними нормальними; відхилення точок від діагоналі сигналізує про відхилення від нормальності. Коробкові діаграми (box plot) дозволяють побачити асиметрію розподілу та можливі аутлаєри.

Кількісна оцінка нормальності часто ґрунтується на показниках асиметрії й ексцесу. Якщо їхні значення невеликі (орієнтовно менше одиниці за модулем) і не перевищують своїх стандартних помилок більш ніж у два рази, розподіл можна вважати прийнятно близьким до нормального, принаймні для практичних цілей. Додатково використовуються статистичні критерії, такі як тест Колмогорова-Смирнова, Шапіро-Вілکا та інші. Проте їх результати потрібно інтерпретувати обережно з огляду на обсяг вибірки: на великих вибірках навіть мінімальні, практично незначущі відхилення від нормальності будуть статистично значущими, тоді як на дуже малих вибірках тест може не виявити суттєвих відхилень через низьку статистичну потужність. Отже, оцінка нормальності є інтегральним процесом, що поєднує візуальні, описові та критерійні показники.

Якщо виявлено, що розподіл суттєво відхиляється від нормального, наступним кроком є аналіз причин такого відхилення. Насамперед варто поставити питання: чи повинен цей показник взагалі бути нормальним. Для

певних психологічних змінних, наприклад рівня суб'єктивного відчуття щастя, задоволеності життям або виражених патологічних симптомів у нормальній популяції, асиметричні розподіли є очікуваними: більшість людей демонструє середні або високі показники благополуччя, тоді як дуже низькі значення зустрічаються рідше; або, навпаки, більшість учасників має низькі бали за шкалами клінічних симптомів, тоді як високі значення – виняток. У таких випадках вимагати нормальності від даних означає накладати на них неприродну модель. Тоді дослідник або приймає використання непараметричних методів, або обирає статистичні моделі, що враховують специфіку розподілу.

Якщо ж розподіл мав би бути близьким до нормального теоретично (наприклад, за підсумковими балами стандартного тесту в однорідній популяції), але виявляється суттєво скошеним або платикуртичним, доцільно зробити ще один крок назад і перевірити наявність аутлаєрів (викидів). Аутлаєри – це спостереження, які суттєво відрізняються від переважної більшості даних. У найпростішому випадку, коли йдеться про одну змінну (univariate outliers) можна виявити за допомогою z -оцінок: значення, що перевищують 3 стандартні відхилення за модулем, зустрічаються надзвичайно рідко, приблизно один раз зі тисячі спостережень, а $z > 4$ або 5 вказує на дуже підозріле спостереження. Коробкові діаграми також дозволяють побачити точки, які виходять далеко за межі «вусів» – тобто за межі діапазону, що охоплює більшість спостережень. Виявлення аутлаєрів може суттєво змінити картину розподілу після їхнього аналізу й, за необхідності, обґрунтованого вилучення.

Проте обмежуватися лише однофакторними аутлаєрами недостатньо, особливо коли дослідження включає кілька взаємопов'язаних змінних. Можлива ситуація, коли за кожною окремою змінною значення респондента не здаються екстремальними, але у багатовимірному просторі їх комбінація виявляється дуже нетиповою. У такому випадку йдеться про мультіваріативні аутлаєри (multivariate outliers), для виявлення яких застосовуються спеціальні показники, зокрема відстань Махаланобіса або інші індекси (наприклад, Cook's distance у регресійному аналізі). Вони оцінюють відстань кожної точки від центроїда вибірки з урахуванням коваріацій між змінними. Якщо значення відстані Махаланобіса для певного випадку є статистично значущим за розподілом χ^2 з числом ступенів свободи, що дорівнює кількості змінних, цей випадок розглядають як потенційний аутлаєр. Проте, як і у випадку з аутлаєрами за однією змінною, механічне видалення таких спостережень без змістовного аналізу є неприпустимим. Необхідно мати розумне пояснення, чому той чи інший респондент або спостереження «не такі», і фіксувати всі рішення в протоколі обробки даних.

Якщо суттєві відхилення від нормальності пов'язані не лише з аутлаєрами (викидами), а й із загальною формою розподілу, можна вдаватися до трансформування даних. Класичними є логарифмічна, квадратнокоренева, обернена та інші трансформації, спрямовані на

зменшення правосторонньої асиметрії, стабілізацію дисперсії або наближення розподілу до нормального. У більш узагальнених підходах використовуються, наприклад, трансформації Вох-Сох або різні варіанти оптимального шкалювання, де відстані між значеннями змінної нелінійно змінюються так, щоб наблизити розподіл до бажаної форми. Однак слід усвідомлювати, що будь-яка трансформація змінює інтерпретацію показника: якщо вихідно змінна мала інтервальний зміст, після складних нелінійних перетворень її інтерпретація може наближатися до порядкової; частково втрачається інтуїтивна зрозумілість числових значень. Тому трансформації слід застосовувати обережно, чітко зазначаючи, які саме перетворення виконано, як вони впливають на інтерпретацію результатів і чому без них аналіз був би менш коректним.

Важливо також пам'ятати, що трансформування даних – не єдиний і не завжди найкращий шлях у випадку відхилення від нормальності. Сучасна статистика пропонує широкий спектр непараметричних та робастних методів, які не вимагають суворої нормальності розподілу й часто є більш доречними, особливо у випадках малих вибірок або виражених асиметрій. Непараметричні критерії, такі як U-критерій Манна-Вітні, критерії Вілкоксона, Краскела-Уолліса, Фрідмана та інші, оперують не самими значеннями, а їх ранговими позиціями, завдяки чому стають менш чутливими до викидів і форми розподілу. Робастні статистики, що використовують, наприклад, усічені середні, медіанний аналіз або спеціальні оцінки дисперсії, також забезпечують більшу стійкість до порушення припущень. Окрім того, методи ресемплінгу, такі як бутстрепінг або джекнайф, дозволяють отримувати оцінки стандартних помилок і довірчих інтервалів для параметрів, розподіл яких складно визначити аналітично або які оцінюються за невеликими вибірками. Бутстрепінг, що ґрунтується на багаторазовому випадковому виборі підвибірок із початкової вибірки з поверненням, надає емпіричний розподіл оцінок, на основі якого можна робити висновки про їх варіативність і точність. Використання таких дозволяє отримувати більш надійні й стійкі результати у складних, «далеких від ідеалу» масивах психологічних даних.

Зрештою, процес підготовки результатів дослідження до статистичного аналізу можна розглядати як послідовність взаємопов'язаних кроків, кожен з яких формує основу для наступного етапу. Обчислення описових статистик і первинна візуалізація даних дозволяють побачити загальну картину й виявити очевидні помилки. Аналіз типів шкал гарантує, що дослідник коректно обиратиме методи аналізу та інтерпретуватиме результати відповідно до природи змінних. Перевірка на пропуски й артефакти, дослідження розподілів, оцінка нормальності, пошук аутлаєрів (викидів), прийняття рішень щодо трансформації чи використання альтернативних статистичних підходів – усе це не просто технічні процедури, а прояв зрілої статистичної культури.

9.2. Алгоритм вибору статистичного критерію

Алгоритм вибору статистичного критерію є одним із ключових етапів у підготовці та проведенні психологічного дослідження, оскільки правильність статистичного висновку безпосередньо залежить від того, наскільки коректно дослідник співвідносить мету аналізу, властивості емпіричних даних та логіку відповідного математичного апарату. На відміну від механічного застосування відомих критеріїв, вибір статистичного методу вимагає глибокого розуміння природи вибірок, типів змінних, вимірювальних шкал, специфіки розподілу та характеру дослідницького завдання.

Передусім необхідно визначити, чи є вибірки, що порівнюються або аналізуються, залежними чи незалежними. Незалежні вибірки характеризуються тим, що спостереження в одній групі не впливають на значення в іншій; це, наприклад, порівняння студентів різних спеціальностей або працівників різних підрозділів. Залежні вибірки, навпаки, передбачають, що дані взято від тих самих респондентів у різні моменти часу або в різних умовах – наприклад, рівень тривожності до та після тренінгу. Класифікація вибірок визначає фундаментальну вісь подальшого вибору статистичного критерію, адже параметричні й непараметричні методи мають різні модифікації для незалежних і залежних даних.

Другим важливим чинником є тип вимірювальної шкали, у якій зафіксовано змінні. Номінальні шкали допускають лише констатацію категорії, порядкові – ранжування, тоді як інтервальні та шкали відношень відкривають можливість здійснювати арифметичні операції, розраховувати середні значення, стандартні відхилення та інші параметри. Відповідно, у випадку номінальних і порядкових змінних можливості параметричних тестів різко обмежені, що робить непараметричні підходи основним інструментом аналізу.

Третім кроком є оцінка характеру розподілу. Нормальність розподілу залежної змінної є критичною умовою застосування параметричних критеріїв, таких як t-тест, ANOVA, кореляція Пірсона чи регресійні моделі з класичними припущеннями. У практиці психологічних досліджень нормальний розподіл зустрічається з різною частотою: у великих вибірках він часто наближається до нормального через роботу центральної граничної теореми, проте у змінних, пов'язаних із суб'єктивним досвідом, розподіл майже завжди скошений. Відхилення від нормальності слід не лише констатувати, але й оцінити їхню величину, адже у випадку великої вибірки незначні відступи можна ігнорувати, тоді як у малих вибірках необхідно застосовувати непараметричні або робастні методи. Якщо розподіл суттєво відхиляється від нормального, можливе застосування математичних трансформацій (логарифмічної, кореневої, оберненої), але лише у тих випадках, коли це узгоджується зі змістом змінної. Якщо ж зберегти нормальність неможливо, необхідно застосовувати непараметричні аналоги параметричних тестів або бутстрепінг,

який дозволяє оцінити статистичні характеристики на основі багатократної генерації вибірок.

Після того як дослідник оцінив типи вибірок, шкалу та розподіл, слід визначити характер завдання. Якщо необхідно порівняти емпіричний розподіл із теоретичним, наприклад із нормальним або рівномірним, використовується одновибірковий критерій – χ^2 Пірсона, критерій Колмогорова–Смірнова. Другий тип завдань – порівняння двох вибірок. Якщо вибірки незалежні, а дані мають інтервальний або відношеневий характер та відповідають нормальному розподілу, застосовується t-критерій Стюдента для незалежних вибірок; у разі порушення нормальності – непараметричний критерій Манна–Уїтні. Для залежних вибірок відповідно застосовується парний t-критерій або критерій Вілкоксона. Третій тип завдань – порівняння трьох і більше груп. Якщо дотримано вимоги параметричності, використовується дисперсійний аналіз (ANOVA), при порушенні нормальності – критерій Краскела–Уолліса; у випадку залежних вимірювань – повторні вимірювання ANOVA або критерій Фрідмана. Четвертий тип статистичних завдань полягає у виявленні зв'язків між змінними. Для інтервальних даних із нормальним розподілом застосовують кореляцію Пірсона, для порядкових або ненормальних – кореляцію Спірмена або Кендалла. Якщо необхідно оцінити причинні моделі, використовується регресійний аналіз – лінійний, множинний або логістичний, залежно від типу залежної змінної. Регресійні моделі вимагають суворої перевірки припущень – нормальності залишків, гомоскедастичності, відсутності мультиколінеарності, лінійності зв'язків.

Однак сучасна психологічна статистика не обмежується лише порівняннями та пошуком кореляцій. Якщо дослідницьке завдання спрямоване на вивчення структури даних, виділення латентних конструктів або типів респондентів, застосовуються методи багатовимірного аналізу. Факторний аналіз дозволяє визначити, які латентні змінні лежать в основі спостережуваних показників, та перевірити, чи відповідає структура опитувальника теоретичній моделі. У дослідженнях особистісних властивостей, мотивації або когнітивних стилів факторний аналіз є ключовим етапом валідації. Якщо мета полягає не в підтвердженні теоретично заданої структури, а в пошуку оптимального простору для представлення даних, застосовується багатовимірне шкалювання, яке дозволяє візуалізувати подібність чи відмінність між об'єктами у вигляді дво- або тривимірних конфігурацій. У випадку, коли необхідно виокремити групи досліджуваних на основі схожості профілів показників, використовується кластерний аналіз – ієрархічний, k-means, аналіз латентних класів. Кластеризація є незамінною в диференціальній психології, психодіагностиці та організаційних дослідженнях, де важливо не лише виміряти рівні змінних, а й виявити природні групування.

Таким чином, вибір статистичного критерію – це багатоступеневий процес, що включає аналіз типу вибірок, шкали вимірювання, характеру розподілу, специфіки дослідницького завдання та вимог до інтерпретації

результатів. Немає універсального статистичного тесту, який був би придатним для всіх ситуацій, тому обґрунтування вибору методу є важливою складовою наукової доброчесності. Коректне застосування статистичних критеріїв є ключовим принципом сучасної доказової психології.

9.3. Статистичні критерії порівняння емпіричного розподілу з очікуваним

У процесі емпіричного дослідження психолог постійно стикається із завданням оцінити, наскільки отримані ним дані узгоджуються з теоретичними уявленнями, статистичними припущеннями або наперед заданими очікуваними значеннями. У психології така перевірка може стосуватися різноманітних аспектів: розподілу частот категорій поведінки, відхилення оцінок респондентів від очікуваного середнього, відповідності результатів нормативним даним, типовим емпіричним закономірностям або теоретичним моделям. Одновибіркові методи перевірки гіпотез дають змогу оцінити, чи є різниця між спостережуваними та очікуваними значеннями статистично значущою, тобто чи може вона бути пояснена випадковістю, чи свідчить про наявність реального ефекту.

Одним із найпростіших, але водночас ефективних інструментів є біноміальний критерій, який застосовується у випадках, коли змінна має лише дві можливі категорії, а дослідник хоче порівняти фактичну частоту однієї з них із частотою, що задається теоретичною гіпотезою. Типовим прикладом у психології може бути оцінка частки респондентів, які дали відповідь «так» або «ні» на запитання про наявність певної поведінкової звички, або перевірка того, чи відсоток учасників із високим рівнем стресу перевищує очікувані 50 %. Математично біноміальний критерій ґрунтується на властивостях біноміального розподілу, який моделює ймовірність появи певної кількості успіхів у N випробуваннях за фіксованої ймовірності p . Якщо очікувана частка становить, наприклад, 0.50, то гіпотеза H_0 стверджує, що варіації спостережуваних частот повністю зумовлені випадковими флуктуаціями, тоді як H_1 передбачає відхилення від цього значення. У психології біноміальний критерій широко використовується, коли дослідження ґрунтується на дихотомічних відповідях – наприклад, для оцінки того, чи більшість підлітків демонструє ризиковану поведінку, чи відсоток студентів, що користуються психологічною допомогою, є більшим за популяційну норму, або чи частка респондентів, які погодилися з твердженням про високий рівень тривоги, статистично значуще перевищує очікувану. Важливим обмеженням цього критерію є його застосовність лише до змінних із двома категоріями та вимога незалежності випробувань. Типовою помилкою є використання біноміального тесту для питання, яке містить більше двох варіантів, або у випадках, коли відповіді респондентів можуть бути залежними – наприклад, при попарних оцінках.

