

ДОСЛІДЖЕННЯ В СОЦІАЛЬНИХ НАУКАХ: ТЕОРІЯ І ПРАКТИКА

А. Бхатгачерджи, Університет Південної Флориди,
abhatt@usf.edu

Н. І. Ситник, Національний технічний університет України
«Київський політехнічний інститут»
natalia_sytnik@ukr.net

**ДОСЛІДЖЕННЯ В СОЦІАЛЬНИХ НАУКАХ:
ТЕОРІЯ І ПРАКТИКА**

АНОЛЬ БХАТТАЧЕРДЖИ

НАТАЛІЯ СИТНИК

**ДОСЛІДЖЕННЯ В СОЦІАЛЬНИХ НАУКАХ:
ТЕОРІЯ І ПРАКТИКА**

А. Бхатгачерджи
Університет Південної Флориди
Тампа Флоріда, США
abhatter@usf.edu

Н. І. Ситник
Національний технічний університет України
«Київський політехнічний інститут»
Київ, Україна
natalia_sytnik@ukr.net

Copyright ©

Опубліковано за ліцензією Creative Commons Attribution-NonCommercial-
ShareAlike 3.0 Unported License

Дослідження в соціальних науках: теорія і практика

А. Бхаттачарджи, Н.І. Ситник

ISBN

ISBN

Опубліковано за ліцензією [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#), відповідно до якої користувачі можуть використовувати, копіювати, передавати іншим, розповсюджувати і посилатися на цю книгу за таких умов:

- посилання на всю книгу чи її частини робляться згідно зі стандартною академічною практикою,
- книга не може бути використана в комерційних цілях,
- користувачі можуть змінити, додати чи модифікувати частини книги, однак розповсюджувати її за такою ж, або схожою ліцензією.

При будь-якому використанні чи поширенні книги умови ліцензії мають бути чітко визначені.

ПЕРЕДМОВА

Наукові дослідження – це потужне джерело нових знань, які трансформуються в новітні технології та продукти наукоємного виробництва, прискорюють зростання продуктивності праці і темпів накопичення суспільного прибутку, сприяють підвищенню конкурентоспроможності суспільного виробництва.

Ця книга має на меті ознайомити студентів-магістрів, аспірантів та молодих науковців із загальноприйнятими на Заході науково-методологічними підходами до проведення наукових досліджень у галузі соціальних наук та сприяти формуванню їх дослідницької компетентності.

Перше та найбільш важливе питання, яке виникає в потенційних читачів цієї книги, полягає в тому, чим вона відрізняється від інших численних навчальних видань, присвячених науковим дослідженням. Існує щонайменше три ключові відмінності. По-перше, цей посібник розглядає не просто науково-дослідні методи, що застосовуються на етапі збору та аналізу емпіричних даних, а цілісний послідовний процес наукового дослідження від початкових до завершальних етапів. Наукова методологія є лише одним із елементів процесу дослідження, причому досить структурованим, а тому легким. Більшість навчальних видань приділяють багато уваги науковій методології, однак не зупиняються детально на найбільш складних, мало структурованих і, водночас, дуже важливих аспектах проведення наукових досліджень, таких як розроблення теорій і наукове мислення. Молоді науковці можуть відносно швидко оволодіти науковою методологією, однак стикаються зі складнощами при генеруванні нових цікавих дослідницьких питань чи побудові наукових теорій. Розділи «Наукове мислення» і «Теорії в наукових дослідженнях» спеціально присвячені формуванню цих складних вмінь, що є необхідними атрибутами емпіричних досліджень.

По-друге, книга є стислою та компактною за дизайном. Щоб не відволікати увагу читача на другорядні речі, у ній розглядаються лише найсуттєвіші концепції, важливі для глибокого та цілісного розуміння процесу наукового дослідження в соціальних науках. Книга написана таким чином, щоб у компактній формі представити сучасні підходи до проведення наукового дослідження, не примушуючи читача постійно звертатися до інших додаткових джерел, які можуть виявитися недоступними.

По-третє, електронна версія книги є безкоштовною. На авторське переконання, наукове знання не повинно обмежуватися такими бар'єрами, як ціна навчального видання та його доступність. Науковий прогрес можливий лише тоді, коли студенти й науковці всього світу мають вільний доступ до досягнень наукової думки, і ця безкоштовна

книга є внеском авторів у забезпечення загального доступу до знань у конкретній галузі – методології проведення наукових досліджень. Однак безкоштовність не означає низьку якість. Найкращі речі на Землі – повітря, вода і сонячне світло, – безкоштовні. Багато ресурсів Google є також безкоштовними і зараз навіть складно уявити життя сучасної людини в добу Інтернету без Google. Деякі дуже складні програмні продукти, доступні сьогодні широкому колу користувачів, такі як Linux і Apache, є також безкоштовними, що не знижує їх цінності.

Книга спрямована на оволодіння теоретико-методологічними основами організації та проведення наукового дослідження. Матеріали книги викладені у 16 розділах, де розглядаються закономірності наукового пізнання, парадигми соціальних наук, сутність емпіричних методів дослідження, методологія розроблення та валідизації конструктів, висвітлені сутність, принципи організації, практичні інструменти та проблемні питання процесу виконання наукового дослідження.

ЗМІСТ

Вступ до дисципліни

1. Наука та наукові дослідження.....	1
2. Наукове мислення.....	12
3. Процес наукового дослідження.....	21
4. Теорії в наукових дослідженнях	31

Основи емпіричних досліджень

5. Дизайн дослідження	44
6. Вимірювання конструктів.....	53
7. Надійність і валідність шкал	68
8. Формування вибірок.....	80

Збір даних

9. Анкетне дослідження	89
10. Експериментальне дослідження.....	100
11. Кейс-дослідження.....	112
12. Інтерпретативне дослідження	124

Аналіз даних

13. Якісний аналіз.....	135
14. Кількісний аналіз: описова статистика.....	141
15. Кількісний аналіз: пояснювальна статистика	149
16. Етика наукових досліджень.....	157

Розділ 1

НАУКА ТА НАУКОВІ ДОСЛІДЖЕННЯ

Що таке дослідження? Залежно від того, хто ставить це начебто просте питання, можна отримати дуже різні відповіді. Одні скажуть, що вони зазвичай досліджують різні он-лайнні сайти, щоб знайти найкраще місце для купівлі товарів чи послуг, які їм потрібні. Канали телевізійних новин начебто проводять дослідження у формі обміну думок стосовно важливих для громадськості тем, таких як наступні вибори чи фінансовані урядом проекти. Студенти, що навчаються в університетах, проводять дослідження в Інтернеті, щоб знайти інформацію, яка їм потрібна для виконання наукових проектів або курсових робіт. Аспіранти, виконуючи дослідження разом зі своїм науковим керівником, можуть їх розглядати як збір та аналіз даних за проектом. Представники бізнесу та бізнес-консультанти досліджують потенційно можливі заходи для вирішення організаційних проблем, таких як «вузькі місця» ланцюга постачання чи ідентифікація трендів покупок споживачами. Однак наведені вище заходи не можуть вважатися науковими дослідженнями допоки: 1) вони не роблять внесок у науку; 2) не використовують наукові методи. Цей розділ досліджує, що саме означають ці терміни.

Наука

Що таке наука? Для одних людей наука асоціюється зі складними шкільними чи університетськими дисциплінами, такими як фізика, хімія чи біологія, що їх можуть опанувати лише дуже здібні учні. Для інших це сфера діяльності науковців, що використовують спеціальне обладнання у своїх лабораторіях. За своєю етимологією (походженням), слово «наука» виникло від латинського *scientia*, що означає «знання». **Наука** – це систематизований масив знань в певній сфері, що здобувається завдяки використанню наукових методів (наукові методи описуються нижче). Окремі науки відносяться до однієї з двох великих категорій: природничі та соціальні науки. **Природничі науки** – це науки, що досліджують природні об'єкти чи феномени, такі як світло, речовини, природні явища, небесні тіла чи людський організм. Природничі науки поділяються на фізичні, геологічні, біологічні науки та інші. До складу фізичних наук входить фізика (наука про фізичні об'єкти), хімія (наука про речовини) та астрономія (наука про небесні тіла). Наука про природні явища складається з окремих дисциплін, таких як геологія (наука про земні надра). Біологічні науки включають, наприклад, такі дисципліни як зоологія (наука про тварин) і ботаніка (наука про рослини). На противагу цьому, **соціальні науки** – це науки про людей та їх об'єднання, такі як групи, організації, суспільства, економіку, і також про їх індивідуальну та спільну поведінку. Соціальні науки можуть бути класифіковані на такі

дисципліни, як психологія (наука про людську поведінку), соціологія (наука про соціальні групи) чи економіка (наука про організації, ринки та національні економіки).

Природничі науки відрізняються від соціальних наук за певними ознаками. Природничі науки є точними, детермінованими й об'єктивними, тобто отримані ними результати не залежать від особи, що проводить наукове дослідження. Наприклад, наукові експерименти у фізиці, такі як вимірювання швидкості звуку в різних речовинах, або індексу рефракції води призведуть до однакових результатів незалежно від часу, місця проведення експерименту та особи дослідника. Якщо два студенти, виконуючи однаковий фізичний експеримент, отримають різні значення досліджуваних фізичних характеристик, це в більшості випадків означатиме, що один або обидва з них припустилися помилки. Однак такий висновок не можна зробити, коли йдеться про дослідження в соціальних науках, які загалом менш точні, менш детерміновані й більш суб'єктивні. Наприклад, якщо вимірюється рівень задоволеності людини своєю працею, використовуючи певний метод, то можна з'ясувати, що один і той же досліджуваний буде відчувати себе більш чи менш задоволеним (чи менш щасливим) у різні дні й навіть у різні години одного й того ж дня. Задоволеність людини може значно змінюватись залежно від поточної інформації, яку вона отримує, або залежно від подій, що сталися протягом робочого дня. Більш того, не існує єдиної методики, що дозволяє точно оцінити задоволеність людини. Отже, одна методика оцінить досліджувану особу як «задоволену», тоді як інша методика може показати, що досліджувана особа є «не зовсім задоволеною». Іншими словами, у соціальних науках існує висока ймовірність помилки вимірювання, значна невизначеність та відсутність консенсусу стосовно політики прийняття рішень. Наприклад, у середовищі дослідників, що працюють у природничих науках, складно знайти розбіжності стосовно швидкості світла чи швидкості руху Землі навколо Сонця, однак, існують суттєві розбіжності серед дослідників, що працюють у соціальних науках, стосовно того, як зменшити рівень безробіття чи вивести економіку країни з рецесії. Кожному студенту, що вивчає соціальні науки, знайомий і звичний високий рівень суб'єктивності, невизначеності та суперечливості, характерний для дисциплін соціального блоку, що насправді є відображенням значної варіативності соціальних об'єктів.

Науки також можуть класифікуватися залежно від своїх цілей. **Фундаментальні науки**, які часто називають «чистими» науками, пояснюють найбільш фундаментальні об'єкти, явища, процеси, сили, взаємозв'язки між ними та закони, що визначають їхню поведінку. Прикладами таких дисциплін є фізика, математика чи біологія. **Прикладні науки**, які також називають практично орієнтованими, використовують наукові знання, отримані фундаментальними науками, у

реальному житті. Наприклад, інженерна справа – це прикладна наука, що спирається на закони фізики та хімії для вирішення практичних завдань, таких як будівництво міцніших мостів чи ефективніших двигунів, тоді як медицина є прикладною наукою, яка використовує закони біології для лікування та збереження здоров'я людини. Як фундаментальні, так і прикладні науки важливі для розвитку суспільства. Однак прикладні науки не можуть у своєму розвитку покладатися виключно на власні досягнення – вони спираються на здобутки фундаментальних дисциплін. Звичайно, промисловість та приватні підприємства більше зацікавлені в прикладних дослідженнях, враховуючи їхню практичну спрямованість, хоча університетська наука проводить як фундаментальні, так і прикладні дослідження.

Наукове знання

Метою науки є створення наукових знань. **Наукове знання** – це сукупність законів і теорій, що пояснюють певний феномен (подію, явище) чи процес, які є об'єктами дослідження, і здобувається завдяки застосуванню наукових методів. **Закони** – це певні сталі й повторювані співвідношення між подіями, явищами чи процесами. **Теорії** можуть бути визначені як системні пояснення досліджуваних феноменів чи процесів. Наприклад, у фізиці закони руху Ньютона описують, що відбувається, коли об'єкт перебуває у стані спокою чи рухається (перший закон Ньютона), які сили потрібні, щоб привести в рух нерухомий об'єкт (другий закон Ньютона), і що трапиться, коли два об'єкта зіткнуться (третій закон Ньютона). Спільно три закони складають базу класичної механіки – теорію об'єктів, що рухаються. Аналогічно теорія оптики пояснює властивості світла й те, як воно поводить себе в різних середовищах, електромагнітна теорія пояснює властивості електричних явищ, а квантова механіка – властивості мікрочастинок. Подібні теорії є і в соціальних науках. Наприклад, теорія когнітивного дисонансу в психології пояснює як люди реагують, коли їхні спостереження за певними подіями чи явищами розходяться з тим, що вони очікували побачити, теорія загального покарання пояснює, чому деякі люди скоюють злочини, використовуючи не ліцензовані версії програмного забезпечення, теорія запланованої поведінки пояснює, як люди у своєму повсякденному житті роблять свідомий вибір із декількох альтернатив і т.п.