Якщо номінальна змінна має три або більше категорій, або якщо досліднику потрібно оцінити відхилення спостережуваного частотного розподілу від теоретичного, застосовується χ^2 -критерій Пірсона для однієї вибірки. Цей критерій дозволяє перевірити, наскільки емпіричні частоти узгоджуються із частотами, очікуваними за певною гіпотезою. Йдеться саме про порівняння частотних розподілів, а не оцінок чи середніх значень, тому χ^2 належить до класичних непараметричних методів. У психологічних дослідженнях його використовують для багатьох цілей: перевірки того, чи розподіл типів темпераменту відповідає популяційному, визначення відмінностей між групами в частоті прояву певної поведінки, оцінки того, чи рівні задоволеності працею розподілені рівномірно або у відповідності до певної моделі. Математично χ^2 визначається як сума квадратів відхилень спостережуваної частоти від очікуваної, поділена на очікувану частоту. Якщо отримане значення перевищує критичне для відповідних ступенів свободи, то гіпотеза про узгодженість емпіричних і теоретичних частот відхиляється. Прикладом психологічної ситуації, де застосовується χ^2 -критерій, є дослідження професійних інтересів: якщо очікується рівномірний розподіл між різними типами, але емпірично спостерігається домінування певних профілів, χ^2 дозволяє статистично підтвердити цю невідповідність. Проте важливим обмеженням є вимога достатності спостережень у кожній категорії – очікувані частоти не повинні бути надто малими, інакше критерій втрачає коректність. Частою помилкою є механічне застосування χ^2 до категорійних розподілів без оцінки їх адекватності для цього методу або ігнорування того, що категорії повинні бути взаємовиключними та повністю вичерпними.

Для змінних, що вимірюються у порядковій або інтервальній шкалі, проблема порівняння емпіричного розподілу з очікуваним найчастіше пов'язана з питанням нормальності або відповідності певному теоретичному закону. Одним із найпоширеніших критеріїв для цього є одновибірковий критерій Колмогорова-Смірнова. На відміну від χ^2 , який працює з частотами, Колмогорова-Смірнова тест аналізує кумулятивні розподіли, порівнюючи максимальне відхилення між емпіричною функцією розподілу та теоретичною. Якщо це відхилення є статистично значущим, дослідник має підстави стверджувати, що вибірка не походить із розподілу певного типу. У психології такий критерій застосовують не лише для перевірки нормальності (хоча саме це є найпоширенішим), але й у випадках, коли розглядаються інші теоретичні розподіли, наприклад у моделях реакційного часу, які часто ґрунтуються на експоненційних або Пуассонівських законах. Прикладом може бути дослідження когнітивних процесів, де час реакції має властивість правостороннього скошення, і дослідник перевіряє, чи емпіричні дані відповідають очікуваному експоненційному розподілу. Суттєвою перевагою тесту Колмогорова-Смірнова є його чутливість до відхилень у всьому діапазоні значень, а не лише у хвостах або центрі розподілу. Проте одним із ключових обмежень є те, що параметри теоретичного розподілу повинні бути задані наперед, а не оцінені з відповідної ж вибірки, що в реальних

психологічних дослідженнях може викликати методологічні складнощі. Також тест Колмогорова-Смірнова тест є надто чутливим у великих вибірках, часто відхиляючи нормальність навіть при мінімальних відхиленнях, що не впливають на практичний аналіз; навпаки, у малих вибірках він може не виявляти суттєвих відхилень. Типовою помилкою є бездумне використання Колмогорова-Смірнова як єдиного критерію нормальності без розгляду графічних методів, таких як Q-Q графіки, або без урахування контексту дослідження, де відхилення від нормальності можуть бути як природними, так і методологічно незначущими.

Окреме місце серед одновибіркових критеріїв належить одновибірковому t-критерію, який дозволяє порівняти середнє значення вибірки з певним теоретичним або конкретним числовим значенням. На відміну від попередніх критеріїв, t-тест працює не з розподілами чи частотами, а з інтервальними даними та припущенням щодо середнього. У психології цей підхід використовується в ситуаціях, коли дослідник має певну гіпотезу щодо середнього значення популяції – наприклад, коли потрібно оцінити, чи середній рівень тривожності групи перевищує нормативне значення за психодіагностичною методикою, чи середній рівень задоволеності життям відрізняється від загальнопопуляційного індикатора, або чи середній показник за шкалою депресії перевищує встановлений клінічний поріг. Математично t-тест оцінює співвідношення між різницею емпіричного та теоретичного середнього і стандартною похибкою, яка визначає ступінь невизначеності у вимірюванні середнього значення. Важливою умовою застосування цього тесту є нормальний розподіл даних або достатньо велика вибірка, для якої працює центральна гранична теорема. Обмеженням є також чутливість до викидів, які можуть суттєво змінити середнє, та до широкої дисперсії, що зменшує статистичну потужність. Типовою помилкою є інтерпретація t-тесту без урахування величини ефекту або без перевірки того, чи значення, з яким порівнюється середнє, має реальну теоретичну або практичну обґрунтованість.

Таким чином, одновибіркові критерії відіграють важливу роль у психологічному аналізі, забезпечуючи можливість оцінки відповідності емпіричних даних теоретичним моделям або попередньо сформульованим очікуванням. Вибір конкретного критерію визначається типом шкали вимірювання, природою змінної, формулюванням гіпотези, вимогами до розподілу даних та логікою дослідження. Біноміальний тест є незамінним при аналізі дихотомічних даних, χ^2 критерій дає змогу оцінити структуру категоріальних розподілів, критерій Колмогорова-Смірнова дозволяє перевіряти складні теоретичні моделі розподілу, а одновибірковий t-тест надає досліднику інструмент для перевірки гіпотез щодо середніх значень. Разом ці критерії утворюють фундаментальний інструментарій для глибокого, статистично обґрунтованого аналізу психологічних даних, забезпечуючи надійну основу для формулювання висновків і подальших інтерпретацій.

9.4. Статистичні критерії порівняння двох емпіричних розподілів

Порівняння двох емпіричних розподілів є одним із найпоширеніших завдань у психологічних дослідженнях, оскільки саме воно дозволяє встановити, чи існують статистично значущі відмінності між групами людей, умовами експерименту або часовими зрізами. У багатьох випадках дослідник прагне з'ясувати, чи відрізняються рівні тривожності у чоловіків та жінок, чи вплинуло втручання на зміну когнітивних показників, чи спостерігається зменшення симптомів депресії після терапії, або ж чи мають люди різного віку відмінні профілі мотивації. Незалежно від того, яка саме психологічна характеристика аналізується, ключовим етапом є вибір адекватного статистичного критерію, який би коректно оцінював різницю між двома вибірками. Невірний вибір статистичного тесту може призвести до хибних висновків, спотвореної інтерпретації, а подекуди – до відтворення ефектів, яких не існує в реальності. Саме тому важливо розуміти логіку та припущення кожного критерію, типи шкал вимірювання, нормальність або ненормальність розподілу даних, а також властивості вибірок – незалежні чи залежні.

Найтипівішими представниками параметричних критеріїв для цього завдання є t-критерій Стьюдента для незалежних та залежних вибірок. Його логіка базується на оцінці різниці між двома середніми значеннями та варіативністю всередині груп. Якщо варіативність велика, то навіть суттєва різниця між середніми може виявитися статистично незначущою; якщо варіативність мала, то навіть невеликі відмінності можуть бути виявлені як достовірні. Психологічні ситуації, де застосовується цей критерій, є надзвичайно численними. Наприклад, дослідник може оцінювати різницю між середнім рівнем стресу у двох групах працівників – тих, хто проходить регулярні тренінги психологічної підтримки, і тих, хто їх не проходить. Якщо дані мають нормальний розподіл, а дисперсії є приблизно рівними, t-критерій забезпечує найбільш ефективну оцінку цієї різниці. Аналогічно, у клінічній психології він використовується для порівняння рівня симптомів депресії до і після терапевтичного втручання – у такому випадку застосовують t-критерій для залежних вибірок, оскільки дані отримані від одних і тих самих людей у двох різних часових точках.

Однак параметричні методи вимагають дотримання припущення нормальності, яке в реальних умовах психологічних досліджень дотримується далеко не завжди. Багато психологічних змінних мають природне скошення: рівень тривожності часто має правосторонній хвіст, час реакції розподілений асиметрично, а кількість конфліктів у сім'ї не може бути нормальною через наявність великої кількості нульових значень. Крім того, параметричні підходи чутливі до викидів, які здатні суттєво змінити середнє значення та збільшити ризик хибних висновків. Тому навіть за бажання використовувати t-критерій дослідник повинен ретельно перевірити властивості розподілу й оцінити варіативність даних.

Коли дані не відповідають параметричним вимогам, застосовують непараметричні критерії. Їхня логіка ґрунтується не на середніх значеннях, а на рангах, тобто на порядкових відносинах між значеннями. Найвідомішим непараметричним тестом для незалежних вибірок є критерій Манна-Уїтні, який визначає, чи можна вважати значення однієї групи статистично вищими або нижчими, ніж значення іншої групи. Фактично цей тест оцінює ступінь перекриття розподілів: якщо один розподіл систематично має вищі значення, то рангова сума цієї групи буде більшою. У психології критерій Манна-Уїтні застосовується у ситуаціях, коли дані вимірюються у порядковій шкалі або мають суттєві відхилення від нормальності. Наприклад, якщо дослідник вивчає рівень агресивності у двох групах підлітків – тих, хто пережив травматичний досвід, і тих, хто не пережив, – то розподіл даних може бути асиметричним або містити викиди, що робить t-критерій непридатним. Критерій Манна-Уїтні дасть змогу адекватно оцінити відмінності між групами, не порушуючи строгих припущень.

Для залежних вибірок непараметричним аналогом є критерій Вілкоксона, який аналізує напрям змін між двома вимірюваннями в одних і тих самих респондентів. Його часто застосовують у клінічних і психотерапевтичних дослідженнях, де розподіл результатів після терапії часто виявляється скошеним, а різниця до/після може бути нерівномірною. Наприклад, у дослідженні впливу когнітивно-поведінкової терапії на рівень соціальної тривожності зміни можуть бути значущими тільки у підгрупи респондентів, у той час як інші майже не демонструють прогресу. Критерій Вілкоксона чутливо фіксує систематичні зрушення рангових значень, що дає змогу оцінити ефективність втручання навіть тоді, коли параметричні тести недоступні.

Окрему роль відіграють критерії для порівняння частотних розподілів. У випадках, коли змінні представлені не числовими значеннями, а категоріями, застосовують критерії порівняння частот. Найпоширенішим серед них є χ^2 -критерій Пірсона, який дає змогу оцінити, чи пов'язані значення двох категоріальних змінних між собою. Він широко застосовується в соціальній і віковій психології, під час аналізу поведінкових особливостей або при порівнянні частоти певних відповідей між групами. Наприклад, можна оцінити, чи відрізняється частота проявів прокрастинації між студентами різних курсів або чи є залежність між рівнем освітньої мотивації та формою навчання. Якщо одна група демонструє суттєво інші частотні характеристики, ніж інша, χ^2 -критерій це зафіксує. Важливо враховувати, що він чутливий до розміру вибірки та до очікуваних частот, тому його застосування потребує мінімальної достатності даних у кожній категорії.

Для парних номінальних спостережень, коли два вимірювання здійснюються у тих самих респондентів, використовується критерій Мак-Немара. Його застосовують у ситуаціях, коли вимірюється зміна статусу респондента за дихотомічною змінною – наприклад, «проявляє симптом» / «не проявляє симптом». У психологічній діагностиці цей критерій є ключовим при

аналізі ефективності інтервенцій. Якщо дослідник прагне з'ясувати, чи змінилася частка дітей із проблемною поведінкою після тренінгової програми, критерій Мак-Немара дозволяє оцінити не загальний рівень, а структурне зміщення часток усередині одних і тих самих респондентів. Він особливо актуальний у дослідженнях, де поведінкові зміни мають категоріальний характер і не можуть бути описані шкалами з багатьма градаціями.

Порівняння двох вибірок є методологічним ядром психологічних досліджень. У практиці психолога майже кожне емпіричне завдання так чи інакше зводиться до порівняння двох груп. Психолог оцінює зміни до і після втручання, співставляє різні соціальні або вікові групи, аналізує гендерні відмінності, порівнює клінічні й неклінічні вибірки, або визначає ефективність навчальних програм. Саме тому важливо не лише знати формули критеріїв, а головне – правильно обирати підхід відповідно до типу змінних, властивостей розподілу та структури вибірки.

На практиці алгоритм вибору тесту виглядає таким чином: якщо змінна кількісна, а групи незалежні та розподіл приблизно нормальний – застосовують незалежний t-критерій. Якщо розподіл ненормальний або є викиди, – критерій Манна–Уїтні. Якщо дані отримані двічі від одних і тих самих респондентів, – залежний t-критерій або критерій Вілкоксона залежно від властивостей розподілу. Якщо змінні категоріальні – χ^2 для незалежних або Мак-Немара для залежних. Усі ці критерії дозволяють відповісти на фундаментальне запитання психологічної статистики: чи є різниця між двома групами настільки великою, щоб її неможливо було списати на випадковість?

9.5. Статистичні критерії порівняння трьох і більше емпіричних розподілів

Статистичне порівняння трьох і більше емпіричних розподілів є однією з найважливіших процедур у психологічних дослідженнях, оскільки більшість реальних експериментальних проєктів передбачає аналіз не двох, а кількох груп одночасно. Це можуть бути вікові підгрупи, рівні експозиції експериментального впливу, різні умови тренінгу, декілька категорій клінічних проявів або різні типи стимулів, що подаються учасникам. Коли кількість порівнюваних умов зростає, класичні підходи, призначені для двох вибірок, перестають працювати, а послідовні парні порівняння призводять до накопичення статистичної помилки першого роду. Саме тому в статистичній практиці була сформована окрема група критеріїв, що дають змогу оцінити наявність загальних відмінностей між кількома вибірками одночасно, а у разі підтвердження таких відмінностей – перейти до аналізу конкретних пар груп із корекцією рівня значущості. Логіка цих методів ґрунтується на тому, що спочатку перевіряється глобальна гіпотеза про рівність усіх середніх або медіан, і лише тоді, коли ця гіпотеза відхиляється, проводиться детальніший аналіз джерел відмінностей.

Перед тим як застосовувати багатовибіркові методи, дослідник повинен установити тип шкали вимірювання, на якій представлені дані, оскільки саме це визначає, які критерії можна використовувати. Якщо дані виміряні в інтервальній або шкалі відношень і відповідають припущенням про нормальний розподіл і однорідність дисперсій, застосовуються параметричні методи, головним з яких є дисперсійний аналіз. Якщо ж дані є ранговими або мають очевидні відхилення від нормальності, використовуються непараметричні аналоги, серед яких найпоширенішим є критерій Крускала-Уолліса. Вибір способу аналізу завжди має ґрунтуватися не лише на формальному визначенні шкали, а й на фактичних властивостях розподілу, перевірених за допомогою графіків, тестів нормальності та аналізу аутлаєрів (викидів). У психологічних вибірках часто спостерігаються асиметричні розподіли, підвищена дисперсія, кластери або інші форми відхилення від нормальності, що робить непараметричні методи корисною альтернативою основним тестам.

Класичним параметричним методом порівняння трьох і більше вибірок є дисперсійний аналіз (ANOVA), який дозволяє оцінити, чи існують статистично значущі відмінності між середніми показниками кількох груп. Його ключовою ідеєю є порівняння двох джерел варіації: варіації між групами, яка відображає систематичні ефекти, і варіації всередині груп, яка характеризує індивідуальні або випадкові відмінності. Якщо різниця між групами перевищує випадкову мінливість, це є підставою для відхилення нульової гіпотези про рівність середніх. Дисперсійний аналіз представлений у різних модифікаціях, залежно від структури дизайну. Найпростішою є однофакторна модель, у якій існує лише один фактор групування. У психологічних дослідженнях такі моделі часто застосовують при порівнянні трьох вікових груп за рівнем когнітивної гнучкості, кількох варіантів тренінгових програм за впливом на рівень стресостійкості або різних типів стимулів за величиною емоційної реакції. Якщо експеримент має складнішу структуру, коли вплив на залежну змінну можуть чинити два або більше факторів, застосовується двофакторний, трьохфакторний або навіть багатофакторний дисперсійний аналіз. Такі моделі дозволяють не лише оцінити окремий вплив кожного фактора, а й перевірити взаємодію факторів, яка є особливо важливою у психології, де ефекти одного чинника часто залежать від значень іншого.

Специфічною формою багатовибіркових порівнянь є дисперсійний аналіз із повторними вимірюваннями, який використовується, коли дані отримано від тих самих учасників у кількох умовах. Це класичні внутрішньогрупові дизайни, що з'являються у дослідженнях пам'яті, уваги, навчання, перцептивних процесів або в інтервенційних експериментах, де вимірювання проводять на різних етапах тренінгу чи лікування. Такий підхід потребує особливої обережності, оскільки дані всередині однієї особи не є незалежними, що зумовлює необхідність використання спеціальних статистичних моделей. Дисперсійний аналіз із повторними вимірюваннями

дає змогу врахувати кореляційну структуру даних, однак потребує дотримання специфічного припущення – сферичності; якщо воно порушене, коефіцієнти F-критерію можуть бути некоректними, і тоді застосовують поправки Грінхауза-Гайссера або Гінфельда-Г'юна. Вибір відповідної поправки є одним із ключових аспектів інтерпретації таких моделей, і недооцінка порушень сферичності часто призводить до хибних висновків.

Якщо дані не відповідають вимогам параметричних методів, стають актуальними непараметричні підходи, серед яких провідним є критерій Крускала-Уолліса. Він ґрунтується на ранговому представленні значень і дозволяє оцінити, чи належать вибірки до однієї генеральної сукупності без припущення нормальності. У психологічних дослідженнях цей критерій часто застосовується у випадках, коли залежна змінна є результатом суб'єктивного рейтингу, шкали Лікерта або коли розподіл має явно виражену асиметрію, медіана краще відображає центральну тенденцію, ніж середнє, або дані містять значну кількість однакових значень. Логіка критерію ґрунтується на порівнянні середніх рангів у групах: якщо ранги суттєво відрізняються, відмінності між групами вважаються статистично значущими. Хоча критерій Крускала-Уолліса не дає змоги безпосередньо оцінити величину ефекту, він дозволяє отримати глобальний висновок про наявність або відсутність відмінностей між трьома чи більше вибірками.

Важливо усвідомлювати, що підтвердження наявності значущих відмінностей лише відкриває шлях до наступного етапу аналізу – порівняння конкретних пар груп. Оскільки багаторазові порівняння різко збільшують ризик помилки першого роду, у статистиці розроблено систему пост-хок тестів, що дозволяють зберегти коректний рівень α . У параметричних моделях найпоширенішими є корекції Тьюкі, Бонферроні та Шеффе, які відрізняються рівнем жорсткості та чутливістю до відхилень у дисперсіях. У непараметричному контексті застосовуються рангові аналоги, такі як критерії Данна або пост-хок порівняння з корекцією Бонферроні. Правильний вибір пост-хок аналізу є критичним, оскільки від цього залежить точність локалізації ефекту, а надто консервативні корекції можуть маскувати реальні відмінності. У психології, де зміни часто є невеликими, але стабільними, баланс між статистичною потужністю та контролем за помилками набуває особливої ваги.

Слід окремо відзначити, що багатовибіркові порівняння значно чутливіші до якості даних, ніж двовибіркові аналізи. Наявність аутлаєрів, пропущених значень, неоднорідність дисперсій або різні розміри груп можуть суттєво вплинути на результати. Тому до початку багатофакторного аналізу дослідник має провести детальну перевірку розподілів, оцінити варіативність у групах, переконатися, що дисперсії не надмірно відрізняються, і за потреби застосувати робастні методи або бутстрепінг. Останній стає дедалі популярнішим у сучасній психологічній статистиці, оскільки дозволяє будувати довірчі інтервали для середніх чи різниць між групами без жорстких припущень про форму розподілу.

У контексті психологічних досліджень застосування багатовибіркових критеріїв є не лише статистичною процедурою, а й методологічним рішенням, яке визначає логіку інтерпретації результатів. Коли дослідник має справу з кількома групами, виникає потреба не просто порівняти їх попарно, а встановити загальну структуру впливу, виявити фактори, що визначають варіацію залежної змінної, і зрозуміти, чи є отримані відмінності випадковими або систематичними. Дисперсійний аналіз дозволяє побачити загальну картину, оцінити внесок кожного фактора та їх взаємодії, а непараметричні методи – забезпечити стійкість висновків у разі складних або асиметричних розподілів.

Таким чином, критерії порівняння трьох і більше емпіричних розподілів утворюють цілісну систему, що дає можливість аналізувати складні психологічні дані з урахуванням їх багатовимірності та варіативності. Вони дозволяють не лише встановити наявність статистично значущих відмінностей, а й зрозуміти структуру цих відмінностей, визначити механізми їх виникнення та забезпечити високий рівень наукової обґрунтованості висновків.

9.6. Статистичні критерії для виявлення взаємозв'язків між змінними

Вивчення взаємозв'язків між психологічними змінними становить одну з ключових аналітичних задач у сучасних емпіричних дослідженнях. Незалежно від того, йдеться про оцінку зв'язку між рівнем тривожності та академічною успішністю, самооцінкою і соціальною активністю чи характеристиками особистості та показниками професійної ефективності, дослідник завжди стикається з необхідністю застосування статистичних процедур, які дозволяють кількісно встановити силу, напрям, значущість та форму зв'язку між досліджуваними показниками. На відміну від процедур порівняння груп, де головним об'єктом аналізу є відмінності між вибірками, статистичні критерії для оцінки зв'язків спрямовані на дослідження взаємної варіації показників і визначення того, наскільки зміна одного параметра є пов'язаною зі зміною іншого. При цьому сам факт наявності зв'язку не завжди означає причинність, і правильна інтерпретація результатів потребує зваженого врахування як теоретичних моделей, так і методологічних припущень статистичних критеріїв.

Першим кроком у дослідженні будь-якого зв'язку є правильний вибір співвідношення між типами шкал вимірювання, оскільки саме вони визначають, які математичні операції і які статистичні критерії можна застосовувати. Зв'язки між номінальними змінними потребують використання критеріїв, заснованих на таблицях спряженості. Коли обидві змінні є категоріальними, застосовується χ^2 -Пірсона або похідні від нього міри асоціації (Cramer's V, коефіцієнт контингенції, коефіцієнт ϕ для дихотомічних змінних). Для взаємозв'язків між порядковими змінними застосовуються

рангові коефіцієнти Спірмена або Кендалла, які не покладаються на припущення нормальності та лінійності і дають можливість працювати з монотонними зв'язками будь-якої форми. У випадку кількісних змінних традиційним інструментом є коефіцієнт лінійної кореляції Пірсона, який дозволяє оцінити силу та напрям лінійного зв'язку між змінними за наявності нормального розподілу й відсутності суттєвих викидів.

Кореляційний аналіз лежить в основі абсолютної більшості психологічних досліджень, однак його можливості достатньо обмежені. Кореляція не дозволяє встановити причинність, не враховує вплив третіх змінних і не дає змоги проаналізувати багатофакторну структуру зв'язків. Крім того, значимість коефіцієнта кореляції залежить від розміру вибірки, а його величина – від варіативності вимірювань. Саме тому кореляційні методи є важливою, але лише початковою ланкою аналітичного процесу.

Для глибшого аналізу взаємозв'язків між змінними у психологічних дослідженнях центральне місце займає регресійний аналіз. Його сутність полягає у побудові рівняння, що дозволяє передбачити значення однієї змінної (залежної) на основі значень однієї чи кількох незалежних (предикторів). Проста лінійна регресія визначає зв'язок між двома інтервальними змінними та дозволяє встановити, як зміна предиктора на одну одиницю впливає на зміну критерію. У цьому випадку регресійний коефіцієнт «В» має зміст прямої кількісної інтерпретації – він показує, на скільки балів у середньому зміниться залежна змінна за зміни предиктора на один бал. Стандартизований коефіцієнт « β » дозволяє порівнювати ефекти різних предикторів між собою, оскільки він відображає зміну залежної змінної в одиницях стандартних відхилень.

Однак у більшості психологічних ситуацій взаємодія між конструктами набагато складніша, ніж може описати проста регресія. Людська поведінка завжди багатофакторна, а психічні явища зумовлені взаємодією цілого набору факторів. Саме тому у реальних дослідженнях основним методом виступає множинна регресія, яка дозволяє одночасно оцінити вплив кількох предикторів на результат. Множинна регресія розділяє загальну дисперсію залежної змінної на частки, кожна з яких пояснюється певним предиктором за умови контролю інших. Завдяки цьому можна визначити унікальний внесок кожної змінної у прогноз залежної змінної.

Одним із найбільш інформативних показників регресійної моделі є коефіцієнт детермінації R^2 , який показує, яка частка дисперсії залежної змінної пояснена сукупністю предикторів. Цей показник виступає аналогом розміру ефекту для всієї моделі, а його значущість перевіряється за допомогою F-критерію в таблиці ANOVA. Однак R^2 не дає інформації щодо внеску окремих предикторів. Для цього використовується аналіз коефіцієнтів β , які показують, наскільки суттєвим є унікальний зв'язок кожного з предикторів із залежною змінною. Важливо, що кореляція змінної з залежною змінною сама по собі не гарантує значущого внеску в регресію: якщо декілька змінних претендують на «один і той самий шматок» дисперсії, він не дістанеться жодній із них. Таким

чином, чим більше предикторів у моделі та чим більше вони корелюють між собою, тим вищий ризик мультиколінеарності – ситуації, коли окремі коефіцієнти стають нестабільними, а інтерпретація моделі – проблематичною.

Для запобігання мультиколінеарності застосовуються діагностичні індикатори, такі як Tolerance та VIF. Значення Tolerance нижче 0.30 або VIF вище 4 свідчать про потенційну небезпеку. У таких випадках доцільно видалити один із предикторів, перетворити змінні або застосувати альтернативні статистичні моделі. З огляду на складність інтерпретації взаємозв'язків у багатофакторних моделях, сучасна психологічна методологія рекомендує уникати автоматичних (покрокових) алгоритмів регресії, які порушують стабільність моделі та можуть призводити до висновків, що не відтворюються на інших вибірках.

Окрему методологічну цінність має ієрархічна регресія, яка дозволяє оцінити внесок нової змінної після контролю вже введених у модель предикторів. Саме вона використовується для аналізу інкрементної валідності – можливості нового психологічного показника покращувати прогноз залежної змінної. Наприклад, дослідник може послідовно додати у модель спочатку демографічні змінні, потім особистісні риси, а далі – показники мотивації, оцінюючи, наскільки кожен рівень покращує якість прогнозу. Це робить ієрархічну регресію важливим інструментом для оцінки унікального внеску нових психодіагностичних методик або конструктів.

Важливою умовою коректного застосування регресійного аналізу є відповідність даних певним статистичним припущенням. Розподіли інтервальних змінних мають бути близькими до нормальних, а зв'язки між змінними – приблизно лінійними. Залишки моделі повинні мати нормальний розподіл, бути незалежними та демонструвати гомоскедастичність – рівність дисперсій на різних рівнях предикторів. Порушення цих припущень знижує надійність моделі й може бути підставою для застосування непараметричних методів, таких як регресія для порядкових змінних або логістична регресія для дихотомічних залежних змінних. Окремо слід зазначити, що деякі психологічні конструкти мають принципово ненормальний розподіл (наприклад, показники щастя або симптоматики), тому зусилля з «нормалізації» даних не завжди є теоретично обґрунтованими.