Мета наукових досліджень полягає в тому, щоб встановлювати закони та формулювати теорії, які можуть пояснити природні й соціальні феномени, або, іншими словами, здобувати наукові знання. Важливо розуміти, що таке знання може бути недосконалим і навіть далеким від істини. Іноді не існує єдиної універсальної істини, а є лише певний баланс «багатьох істин». Теорії, на яких базується наукове знання, є тільки поясненням окремого феномену, що запропонований вченими. Отже, можуть існувати кращі та гірші пояснення, залежно від того, наскільки

добре вони відповідають реаліям і, відповідно, можуть бути кращі та гірші теорії. Прогрес науки відзначається рухом у часі від недосконалих теорій до більш досконалих у результаті ретельнішого спостереження, використання більш точного інструментарію та більш інформативних логічних суджень.

Науковці встановлюють наукові закони і створюють нові теорії завдяки процесу логічного мислення та доказів. Логіка (теорія) та практичні докази (спостереження) є тими двома стовпами, на яких базується наукове знання. В науці теорія і спостереження є взаємопов'язаними й не можуть існувати одне без одного. Теорії надають сенс і значення тому, що спостерігається, тоді як спостереження допомагають підтвердити чи уточнити існуючу теорію, або сконструювати нову. Будь-які інші засоби отримання знання (наприклад, релігія чи міфологія) не можуть вважатися наукою.

Наукове дослідження

Вважаючи, що теорії та спостереження є двома стовпами науки, наукове дослідження може виконуватися на двох рівнях: теоретичному та емпіричному. Завданнями теоретичного рівня дослідження є розвиток абстрактних концепцій стосовно природних чи соціальних подій чи явищ, а також встановлення зв'язків між цими концепціями, тобто, створення теорій. Завданнями емпіричного рівня дослідження є тестування теоретичних концепцій і зв'язків, щоб з'ясувати, наскільки добре вони відображають реалії навколишньої дійсності, відомі завдяки практиці, і подальше удосконалення на цій основі існуючих теорій. Із часом теорія стає все більш і більш точною (тобто все краще відповідає реаліям, які спостерігаються), і наука досягає зрілості. Наукове дослідження означає постійний рух уперед і назад між теорією та спостереженням. Отже, якщо покладатися виключно на спостереження при поясненні причин якогось явища й ігнорувати теорії, не можна розраховувати на проведення валідного наукового дослідження.

Залежно від кваліфікації та інтересів науковця, наукове дослідження може приймати одну з двох форм: індуктивне чи дедуктивне. В **індуктивному дослідженні** метою науковця є побудова теоретичних концепцій і виявлення закономірностей, спираючись на дані спостереження. У **дедуктивному дослідженні** метою науковця є тестування концепцій і виявлених закономірностей, відомих із теорії, спираючись при цьому на нові емпіричні дані. Отже, індуктивне дослідження також правомірно назвати дослідженням із *побудови теорії*, тоді як дедуктивне дослідження є дослідженням із *тестування теорії*. Зверніть увагу, що метою тестування теорії є не тільки тестування, але й можливе уточнення, розвиток чи удосконалення теорії. Рис 1.1. ілюструє комплементарну природу індуктивного та дедуктивного досліджень. Відзначимо, що індуктивне та дедуктивне дослідження є двома

половинами дослідницького циклу, що постійно повторюється між теорією та спостереженнями. Дослідник не може виконувати індуктивне чи дедуктивне дослідження, якщо він не обізнаний як з теоретичними, так і з емпіричними компонентами дослідження. Зрозуміло, що зрілий науковець – це той, хто може охопити весь цикл дослідження і може виконувати як індуктивне, так і дедуктивне дослідження.

Важливо розуміти, що побудова теорії (індуктивне дослідження) і тестування теорії (дедуктивне дослідження) є необхідними для розвитку науки. Елегантні теорії нічого не варті, якщо вони не відповідають дійсності. Аналогічно, великі масиви даних будуть не потрібні, якщо вони не допомагають у конструюванні змістовних теорій. Замість того, щоб розглядати ці два процеси в круговому зв'язку, як показано на рис.1.1., більш адекватним є їх відображення у вигляді спіралі. При цьому кожна нова ітерація між теорією та емпіричними даними поглиблює й уточнює існуюче пояснення досліджуваного феномену та «шліфує» теорію. Хоча як індуктивне, так і дедуктивне дослідження важливі для прогресу науки, здається, що індуктивне дослідження (побудова теорій) є ціннішим, коли існує небагато теорій чи пояснень, тоді як дедуктивне дослідження (тестування гіпотез) є більш продуктивним, коли існує багато конкуруючих теорій стосовно одного й того ж феномену та необхідно дізнатися, яка з них краще «працює» й за яких обставин.

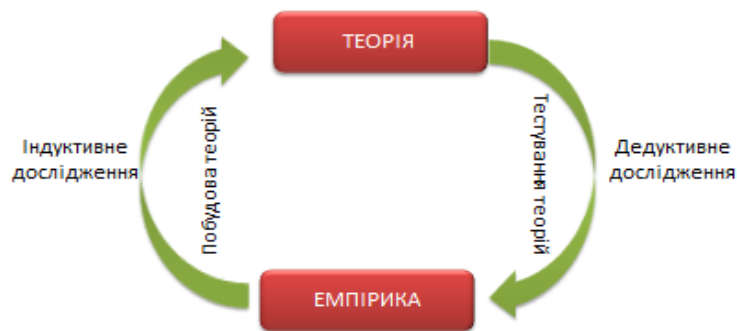


Рис. 1.1. Цикл наукового дослідження

Процес побудови й тестування гіпотез у соціальних науках ускладнюється невизначеною природою теоретичних концепцій, часто неадекватним інструментарієм, що використовується для збору даних, і наявністю багатьох факторів, які впливають на досліджуваний феномен і дію яких складно врахувати. У соціальних науках також досить складно виявити хибні теорії, що не працюють. На відміну від теорій природничих наук, теорії соціальних наук рідко бувають довершеними, що відкриває багато можливостей для дослідників щодо їх покращення і побудови альтернативних.

Отже, виконання наукового дослідження вимагає наявності двох вмінь – теоретичних і методологічних, – які необхідні для роботи на

теоретичному та емпіричному рівнях. Методологічні вміння («знати як») відносно стандартні, мало варіюють між дисциплінами і відносно легко набуваються в процесі виконання наукових проектів. Однак теоретичні вміння («знати що») значно важче набути, вони вимагають років спостережень та осмислення, є неявними, тобто такими, що не можуть бути засвоєні через навчання, а скоріше набуваються через досвід. Усі великі науковці в історії людства, такі як Галілео Галілей, Ісаак Ньютон, Альберт Ейнштейн, Нільс Бор, Адам Сміт і Чарльз Дарвін були великими теоретиками й увійшли в історію завдяки тому, що запропоновані ними теорії змінили поступ науки. Методологічні вміння потрібні, щоб стати звичайним дослідником, тоді як теоретичні вміння потрібні, щоб стати видатним дослідником!

Науковий метод

Вище був наведений опис науки як здобуття знань за допомогою наукового методу. Отже, що конкретно означає «науковий метод»? **Науковий метод** – це стандартизований набір технік для побудови наукового знання, таких як проведення валідних спостережень, інтерпретації результатів дослідження і формулювання висновків стосовно результатів. Науковий метод дозволяє дослідникам незалежно та цілісно тестувати існуючі теорії та результати досліджень і робити їх доступними для відкритих обговорень, модифікацій та удосконалень. Науковий метод повинен відповідати чотирьом характеристикам:

- *Відтворюваність*: інші дослідники повинні мати змогу незалежно відтворити або повторити дане наукове дослідження й отримати при цьому схожі, якщо не ідентичні результати.
- *Точність*: теоретичні концепти, які часто важко виміряти, повинні бути визначені настільки точно, щоб інші дослідники могли їх виміряти і протестувати теорію, яка на них будується.
- *Простота*: коли існують численні пояснення певного феномену, науковці повинні завжди приймати найпростіше й найбільш «економне» з точки зору логіки пояснення. Цей концепт називається простотою, або «бритвою Оккама». Ця характеристика утримує науковців від побудови занадто складних і зарозумілих теорій із нескінченним числом концептів і зв'язків, які можуть пояснити все потроху й нічого конкретно.
- *Можливість спростування*: теорія повинна формулюватися в такий спосіб, щоб її можна було спростувати. Теорії, які не можуть бути перевірені чи спростовані, не є науковими теоріями. Теорія, що викладена з використанням неточних термінів і концепти якої неможливо виміряти, не є науковою теорією. Наприклад, ідеї Зігмунда Фрейда щодо психоаналізу потрапляють у цю категорію і, відповідно, не можуть вважатися теорією, хоча психоаналіз може

мати практичне застосування в лікуванні деяких типів захворювань.

Будь-яка галузь досліджень, яка не дозволяє за допомогою наукового методу протестувати свої базові закони чи теорії, не може називатися наукою. Наприклад, теологія (дослідження релігії) не є наукою, тому що теологічні ідеї (такі, як існування Бога) не можуть бути протестовані незалежними спостерігачами, використовуючи відтворений, точний і економний метод, який можна спростувати. Аналогічно, мистецтво, музика, література чи юриспруденція не є науками, хоча вони є творчими видами діяльності.

Науковий метод по відношенню до соціальних наук включає численні наукові підходи та техніки, такі як якісні і кількісні дані, статистичний аналіз, експерименти, виробничі дослідження, кейси і т.п. Значна частина посібника присвячена вивченню цих досить складних методів. Однак потрібно розуміти, що науковий метод діє, головним чином, на емпіричному рівні досліджень, тобто стосується того як проводити спостереження, аналізувати та інтерпретувати отримані результати. Лише незначною мірою науковий метод безпосередньо застосовується на теоретичному рівні, який є найбільш складною частиною наукового дослідження.

Типи наукових досліджень

Залежно від цілей дослідження наукові проекти можуть бути віднесені до однієї з трьох груп: пошукові, описові та пояснювальні. **Пошукові дослідження** часто виконуються в нових напрямках досліджень, де їхніми цілями є: 1) оцінити значення чи ступінь розвитку певного феномену, проблеми чи процесу, 2) запропонувати деякі початкові ідеї стосовно феномену, 3) протестувати необхідність проведення більш глибокого дослідження стосовно феномену. Наприклад, громадяни країни зазвичай незадоволені політикою уряду в період економічного спаду. Пошукове дослідження може бути спрямоване на вимірювання рівня незадоволеності громадян, з'ясування того, у яких показниках це незадоволення може бути оцінене (наприклад, частотою публічних протестних акцій), і можливі причини незадоволення, наприклад неефективна політика уряду в подоланні інфляції, формуванні відсоткових ставок, боротьбі з безробіттям або високими податками. Таке дослідження може включати вивчення інформації з офіційних джерел (оцінка економічних індикаторів розвитку, таких як ВВП, рівень безробіття, індекс споживчих цін і т.п.), дані з інших джерел інформації, отримані шляхом інтерв'ювання експертів, авторитетних економістів або офіційних представників уряду і/або вивчення історичного досвіду стосовно того, як схожі проблеми вирішувалися в минулому. Таке дослідження не дасть дуже точного розуміння проблеми, але може бути корисним у встановленні природи й

ступеня прояву проблеми та слугувати початковим етапом, що передує більш глибокому та комплексному дослідженню.

Описові дослідження спрямовані на проведення ретельних спостережень і детальне документування досліджуваного феномену. Такі спостереження повинні базуватися на наукових методах (що відповідають чотирьом критеріям науковості) і, відповідно, бути більш надійними, ніж звичайне спостереження невідготовлених осіб. Прикладами описових досліджень є складання хронологічних таблиць для оцінки рівня зайнятості за секторами економіки або регіонами за даними державної служби зайнятості. Якщо в методиці розрахунків робляться якісь зміни, оцінки надаються й зі змінами, й без, щоб читач зміг провести порівняння «до/після» відносно наведених статистичних показників.

Пояснювальне дослідження проводиться для того, щоб пояснити досліджуваний феномен, проблему чи процес. Якщо описове дослідження встановлює факти стосовно певного феномену (відповідаючи на питання «що?», «де?», «коли?»), пояснювальне дослідження намагається знайти відповіді на питання «чому?» і «як?». Воно спрямоване на ідентифікацію причин і наслідків даного феномену. Більшість академічних досліджень належать саме до категорії пояснювальних, хоча елементи пошукових та описових досліджень можуть також знадобитися на початкових етапах академічного дослідження. Пошук пояснень досліджуваних подій чи явищ вимагає від дослідника розвинутих навичок теоретизації та інтерпретації, а також інтуїції, креативності та особистого досвіду. Ті, кому вдається робити це добре, стають найбільш авторитетними науковцями в певних галузях знань.