Сфера застосування регресійного аналізу у психології є надзвичайно широкою. Він дозволяє дослідникам порівнювати відносну важливість різних психологічних чинників, аналізувати вплив особистісних рис на поведінку, виявляти структурні закономірності у процесах навчання, спілкування або емоційної регуляції. У прикладних дослідженнях регресія використовується для побудови моделей прогнозування, наприклад, оцінки ризику професійного вигорання, прогнозу академічної успішності або ефективності психологічних втручань. У таких випадках доцільно застосовувати регресію для перевірки якості прогнозу на нових вибірках, що значно підвищує екологічну валідність отриманих результатів.

Таким чином, статистичні критерії для аналізу взаємозв'язків між змінними становлять методологічну основу психологічних досліджень. Вони дозволяють не лише оцінити силу і напрям зв'язку між окремими змінними, але й побудувати складні багатofакторні моделі, що відображають реальну структурну організацію психологічних конструктів. Регресійний аналіз при цьому відіграє центральну роль, забезпечуючи гнучкий і точний інструмент для вивчення багатовимірних процесів людської поведінки та психіки. Застосування цих методів вимагає як високої статистичної компетентності, так і уважного ставлення до теоретичних моделей та психометричних характеристик даних. І саме завдяки такому поєднанню статистичної точності та теоретичної рефлексії дослідник отримує можливість робити обґрунтовані висновки щодо складних психологічних закономірностей та їх практичного значення.

9.7. Оцінка розміру ефекту в психологічних дослідженнях

Розмір ефекту в психологічних дослідженнях є одним із ключових показників, який дозволяє оцінити не лише факт наявності певного статистичного зв'язку або відмінності, а й практичну силу та значущість виявленого явища. На відміну від *p-value*, який лише інформує про те, чи сумісні дані з нульовою гіпотезою, розмір ефекту вказує на те, наскільки великою є різниця або наскільки тісним є зв'язок між змінними. Саме тому сучасні стандарти наукової звітності у психології наполягають на обов'язковому поданні оцінок розміру ефекту у кожному емпіричному дослідженні. Це не просто технічна вимога, а істотний компонент доказовості, адже статистично значущі результати зазвичай досягаються навіть за дуже малих ефектів, коли вибірка є великою; навпаки, у малих вибірках навіть помітні ефекти можуть виявитися статистично незначущими. Звідси випливає, що *p-value* без контролю за розміром вибірки не дає уявлення про силу впливу, тоді як розмір ефекту забезпечує універсальний, стандартизований спосіб оцінити змістовність результатів, незалежно від вибірки й особливостей конкретного статистичного методу.

Для ілюстрації різниці між значущістю та величиною ефекту часто наводять приклади, у яких один і той самий дослідницький задум реалізується через різні статистичні підходи. Якщо порівнюються дві групи за середніми значеннями, можна застосувати *t*-критерій і подати результат у вигляді різниці між середніми, або ж можна розглядати групову змінну як бінарний предиктор та оцінити зв'язок із кількісним показником через точково-бісеріальну кореляцію, що є окремим випадком кореляції Пірсона. Попри відмінність у математичному апараті, обидва підходи оцінюють по суті одне й те саме – ступінь розбіжності між двома вибірками. Це дозволяє зрозуміти важливе методологічне положення: у психології різні статистичні інструменти часто приводять до еквівалентних результатів, якщо вони вимірюють одну й ту саму концептуальну величину. Це підкреслює універсальність стандартних

показників розміру ефекту, насамперед d Коена та r Пірсона, які утворюють два взаємопов'язані «сімейства» індикаторів – перше описує різницю між групами, друге характеризує силу зв'язку між змінними

У психологічних дослідженнях найбільш поширеною є оцінка d Коена, яка виражає різницю між середніми значеннями груп у одиницях стандартного відхилення. Це дозволяє абстрагуватися від масштабу вимірювання та отримати універсальну величину, яку можна порівнювати між дослідженнями, навіть якщо вони застосовують різні тести чи шкали. Наприклад, якщо одна група має середнє значення 80, а інша – 74, то сама по собі різниця у 6 балів мало що говорить нефахівцеві; однак якщо ця різниця становить 0,75 стандартного відхилення, це вже дає змогу дійти висновку про досить виражений ефект. Коефіцієнт r у свою чергу відображає силу лінійного зв'язку між двома змінними й може варіювати від -1 до 1 . Нерідко дослідники запитують, як співвідносяться d і r , і чи можна перевести один індикатор в інший. Відповідь позитивна, і для нормальних розподілів, рівних дисперсій та симетричних груп існують точні формули конвертації. Такі співвідношення дозволяють наводити різні ефекти у єдиній стандартизованій формі та оцінювати їх відносну інтенсивність. У презентації наведено класичні відповідності: $r = 0,10$ приблизно відповідає $d = 0,20$; $r = 0,30$ – $d \approx 0,63$; $r = 0,50$ – $d \approx 1,15$. Це демонструє, що навіть відносно невелика кореляція може відповідати відчутному розриву між групами, і навпаки

Проте важливо пам'ятати, що будь-які «загальні» рекомендації щодо інтерпретації величин ефекту є умовними і залежать від контексту. Ще Джейкоб Коен наголошував, що його класифікації «малий», «середній» і «великий» ефект слід застосовувати з обережністю, бо в різних галузях психології різні ефекти можуть бути цінними. Наприклад, у клінічній психології навіть невелике підвищення благополуччя може мати велике практичне значення, тоді як у когнітивній психології слабкі ефекти часто є методологічним шумом. У презентації наведено показовий приклад з метааналізу ефективності психотерапії, де середній ефект становив близько $r = 0,32$, що означає 68% шансів того, що випадково обрана людина, яка отримує терапію, почуватиметься краще, ніж випадково обрана людина з тією ж проблемою, але без терапії. Хоча $r = 0,32$ виглядає помірним, його практична значущість у контексті допомоги дуже велика. У цьому полягає фундаментальна різниця між статистичною та прикладною значущістю ефектів: навіть «слабкий» ефект може мати вагоме соціальне чи клінічне значення, якщо він стосується важливих життєвих результатів або великої популяції. Навпаки, високі значення ефектів у лабораторних умовах можуть втрачати своє прикладне значення у реальних контекстах

Ще одним показником, який набуває популярності, є CLES – «common language effect size», або «ефект звичайної мови». Він відображає імовірність того, що випадково обрана людина з однієї групи матиме більший показник, ніж випадково обрана з іншої. Цей підхід робить розмір ефекту значно зрозумілішим для нефахівців, оскільки перекладає статистичні величини у

форму практичної, інтуїтивно зрозумілої імовірності. Наприклад, якщо $CLES = 0,70$, це означає, що у 70% випадків людина з експериментальної групи демонструватиме вищий показник, ніж людина з контрольної. Інтерпретація розміру ефекту в термінах CLES часто полегшує пояснення результатів у міждисциплінарних дослідженнях, у звітах для освітніх установ чи при роботі з практиками, які не володіють статистичним апаратом. Аналогом цього підходу для кореляцій є $CL(r)$, запропонований Данлапом, який дозволяє оцінити, з якою ймовірністю людина з вищим значенням однієї змінної матиме також вище значення іншої. Такі інтерпретації особливо корисні в соціальній психології, педагогічних дослідженнях і прикладних програмах оцінки ефективності втручань.

Розмір ефекту відіграє ключову роль і в метааналітичних дослідженнях. Метааналіз передбачає узагальнення результатів багатьох досліджень, а для цього потрібні стандартизовані індикатори, які дозволяють порівнювати, підсумовувати або зважувати окремі ефекти. Використання лише p -value унеможлиблює такі порівняння, натомість d , r , η^2 , R^2 та інші індикатори забезпечують уніфікований підхід. Ба більше, у складних багатофакторних моделях (наприклад, у дослідженнях із застосуванням множинної регресії, моделювання структурними рівняннями чи аналізу дисперсії) розмір ефекту дозволяє оцінити внесок окремих предикторів у варіацію залежної змінної. Це надзвичайно важливо, коли модель містить багато змінних, і необхідно визначити, які з них є справді впливовими, а які – незначущими у практичному сенсі.

Нарешті, розмір ефекту є базовим параметром при визначенні статистичної потужності й плануванні вибірки. Щоб дослідження мало достатню ймовірність виявити ефект, якщо він справді існує, потрібно заздалегідь оцінити очікувану величину ефекту, часто ґрунтуючись на попередніх дослідженнях або метааналізах. Це дозволяє встановити мінімальний необхідний обсяг вибірки, уникнути недопотужних досліджень і підвищити відтворюваність результатів. У вже проведених дослідженнях розмір ефекту дозволяє оцінити, наскільки результати є стійкими та чи могли вони бути виявлені за наявного обсягу вибірки.

Таким чином, оцінка розміру ефекту у психологічному дослідженні є не просто додатковим параметром у статистичному аналізі, а невід'ємною частиною сучасної наукової культури. Вона забезпечує можливість адекватно інтерпретувати результати, порівнювати їх між собою, застосовувати їх у практиці, планувати дослідження та робити обґрунтовані висновки щодо сили та значущості виявлених явищ. У психології найчастіше застосовують d Коена, r Пірсона та індикатори поясненої дисперсії R^2 або η^2 , і важливо пам'ятати, що їхня цінність визначається не лише числовим значенням, а й контекстом, у якому вони інтерпретуються. Розмір ефекту є тим інструментом, який дозволяє перейти від статистичної інтерпретації до змістовної, від формальної значущості – до реальної, від чисел – до пояснення реальних

психологічних процесів, що лежать в основі поведінки, емоцій чи когнітивних механізмів людини.

Завдання на самостійну підготовку

1. Чому підготовка емпіричних даних до статистичного аналізу розглядається як центральна ланка наукового процесу, а не як суто технічний етап?
2. Які показники центральної тенденції та мінливості є мінімально необхідними для первинного опису психологічних даних і які висновки можна зробити на їх основі?
3. Які типові помилки введення та технічні артефакти можна виявити на етапі аналізу описових статистик і крайніх значень?
4. Як правильне визначення типу вимірювальної шкали (номінальна, порядкова, інтервальна, відношень) впливає на вибір подальших статистичних процедур?
5. Чому при аналізі пропущених значень важливо не лише зафіксувати їхню кількість, але й зрозуміти механізм їх виникнення (випадкові, систематичні тощо)?
6. Чому оцінка нормальності розподілу має спиратися на поєднання графічних методів (гістограми, Q–Q, box-plot) та кількісних індикаторів (асиметрія, ексцес, критерії нормальності)?
7. Що таке аутлаєри (викиди), якими методами їх виявляють і за яких умов їх видалення може бути обґрунтованим?
8. У яких ситуаціях доцільніше не трансформувати дані, а обрати непараметричні чи робастні статистичні методи (рангові критерії, бутстрепінг тощо)?
9. Яку послідовність кроків можна вважати базовим алгоритмом вибору статистичного критерію?
10. У чому полягає концептуальна відмінність між одновибірковими критеріями (біноміальний тест, χ^2 для однієї вибірки, Колмогорова–Смірнова, одновибірковий t-тест) щодо завдань, які вони розв'язують?
11. Які обмеження мають критерії χ^2 Пірсона при аналізі частотних розподілів у психології?
12. За яких умов t-критерій Стюдента (для незалежних і залежних вибірок) може виявитися некоректним, навіть якщо формально змінна є інтервальною?
13. У чому полягає принципова різниця між параметричними критеріями порівняння груп (t-тест, ANOVA) та їх непараметричними аналогами (Манна–Уїтні, Вілкоксона, Краскела–Уолліса, Фрідмана)?
14. Яку роль відіграє дисперсійний аналіз (ANOVA) при порівнянні трьох і більше вибірок і чому просте множення парних t-тестів є методологічно хибним?
15. Що таке пост-хок аналіз у рамках порівняння трьох і більше груп і чому необхідні корекції рівня значущості (Тьюкі, Бонферроні тощо)?

16. Які припущення лежать в основі регресійного аналізу (лінійність, нормальність залишків, гомоскедастичність, відсутність мультиколінеарності) і як їх порушення впливає на інтерпретацію результатів?

17. У чому зміст відмінності між простою та множинною регресією в психологічних дослідженнях і що дає аналіз коефіцієнта детермінації R^2 ?

18. Як ієрархічна регресія використовується для оцінки інкрементної валідності нових психологічних показників?

19. Чому p -value не можна розглядати як показник сили ефекту і яку роль у сучасних дослідженнях відіграють стандартизовані показники розміру ефекту (d Коена, r , η^2 , R^2)?

20. Як інтерпретація розміру ефекту в термінах «мови ймовірностей» (CLES, CL(r)) допомагає пояснювати результати психологічних досліджень нефахівцям?

РОЗДІЛ 10 ПРЕДСТАВЛЕННЯ РЕЗУЛЬТАТІВ НАУКОВОГО ДОСЛІДЖЕННЯ

10.1. Значення візуалізації даних в науковому дослідженні

Значення візуалізації в науковому дослідженні важко переоцінити, особливо в галузях, де об'єктом аналізу є складні, багатовимірні та контекстуально зумовлені явища, як-от у психології. У процесі проведення експериментального дослідження психолог отримує масив емпіричних даних, які підлягають статистичній обробці. Одним із найпоширеніших початкових етапів такої обробки є обчислення описових статистик – середнього арифметичного, медіани, моди, дисперсії, стандартного відхилення, діапазону та інших показників центральної тенденції й варіативності. Ці індикатори справді дозволяють описати загальні характеристики вибірки, сформулювати перше уявлення про розподіл результатів і створити базовий фон для подальшого аналізу. Однак самі по собі вони є лише грубим наближенням до реальної структури даних. Описові статистики агрегують інформацію, «спресовуючи» її в кілька числових значень, які приховують важливі аспекти розподілу – асиметрію, багатoverшинність, наявність викидів, кластери, різні профілі у підгрупах. У психологічних дослідженнях, де аналізуються індивідуальні відмінності, складні патерни поведінки, багатовимірні конструкти або суб'єктивні переживання, ігнорування цих аспектів може призводити до грубого спрощення й, зрештою, до хибних висновків.