Історія наукової думки

Перед тим, як закінчити цей розділ, буде цікаво зробити невеликий історичний екскурс, щоб побачити основні етапи еволюції науки та згадати найвидатніші уми людства. Хоча приклади наукового прогресу накопичувалися протягом багатьох століть, терміни «наука», «науковці» та «науковий метод» з'явилися лише в ХІХ ст. До цього часу наука вважалася частиною філософії та співіснувала з іншими галузями філософії, такими як логіка, метафізика й естетика, хоча кордони між цими галузями були досить розмитими.

На початкових етапах розвитку людства знання зазвичай розглядалося в термінах теологічних догм, що базувалися на істині. Такі уявлення були піддані сумніву давньогрецькими філософами, зокрема Платоном, Аристотелем і Сократом у ІІІ ст. до н.е. Вони припустили, що для більш глибокого розуміння фундаментальних законів існування природи та світу необхідно звернутися до процесу систематизованого логічного мислення, який називається **раціоналізмом**. Зокрема, класична робота Аристотеля «Метафізика» (що буквально означає «за межами

фізичного існування») розмежувала теологію (вивчення Бога), онтологію (вивчення буття та існування) й універсальну науку (вивчення фундаментальних принципів, на яких базується логіка). Раціоналізм (який потрібно відрізнити від раціональності) розглядає причину як джерело знань й обґрунтування, вважаючи, що критерій істини не є чуттєвим, а скоріше інтелектуальним чи дедуктивним і часто виводиться з набору первинних принципів чи аксіом (такого як Аристотелівський закон несуперечливості).

Наступний важливий крок у розвитку наукової думки стався в XVI ст., коли британський філософ Френсіс Бекон (1561-1626) припустив, що знання може бути отримане лише через спостереження за подіями та процесами реального світу. Базуючись на цій передумові, Бекон розглядав набуття знань скоріше як емпіричну діяльність (а не мислення) і розвинув емпіризм у потужну галузь філософії. Роботи Бекона мали наслідком популяризацію індуктивних методів наукового дослідження та розвиток «наукового методу». Науковий метод, який спочатку називався «методом Бекона», включав систематичне спостереження, вимірювання та експериментування, і, вірогідно, посіяв зерна атеїзму шляхом відхилення теологічних догм як таких, що не піддаються спостереженню.

Емпіризм продовжував боротися з раціоналізмом в епоху Середньовіччя, коли філософи шукали найбільш ефективні способи отримання валідного знання. Французький філософ Рене Декарт підтримував раціоналістів, тоді як британські філософи Джон Лок і Девід Г'юм поділяли погляди емпіриків. Інші науковці, такі як Галілео Галілей та Ісаак Ньютон, намагалися об'єднати обидві ідеї в природничу філософію, зосередившись, головним чином, на розумінні природи та фізичного всесвіту, який вважався попередником природничих наук. Галілео Галілей (1564-1642) був, можливо, першим, хто вважав, що закони природи можуть бути описані математично, і зробив внесок у розвиток астрономії завдяки інноваційному поєднанню експериментування та математики.

У XVIII ст. німецький філософ Іммануїл Кант намагався вирішити суперечку між емпіризмом і раціоналізмом у своїй книзі «Критика чистого розуму», висловивши думку, що досвід є повністю суб'єктивним. Отже, застосування емпіризму та раціоналізму, спираючись на чистий розум без попереднього урахування суб'єктивної природи досвіду, може призвести лише до теоретичних ілюзій. Ідеї Канта започаткували розвиток німецького ідеалізму, який дав поштовх для подальшого розвитку інтерпретативних методик, таких як феноменологія, герменевтика і критична соціальна теорія.

Приблизно в той же час французький філософ Огюст Конт (1798 - 1857), відомий як засновник соціології, зробив спробу поєднати раціоналізм й емпіризм у новій доктрині, що отримала назву позитивізм.

Він припустив, що теорія й емпірика(спостереження) циклічно залежать одне від одного. Хоча теорії можуть бути створені в процесі мислення, вони тільки тоді будуть справжніми теоріями, коли їх можна перевірити через спостереження. Наголос на перевірці (верифікації) означав перехід від філософії й метафізики до сучасної науки й подальший розвиток «наукового методу» як первинного засобу валідації надбань науки. Еміль Дюркгейм, спираючись на ідеї Конта, започаткував соціологічний позитивізм, а Людвіг Вітгенштайн – логічний позитивізм.

На початку ХХ ст. сильні позиції позитивізму похитнулися завдяки зусиллям інтерпретативних соціологів–антипозитивістів, що належали до школи німецьких ідеалістів. Позитивізм зазвичай асоціювався з кількісними методами дослідження, такими як експеримент чи анкетування, тоді як антипозитивізм користувався якісними методами, такими як неструктуроване інтерв'ю чи спостереження. Навіть практики позитивізму, такі як американський соціолог Пауль Лазарсфельд, що започаткував масштабні анкетні дослідження та статистичні методи аналізу результатів анкетування, визнавав потенційні проблеми, пов'язані з помилками спостерігача й структурними обмеженнями позитивістського дослідження. У відповідь антипозитивісти підкреслювали, що соціальна діяльність повинна вивчатися інтерпретативними засобами, базуючись на розумінні значення та цілі, які індивіди надають своїм особистим діям, що надихнуло Джоржа Сіммела на дослідження символічного інтеракціонізму, Макса Вебера на дослідження ідеальних типів, а Едмунда Хассерлана дослідження феноменології.

В середині та наприкінці ХХ ст. школи позитивістського й антипозитивістського мислення були об'єктами критики та подальшого удосконалення. Британський філософ Карл Поппер припустив, що людське знання базується не на твердому й непохитному фундаменті, а скоріше на наборі попередніх гіпотез, які не можуть бути доведені остаточно, однак можуть бути спростовані. Емпіричні докази є основою для відхилення цих гіпотез або «теорій». Цей метатеоретичний підхід, названий **постпозитивізм** чи постемпіризм, розвиває позитивізм, припускаючи, що неможливо довести істину, хоча можливо відхилити неправильні уявлення. Підхід зберігає позитивістські погляди про об'єктивну істину та важливість наукового методу.

Аналогічно, антипозитивісти також зазнали критики, оскільки вони намагалися зрозуміти суспільство, а не виявляти його вади чи змінювати на краще. Корені останньої думки лежать в «Капіталі» Карла Маркса та Фрідріха Енгельса, які критикували капіталістичне суспільство як таке, що є соціально нерівним, і рекомендували подолати цю нерівність шляхом класового конфлікту та пролетарської революції. Марксизм надихнув робітничий клас на соціальні революції в таких

країнах як Німеччина, Італія, Росія і Китай, однак у цілому не зміг забезпечити соціальну рівність, яку декларував.

Напрямок критичних досліджень (який також називається критичною теорією) був заснований у ХХ ст. Максом Хорхеймером і Юргеном Хабермасом. Вони дотримуються схожих ідей критики та вирішення проблем соціальної нерівності, водночас додаючи тезу, що люди можуть і повинні свідомо діяти для зміни свого соціального та економічного життя, хоча їхні можливості зробити це обмежені різними формами соціального, культурного чи політичного панування. Прихильники критичних досліджень намагаються викрити обмежувальні умови статусу кво, аналізуючи опозиційні погляди, конфлікти і протиріччя в сучасному суспільстві, а також усунути причини класового панування, щоб звільнити клас, який пригнічують.

Більше інформації стосовно різних філософських поглядів і підходів буде надано в наступних розділах цієї книги.

Розділ 2

НАУКОВЕ МИСЛЕННЯ

Виконання хорошого дослідження означає, перш за все, перехід розумової діяльності дослідника в режим проведення дослідження. Це вимагає візуалізації абстрактної картини на основі реальних спостережень та асоціативного мислення для виявлення неочевидних концептів та паттернів, а також їх синтезу в загальні закони й теорії, що можуть бути застосовані для пояснення інших ситуацій, які відрізняються від сфери первинного спостереження. Дослідження передбачає постійний рух уперед і назад від емпіричного рівня, де виконуються спостереження, до теоретичного рівня, де ці спостереження отримують абстрактну форму у вигляді узагальнених законів і теорій. Дослідницькі навички потребують багато років для свого розвитку, і не є тим, чому можна навчитися в магістратурі чи аспірантурі, або під час навчання на виробництві, і є, можливо, найбільш дефіцитними серед тих, хто пише кандидатські дисертації. Серед абстракцій мислення, якими повинен володіти справжній дослідник, особливо важливими є предмет дослідження, конструкти, гіпотези, операціоналізація, теорії, моделі, індукція, дедукція і т.п., які будуть розглядатися в цьому розділі.

Предмет дослідження

Одне з перших рішень у соціальних дослідженнях стосується визначення предмету наукового дослідження. **Предмет дослідження** – це окрема людина, колектив, організація, країна чи інший об'єкт, який є мішенню дослідження і на який спрямована пізнавальна діяльність дослідника. Предмет дослідження визначається, насамперед, метою дослідження. Наприклад, якщо вивчається поведінка споживачів, результати навчання працівників або їхнє ставлення до нових технологій, тоді предметом дослідження будуть окремі особи, споживачі або працівники. Якщо вивчається командна робота в організаціях чи формування груп волонтерів, тоді предметом дослідження буде група. Якщо мета дослідження полягає в тому, щоб зрозуміти, як компанії можуть підвищити свою прибутковість або приймати виважені управлінські рішення, тоді предметом дослідження буде компанія. У цьому випадку, хоча рішення приймаються людьми, які працюють у компаніях, вважається, що вони представляють більшою мірою свої організації, а не самих себе. Якщо дослідження спрямоване на розуміння відмінностей у національних культурах різних країн, тоді предметом дослідження будуть країни. Неживі об'єкти теж можуть бути предметом дослідження соціальних наук. Наприклад, якщо дослідник зацікавлений у розумінні того, як зробити веб-сторінки в Інтернеті більш привабливими для користувачів, тоді предметом дослідження будуть веб-сторінки, а не

користувачі. Якщо ж вивчається процес передачі знань між компаніями, тоді предметом дослідження буде сукупність компаній, що передають й отримують знання.

Розуміння предмету дослідження може бути достатньо складним процесом. Наприклад, якщо необхідно дослідити, чому певні інновації є більш успішними, ніж інші, тоді предметом аналізу стає інновація. Однак, якщо необхідно дізнатися, чому деякі організації створюють інноваційні продукти частіше, ніж інші, тоді предметом дослідження стає організація. Отже, два пов'язаних дослідницьких питання в межах одного і того ж дослідження можуть мати два різних предмета дослідження.

Розуміння предмету дослідження є важливим, оскільки це визначає, якого типу дані потрібно зібрати для конкретного дослідження й де саме потрібно їх збирати. Якщо предметом аналізу є організація, то необхідно вимірювати змінні організаційного рівня, такі як розмір організації, прибуток, рентабельність чи організаційну структуру. Ці дані можуть надійти з різних джерел, таких як фінансові звіти чи інтерв'ювання керівного складу організації. Деякі змінні, наприклад, зарплата директора, можуть виглядати як змінна індивідуального рівня, однак фактично це може бути змінною організаційного рівня. Іноді можливо збирати дані з нижчих рівнів аналізу й потім переводити їх на більш високий рівень. Наприклад, щоб дослідити командну роботу в організаціях, можна провести анкетування учасників команд у різних командах, розрахувати середні значення визначених змінних, щоб створити узагальнений показник командного рівня (наприклад, групової солідарності чи конфліктності).

Концепти, конструкти та змінні

Як відзначалося в першому розділі, хоча дослідження може бути пошуковим, описовим чи пояснювальним, більшість наукових досліджень намагаються бути пояснювальними, у тому сенсі, що вони шукають потенційні пояснення природних чи соціальних феноменів, які вони досліджують. Пояснення вимагають розвитку **концептів**, або узагальнених властивостей чи характеристик, пов'язаних із об'єктами, подіями чи індивідами. Концепт (за значенням близький до поняття) є мисленнєвим утворенням, яке в процесі мислення заміщує уявлення науковця про реальний об'єкт. Об'єкти на зразок людини, компанії чи країни не є концептами, однак їхні специфічні характеристики чи поведінка, такі як ставлення до інновацій, стратегія розвитку чи політика зайнятості, можуть розглядатися як концепти.

Свідомо чи ні, ми використовуємо різні види концептів у своїх щоденних комунікаціях. Деякі з цих концептів розвивалися протягом тривалого часу завдяки використанню спільної мови. Іноді дослідники запозичують концепти з інших дисциплін чи мов для пояснення

досліджуваного феномену. Наприклад, ідея гравітації, запозичена з фізики, може знайти застосування в роздрібній торгівлі, щоб описати чому споживачі «тяжіють» до певних магазинів, які вони вподобали. Аналогічно, концепт відстані може бути використаний, щоб пояснити ступінь соціальної відчуженості між двома працівниками, які працюють в одному відділі. Іноді ми створюємо наші власні концепти для опису характеристик, які не були раніше описані. Наприклад, технострес – це новий концепт, що стосується розумового стресу, із яким стикається той, хто в стислі терміни має опанувати нову технологію.