Нерідко можна спостерігати ситуацію, коли дослідники обмежуються поданням таблиці з описовими статистиками та одним–двома коефіцієнтами кореляції або результатами регресійного аналізу, вважаючи, що цього достатньо для повного розкриття емпіричних даних. Такий підхід формально відповідає вимогам до статистичної звітності, але фактично залишає читача наодинці з абстрактними цифрами, позбавленими наочного виміру. Таблиця може містити середні бали для двох груп, стандартні відхилення та значення *p-value*, але без візуального подання розподілу стає неможливо оцінити, чи справді різниця є змістовною, чи не зумовлена вона кількома екстремальними спостереженнями, чи не приховуються за «середнім» значенням дві різні підгрупи. Проблема полягає в тому, що людина як пізнавальний суб'єкт краще сприймає структуру, ніж ізольовані числа: наш когнітивний апарат еволюційно пристосований до виявлення форм, патернів, пропорцій, а не до оперування абстрактними величинами. Саме візуалізація даних дозволяє перевести сухі числові ряди у форму, де структура стає очевидною: розподіл набуває форми, зв'язки між змінними – контуру, а відхилення – помітних сигналів, що потребують пояснення.

Класичним прикладом того, наскільки небезпечно спиратися лише на описові статистики без візуалізації, є знаменитий квартет Енскомба, створений

статистиком Френсісом Енскомбом у 1973 році. Він складався з чотирьох штучно сконструйованих наборів даних, які мали абсолютно ідентичні значення середнього, дисперсії, коефіцієнта кореляції між змінними, а також однакові результати лінійної регресії.

Таблиця 10.1 – Дані для квартету Енскомба

№	А		Б		В		Г	
	Х	У	Х	У	Х	У	Х	У
1	10	8,04	10	9,14	10	7,46	8	6,58
2	8	6,95	8	8,14	8	6,77	8	5,76
3	13	7,58	13	8,74	13	12,74	8	7,71
4	9	8,81	9	8,77	9	7,11	8	8,84
5	11	8,33	11	9,26	11	7,81	8	8,47
6	14	9,96	14	8,1	14	8,84	8	7,04
7	6	7,24	6	6,13	6	6,08	8	5,25
8	4	4,26	4	3,1	4	5,39	19	12,5
9	12	10,84	12	9,13	12	8,15	8	5,56
10	7	4,82	7	7,26	7	6,42	8	7,91
11	5	5,68	5	4,74	5	5,73	8	6,89
M_x, M_y	9,0	7,5	9,0	7,5	9,0	7,5	9,0	7,5
SD_x, SD_y	3,32	2,03	3,32	2,03	3,32	2,03	3,32	2,03
r	0,82		0,82		0,82		0,82	

Якщо дивитися лише на таблицю числових показників, усі чотири набори здаються повністю еквівалентними і, з погляду традиційної статистичної звітності, взаємозамінними. Однак варто побудувати графіки розсіювання – і відкривається радикально інша картина. У першому наборі точки лягають майже ідеально вздовж прямої, ілюструючи класичну лінійну залежність. У другому дані утворюють криву, що свідчить про виражений нелінійний зв'язок, який цілком невидимий на рівні коефіцієнта кореляції. У третьому більшість точок розміщені близько до прямої, але один сильний викид суттєво впливає на параметри моделі, і візуальний огляд негайно ставить питання про доцільність включення цього спостереження. У четвертому наборі майже всі точки розташовані майже вертикально, а одна є екстремальною і фактично створює ілюзію кореляції там, де її по суті немає. Без візуалізації ми мали б справу з формально однаковими наборами даних, але з точки зору змістового аналізу – це чотири абсолютно різні ситуації, які вимагають різних висновків, різних моделей і різних інтерпретацій.

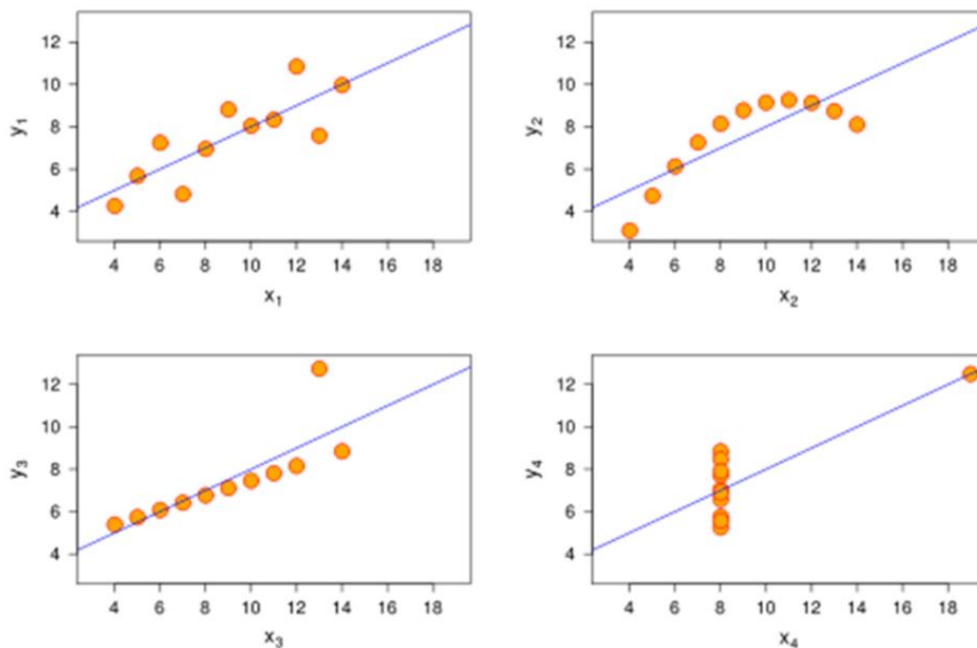


Рисунок 10.1 – Квартет Енскомба

Цей приклад особливо промовистий для психологічних досліджень, де структура даних часто є складною, багатовимірною і далекою від ідеалізованих припущень класичної статистики. Візуалізація даних у психології – це не косметичне оформлення результатів, а інструмент виявлення прихованих закономірностей, які не видно в числах; відхилень і викидів, що можуть істотно спотворювати загальну картину; порушень припущень статистичних моделей, таких як лінійність, нормальність розподілу, гомогенність дисперсій; а також аномальних або парадоксальних результатів, які вимагають додаткового тлумачення або повторної перевірки. Наприклад, гістограми, діаграми розсіювання, «ящики з вусами», щільності розподілу або профільні графіки можуть миттєво показати, що дані далекі від нормального розподілу, містять кілька кластерів або різко відрізняються між групами. У багатьох випадках саме графік змушує дослідника переглянути початкові гіпотези, відмовитися від стандартних параметричних тестів, застосувати робастні методи або інші моделі, які краще відповідають емпіричній структурі.

Особливе значення візуалізація має в контексті реплікаційної кризи та проблеми недостатньої прозорості даних у сучасній психології. Однією з ключових претензій до значної частини опублікованих робіт є те, що вони подають тільки стислий набір числових показників, без повноцінної демонстрації структури даних. Такий підхід унеможливорює незалежну оцінку якості вимірювання, стабільності ефектів і адекватності застосованих статистичних процедур. Коли дані не візуалізуються, рецензенти й читачі позбавлені можливості побачити, чи не зумовлений «значущий» результат кількома крайніми спостереженнями, чи не є різниця між групами наслідком нерівномірного розподілу, чи не суперечить емпірична картина припущенням,

на яких базується вибраний статистичний тест. Візуалізація, таким чином, є важливим елементом наукової прозорості: добре побудований графік часто говорить більше, ніж сторінка таблиць.

У дискусіях про відкриті дані візуалізація часто розглядається як проміжна ланка між повним доступом до сирих даних і кінцевими статистичними висновками. З одного боку, сам по собі графік не замінює відкриття датасету – без вихідних даних неможливо відтворити аналіз, перевірити розрахунки, застосувати альтернативні моделі. З іншого боку, навіть за умови наявності відкритих даних, відсутність якісних візуалізацій у публікації робить результати менш прозорими й менш зрозумілими для широкої аудиторії. Візуалізація виступає у ролі «вікна» в структуру даних, забезпечуючи базову можливість оцінити, чи узгоджуються числові показники із загальною картиною. У середовищі відкритої науки важливо, щоб разом із датасетами публікувалися також коди побудови графіків, оскільки це дозволяє відтворити не лише аналітичні, а й візуальні процедури. Відсутність таких матеріалів робить статтю менш відтворюваною навіть у разі формальної наявності даних.

У психологічних дослідженнях візуалізація виконує ще одну важливу функцію – функцію комунікації складних ідей і результатів для різних аудиторій. Графіки, побудовані з дотриманням наукових стандартів, роблять результати доступними не лише для статистиків чи досвідчених дослідників, але й для здобувачів освіти, практичних психологів, лікарів, освітян, управлінців, а також представників суміжних галузей, які не володіють поглибленими знаннями в галузі математичної статистики. Добре продумана візуалізація дозволяє інтуїтивно зрозуміти основні тенденції, оцінити розмір ефекту, побачити перекриття між групами, проілюструвати динаміку змін, порівняти різні моделі. Вона створює спільну «мову», яка полегшує міждисциплінарну взаємодію і сприяє поширенню наукових результатів за межі вузького академічного кола.

Важливо підкреслити, що візуалізація даних не повинна обмежуватися простим відтворенням стандартних шаблонів на кшталт стовпчикових діаграм із помилками середнього. Психологічні дані часто мають багатшарову структуру, і завдання дослідника полягає в тому, щоб обрати такі форми візуалізації, які дозволять максимально повно відобразити складність явища. Наприклад, при вивченні особистісних профілів доцільно використовувати радарні діаграми, профільні графіки чи кластерні візуалізації, які демонструють групи з різними поєднаннями рис, а не лише середні значення по кожній шкалі. При аналізі динаміки станів упродовж часу корисними є лінійні графіки з індивідуальними траєкторіями або агреговані графіки із зазначенням інтервалів довіри. У дослідженнях, де важливу роль відіграють латентні змінні, візуалізація результатів факторного аналізу, багатовимірного шкалювання, структурного моделювання дозволяє показати, як «розташовуються» об'єкти й змінні у латентному просторі. Таким чином, візуалізація стає частиною теоретичної інтерпретації, а не лише інструментом ілюстрації.

Разом із тим візуалізація може стати і джерелом маніпуляцій, якщо застосовується безвідповідально. Неправильно обрана шкала, усічена вісь, навмисне збільшення або зменшення візуальних відмінностей, вибіркове подання лише частини даних можуть призводити до спотворення реальної картини. Графіки, покликані підвищувати прозорість, іноді використовуються як засіб маскування слабких результатів або перебільшення ефектів. Щоб цього уникнути, необхідні чіткі стандарти візуальної звітності, які б регламентували не лише технічні аспекти побудови графіків, а й принципи етичного подання даних: повнота, коректність масштабів, відсутність свідомих викривлень, пояснення вибору тих чи інших способів візуалізації. В ідеалі наукова стаття повинна містити лаконічні, але інформативні візуалізації, а разом з тим – можливість доступу до повної версії графіків і вихідних даних у відкритих репозиторіях.

Значення візуалізації стає особливо очевидним у ситуаціях, коли результати виявляються неочікуваними, суперечливими або слабкими. У таких випадках саме графічне подання може показати, чи йдеться про справжній феномен – наприклад, про складну нелінійну залежність або про наявність прихованих підгруп, – чи про артефакт, пов'язаний із викидами, помилками вимірювання або порушеннями припущень моделі. Візуалізація тут виконує роль діагностичного інструменту, що дає змогу відокремити змістовні ефекти від шуму. У результаті публікуються ефекти, які згодом не відтворюються й підіриває довіру до психологічної науки.

Таким чином, значення візуалізації в науковому дослідженні виходить далеко за межі декоративного оформлення результатів. Вона є ключовим елементом якісного аналізу даних, засобом виявлення структурних особливостей вибірки, інструментом перевірки припущень статистичних моделей, каналом наукової комунікації та важливим компонентом відкритої науки. В умовах реплікаційної кризи візуалізація виступає не лише технічним, а й етичним ресурсом, який може або сприяти підвищенню прозорості та відтворюваності, або, навпаки, за відсутності стандартів і належної культури використання ставати інструментом маскування проблем. Справжня наукова відповідальність вимагає від дослідника не просто «додати кілька графіків» до статті, а вибудовувати візуальне подання даних як органічну частину дослідницького процесу, де кожен графік є аргументом, відкритим для критичного розгляду, перевірки та реплікації. Лише за таких умов візуалізація стає тим, чим вона покликана бути в сучасній науці, – інструментом не прикрашання, а поглиблення розуміння реальності.

10.2. Принципи та правила якісного графічного представлення даних

Однією з ключових вимог до наукової графіки є відмова від надмірної декоративності та візуальних ефектів, які спотворюють сприйняття кількісних співвідношень. Зокрема, тривимірні (3D) графіки, популярні у масових

презентаціях і бізнес-візуалізації, у науковому контексті вважаються неприйнятними, оскільки вони викривляють площу й об'єм, створюючи оманливе враження щодо величини або динаміки ефекту. Підтінення, перспективні викривлення, псевдооб'ємність, додаткові тіні – усі ці елементи можуть привнести неправдиві акценти у сприйняття графіка. Наприклад, у стовпчастій діаграмі навіть невеликі зміни кута огляду можуть призвести до того, що один стовпець виглядатиме значно більшим за інший, хоча різниця у значеннях є мінімальною. У галузях, де точність передавання результатів дослідження є критично важливою, наприклад, у клінічній чи освітній психології, психології праці або психодіагностиці, подібні викривлення можуть мати реальні наслідки, оскільки результати можуть бути неправильно інтерпретовані практиками, адміністраторами чи науковцями.

Графік повинен містити лише ті елементи, які необхідні для розуміння інформації, і не повинен бути ані перевантаженим текстом та декоративними елементами, ані надмірно «повітряним», де простір заповнений порожнечою. Розміщення підписів, заголовків, маркерів осей, легенд, числових значень повинно бути збалансованим і читабельним. Занадто дрібні написи роблять графік нечітким, а надто великі – справляють враження непропорційності та руйнують загальну структуру представлення. Графік має бути створений таким чином, щоб читач без додаткових пояснень міг зрозуміти, яку змінну він відображає, яка шкала використовується, які одиниці вимірювання застосовано, який час, умови або вибірки порівнюються. Відсутність підписів осей або одиниць вимірювання перетворює графік на формальну картинку, що може бути інтерпретована довільним чином. Якщо на осі позначено тільки цифри без зазначення того, що саме вони означають, то графік втрачає функцію наукового аргументу та досягає рівня декоративності.