Концепти можуть мати різний рівень абстрактності. Деякі концепти, наприклад, розмір організації, дуже точні й об'єктивні, тоді як інші, наприклад, креативність організації є більш абстрактними та складними для уявлення. **Конструкт** – це абстрактний концепт, який спеціально обраний чи створений, щоб пояснити даний феномен. Конструкт може бути простим концептом, (наприклад, розмір організації), або комбінацією з декількох концептів. Наприклад, конструкт письмова комунікативна здібність може складатися з декількох концептів, що лежать в його основі – словниковий запас, орфографія і синтаксис. Якщо розмір організації є одновимірним, то комунікативна здібність є багатовимірним конструктом, тобто таким, що складається з багатьох концептів, які лежать у його основі. Відмінність між конструктами й концептами є зрозумілішою в багатовимірних конструктах, де абстракція вищого рівня називається конструктами, тоді як абстракція нижчого рівня називається концептами. Однак ця відмінність стає майже непомітною у випадку одновимірних конструктів.

Конструкти, що використовуються в наукових дослідженнях, повинні мати точні й однозначні визначення, які можуть бути використані іншими дослідниками, щоб точно розуміти, що вони означають. Наприклад, конструкт дохід організації, який здається начебто простим, насправді може означати місячний чи річний дохід, дохід до чи після оподаткування, а тому не є ні точним, ні однозначним.

Існує два типи визначень: словникові і операціональні. Згідно з більш відомим словниковим визначенням, конструкт часто визначається через синоніми. Наприклад, ставлення може бути визначене як почуття, уявлення про цінності, переконання, погляди, а також конкретні пристрасті та інтереси, що поділяються людиною. У свою чергу почуття можуть визначатися як ставлення і т.п. Такі визначення циклічної природи не є особливо корисними в наукових дослідженнях для встановлення значення і сутності конструкту. Наукове дослідження вимагає **операціональних (робочих) визначень**, які характеризують конструкти стосовно того, як вони будуть вимірюватися на емпіричному рівні. Наприклад, операціональне визначення такого конструкту як температура повинно точно вказати, у якій шкалі – за Цельсієм, Фаренгейтом чи Кельвіном – буде вимірюватися температура. Такий

конструкт як дохід має бути визначений стосовно того, який саме дохід буди досліджуватися – місячний, річний, до чи після оподаткування. Очевидно, що такі конструкти як компетентність, мотивація чи творчий потенціал визначити на операціональному рівні досить проблематично.

Термін, який часто пов'язаний та іноді неправильно вживається замість конструкта, – це змінна величина. Якщо говорити з позицій етимології, змінна величина є певною кількістю, яка може змінюватися (наприклад, від низьких до високих чи від негативних до позитивних значень і т.п.), на відміну від констант, що не змінюються. Однак у наукових дослідженнях змінна є вимірюваним показником абстрактного конструкту. Як абстрактні величини, конструкти не можна виміряти безпосередньо, отже необхідно звертатися до величин, які їх представляють, тобто до змінних. Наприклад, інтелект індивідуума часто вимірюють за шкалою коефіцієнта інтелекту. У цьому випадку інтелект є конструктом і бал IQ є змінною, що вимірює цей конструкт. Залежно від того, наскільки добре IQ вимірює інтелект, бал IQ може бути гарним чи поганим показником цього конструкту. Як показано на рис. 2.1., наукове дослідження виконується у двох площинах: теоретичній та емпіричній. Конструкти визначаються в теоретичній (абстрактній) площині, тоді як змінні є операціональними і вимірюються в емпіричній площині. Для наукового мислення потрібно мати здатність легко переходити з однієї площини в іншу.



Рис. 2.1. Теоретичний та емпіричний рівні дослідження

Залежно від мети застосування, змінні величини можна класифікувати на незалежні, залежні, змінні-посередники, модератори і контрольовані. Змінні, що пояснюють інші змінні, називаються **незалежними**. Змінні, що пояснюються іншими змінними, називаються **залежними**. Змінні, які пояснюються незалежними змінними і водночас самі пояснюють залежні змінні, називаються **змінними-посередниками** (або медіаторами). Змінні, які впливають на зв'язок між незалежними і залежними змінними, називаються **модераторами**. Наприклад, якщо стверджується, що вищий рівень інтелекту є причиною кращих успіхів у

навчанні студентів, тоді інтелект є незалежною, а успішність навчання залежною змінною. Також можуть бути інші сторонні змінні, які не потрібні у поясненні даної залежної величини, однак мають певний вплив на залежну величину. Такі змінні в науковому дослідженні мають контролюватися, а тому називаються **контрольованими**.

Щоб зрозуміти відмінності між цими типами змінних, розглянемо приклад, наведений на рис 2.2. Якщо ми вважаємо, що рівень інтелекту впливає на (чи пояснює) академічні успіхи студентів, тоді такий показник інтелекту як IQ є незалежною змінною, а показник академічних досягнень (середній бал диплому) є залежною змінною. Якщо ми вважаємо, що ефект інтелекту на академічні досягнення також залежить від зусиль, які докладає студент у процесі навчання (тобто серед двох студентів із однаковим рівнем інтелекту студент, який докладає більше зусиль, досягає кращих результатів у навчанні, ніж той, хто докладає менше зусиль), тоді зусилля стають змінною-модератором. Якщо академічні досягнення вважаються проміжним кроком до отримання більш високих доходів у майбутньому, тоді академічні досягнення стають змінною-посередником у зв'язку між інтелектом і потенціалом доходів. Отже, змінні визначаються як незалежні, залежні, посередники чи модератори залежно від природи зв'язків між ними. Загальна мережа зв'язків між набором пов'язаних конструктів називається **номологічною мережею** (див. рис. 2.2). Дослідницьке мислення вимагає не тільки здібності встановити конструкти на основі емпіричних спостережень, але й здібності створити мислену номологічну мережу, що поєднує абстрактні конструкти.

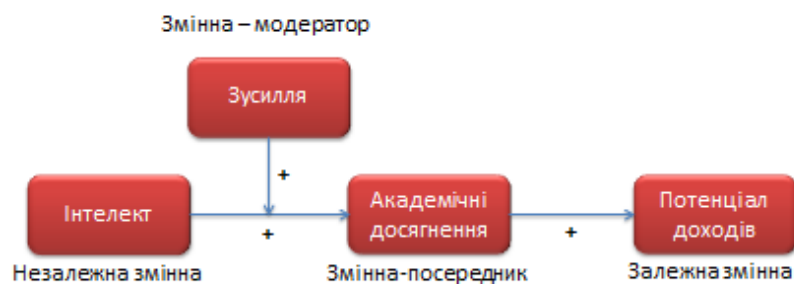


Рис.2.2. Номологічна мережа конструкту «інтелект»

Теоретичні положення й гіпотези

Рис. 2.2. показує, як теоретичні конструкти, такі як інтелект, зусилля, академічні досягнення та потенціал доходів пов'язані між собою в номологічній мережі. Кожний із цих зв'язків становить певне теоретичне положення. Шукаючи пояснення досліджуваного феномену чи процесу, не буде адекватним лише ідентифікувати ключові концепти й конструкти, що їх зумовлюють. Потрібно ідентифікувати та встановити характер, або паттерни зв'язків між цими конструктами. Такі паттерни

зв'язків називаються теоретичними положеннями. **Теоретичне положення** – це попередній і гіпотетичний зв'язок між конструктами, який формулюється в декларативній формі. Наведемо приклад теоретичного положення: «зростання інтелекту студентів призведе до зростання їхніх академічних досягнень». Таке декларативне твердження не обов'язково має бути істинним, однак воно обов'язково підлягає емпіричній перевірці з використанням реальних даних, щоб можна було встановити його істинність. Теоретичні положення зазвичай отримують, базуючись на логіці (дедукції) чи емпіричних спостереженнях (індукції).

Оскільки теоретичні положення є зв'язками між абстрактними конструктами, вони не можуть бути перевірені безпосередньо. Натомість, вони тестуються опосередковано шляхом дослідження зв'язку між відповідними змінними, які їх представляють. Емпіричне формулювання теоретичних положень стосовно зв'язку між змінними називається **гіпотезою** (див. рис. 2.1). Оскільки бали IQ і середній бал за дипломом є операціональними показниками інтелекту та академічних досягнень відповідно, сформульоване вище теоретичне положення може бути виражене у формі гіпотези: «Збільшення балу IQ студента призводить до збільшення середнього балу диплому». Теоретичні положення формулюються в теоретичній площині, тоді як гіпотези – в емпіричній. Отже, гіпотези підлягають емпіричній перевірці, щоб встановити чи є відповідне теоретичне положення валідним.

Гіпотези можуть бути сильними і слабкими. Гіпотеза «Бали IQ студентів пов'язані з їхніми академічними успіхами» є прикладом слабкої гіпотези, оскільки вона не вказує ні на напрям гіпотези (зв'язок має бути позитивним чи негативним), ні на її причини (тобто, інтелект є причиною академічних досягнень чи, навпаки, академічні досягнення зумовлюють інтелект). Сильнішою гіпотезою буде така: «Бали IQ студентів позитивно пов'язані з їхніми академічними досягненнями», яка вказує на напрям, але не на причину. Кращою гіпотезою буде наступна: «Бали IQ студентів мають позитивний ефект на їхні академічні досягнення», яка вказує і на напрям зв'язку, й на причину, тобто інтелект зумовлює академічні досягнення, а не навпаки. Позначки на рис. 2.2 вказують на напрям відповідних гіпотез.

Також необхідно відзначити, що наукові гіпотези повинні чітко вказувати незалежні та залежні змінні. Із останнього формулювання гіпотези «Бали IQ студентів мають позитивний ефект на їхні академічні досягнення» зрозуміло, що інтелект є незалежною змінною («причиною»), а академічні досягнення є залежною змінною («ефектом»). Далі також зрозуміло, що ця гіпотеза може бути оцінена щодо її істинності: вона є істинною, якщо вищий рівень інтелекту призводить до вищих академічних досягнень, і вона є хибною, якщо високий рівень інтелекту не має впливу або веде до гірших академічних досягнень. У подальшому буде показано, як емпірично тестувати такі причинно-

наслідкові зв'язки. Формулювання на кшталт «студенти зазвичай інтелектуально обдаровані» або «усі студенти можуть досягти високих академічних досягнень» не є науковими гіпотезами, оскільки вони не вказують на залежні та незалежні величини, як і на напрям зв'язку, а отже, і не можуть бути оцінені як істинні чи хибні.

Теорії та моделі

Теорія – це набір упорядкованих взаємозалежних конструктів і теоретичних припущень, який має на меті пояснити і спрогнозувати поведінку досліджуваного феномену чи процесу за певних умов і обмежень. Фактично, теорія є упорядкованим набором пов'язаних теоретичних припущень. Тоді як теоретичні положення зазвичай пов'язують два чи три конструкти, теорії представляють собою систему численних конструктів і теоретичних припущень. Отже, теорії можуть бути значно складнішими, більш абстрактними й масштабними, ніж теоретичні положення чи гіпотези.

Люди, незнайомі з науковими дослідженнями, часто думають, що теорія є певним спекулятивним припущенням, або щось протилежне фактам. Наприклад, люди часто кажуть, що навчання повинно бути менш теоретичним і більш практичним. Однак практика чи факти не є альтернативою теорії, а необхідними, суттєвими компонентами для перевірки валідності теорій. Хороша наукова теорія повинна бути добре обґрунтованою фактичним матеріалом і мати практичне застосування, тоді як погано визначеним теоріям зазвичай бракує цих рис. Відомий дослідник організації Курт Левін сказав: «Теорія без практики стерильна, а практика без теорії сліпа». Отже, як теорія, так і факти (чи практика) важливі для наукових досліджень.

Теорії надають пояснення явищам природи чи суспільства. Як підкреслювалось в розділі 1, ці пояснення можуть бути хорошими чи поганими. Отже, можуть бути хороші чи погані теорії. Розділ 3 описує деякі критерії, що можуть бути використані, щоб оцінити наскільки гарною насправді є теорія. Тим не менш, для науковців важливо розуміти, що теорія не є «істиною», і в ній немає нічого сакрального, вони не можуть бути прийнятими лише тому, що їх запропонував авторитетний науковець. У міру наукового прогресу погані теорії замінюються кращими, із більшими пояснювальними можливостями. Для дослідників важливим викликом є побудова кращих і більш універсальних теорій, які можуть пояснити досліджуваний феномен краще, ніж вже існуючі.

Модель – це термін, який часто використовується у зв'язку з теорією. **Модель** – це презентація всієї системи або її певної частини, яка сконструйована для вивчення цієї системи (тобто, як система працює, і що на неї впливає). Тоді як теорія прагне пояснити феномен, модель прагне представити феномен. Моделі часто використовуються тими, хто

приймає рішення, щоб обґрунтувати важливі рішення з урахуванням визначеного набору вхідних параметрів. Наприклад, маркетологи можуть використовувати моделі, щоб вирішити, скільки грошей потрібно витратити на рекламу різних ліній продукції на основі певних параметрів, таких як витрати на рекламу в попередньому періоді, обсяг продажів, динаміка ринку й конкурентна продукція. Аналогічно, синоптики можуть використовувати моделі, щоб спрогнозувати динаміку змін погоди на підставі таких параметрів, як швидкість вітру, його напрям, температура та вологість повітря. Хоча такі моделі є корисними, вони можуть мало що давати в сенсі пояснення витрат на рекламу чи прогнозів погоди.