Значною мірою культура наукової графіки формується через дотримання принципу економії виразних засобів, відомого як концепція «максимум інформації при мінімумі чорнила», що була сформульована Едвардом Тафте. Сенс цього принципу полягає в тому, що кожен елемент графіка повинен нести смислове навантаження. Усе, що не є необхідним для розуміння змісту даних – декоративні сітки, надлишкові рамки, тіні, градієнти, зайва кількість шрифтів, кольорові ефекти, – слід усувати. Графік має бути настільки інформативним, наскільки це можливо, і водночас позбавленим візуального шуму, який відволікає увагу від суті. У науковому контексті зайва декоративність не просто ускладнює інтерпретацію, а й може створювати небезпеку маніпуляції даними. Наприклад, вибір кольорів, які привертають увагу до другорядних елементів, може змістити акцент з основних результатів.

Обов'язковою вимогою до будь-якого наукового графіка є наявність чітко підписаних осей. Кожна вісь повинна містити назву змінної, яку вона представляє, а також одиницю вимірювання. Це дозволяє інтерпретувати значення не лише в контексті «більше-менше», а й у реальних параметрах. Наприклад, якщо графік демонструє рівень тривожності, потрібно вказати, що шкала вимірювання – бали або умовні одиниці (у.о.), якщо шкала авторська.

Відсутність одиниць вимірювання створює простір для помилкових або вільних тлумачень результатів, чого слід уникати.

Також важливою є єдність стилю оформлення. Зокрема, розмір шрифту, який використовується у графіках, повинен відповідати основному тексту документа. Допускається зменшення розміру на 1-2 пункти, але не більше. Занадто дрібний шрифт робить графік важким для читання, особливо якщо публікація чи презентація передбачає демонстрацію на екрані або в друкованому форматі.

Кожен графік має виконувати конкретну функцію. Якщо метою є показати динаміку доцільно використовувати лінійний графік; якщо потрібно продемонструвати розподіл – гістограму або коробковий графік; якщо необхідно порівняти середні значення – стовпчасту діаграму з позначенням довірчих інтервалів. Використання невідповідного типу графіка призводить до спотворення змісту.

Наприклад, застосування секторної діаграми для відображення змін за часом є помилкою, адже вона не передає хронологічної послідовності. Так само використання стовпчастого графіка для кореляційних зв'язків не дає змоги побачити форму залежності. Функціональність полягає у відповідності графічного засобу типу даних і дослідницькому завданню.

Відомо, що стовпчасті діаграми приховують варіативність, затирають інформацію про форми розподілу та роблять невидимими викиди, які можуть істотно впливати на статистичні висновки. У психологічних дослідженнях, де результати часто ґрунтуються на невеликих вибірках і чутливих показниках, така втрата інформації є особливо небезпечною. Натомість графіки з позначенням довірчих інтервалів або інтервальних оцінок, візуалізація індивідуальних значень, графіки щільності розподілу або точкові діаграми дозволяють демонструвати як середні тенденції, так і структуру індивідуальних відмінностей. Такі графіки дозволяють одночасно оцінити і середнє значення, і надійність цього показника, і ступінь варіативності. Це особливо важливо для міжгрупових порівнянь або в дослідженнях, де перевіряється статистична значущість ефекту.

Особливого значення набуває питання правильного вибору шкал. Якщо шкала починається не з нуля, це обов'язково має бути чітко позначено, і необхідно пояснити, чому саме таке рішення було прийнято. Маніпулятивне обрізання шкали здатне штучно збільшити або зменшити різницю між групами, і цим активно користуються в мас-медіа, але такі прийоми є неприйнятними в науковій практиці. Наприклад, різниця між 93 і 96 балами на графіку, що починається з 90, виглядатиме як різниця між 30 і 60, хоча насправді вона лише 3 бали. У психології, де різниця між групами може бути дуже тонкою, некоректне використання шкал може повністю змінити зміст висновків.

Формат зображення також має важливе значення. У наукових роботах краще не використовувати графіки у форматі .jpeg, оскільки алгоритм стискання спотворює дрібні деталі, призводить до появи артефактів і знижує чіткість ліній та підписів. Доцільно використовувати формати, що зберігають

високу якість, – насамперед .png. Це забезпечує чіткість ліній, правильне відтворення шрифтів і збереження кольорової палітри. У випадку схем, які можуть масштабуватися без втрати якості, оптимальними є векторні формати. Питання формату набуває особливої ваги в умовах публікації даних у відкритих репозиторіях, які є важливою складовою сучасної наукової культури: погана якість графіків унеможлиблює вторинний аналіз результатів і робить дослідження менш прозорим.

Не менш важливим аспектом графічного представлення даних є вибір кольорової гами. Кольори повинні бути не лише інформативними, але й доступними для людей з різними формами кольоросприйняття. Кольори мають бути контрастними, але не надто яскравими. Нейтральні палітри або комбінації синього й оранжевого, чорного й червоного, сірої шкали з акцентним кольором забезпечують і читабельність, і наукову стриманість. Неправильно підібрана кольорова схема може змістити увагу з важливих елементів, а також створити ілюзію наявності або відсутності закономірності.

Отже, графічне представлення даних у науковому дослідженні є не лише технічним, а й концептуальним завданням. Від дотримання принципів чіткості, стриманості, інформативності, коректності шкал, адекватного вибору кольорів і форматів залежить не лише якість візуального подання результатів, а й рівень прозорості, достовірності та відтворюваності всього дослідження. У сучасній науці графік є не просто додатком до тексту, а цілісним інтелектуальним висловлюванням, яке повинно відповідати високим стандартам наукової культури. Саме через якісну візуалізацію здійснюється перехід від цифр до розуміння, від результатів до інтерпретації, від індивідуального дослідження до колективної наукової дискусії.

10.3. Графічне представлення даних залежно від типу шкали вимірювання

Вибір типу графіка безпосередньо залежить від характеру емпіричних даних і шкали вимірювання, яка використовується під час дослідження. Невідповідність між типом шкали та способом графічного представлення є однією з найпоширеніших помилок у наукових роботах, що часто призводить до спотворення змісту результатів. Тому кожен дослідник повинен добре орієнтуватися у видах шкал і розуміти, які графічні засоби найадекватніше відображають інформацію, отриману в межах певного вимірювального підходу.

У межах класичної типології, запропонованої С. Стівенсом, емпіричні показники психологічних досліджень можна віднести до чотирьох основних шкал: номінальної, порядкової, інтервальної та шкали відношень. Кожна з цих шкал характеризується власним набором математичних операцій, які є допустимими й інтерпретаційно коректними, а отже, й власною логікою графічного представлення. Ігнорування специфіки шкали, наприклад, використання лінійного графіка для номінальних даних або подання порядкових змінних як рівноінтервальних є методологічною помилкою.

Для номінальних змінних характерна категоріальність, тобто відсутність логічного порядку між значеннями. У таких випадках головним завданням візуалізації є демонстрація частот або відсоткових співвідношень між категоріями. Найчастіше використовуються стовпчасті (барові). Кожен стовпець представляє окрему категорію, а його висота відображає кількість або відсоток респондентів. Наприклад, при дослідженні типів темпераменту у студентів можна зобразити, скільки осіб належить до групи сангвініків, холериків, флегматиків і меланхоліків (рис. 10.2).

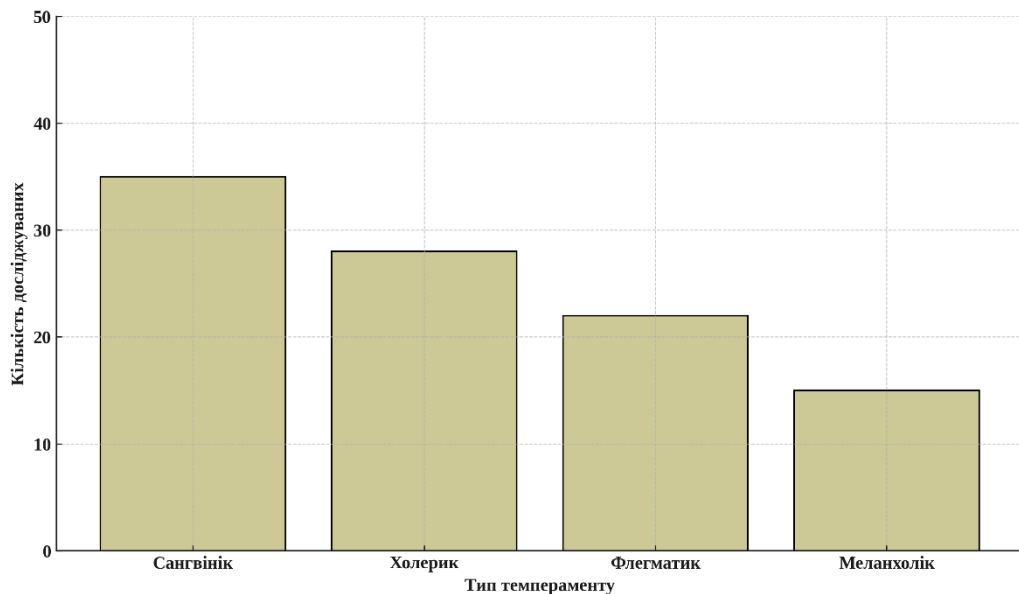


Рисунок 10.2 – Розподіл типів темпераменту у здобувачів вищої освіти (n = 100)

Складені стовпчасті діаграми застосовуються, коли потрібно показати взаємозв'язок між двома номінальними змінними. Наприклад, можна подати співвідношення типів темпераменту в чоловіків і жінок у межах однієї діаграми, де кожен стовпець поділено на кольорові сегменти, що представляють категорії іншої змінної.

Секторні діаграми (pie charts) – ілюструють частки кожної категорії в загальній сукупності (рис. 10.3). Вони зручні, коли кількість категорій не перевищує чотирьох-п'яти, і коли важливо показати структуру розподілу. Водночас у науковій літературі секторні діаграми часто критикують за низьку точність зчитування, тому їх варто використовувати лише як допоміжний інструмент. Основна вимога до візуалізації номінальних змінних – чіткість та унікальність кольорових позначень. Кожна категорія має бути легко відрізнена від інших, а кольори повинні бути нейтральними та контрастними, без надмірної декоративності.

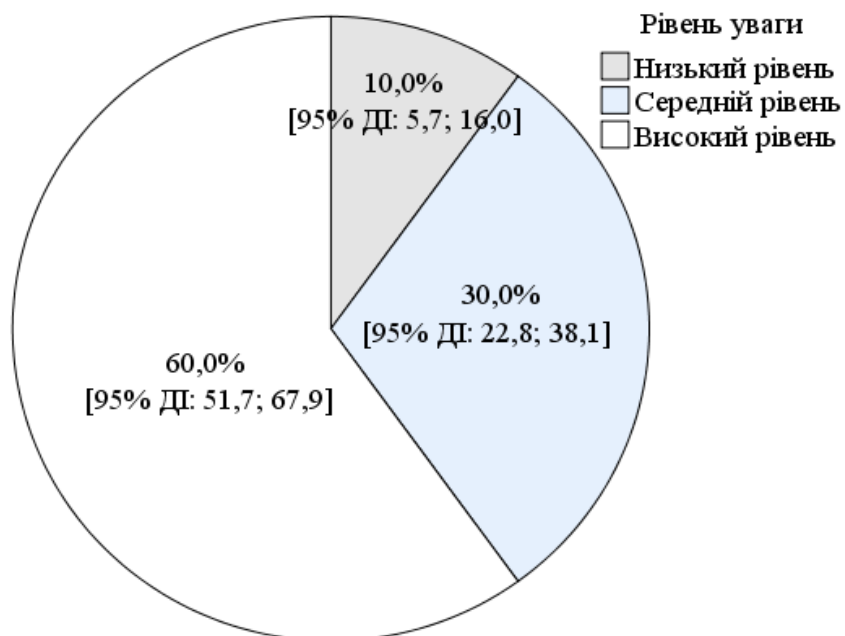


Рисунок 10.3 – Розподіл рівнів уваги у рятувальників (n = 100)

Порядкові шкали дозволяють ранжувати об'єкти, але не визначають точну відстань між рівнями. Тому для таких змінних важливо не передавати хибне враження про рівномірність інтервалів. Найкраще підходять такі види графіків: стовпчасті діаграми з упорядкованими категоріями – розміщення рівнів змінної (наприклад, «низький», «середній», «високий») повинно бути послідовним за зростанням або спаданням. Це дає змогу швидко побачити тенденцію, не порушуючи логіки шкали.

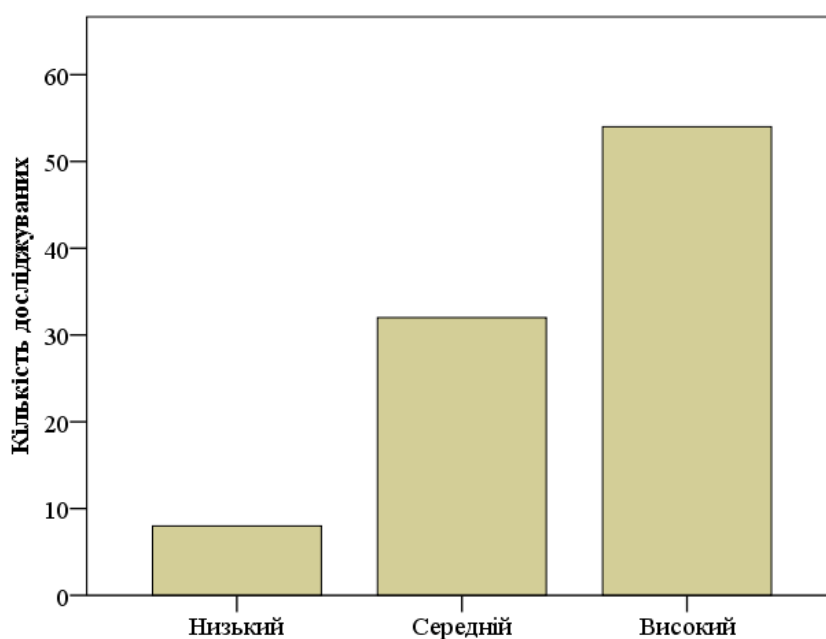


Рисунок 10.4 – Розподіл рятувальників за рівнями психологічного благополуччя (n = 97)

Полігони частот – використовуються для відображення розподілу рангових оцінок. Полігон дозволяє побачити форму розподілу (симетрію, асиметрію, наявність піків або викидів), що може мати важливе інтерпретаційне значення.