Моделі відрізняються за своїми видами й можуть бути математичними, мережевими чи покроковими. Моделі можуть бути також описовими, прогнозними чи нормативними. Описові моделі часто використовуються для презентації складних систем, візуалізації змінних і зв'язків, що існують у таких системах. Модель витрат на рекламу може належати до цього виду моделей. Прогнозні моделі (наприклад, регресійна модель) дозволяють зробити прогноз майбутніх подій. Прогноз погоди є прикладом такої моделі. Нормативні моделі використовують, щоб скерувати якусь діяльність згідно з загальноприйнятими нормами чи практикою. Моделі також можуть бути статичними, якщо вони представляють стан системи на даний час, або динамічними, якщо вони представляють еволюцію системи за певний період часу.

Процес розвитку теорії чи моделі може включати індуктивне та дедуктивне мислення. Згадайте з розділу 1, що **дедукція** є процесом, при якому висновки про якийсь феномен чи процес здійснюються на підставі теоретичних чи логічних причин і первинного набору передумов. Наприклад, якщо банк вводить строгий код етики для своїх працівників (передумова 1) і Петренко є працівником цього банку (передумова 2), то Петренко дотримується коду етики банку (висновок). При застосуванні дедукції висновки мають бути вірними, якщо вірними є передумови та причини.

Індукція є процесом, при якому висновки базуються на фактах чи даних спостереження. Наприклад, якщо фірма витрачає багато грошей на кампанію з просування товару (спостереження 1), але продажі не зростають (спостереження 2), тоді, вірогідно, кампанія з просування була проведена погано (висновок). Однак можуть бути альтернативні пояснення поганих продажів, такі як економічний спад, вихід на ринок конкуруючих продуктів чи брендів, або проблеми з ланцюгом постачань. Індуктивні висновки є фактично тільки гіпотезами й можуть бути відхилені. Дедуктивні висновки загалом більш сильні, ніж індуктивні, однак дедуктивний висновок, що будується на неправильних початкових умовах, також буде неправильним.

Як показано на рис. 2.3., індуктивне та дедуктивне мислення нерозривно пов'язані зі створенням теорій і моделей. Індукція вступає в дію, коли спостерігається певний факт і виникає питання «Що відбувається?». Відповідаючи на це питання, дослідник пропонує одне чи декілька вірогідних пояснень (гіпотез). Потім, керуючись своїм розумінням досліджуваного феномену, звертається до дедукції, щоб обмежити кількість вірогідних пояснень найбільш правдоподібними, базуючись на логіці й розумних передумовах. Дослідники повинні бути здатні легко переходити від індуктивного до дедуктивного мислення, якщо вони хочуть удосконалити або розвинути певну теорію чи модель, або запропонувати інші, що, фактично, і є змістом наукового дослідження.



Рис. 2.3. Процес побудови моделі

Розділ 3

ПРОЦЕС НАУКОВОГО ДОСЛІДЖЕННЯ

В розділі 1 ми з'ясували, що наукове дослідження є процесом отримання наукового знання шляхом застосування наукових методів. Але як проводиться дослідження? Цей розділ присвячений розгляду процесу наукового дослідження, передумовам і результатам наукового процесу.

Парадигми в соціальних дослідженнях

Дизайн і проведення дослідження формуються на основі ментальних моделей, які ми використовуємо, щоб організувати процеси власного мислення і спостереження. Ці ментальні моделі або межі (системи переконань) називаються **парадигмами**. Слово «парадигма» було популяризоване Томасом Куном (1962) у його книзі «Структура наукових революцій», де він досліджував історію природничих наук для того, щоб виявити ті види діяльності, які призвели до прогресу в науці. Схожі ідеї можна застосувати по відношенню до соціальних наук, де соціальні реалії по-різному сприймаються різними людьми, що може обмежувати їхнє мислення й уявлення відносно досліджуваних феноменів. Наприклад, представники різних політичних партій можуть мати різні уявлення стосовно ролі уряду в житті народу і, відповідно, пропонувати різні шляхи щодо вирішення існуючих в країні соціальних проблем. Західні суспільства роблять акцент на правах особистості, таких як право приватної власності чи свобода слова. Східні суспільства, навпаки, намагаються встановити баланс прав особистості та прав родини, організацій і уряду, а тому в своїй політиці намагаються більше орієнтуватися на колектив, ніж на окрему людину. Такі відмінності в перспективі часто призводять до того, що західні політики критикують уряди східних країн за авторитаризм, тоді як представники східних країн критикують західні суспільства за їх жагу до наживи, високий рівень злочинності і створення культу особистості. Наші особисті парадигми нагадують «кольорові окуляри», які визначають, у яких кольорах ми бачимо навколишній світ, і як ми осмислюємо те, що бачимо.

Парадигми не завжди легко розпізнати, оскільки вони неявні, неочевидні і сприймаються на віру. Однак виявлення парадигм є ключем до розуміння й подолання розбіжностей в уявленнях людей стосовно одних і тих же подій чи явищ суспільства. Наприклад, у дослідженнях із соціальних наук для того, щоб зрозуміти, чому впровадження нової технології було успішним в одній організації і зазнало поразки в іншій, дослідник, що дивиться на світ через «раціональні лінзи», буде шукати раціональне пояснення проблеми, наприклад, в неадекватній організації процесу впровадження чи низькій відповідності між технологією та контекстом завдання, для якого ця технологія призначалася. Інший

дослідник, що дивиться на ту ж проблему через «соціальні лінзи», буде шукати соціальні проблеми, такі як недостатній рівень підготовки працівників, низька підтримка менеджменту, або невідповідна організаційна культура. Дослідник, який дивиться на проблему через «політичні лінзи», сконцентрується на особливостях організаційної політики, яка могла загальмувати процес упровадження технології. Отже, неусвідомлювані парадигми часто обмежують концепти, які дослідники обирають для виміру, коло їхніх спостережень і, відповідно, інтерпретацію досліджуваних явищ чи подій. Однак, ураховуючи складну природу соціальних феноменів, можливо, що всі вищевказані парадигми є тією чи іншою мірою правильними, і більш глибоке розуміння проблеми може потребувати розуміння й застосування численних парадигм.

Сьогодні серед дослідників соціальних наук є популярними дві парадигми: позитивізм і постпозитивізм. **Позитивізм**, що базується на роботах французького філософа Огюста Конта, залишався домінуючою парадигмою до середини ХХ ст. Згідно з цією парадигмою, наука, або процес генерування знання, має бути обмеженою тим, що можна спостерігати та вимірювати. Позитивізм прагне покладатися виключно на теорії, які можуть бути перевірені безпосередньо. Хоча спочатку позитивізм був спробою відокремити наукове пізнання від релігії (де передумови не можуть спостерігатися об'єктивно), позитивізм призвів до розвитку *емпіризму* чи сліпої віри в результати, отримані через спостереження, і відмови від будь-яких спроб розширити висновки на інші факти, окрім досліджених. Оскільки людські думки й емоції не піддаються безпосередньому вимірюванню, їх вважали некоректними об'єктами наукового дослідження. Розчарування в жорсткій емпіричній природі позитивістської філософії призвело до розвитку **постпозитивізму**, чи постмодернізму протягом середини-кінця ХХ ст. Важливим положенням постпозитивізму є те, що обґрунтовані пояснення соціальних явищ і подій можуть бути зроблені на основі поєднання емпіричних спостережень із логічним мисленням. Постпозитивісти розглядають наукові закономірності не як щось однозначно визначене, а скоріше ймовірне (тобто те, що формується під впливом численних обставин), і часто намагаються виявити ці обставини, щоб краще зрозуміти реалії соціального світу. У подальшому постпозитивізм розпався на два окремих напрями – *суб'єктивістів*, які розглядають світ скоріше як суб'єктивну конструкцію суб'єктивного людського мислення, ніж об'єктивну реальність; і *критичних реалістів*, які вважають, що є одна реальність, незалежна від мислення людини, однак її неможливо пізнати з високим ступенем визначеності.

Баррел і Морган (1979) у своїй епохальній книзі «Соціальні парадигми й організаційний аналіз» припустили, що спосіб, за допомогою якого дослідники соціальних наук уявляють і досліджують соціальні феномени, формується двома наборами філософських припущень:

онтологія і епістемологія. **Онтологія** відноситься до наших припущень стосовно того, як ми бачимо світ, а саме, чи складається світ, головним чином, із соціального порядку або постійних змін. **Епістемологія** відноситься до наших припущень стосовно найкращого способу пізнання світу, а саме, чи звертаємось ми до об'єктивного або суб'єктивного підходу в пізнанні соціальної реальності. Використовуючи ці два набори припущень, можна виділити чотири категорії досліджень соціальних наук (рис. 3.1).

Якщо для дослідника світ складається, головним чином, із соціального порядку (онтології), і він націлений на вивчення паттернів впорядкованих подій чи процесів, вважаючи найкращими способами дослідження дійсності об'єктивні підходи, які є незалежними від особи дослідника, (наприклад, використовуючи стандартизовані інструменти збору даних – анкети), то це означає, що він звертається до парадигми **функціоналізму**. Однак, якщо науковець вважає, що найкращим способом дослідження соціального порядку є суб'єктивна інтерпретація соціальних феноменів учасниками дослідження (наприклад, інтерв'ювання різних учасників і усунення відмінностей між їхніми відповідями шляхом використання власних суб'єктивних суджень), тоді науковець звертається до **інтерпретативної** парадигми. Якщо дослідник вважає, що світ складається з радикальних змін, намагається їх зрозуміти і прийняти, використовуючи об'єктивний підхід, то він дотримується парадигми **радикального структуралізму**. Якщо ж дослідник прагне зрозуміти соціальні зміни, спираючись на суб'єктивні прогнози учасників дослідження, тоді він стоїть на позиціях парадигми **радикального гуманізму**.



Рис. 3.1. Чотири парадигми досліджень у соціальних науках
(джерело: Баррел і Морган, 1979)

На сьогодні більшість соціальних досліджень наслідують природничі науки і дотримуються функціональної парадигми. Функціоналісти вважають, що соціальний устрій чи розвиток можна пояснити через властивості їхніх функціональних компонентів. Вони намагаються розділити проблему на окремі невеликі компоненти і детально досліджують один чи декілька з них, спираючись на об'єктивістські методи, такі як анкетні чи експериментальні дослідження. Однак із появою постпозитивістського мислення невелика, хоча і зростаюча чисельність дослідників намагається пояснити соціальний устрій, використовуючи суб'єктивістські методи, такі як інтерв'ювання й етнографічні дослідження. Радикальний гуманізм і радикальний структуралізм продовжують складати лише незначну частку соціальних досліджень, оскільки науковці більше зацікавлені в розумінні загальних закономірностей розвитку подій, явищ чи процесів, ніж випадкових феноменів. Разом із тим, при дослідженні причин соціальних змін на рівні окремих країн, можливо, радикальний гуманізм буде найбільш адекватним. Зазвичай соціальні й організаційні феномени включають як елементи стабільності, так і елементи змін. Наприклад, організаційний успіх у бізнес-середовищі залежить від формалізованих бізнес-процесів, робочих процедур та посадових обов'язків і водночас обмежений впливом динамічного комплексу факторів, таких як конкуренти, конкуруючі продукти та клієнтська база. Отже, цілісне й багатопланове розуміння соціальних феноменів вимагає визнання та застосування мультипарадигмового підходу до проведення досліджень.

Загальна характеристика процесу дослідження

Як прийняті дослідником парадигми мислення визначають процес соціального дослідження? По суті, усі академічні дослідження є покроковим процесом спостереження, осмислення та валідації. На стадії спостереження дослідники спостерігають за соціальним феноменом, що їх цікавить. На стадії осмислення чи раціоналізації вони намагаються зрозуміти сутність досліджуваного феномену шляхом встановлення логічних зв'язків між різними частинами картини, що спостерігалася. Іноді це може призвести до створення теорії. І, нарешті, на стадії валідації дослідники перевіряють теорії, що може мати результатом внесення змін чи розвиток теорії, яка була сформульована раніше.

Дизайни наукових досліджень відрізняються залежно від того, чи дослідник починає зі стадії спостереження і потім намагається його осмислити (індуктивне дослідження), чи він починає безпосередньо з раціоналізації або теорії та намагається перевірити цю теорію. (дедуктивне дослідження). Отже, цикл спостереження – осмислення – валідація дуже схожий із циклом індукція – дедукція, що був описаний в розділі 1.

Більшість традиційних досліджень за своєю природою відносяться до дедуктивних і базуються на парадигмі функціоналізму. Рис. 3.2. надає схематичне уявлення про такий дослідницький проект. Рисунок відображає види діяльності, які потрібно виконувати в дослідженні на основі парадигми функціоналізму. Ці види діяльності вкладаються в три фази: пошукове дослідження, дизайн дослідження й проведення дослідження. Зверніть увагу, що така схема не є дорожньою картою чи блок-схемою для всіх досліджень. Вона застосовується тільки при проведенні досліджень, що базуються на парадигмі функціоналізму, а тому може й повинна видозмінюватися залежно від мети конкретного наукового проекту.



Рис. 3.2. Процес дослідження на основі парадигми функціоналізму

Перша фаза дослідження – це **пошукове дослідження**. Вона включає пошук і вибір проблемних питань для подальшого дослідження, а також вивчення опублікованих літературних джерел за певним напрямом, щоб з'ясувати поточний стан знань, та ідентифікацію теорій, що можуть бути корисними у пошуку відповідей на питання даного дослідження.

Першим кроком пошукової фази є ідентифікація одного чи декількох дослідницьких питань стосовно подій, явищ чи процесів, що становлять інтерес для дослідника. Прикладами таких питань можуть бути фактори, що мотивують споживачів купувати товари та послуги в Інтернет-магазинах, або що спонукає працівників до більш продуктивної праці і т.п. Дослідницькі питання стосуються, головним чином, того що, чому, як саме і коли відбувається. Більш цікавими є питання, що вимагають залучення великих досліджуваних груп та їх порівняння з

урахуванням конкретних обставин і дії різних факторів. Наприклад, питання «Що дозволяє компаніям бути інноваційними?» стосується складних проблем реального бізнесу та є більш цікавим, ніж гіпотетичні надумані проблеми з неочевидними відповідями. Вузько сфокусовані дослідницькі питання (часто такі, що передбачають альтернативну відповідь «так» чи «ні») загалом менш цікаві та інформативні, коли йдеться про виявлення суттєвих властивостей соціальних феноменів. Нецікаві наукові питання зазвичай мають наслідком нецікаві наукові результати, які важко опублікувати.

Наступним кроком є огляд літератури в досліджуваній галузі знань. Мета огляду літератури триєдина: 1) з'ясувати поточний стан знань в досліджуваній галузі; 2) встановити ключових авторів, публікації, теорії та наукові надбання в даній галузі; 3) виявити існуючі прогалини в знаннях. У наш час огляд літератури зазвичай виконується із використанням комп'ютеризованого пошуку за ключовими словами в он-лайн баз даних. Ключові слова можуть бути об'єднані з використанням операцій «і» та «або», щоб звузити чи розширити результати пошуку. Коли виявлено список найбільш важливих публікацій, дослідник повинен переглянути текст чи хоча б резюме кожної з них, щоб визначити потребу в їх більш детальному аналізі. Огляд літератури має бути достатньо повним. Його результати можуть бути узагальнені у вигляді таблиці та структуровані більш детально з використанням схем, таких як матриця концептів. Добре проведений огляд літератури має виявити, чи були певні наукові питання предметом попередніх досліджень (що може зняти потребу в їх повторному дослідженні), чи не з'явилися нові й більш цікаві дослідницькі питання, чи, можливо, ці питання потрібно видозмінити або переформулювати з урахуванням даних літературного огляду. Аналіз літератури також може наштовхнути на нові ідеї чи потенційно можливі відповіді на питання, які становлять науковий інтерес дослідника, і/або допомогти ідентифікувати теорії, які раніше були корисними при дослідженні схожих питань.

Оскільки дослідження на основі парадигми функціоналізму включає перевірку гіпотез, третім кроком є ідентифікація однієї чи декількох теорій, які можуть допомогти в конкретному дослідженні. Хоча огляд літератури може виявити широкий спектр концептів чи конструктів, потенційно пов'язаних із предметом дослідження, теорія допоможе ідентифікувати, які з цих конструктів знаходяться в логічному зв'язку з досліджуваним феноменом і яким саме чином. Невідповідні теорії можуть мати результатом вимірювання широкого спектру менш релевантних, або і не релевантних конструктів, водночас мінімізуючи шанси отримання змістовних і невідповідних результатів. У дослідженнях на основі парадигми функціоналізму теорії можуть використовуватися як логічна основа формулювання гіпотез для їх емпіричної перевірки. Очевидно, не всі теорії однаково добре підходять для вивчення всіх

соціальних феноменів. Теорії мають бути ретельно відібрані на основі відповідності їхніх положень проблемі дослідження. Теорії та процеси їх розробки будуть детально розглянуті в наступному розділі.

Дизайн дослідження є наступною фазою дослідницького процесу. Цей процес стосується створення плану тих видів діяльності, які необхідно виконати, щоб успішно знайти відповіді на питання, поставлені на стадії пошукового дослідження. Це включає відбір наукових методів, операціоналізацію конструктів, що становлять інтерес даного дослідження, і вибір адекватної стратегії формування вибірки.

Операціоналізація є процесом дизайну точних показників для виміру абстрактних теоретичних конструктів. Це головна проблема в соціальних дослідженнях, урахувавши те, що багато конструктів, таких як мотивація чи інноваційність важко визначити, не кажучи вже про їх точне вимірювання. Операціоналізація починається з конкретизації операціональних визначень (або концептуалізації) конструктів даного дослідження. Далі науковець має проаналізувати літературу, щоб з'ясувати, чи встановлені показники відповідають даному операціональному визначенню і можуть використовуватися безпосередньо або в модифікованому вигляді для виміру конструктів даного дослідження. Якщо таких показників не існує, або вони незадовільні, чи відображають інший конструкт, тоді, можливо, потрібно розробити нові показники для вимірювання конструктів, які планується досліджувати. Це означає, що потрібно точно визначити, як бажані конструкти будуть вимірюватися (тобто кількість показників, які саме і т.п.). Таке визначення може перетворитися на довгий і складний процес із численними перевірками та модифікаціями, допоки новий інструмент зможе бути прийнятий як «науково валідний». У розділі про вимірювання ми обговоримо операціоналізацію конструктів детальніше.

Одночасно з операціоналізацією дослідник має прийняти рішення стосовно того, які методи дослідження він збирається використати для збору даних. Ці методи можуть включати кількісні методи, такі як експерименти чи анкетування, і якісні методи, такі як кейс-дослідження чи дослідження дією, або, можливо, їх комбінацію. Якщо планується провести експеримент, тоді яким має бути його дизайн? Якщо анкетування, тоді яким буде його вид – опитування по телефону, Інтернету чи їх комбінація? Для складних, невизначених і багатопланових соціальних феноменів більш доречно використання комплексу методів, що може допомогти посилити унікальні переваги кожного методу й отримати нові ідеї й результати, які не можна отримати, спираючись лише на один метод.

Дослідники також повинні ретельно обирати цільову групу, що буде брати участь у дослідженні та стратегію **формування вибірок**. Наприклад, будуть досліджуватися окремі особи, компанії чи робочі

групи в компаніях? Стратегія формування вибірок тісно пов'язана з вибором предмету дослідження. Добираючи вибірку, потрібно приділити достатню увагу тому, щоб уникнути «неправильних» вибірок, сформованих, наприклад, за принципом доступності, що може призвести до отримання хибних даних. Формування вибірок буде досліджено в наступному розділі.

На цій стадії може бути корисним написати **дослідницькі пропозиції**, конкретизуючи й обґрунтовуючи ті рішення, які були зроблені на попередніх стадіях дослідження. Така багатопланова пропозиція має містити інформацію стосовно проблемних питань дослідження, обґрунтування того, чому саме ці питання були обрані для дослідження, аналіз попередніх наукових робіт; теорій, на які буде спиратися дослідник; гіпотез, що будуть тестуватися; показників оцінки конструктів, опису й обґрунтування методів дослідження, які планується застосувати; обраної стратегії формування вибірок. Агенції з надання наукових грантів зазвичай вимагають такі пропозиції, щоб обрати найкращі з них для фінансування. Навіть якщо фінансова підтримка не очікується, пропозиція може бути рушійною силою для отримання зворотного зв'язку від інших дослідників і виявлення потенційних проблем, що можуть виникнути під час виконання дослідницького проекту (наприклад, відсутність деяких важливих конструктів), ще до його початку. Такий зворотний зв'язок на початковому етапі дослідження може виявитися дуже цінним, оскільки коригувати проблеми після того, як дані зібрані, буде запізно.

Визначивши, кого досліджувати (вибірка), що оцінювати (концепти) і як збирати дані (методи дослідження), дослідник готовий переходити до стадії **проведення досліджень**. Вона включає пілотне тестування інструментів виміру, збір даних та їх аналіз.

Пілотне тестування часто випускається з уваги, однак це надзвичайно важлива частина дослідницького процесу. Воно допомагає виявити потенційні проблеми дослідження чи інструментарію (наприклад, з'ясувати, чи відповідають питання анкети інтелектуальному рівню досліджуваної групи). Водночас, воно дозволяє переконатися, що інструменти вимірювання надійні і є валідними показниками досліджуваних конструктів. Після успішного пілотного тестування дослідник переходить до **збору даних** в обраній вибірці. Залежно від методів дослідження, дані можуть бути кількісними та якісними.

Отримані дані аналізуються та інтерпретуються, щоб зробити висновки стосовно поставлених у дослідженні питань. Залежно від типу отриманих даних (кількісні чи якісні), виконується кількісний аналіз даних (із використанням статистичних методів аналізу, як, наприклад, регресійний аналіз) чи якісний аналіз (наприклад, кодування чи контент аналіз).

Заключна стадія дослідження містить підготовку **звіту**, який документує сам процес дослідження і його основні результати у формі наукової статті, дисертації чи монографії. У звіті докладно описуються особливості процесу дослідження (теоретичні положення, обрані конструкти, методи дослідження, формування вибірки тощо) і надається їх обґрунтування, а також наводяться результати кожної стадії. Дослідницький процес повинен бути описаний із достатнім ступенем деталізації, щоб інші дослідники могли його відтворити, перевірити результати й оцінити, наскільки прийнятними з наукової точки зору є отримані висновки. Наявність початкової дослідницької пропозиції значно полегшує та прискорює процес підготовки завершального звіту. Відзначимо, що дослідження не буде мати цінності, допоки його процес і результати не задокументовані для майбутніх дослідників. Таке документування є життєво важливим для прогресу науки.

Поширені помилки при проведенні дослідження

Проведення дослідницького процесу пов'язане з певними проблемами і підводними каменями. У результаті дослідники-початківці після значних витрат часу та зусиль раптом з'ясовують, що не змогли надати переконливі відповіді на питання дослідження, або що результати не представляють загального інтересу. Такі проблеми зазвичай мають наслідком відхилення статей науковими журналами. Деякі з найбільш розповсюджених помилок описуються нижче.

Недостатньо актуальні питання дослідження. Іноді дослідник обирає для дослідження проблеми, що цікаві для нього, але не для наукової спільноти, тобто дослідження не додає нового знання чи уявлення стосовно досліджуваного феномену. Оскільки дослідницький процес вимагає від дослідника значних інвестицій часу та зусиль, він повинен бути переконаний (і здатний переконати інших), що питання дослідження, на які він намагається знайти відповіді, стосуються реальних і мало досліджених у минулому проблем, із якими стикаються достатньо багато людей.

Дослідження феноменів із коротким терміном життя. Інша поширена помилка – це фокусування на феноменах-одноденках. Типовий приклад – дослідження деяких швидкосмінних технологій чи практик, що набули розповсюдження в наші дні. Оскільки може пройти декілька років, перш ніж дослідження буде закінчене, а його результати будуть підготовлені до публікації, може статися, що загальний інтерес до цих новинок щезне раніше. Найкращою стратегією може бути вивчення «невмираючих» тем, які роками залишаються актуальними.

Проблеми, що не піддаються дослідженню. Деякі наукові проблеми не можуть бути адекватно вирішені лише на основі доказів, які отримуються шляхом спостереження або через обмеженість тих методів і

процедур дослідження, якими прийнято в даний час користуватися. Таких проблем краще уникати. Однак деякі амбітні проблеми, що не піддаються дослідженню, можуть бути видозмінені або переформульовані на добре зрозумілі, актуальні проблеми, які піддаються дослідженню.

Улюблені методи дослідження. Багато дослідників тяжіють до використання деяких улюблених ними методів, (наприклад, анкетування), і підбирають під них дослідницькі проблеми. Це неправильний підхід. Методи дослідження повинні обиратися, щоб найкраще відповідати проблемі дослідження, а не навпаки.

Збір даних навмання. Деякі дослідники схильні спочатку збирати дані (користуючись уже відомими інструментами), а потім вирішувати, що з ними робити. Зверніть увагу, що збір даних є тільки одним з кроків у довгому та складному процесі планування, дизайну й виконання дослідження. Якщо дослідники одразу перестрибують до збору даних без ретельного планування, то отримані дані можуть виявитися хибними, а зусилля, витрачені на їх збір, марними. Велика кількість даних не може компенсувати дефіцит у плануванні та дизайні досліджень, а особливо відсутність цікавих наукових питань.

Розділ 4

ТЕОРІЇ В НАУКОВИХ ДОСЛІДЖЕННЯХ

Як ви знаєте з попередніх розділів, наука – це знання, представлене як сукупність «теорій», що виникли шляхом застосування наукових методів. У цьому розділі ми розглянемо, що являє собою теорія, для чого в дослідженнях потрібні теорії, як вони оцінюються і як їх можна застосовувати в дослідженнях. Також будуть представлені наочні приклади п'яти теорій, що найчастіше використовуються в соціальних дослідженнях.