Боксплоти (коробкові графіки) – особливо корисні для порядкових змінних коли потрібно показати медіану та варіативність рангових оцінок. Боксплот дозволяє візуально оцінити розмах значень і наявність аномальних спостережень (рис. 10.5).

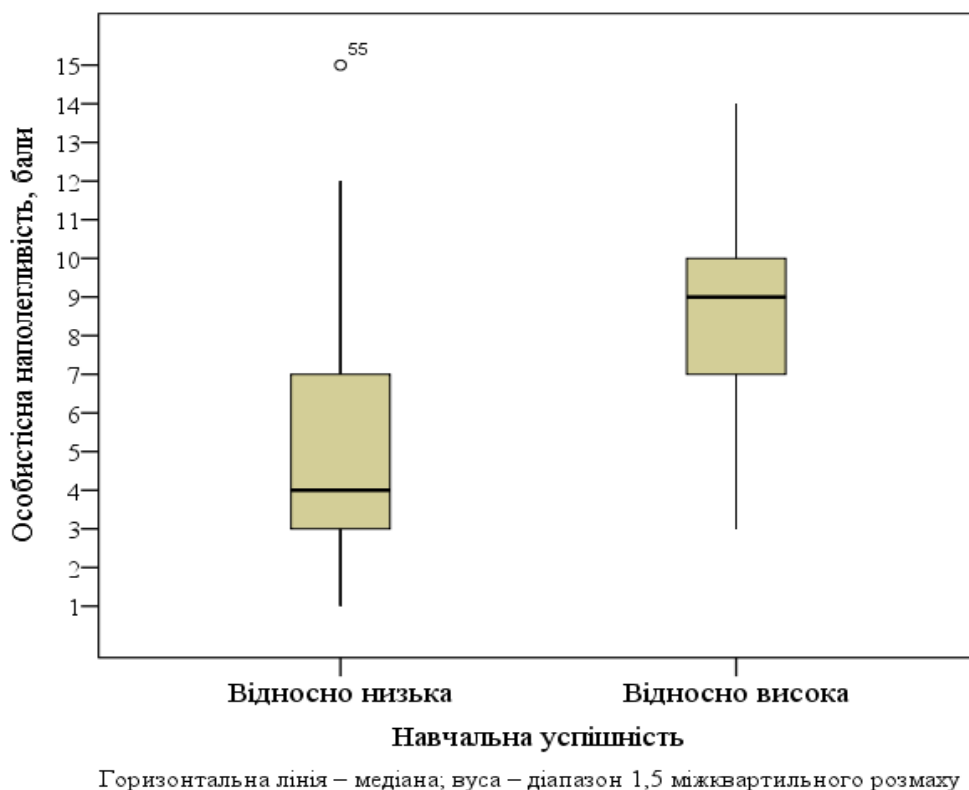


Рисунок 10.5 – Особливості наполегливості здобувачів вищої освіти з різним рівнем навчальної успішності ($n_1 = 61$; $n_2 = 44$)

Шкали інтервалів відношень дозволяють проводити найповніший статистичний аналіз, тому і спектр графічних методів тут є найширшим. До основних належать.

Лінійні графіки (line plots) – застосовуються для демонстрації динаміки змін у часі або залежності між кількісними показниками. Наприклад, можна зобразити зміну рівня емоційного вигорання працівників протягом року чи динаміку показника тривожності після серії тренінгів. Вісь X у таких графіках відображає часову або незалежну змінну, а вісь Y – залежну змінну (рис. 10.6).

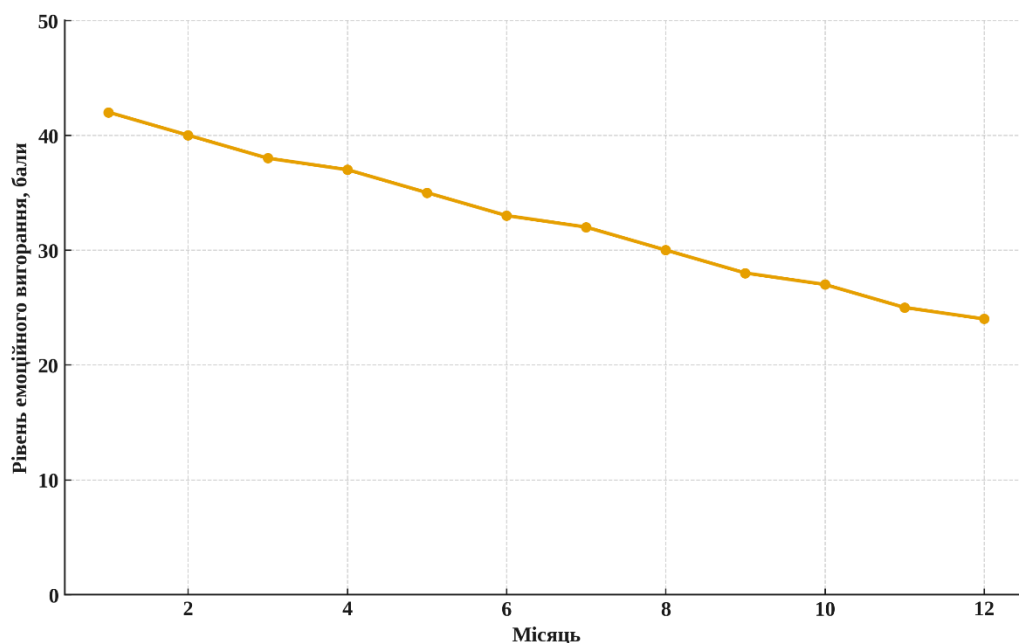


Рисунок 10.6 – Динаміка емоційного вигорання працівників протягом року (n = 20)

Графіки розсіювання (scatter plots) – використовуються для аналізу зв'язку між двома кількісними змінними. Кожна точка на площині відповідає парі значень X і Y для окремого респондента (рис. 10.7). Такий графік дозволяє візуально оцінити характер залежності (лінійний, криволінійний, відсутній), силу зв'язку та наявність викидів.

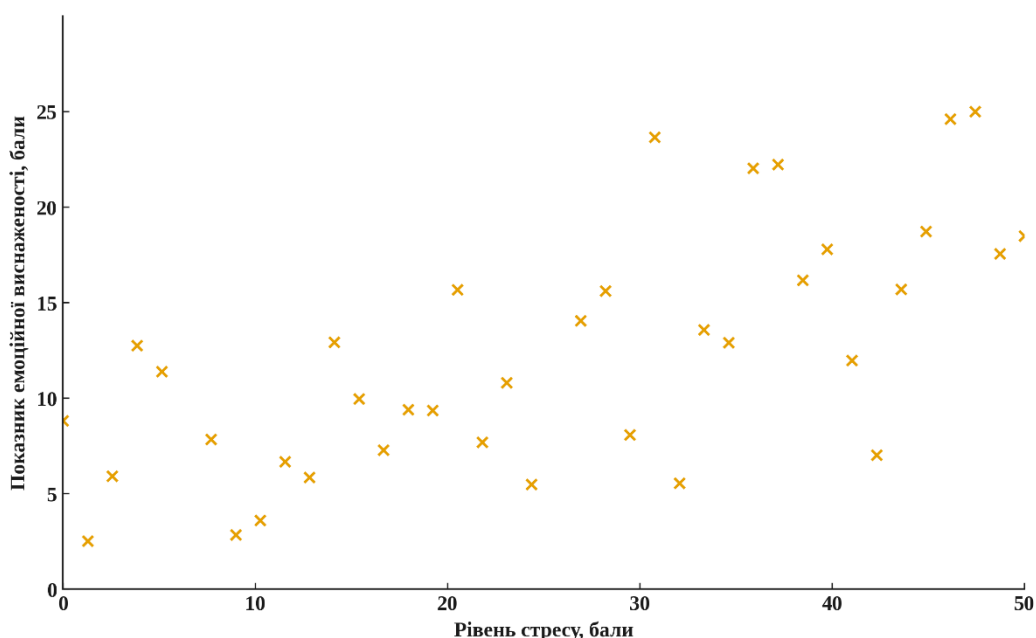
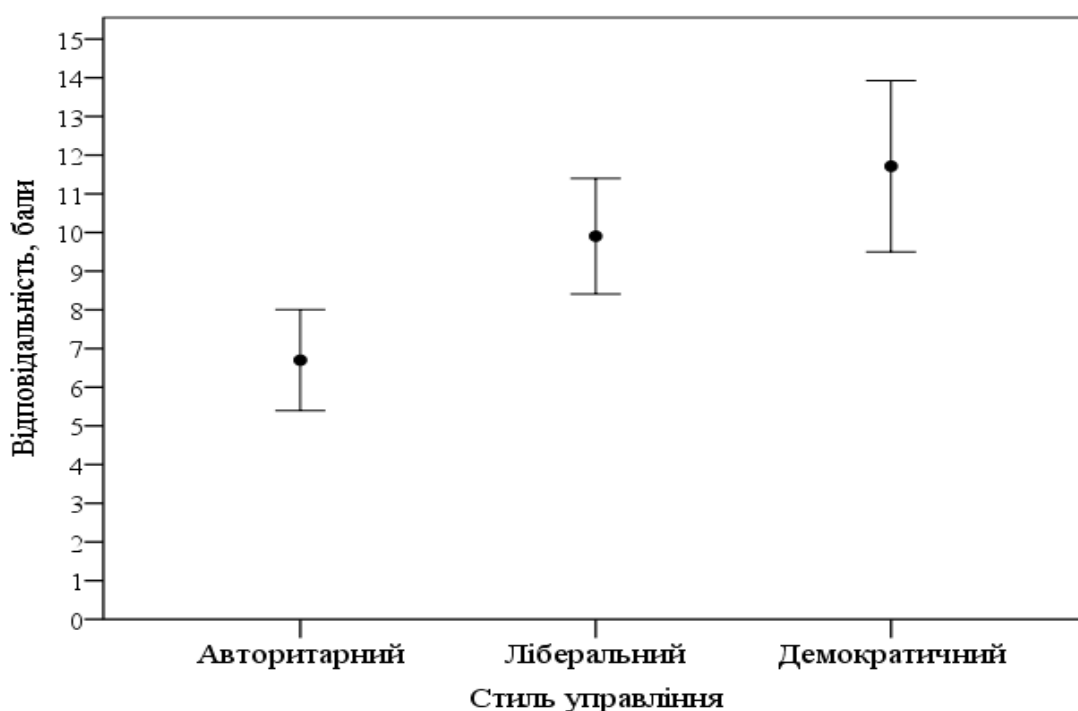


Рисунок 10.7 – Залежність емоційної виснаженості особистості від сприйняття рівня стресу (n = 40)

Гістограми – відображають частоту появи певних значень або інтервалів. Вони особливо корисні для оцінки розподілу даних – чи є він

нормальним, симетричним, чи має скошення. Це важливо для вибору адекватних статистичних тестів, оскільки багато з них базуються на припущенні нормальності розподілу.

Графіки середніх із довірчими інтервалами (mean plots with confidence intervals) – сучасний стандарт у наукових публікаціях. Традиційні стовпчасті діаграми середніх давно зазнали критики за те, що вони приховують варіативність даних і вводять читача в оману відносно точності оцінок. На рис. 10.8. наведено приклад графіка, який демонструє середні значення трьох груп разом із 95% довірчими інтервалами. Такий формат подання суттєво підвищує інформативність даних, демонструючи не лише середні значення, але й варіативність результатів. Він дозволяє швидко оцінити, чи існує перекриття довірчих інтервалів між групами, що є важливою інформацією під час інтерпретації.



Стовпці помилок – 95% довірчий інтервал середнього значення

Рисунок 10.8 – Особливості особистісної відповідальності у працівників організацій з різним стилем керівництва ($n_1 = 35$; $n_2 = 45$; $n_3 = 37$)

Таким чином, графічне представлення даних повинно ґрунтуватися на поєднанні математичної структури шкали та цілей аналізу. Кожен тип шкали вимагає відповідного типу графічного подання. Якщо порівнюється більше трьох змінних одночасно, краще використовувати багатопанельні графіки або інтерактивні інструменти (наприклад, у програмі JASP чи R), щоб уникнути перевантаження одного зображення. Дотримання цього принципу не лише підвищує наукову коректність результатів, а й робить їх зрозумілими для

аудиторії. Візуалізація, що враховує природу даних, є невід'ємною ознакою професійності дослідника та показником високої культури

10.4. Особливості представлення результатів дослідження в таблицях

Таблиці в науковому дослідженні посідають особливе місце, оскільки саме вони забезпечують найбільш структурований, логічний і точний спосіб подання числової інформації, який дозволяє уникнути неоднозначності, властивої суто текстовому опису або навіть окремим видам графічних візуалізацій. На відміну від графіків, що фокусуються передусім на демонстрації тенденцій, співвідношень або загальних закономірностей, таблиці виконують іншу, не менш важливу функцію – вибудовують фактологічну основу дослідження, фіксуючи точні числові значення, що лягають в основу статистичних висновків. У цьому полягає їхня наукова цінність: таблиця не лише відображає результати, а й засвідчує суворість проведених обчислень, дає змогу перевірити аналітичні процедури і, за необхідності, повторити їх. Саме в точності й структурованості таблиці полягає її сила, адже вона дозволяє уникнути спрощень або довільних інтерпретацій, які нерідко виникають під час побудови графіків.

Потреба у табличному поданні даних особливо зростає в тих випадках, коли дослідник працює зі складними структурами даних, багатовимірними моделями або багатофакторними експериментальними дизайнами. Хоча графік може ефектно підкреслити загальний тренд, лише таблиця дозволяє відтворити повну числову картину. У психології такі ситуації трапляються надзвичайно часто, адже ця наука оперує як кількісними, так і якісними змінними, а також різноманітними показниками, що описують поведінкові, емоційні, когнітивні та мотиваційні характеристики. У разі проведення дисперсійного аналізу, регресійного моделювання, факторного аналізу або будь-якого іншого комплексного методу саме таблиці забезпечують можливість компактного та водночас детального подання результатів, без якого науковий текст перетворився б на громіздкий, перевантажений послідовний опис чисел, що значно ускладнює його сприйняття.