Теорії

Теорії – це пояснення природних чи соціальних процесів, подій чи явищ. Більш формально **теорія** є системою конструктів (концептів) і теоретичних положень (зв'язків між цими конструктами), що спільно представляють логічне, систематичне і всебічне пояснення досліджуваного феномену в межах певних припущень й умов (Бакарах, 1989)¹.

Теорії повинні скоріш пояснювати, чому щось відбувається, ніж описувати чи прогнозувати. Взагалі можливо прогнозувати події чи процеси, використовуючи набір прогностичних факторів (предикторів) без пояснення того, чому ці події сталися. Наприклад, маркетингові аналітики передбачають коливання на біржі, базуючись на ринкових повідомленнях, звітах про доходи провідних компаній, офіційних статистичних даних і звертаючись до раніше встановлених кореляційних зв'язків. Прогноз вимагає тільки кореляцій, тоді як пояснення вимагає казуїстики, або розуміння причинно-наслідкових зв'язків. Для їх виявлення потрібно дотримання трьох умов: 1) кореляція між двома конструктами; 2) часова послідовність (причина передуює ефекту в часі); 3) відхилення альтернативної гіпотези шляхом тестування. Наукові теорії відрізняються від теологічних, філософських та інших пояснень тим, що вони можуть бути емпірично протестовані з використанням наукових методів досліджень.

Пояснення можуть бути випадковими та закономірними. **Випадковими** (ідеографічними) називаються ті пояснення, що пояснюють окрему ситуацію чи подію дією нерегулярних, випадкових факторів. Наприклад, якщо студент отримав погану оцінку на екзамені, то це могло статися тому, що: 1) він забув, що у нього сьогодні екзамен; 2) він спізнився на екзамен через затори на дорогах; 3) він не справився з хвилюванням під час екзамену; 4) він не підготувався як слід в т.п. Такі

¹Bacharach S. B. Organizational Theories: Some Criteria for Evaluation / S. B. Bacharach //Academy of Management Review . – 1989. – Vol. 14:4.– P. 496-515.

пояснення можуть бути детальними, точними і валідними, однак вони не будуть пояснювати інші подібні ситуації, навіть ті, що стосуються тієї ж особи, а отже не є загальними. **Закономірні** (номотетичні) пояснення, навпаки, прагнуть пояснити велику групу ситуацій чи подій, а не лише окремі з них. Наприклад, студенти погано склали екзамени тому, що вони не приділили достатньо часу підготовці до них або через систематичні пропуски занять. Оскільки закономірні пояснення конструюються таким чином, що мають узагальнений характер для певних подій чи осіб, вони менш точні, менш повні й детальні. Водночас, вони пояснюють події «економно», застосовуючи тільки декілька змінних. Оскільки теорії також створюють, щоб надати узагальнені пояснення паттернам подій, процесів чи явищ, теоретичні пояснення загалом закономірні за своєю суттю.

Щоб зрозуміти, що таке теорія, важливо розуміти, що не є теорією. Теорія не є емпіричними даними, фактами, класифікацією чи таксономією. Набір даних – це не теорія, як купа каміння не є будинком. Аналогічно, набір конструктів не є теорією, оскільки теорії, окрім побудови конструктів, повинні включати теоретичні положення, пояснення та граничні умови. Дані, факти й результати відносяться до емпіричного рівня, або до рівня спостережень, тоді як теорії діють на концептуальному рівні й ґрунтуються скоріше на логіці, ніж на спостереженнях.

Використання теорій у дослідженнях відкриває багато переваг. По-перше, теорії відкривають внутрішню логіку, що лежить в основі феноменів природи чи суспільства, пояснюючи що і чому є ключовими рушійними силами та ключовими результатами досліджуваного феномену, які процеси обумовлюють його виникнення й розвиток. По-друге, вони допомагають зрозуміти сутність феномену, синтезуючи попередні емпіричні дані в теоретичній структурі і погоджуючи суперечливі результати шляхом виявлення можливих факторів, що впливають на зв'язок між двома конструктами в різних дослідженнях. По-третє, теорії задають напрям подальших досліджень, допомагаючи встановити конструкти та зв'язки, що заслуговують на більш детальне вивчення. По-четверте, теорії можуть зробити внесок в побудову спільного знання, закриваючи прогалини між іншими теоріями і спонукаючи до переоцінки існуючих теорій у новому світлі.

Однак теорії теж мають свої обмеження. Як спрощені пояснення дійсності, вони не завжди адекватно розкривають причини конкретних досліджуваних феноменів, спираючись на обмежений набір конструктів і зв'язків. Теорії створюються для простих і економних пояснень, тоді як дійсність може бути набагато складнішою. Окрім того, теорії можуть накладати свого роду «шори», обмежуючи поле зору дослідника. У результаті залишаються непоміченими важливі концепти, які не визначені конкретною теорією.

Побудова блоків теорії

Девід Уеттен (1989) припустив, що теорія складається з чотирьох блоків: конструкти, теоретичні положення, логіка й обмежувальні умови (припущення). Конструкти вказують, що саме відображають теорії, тобто які концепти важливі для пояснення досліджуваного феномену. Положення відображають, як саме концепти пов'язані між собою. Логіка відповідає на питання, чому саме ці концепти пов'язані, а обмежувальні умови (припущення) з'ясовують обставини, за яких концепти та зв'язки між ними працюють. Хоча загальна характеристика конструктивів і теоретичних положень була розглянута раніше в розділі 2, зараз вони будуть описані детальніше.

Конструкти – це абстрактні концепти, визначені на високому рівні абстрагування, які спеціально обрані для пояснення досліджуваного феномена. Згадайте з розділу 2, що конструкти можуть бути одновимірними, тобто містити лише один концепт, як, наприклад, вага чи вік, або багатовимірними, тобто містити декілька концептів, що лежать в їх основі, наприклад, особистість чи культура. Якщо одні конструкти легко зрозуміти, інші, такі як креативність чи інноваційний потенціал є більш складними й абстрактними, водночас, такі конструкти як цінності, ставлення, навченість можуть представляти швидше тимчасові тенденції, ніж постійний стан. Тим не менш, всі конструкти повинні мати точні й однозначні операційні визначення, які однозначно визначають, як саме конструкт буде вимірюватися і на якому рівні аналізу – індивідуальному, груповому, організаційному і т.п. Змінні є вимірюваними представниками абстрактних конструктивів. Наприклад, бал за шкалою інтелекту є змінною, що була запропонована для вимірювання абстрактного конструкту, який називається інтелект. Як відзначалось раніше, наукове дослідження виконується у двох площинах: теоретичному й емпіричному. Конструкти виділяються в теоретичній площині, тоді як змінні встановлюються та вимірюються на емпіричному рівні (спостереження). Відмінності між конструктами (концептуалізованими в теоретичній площині) і змінними (вимірюваними в емпіричній площині), показані на рис. 4.1.

Теоретичні положення – це асоціації, що постулюються між конструктами, базуючись на дедуктивній логіці. Вони формулюються в декларативній формі і в ідеалі повинні визначати причинно-наслідковий зв'язок (на кшталт, якщо сталася подія X, тоді її наслідком буде подія Y). Хоча пропозиції можуть бути гіпотетичні, вони обов'язково підлягають тестуванню і відхиляються, якщо не підтримуються емпіричними спостереженнями. Однак, подібно до конструктивів, положення визначаються на теоретичному рівні, і вони можуть бути протестовані вивченням відповідного зв'язку між вимірюваними змінними конструктивів. Емпіричне формулювання теоретичних положень, що має форму зв'язку між змінними, називається **гіпотезою**. Відмінність між

положеннями, що формулюються на теоретичному рівні, і гіпотезами, що формулюються на емпіричному рівні, проілюстрована на рис. 4.1.

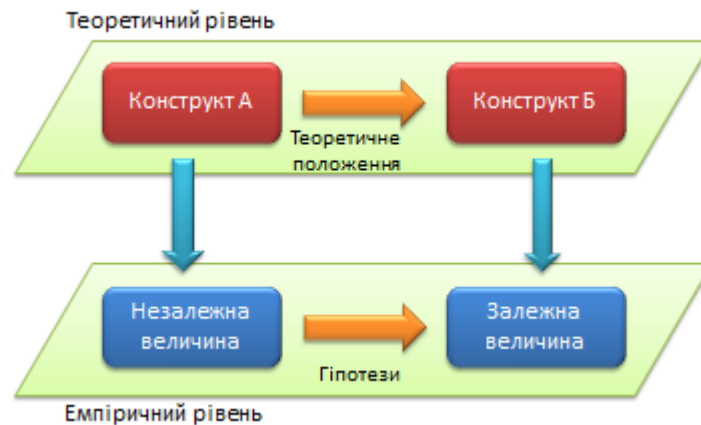


Рис.4.1. Відмінності між теоретичними та емпіричними концептами

Третій будівельний блок теорії – це **логіка**, яка забезпечує основу для підтвердження постульованих положень. Логіка діє як «клей», що зв'язує теоретичні конструкти і забезпечує змістовність та релевантність зв'язків між конструктами. Логіка також представляє пояснення, що лежить в основі теорії. Без логіки теоретичні положення будуть випадковими, суб'єктивними та беззмістовними, а отже, не можуть бути поєднані в єдину «систему теоретичних положень», що є серцем теорії.

Нарешті, все теорії обмежені **припущеннями** стосовно цінностей, часу та форми, а також умовами, які визначають межі застосування теорії. Наприклад, багато економічних теорій припускають, що люди є раціональними (чи досить раціональними істотами), а тому, щоб зрозуміти поведінку людей, теорії звертаються до максимізації корисності, базуючись на співвідношенні витрат і результату. Політичні науки, навпаки, вважають, що люди більш політичні, ніж раціональні, і тому люди намагаються позиціонувати себе в професійному чи особистому оточенні в такий спосіб, щоб максимізувати власну владу й контроль над іншими людьми. Враховуючи природу їхніх базових припущень, економічні та політичні теорії не можуть порівнюватися безпосередньо, і дослідники не повинні звертатися до економічних теорій, якщо їхня мета полягає в розумінні структури влади чи її еволюції в організації. Аналогічно, теорії можуть спиратися на неявні культурні припущення (наприклад, що вони застосовуються до особистої чи колективної культури), просторові припущення (наприклад, що вони застосовуються до певних громад, і не застосовуються до інших) і т.п. Якщо теорія застосовується чи перевіряється належним чином, всі неявні припущення, які формують межі використання теорії, мають бути належним чином зрозумілі. На жаль, теоретики рідко надають точні

формулювання своїх неявних припущень, що веде до невідповідного застосування теорій до проблемних ситуацій у наукових дослідженнях.

Ознаки хорошої теорії

Теорії представляють собою спрощене і часто неповне пояснення складних соціальних реалій, а тому можуть надавати краще чи гірше пояснення соціальних явищ і, відповідно, бути «хорошими» чи «поганими». Як оцінити, наскільки хорошою є теорія? Для цього були запропоновані різні критерії, найбільш важливі з яких наведені нижче:

- **Логічна узгодженість.** Чи не існує логічних протиріч між конструктами, теоретичними положеннями, обмежувальними умовами і припущеннями? Якщо деякі з цих будівельних блоків теорії не узгоджені між собою (наприклад, якщо теорія передбачає раціональність, однак деякі конструкти представляють нераціональні концепти), тоді теорія є недосконалою.
- **Пояснювальна потужність.** В якій мірі дана теорія пояснює чи прогнозує реальність? Хороші теорії, вочевидь, пояснюють досліджуваний феномен краще, ніж погані, що часто вимірюється коефіцієнтом детермінації в регресійному аналізі.
- **Можливість спростування.** Британський філософ Карл Поппер у 1940-х рр. стверджував, що теорії валідні, якщо вони можуть бути спростовані. Можливість спростування забезпечує, що теорія потенційно відхиляється, якщо емпіричні дані не узгоджуються з теоретичними положеннями, які можуть бути емпірично перевірені. Тавтології, на кшталт «день, коли температура повітря висока, є спекотним днем», не можуть бути перевірені емпірично, тому що спекотний день визначається і оцінюється як день з високою температурою повітря, отже, такі твердження не можуть вважатися теоретичними положеннями. Можливість спростування вимагає наявності альтернативних пояснень, забезпечує адекватну вимірюваність конструктів і т.п. Однак відзначимо, що твердження «теорія може бути спростованою» не означає, що теорія повинна бути спростованою. Якщо теорія дійсно спростовується на основі емпіричних доказів, тоді, вірогідно, проблема в поганій теорії!
- **Економічність.** Економічність досліджує, наскільки феномен пояснюється певною кількістю змінних. Концепт був запропонований англійським логіком Вільямом Оккамом у XIV ст. і тому часто називається «брита Оккама». Він стверджує, що серед конкуруючих пояснень, що задовільно пояснюють досліджуваний феномен, найпростіша теорія, яка використовує найменшу кількість змінних, або найменшу кількість обмежень, є найкращою. Пояснення складних соціальних феноменів завжди може бути ускладнене введеннями все нових і нових конструктів. Однак такий підхід суперечить меті отримання такої теорії, яка є

найпростішим і найбільш загальним поясненням дійсності. Економічність пов'язана зі ступенем свободи даної теорії. Економічні теорії мають вищий ступінь свободи, що дозволяє узагальнювати їх до рівня інших контекстів, умов і досліджуваних об'єктів.