Водночас важливо розуміти, що якість таблиці є не технічною подробицею оформлення, а показником загальної дослідницької культури автора. Наукова таблиця не може бути довільною композицією чисел або вставкою, що виконує лише декоративну функцію. Вона має бути максимально змістовною, завершеною та такою, що дозволяє читачеві самостійно зрозуміти представлену інформацію без постійного звернення до основного тексту. Тому вимоги до правильного оформлення таблиць у психологічних дослідженнях є надзвичайно високими. Одна з найпоширеніших помилок полягає у відсутності змістовного заголовка. Таблиця без чітко сформульованої назви перетворюється на масив даних без чіткої інтерпретації, а читач змушений самостійно здогадуватися, що саме

автор намагався продемонструвати. Наукова культура вимагає, щоб заголовок відображав ключову ідею таблиці, пояснював, що саме вимірювалось, у яких умовах, у яких групах або за якими методами проведено аналіз.

Не менш серйозною проблемою є відсутність підписів до стовпців і рядків. Таблиця, в якій наведено числові значення без позначення змінних, одиниць вимірювання або груп, фактично позбавляється сенсу й не може виконувати свою функцію. Читач повинен без утруднень і без додаткових пояснень у тексті зрозуміти, що означає кожен рядок і кожен стовпець. Завдання таблиці – бути самодостатнім інформаційним блоком у структурі наукової праці. Якщо для її розуміння потрібно звертатися до кількох попередніх розділів або здогадуватися про значення символів, це свідчить про суттєві помилки оформлення та порушення принципів наукової доброчесності.

Надзвичайно важливою є також проблема логічного структурування інформації. Досить часто автори, особливо початківці, намагаються розмістити якомога більше числових значень в одній таблиці, не замислюючись про когнітивне навантаження на читача. У результаті таблиця стає не лише важкою для сприйняття, але й фактично втрачає свою функцію, оскільки перетворюється на хаотичний масив інформації, який складно інтерпретувати. Науковий стиль подання вимагає логічного впорядкування: групування зразків за змінними, факторами, групами респондентів, рівнями значущості або шкалами вимірювання. Якщо структура таблиці не відображає логіку дослідження, вона буде неефективною, незалежно від кількості даних.

У наукових текстах нерідко спостерігається ще одна помилка – надмірне використання скорочень, аббревіатур або символів, які автор не пояснює. Для досвідченого дослідника, що регулярно працює з певним комплексом статистичних позначень, такі маркери можуть бути звичними, але для широкої аудиторії, до якої належить більшість читачів психологічних праць, вони можуть залишатися незрозумілими. Коли у таблиці наведено, наприклад, M, SD, SE, p, η^2 , β і жодного пояснення не подано, інтерпретація перетворюється на інтелектуальну гру в здогадування. Примітки внизу таблиці – важлива частина її структури і слугують гарантією того, що вона буде однаково зрозумілою для всіх читачів.

Правильне форматування таблиці також є ключовим параметром її наукової якості. Таблиця повинна бути лаконічною, симетрично структурованою, з рівномірними відступами, вирівняними стовпцями та однорідним шрифтом. Усі числові дані мають бути вирівняні. Недопустимим є використання надмірної декорації, кольорових заливок, декоративних рамок або тіней – усе це суперечить принципам академічного оформлення й може створити враження непрофесійності.

Важливим аспектом подання таблиць у психологічному дослідженні є їх взаємозв'язок з основним текстом. Таблиця не може існувати окремо, як декоративний елемент. Вона має бути інтегрована у структуру наукової аргументації, а текст повинен містити пряме посилання на неї, пояснюючи, що саме в ній представлено та які висновки можна зробити. Недостатньо просто

зазначити «див. Таблицю 1» – потрібно описати, які закономірності підтверджуються цифрами, які тенденції виявляються, чи підтримують дані основну гіпотезу або висувають нові інтерпретації. Таблиця має працювати разом із текстом, утворюючи єдину логічну систему, що забезпечує переконливість, прозорість та відтворюваність дослідження.

Не менш важливим є розуміння того, що таблиці та графіки виконують різні функції і не можуть повністю замінювати одна одну. Таблиця доречна тоді, коли потрібні точні числові значення, детальний аналіз або багатовимірне представлення даних. Графік, навпаки, акцентує увагу на загальних трендах, динаміці, співвідношеннях між групами або візуальних контрастах. У професійній психологічній літературі обидва способи подання даних використовуються разом: таблиці забезпечують фактологічну базу, а графіки – інтуїтивну інтерпретацію. Ігнорування будь-якого з цих інструментів призводить до збіднення дослідження та обмеження його аналітичних можливостей.

У підсумку таблиця в психологічному дослідженні є не просто формою подання даних, а індикатором професійної компетентності автора, рівня його наукової культури та уваги до принципів репрезентативності, точності та прозорості. Правильно оформлена таблиця не лише фіксує результати, а й дозволяє іншим дослідникам відтворити аналіз, перевірити статистичні обчислення або порівняти дані з власними результатами. Уміння створювати і грамотно використовувати таблиці одна з фундаментальних компетентностей науковця, яка визначає якість дослідження, його достовірність і внесок у розвиток психологічної науки.

Нижче представлені типові приклади оформлення таблиць згідно вимог Американської психологічної асоціації (APA).

Таблиця 10.2 – Суб'єктивна оцінка стрес-факторів пандемії COVID-19

Стрес-фактори	Градації відповідей	Всього (n = 115)	Суб'єктивна оцінка впливу пандемії на життя			
			вплинула (n = 85)	не вплинула (n = 30)	χ^2	V
Тривога зараження вірусом	сильно переживаю	26 (22,6%)	21 (24,7%)	5 (16,6%)	8,10*	0,13
	помірно переживаю	46 (40,0%)	35 (41,2%)	11 (36,7%)		
	не переживаю	43 (37,4%)	29 (34,1%)	14 (46,7%)		
Тривога економічних наслідків	сильно переживаю	45 (39,1%)	36 (42,4%)	9 (30,0%)	11,70**	0,24
	помірно переживаю	30 (26,1%)	23 (27,1%)	7 (23,3%)		
	не переживаю	40 (34,8%)	26 (30,5%)	14 (46,7%)		

Примітка: * p < 0,05; ** p < 0,001; V – індекс розміру ефекту V-Крамера.

Таблиця 10.3 – Відмінності в прояві життєстійкості рятувальників з різним рівнем професійної успішності ($n_1 = 54$; $n_2 = 42$)

Змінні	Успішні рятувальники M (SD)	Відносно успішні рятувальники M (SD)	t	d-Коена
Життєстійкість	48,78 (11,35)	41,30 (11,01)	2,03*	0,67
Залученість	26,12 (5,81)	17,10 (4,74)	4,53***	1,70
Контроль	18,72 (4,53)	14,09 (4,78)	1,99*	0,99
Прийняття ризику	9,18 (2,12)	9,90 (2,24)	-0,71	0,33

Примітка: M – середнє значення; SD – стандартне відхилення; * $p < 0,05$, *** $p < 0,001$; d-Коена – індекс розміру ефекту.

Таблиця 10.4 – Особливості стильової саморегуляції рятувальників з різним рівнем стресостійкості ($n_1 = 33$; $n_2 = 37$; $n_3 = 45$)

Регуляторні шкали	Рівень стресостійкості			F	p	η^2
	Низький M (SD)	Середній M (SD)	Високий M (SD)			
Планування	7,23 (1,14)	2,52 (0,17)	2,49 (0,21)	8,361	0,000	0,052
Моделювання	2,39 (1,05)	8,14 (1,54)	8,08 (1,43)	10,941	0,000	0,071
Програмування	8,95 (2,06)	3,18 (0,11)	8,40 (1,95)	0,974	0,480	0,003
Гнучкість	3,29 (0,22)	6,10 (1,51)	5,31 (1,83)	1,344	0,124	0,008
Самостійність	2,20 (0,61)	6,17 (1,98)	7,55 (2,03)	5,671	0,000	0,045
Загальна CP	28,10 (5,15)	30,28 (6,67)	34,61 (5,54)	2,736	0,273	0,007

Примітка: M – середнє значення; SD – стандартне відхилення; F – статистика дисперсійного аналізу Фішера; p – статистична значущість; η^2 – розмір ефекту.

Таблиця 10.5 – Описова статистика та кореляція між суб'єктивним сприйняттям стресу, психологічним капіталом та задоволеністю життям особистістю ($n = 160$)

Змінні	1	2	3	M	SD	S	K
1. Суб'єктивне сприйняття стресу	1,00			26,73	6,24	0,04	-0,19
2. Психологічний капітал	-0,54***	1,00		100,91	21,24	-0,21	-0,51
3. Задоволеність життям	-0,34***	0,32***	1,00	22,61	5,13	-0,13	-0,39

Примітка: M – середнє значення; SD – стандартне відхилення; S – асиметрія; K – ексцес; *** – кореляція статистично значуща на рівні 0,001.

Завдання на самостійну підготовку

1. Чому візуалізація даних у науковому дослідженні, зокрема психологічному, не може розглядатися як суто «декоративний» елемент?
2. Які обмеження мають описові статистики без візуального подання розподілу даних і до яких помилок це може призвести?
3. Як приклад квартету Енскомба демонструє небезпеку інтерпретації результатів лише за середніми, дисперсіями та кореляціями без графіків?
4. У чому полягає епістемологічна цінність графіків розсіювання (scatter plots) для аналізу структури даних у психології?
5. Яким чином візуалізація сприяє виявленню порушень припущень статистичних моделей (лінійність, нормальність, гомогенність дисперсій)?
6. Які основні функції візуалізації в комунікації результатів дослідження для різних аудиторій (науковці, практики, студенти, управлінці)?
7. Чому стереотипні стовпчасті діаграми середніх без відображення варіативності вважаються недостатніми для сучасної наукової звітності?
8. Які потенційні маніпуляції можливі при некоректному використанні візуалізації (обрізані шкали, 3D-ефекти, вибіркове відображення даних)?
9. У чому полягає сутність принципу «максимум інформації при мінімумі чорнила» і як його застосувати до психолого-статистичної графіки?
10. Які обов'язкові елементи повинен містити коректно оформлений науковий графік (підписи осей, одиниці вимірювання, легенда тощо)?
11. У чому полягає небезпека неправильного вибору шкали на осях (наприклад, обрізка осі Y не з нуля) для інтерпретації психологічних результатів?
12. Як тип шкали вимірювання (номінальна, порядкова, інтервальна, відношень) визначає вибір оптимального типу графіка?
13. Які особливості має візуалізація номінальних змінних і чому стовпчасті діаграми частіше є кращим вибором, ніж секторні?
14. Як доцільно візуалізувати порядкові змінні, щоб не створювати ілюзії рівності інтервалів між рівнями шкали?
15. Які переваги мають боксплоти у порівнянні зі стовпчастими діаграмами при представленні результатів психологічних вимірювань?
16. У яких випадках лінійні графіки, гістограми та графіки розсіювання є найадекватнішими інструментами для візуалізації інтервальних/шкал відношень?
17. Чому графіки середніх із довірчими інтервалами вважаються більш інформативними, ніж «класичні» стовпчасті діаграми середніх?
18. Які головні принципи якісного табличного представлення даних у психологічному дослідженні (заголовок, підписи змінних, примітки, логічна структура)?
19. У чому відмінність функцій таблиць і графіків у науковому тексті та чому вони мають доповнювати, а не замінювати одне одного?
20. Які типові помилки оформлення таблиць (надмір даних, неструктурованість, незрозумілі скорочення, відсутність одиниць вимірювання) знижують прозорість дослідження?

СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Бедан В.Б. Експериментальна психологія: курс лекцій (для студентів психологів факультету психології, політології та соціології). Одеса: ОЮА, 2020. 102 с.
2. Боснюк В.Ф. Експериментальна психологія: курс лекцій. Харків: НУЦЗУ, 2018. 192 с.
3. Боснюк В.Ф. Математичні методи в психології: курс лекцій. Мультимедійне навчальне видання. Х.: НУЦЗУ, 2020. 141 с.
4. Важинський С.Е., Щербак Методика та організація наукових досліджень: навч. посіб. Суми: СумДПУ імені А. С. Макаренка, 2016. 260 с.
5. Васюк К.М., Шпортун О.М. Експериментальна психологія: теоретичні та практичні аспекти: навчальний посібник. Вінниця, 2020. 372 с.
6. Горбунова В.В. Експериментальна психологія в схемах і таблицях: Навчальний посібник. К.: «ВД «Професіонал», 2017. 208 с.
7. Грибенюк Г. С. Експериментальна психології: курс лекцій. Черкаси, 2017. 261 с.
8. Колтунович Т.А., Чуйко Г.В., Чаплак Я.В. Соціально-психологічні експерименти: навч. посібник. Чернівці: Чернівець. нац. ун-т ім. Ю. Федьковича, 2022. 192 с.
9. Максименко С.Д., Носенко Е.Л. Експериментальна психологія. К.: Центр учбової літератури, 2021. 360 с.
10. Малхазов О.Р. Експериментальна психологія: практичний посібник. К.: Талком, 2020. 321 с.
11. Романенко О.В. Експериментальна психологія. Київ: Нац. акад. внутр. справ, 2017. 146 с.
12. Савченко О.В. Експериментальна психологія: методичні рекомендації щодо проведення практичних робіт. Херсон: ПП Вишемирський В.С., 2012. 212 с.
13. Сімко Р.Т. Експериментальна психологія: навчально-методичний посібник. Кам'янець-Подільський: Видавець Ковальчук О.В., 2021. 198 с.
14. Туриніна О.Л. Методологія та методи психологічного дослідження: навч.-метод. посіб. Київ: ДП «Вид. дім «Персонал», 2018. 206 с.
15. Фесун Г.С., Радчук В. М., Канівець Т. М. Методи психологічних досліджень: навч.-метод. посібник. Чернівці: Чернівецький національний університет, 2018. 192 с.
16. Шатирко Л.О. Експериментальна традиція психології навчання в Україні. Київ: Видавничий Дім «Слово», 2017. 324 с.
17. American Psychological Association. Publication manual of the American Psychological Association: 7-th ed., Waschinton, DC: Author, 2020. 428 p.
18. Morling B. Research Methods in Psychology: Evaluating a World of Information. 4th ed. New York: W. W. Norton, 2021. 663 p.

Навчальне видання

**Боснюк Валерій Федорович
Грибенюк Геннадій Сергійович**

ЕКСПЕРИМЕНТАЛЬНА ПСИХОЛОГІЯ

Навчальний посібник

Формат 60x84 1/16. Ум.друк.арк. 16,77. Тираж ____ пр. Зам. № ____.
Національний університет цивільного захисту України.
18034, м. Черкаси, вул. Онопрієнка, 8.