Підходи до створення теорій

Як дослідники створюють теорії? Стейнфельд і Фальк (1990)² рекомендують чотири підходи до створення теорій. Перший полягає в побудові теорій індуктивно, спираючись на паттерни подій чи процесів, що спостерігаються в емпіричних дослідженнях. Такий підхід називається «обґрунтованою побудовою теорії», тому що теорія ґрунтується на емпіричних спостереженнях. Підхід значно залежить від здібностей дослідника до спостереження та інтерпретації, а отримана теорія може бути суб'єктивною й такою, яку неможливо спростувати. Крім того, спостереження окремих паттернів подій не обов'язково призводить до побудови теорії, якщо дослідник не здатний надати адекватне пояснення подій, що спостерігаються. Цей підхід буде розглянутий детальніше в матеріалах цього розділу, присвячених якісному аналізу.

Другий підхід до побудови теорій полягає в проведенні концептуального аналізу знизу вгору для ідентифікації різних наборів факторів, що впливають на досліджуваний феномен, використовуючи визначені структури. Однією такою структурою може бути проста структура типу «вхід-процес-вихід», де дослідник може вивчити різні категорії вхідних даних, таких як особистісні, організаційні і/чи технологічні фактори, що потенційно можуть бути пов'язані з досліджуваним феноменом, і описати процеси, що пов'язують ці фактори з досліджуваним феноменом. Це також індуктивний підхід, який значно залежить від індуктивних здібностей дослідника, а інтерпретація – від його попередніх знань щодо досліджуваного феномену.

Третій підхід до побудови теорій – це поширення чи модифікація існуючих теорій для пояснення нового контексту, наприклад, поширення теорії індивідуального навчання для пояснення організаційного навчання. Коли робиться таке поширення, одні концепти, положення і/чи обмежувальні умови старої теорії можуть залишитися без змін, тоді як інші модифікуються відповідно до нового контексту. Такий дедуктивний підхід примножує широкий набір соціальних теорій, що були розвинуті попередніми дослідниками, і є ефективним шляхом побудови нових теорій на основі існуючих.

²Steinfeld C.W., Fulk J. The Theory Imperative / Organizations and Communications Technology, C.W.Steinfeld, J.Fulk (eds.). –Newbury Park, CA: Sage Publications. – 1990.

Четвертий підхід полягає в застосуванні існуючих теорій у зовсім новому контексті, на підставі структурної схожості між двома контекстами. Цей підхід покладається на мислення за аналогією і, можливо, є найбільш креативним шляхом теоретизування з використанням дедуктивного підходу. Наприклад, Маркус (1987)³ запропонувала теорію критичної маси в зростанні мереж, використовуючи схожість за аналогією між ядерним вибухом і неконтрольованим ростом мереж чи заснованих на мережах бізнес-структурах. Як ядерний вибух вимагає критичної маси радіоактивної речовини, так і, за теорією Маркус, мережа вимагає критичної маси користувачів для забезпечення її росту, і без такої критичної маси користувачі можуть залишити мережу, спричинивши можливу смерть мережі.

Приклади теорій соціальних наук

У цьому підрозділі надається короткий огляд деяких ілюстративних теорій із різних соціальних дисциплін. Вони пояснюють різні типи соціальних процесів, використовуючи набір конструктивів, теоретичних положень, логіки й обмежувальних умов (припущень). Наступний матеріал є лише спрощеним вступом до цих теорій, і читачам рекомендується звернутися до першоджерел, щоб ознайомитися з усіма деталями й особливостями кожної теорії.

Теорія агентських відносин. Теорія агентських відносин, яка також називається теорією принципала-агента, є класичною теорією в економіці організацій. Вона була запропонована Россом (1973)⁴ для пояснення двосторонніх зв'язків (наприклад, між роботодавцем і працівником, між менеджментом організації й акціонерами, між продавцями і покупцями), чії цілі не співпадають між собою. Мета теорії агентських відносин полягає в тому, щоб конкретизувати оптимальні контракти й умови, за яких ці контракти допоможуть мінімізувати ефект невідповідності цілей сторін. Ключова передумова цієї теорії в тому, що люди є егоїстичними, достатньо раціональними і не схильними до високого ризику. Теорія може застосовуватися на рівні як індивідуальних, так і організаційних відносин.

У ній двома сторонами виступають принципал і агент. Принципал наймає агента для виконання певних завдань. Якщо мета принципала полягає у швидкому й ефективному завершенні завдання, то мета агента може бути в тому, щоб працювати в зручному для себе темпі, уникати ризиків і ставити власні інтереси (як, наприклад, заробітна плата) вище за

³Markus M. L. Toward a 'Critical Mass' Theory of Interactive Media: Universal Access, Interdependence, and Diffusion / M. L. Markus // *Communication Research*. – 1987. – Vol. 14:5. – P. 491-511.

⁴Ross S. A. The Economic Theory of Agency: The Principal's Problem / S. A. Ross // *American Economic Review*. – 1973. – Vol. 63:2. – 134-139.

корпоративні. Отже, цілі двох сторін не співпадають. Складна й суперечлива природа проблеми може призвести до проблеми асиметричної інформації, яка виникає тому, що принципал не може постійно слідкувати за поведінкою агента або точно оцінити його професійні вміння. Така асиметрія може призвести до того, що агент не буде докладати необхідних зусиль для виконання поставленого завдання (проблема моральної шкоди) або завищить оцінку своїх компетенцій, щоб отримати роботу, однак не зможе її виконувати належним чином (проблема неправильного відбору). Типові контракти, що базуються на поведінці працівника, такі як щомісячна заробітна плата, не можуть подолати ці проблеми. Отже, теорія агентських відносин рекомендує використання контрактів, що засновані на результаті, таких як комісії чи винагорода за результатом, або комбіновані контракти, що поєднують мотивацію, засновану на поведінці та на отриманих результатах. План придбання акцій для працівника є прикладом контракту за результатами діяльності і гарною альтернативною контрактів, заснованих на поведінці працівника. Теорія агентських відносин також рекомендує набір інструментів, що можуть бути застосовані роботодавцями для підвищення ефективності контрактів, заснованих на поведінці. Прикладами таких інструментів є інвестування в механізм моніторингу за роботою працівників шляхом наймання контролерів для подолання інформаційної асиметрії, дизайн поновлюваних контрактів, що залежать від результатів діяльності агента (оцінка діяльності робить контракт частково залежним від результатів), покращення структури робочих завдань, щоб зробити їх такими, що піддаються контролю й оцінці і т.п.

Теорія запланованої поведінки. Поширена в літературі з соціальної психології, теорія запланованої поведінки (ТЗП) була розроблена Еженом (1991)⁵. Вона є узагальненою теорією людської поведінки й може бути використана для вивчення широкого діапазону організаційних ситуацій. Теорія передбачає, що індивідуальна поведінка є свідомим вибором людей і формується внаслідок когнітивного мислення і соціального тиску. Теорія постулює, що поведінка базується на намірах людини щодо своєї поведінки, яка, в свою чергу, є функцією ставлення людини до своєї поведінки, суб'єктивних норм і очікуваного контролю (рис. 4.2). Ставлення визначається як загальні позитивні чи негативні відчуття людини щодо вибору того чи іншого стилю поведінки, яке може бути оціненим сумою переконань людини стосовно різних наслідків поведінки, зваженою на бажаність цих наслідків. Суб'єктивні норми відносяться до індивідуальних очікувань стосовно того, чи люди, важливі для даної особи, очікують від неї певної поведінки. Суб'єктивні норми представлені як зважена комбінація очікуваних норм різних референтних груп, таких як колеги, друзі чи керівники. Поведінковий контроль є

⁵Ajzen I. The Theory of Planned Behavior / I. Ajzen //Organizational Behavior and Human Decision Processes.— 1991. — Vol. 50. — P. 179-211.

індивідуальним очікуванням внутрішнього чи зовнішнього контролю, що обмежує поведінку особи. Внутрішній контроль може включати здатність людини дотримуватися обраного стилю поведінки, тоді як зовнішній контроль відноситься до наявності зовнішніх ресурсів, потрібних для підтримки поведінки. ТЗП також передбачає, що іноді люди мають намір дотримуватися певної поведінки, однак не мають для цього ресурсів, і тому вважає, що поведінковий контроль може мати прямий ефект на поведінку, на додаток до непрямого ефекту, опосередкованого наміром.

ТЗП є розвитком попередньої теорії причинної дії, яка включала намір і суб'єктивні норми як ключові рушійні сили наміру, але не поведінкового контролю. Останній конструкт був доданий Еженом в ТЗП, щоб урахувати ту обставину, що люди можуть мати неповний контроль над своєю власною поведінкою (коли, наприклад, відсутність швидкісного Інтернету обмежує можливості доступу до певних веб-ресурсів).



Рис.4.2. Теорія запланованої поведінки

Теорія дифузії інновацій. Визначна для наукової літератури з комунікацій, теорія дифузії інновацій (ТДІ) пояснює, як інновації поширюються в різних культурах і суспільствах. Концепт був спочатку досліджений французьким соціологом Габріелем Тардом, однак сама теорія розвинута Евереттом Роджерсом в 1962 р. на основі синтезу даних 508 досліджень з дифузії. Чотири ключові елементи теорії – інновація, канали комунікації, час і соціальна система. Інновації можуть включати нові технології, практичну діяльність чи ідеї, а сприйматися як окремими особами, так і організаціями. На макрорівні ТДІ розглядає дифузії інновацій як процес комунікації, у якому члени соціальної системи дізнаються про інновацію та її потенційні переваги через певні канали комунікації (такі як засоби масової інформації чи перші користувачі) і готові її впровадити. Дифузія є процесом, що розгортається в часі. Процес дифузії повільно починається в середовищі небагатьох користувачів-послідовників, досягає піку, коли інновація приймається загальним потоком населення й уповільнюється, коли кількість послідовників досягає насичення. Кумулятивний паттерн сприйняття інновації має S-подібну форму, як показано на рис 4.3., а розподіл користувачів-

послідовників представляє собою нормальний розподіл. Користувачі відрізняються між собою, і можуть бути класифіковані в групи залежно від часу на сприйняття інновації. Швидкість дифузії також залежить від характеристик соціальної системи, таких як наявність лідерів, що формують громадську думку, й агентів змін, які впливають на поведінку інших.

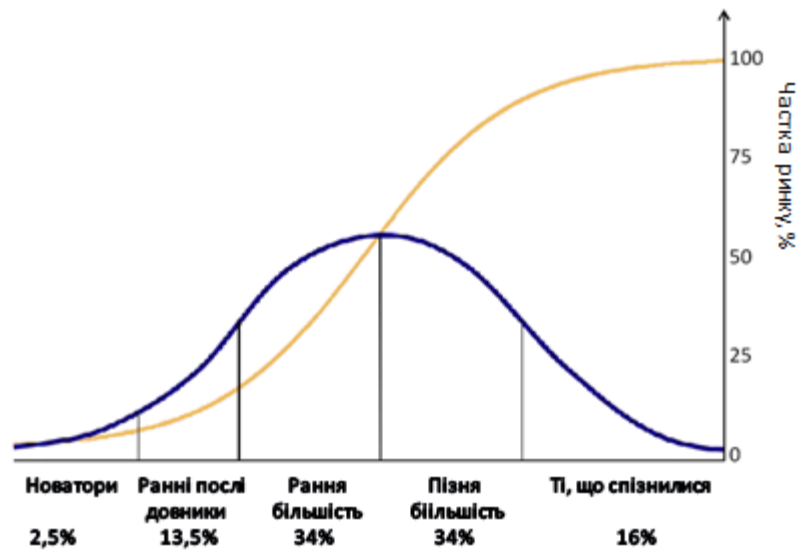


Рис. 4.3. S-подібна крива дифузії

Роджерс (1995)⁶ припустив, що на мікрорівні (рівні користувача) адаптація інновацій є процесом, який складається з п'яти етапів: 1) знання, коли користувачі вперше дізнаються про інновацію через засоби масової інформації чи канал міжособистісних комунікацій; 2) переконання: коли перші користувачі переконують інших спробувати інновацію; 3) рішення: робиться рішення щодо прийняття інновації; 4) впровадження: початкове використання інновації; 5) підтвердження: робиться рішення продовжити користуватися інновацією повною мірою (рис. 4.4.). П'ять характеристик інновації впливають на рішення користувача щодо адаптації інновації: 1) очікувані переваги інновації порівняно з попередніми інноваціями; 2) сумісність, тобто наскільки інновація сумісна з робочими звичками користувача, його переконаннями та цінностями; 3) складність, тобто наскільки складно ознайомитися та використовувати інновацію; 4) можливість перевірки, тобто якою мірою інновація може бути апробована до початку постійного використання; 5) наочність, тобто в якій мірі можна відслідкувати результати використання інновації. Останні дві характеристики часто опускаються в дослідженнях із інновацій. Складність негативно корелює з прийняттям інновацій, тоді як інші чотири фактори корелюють позитивно. Прийняття інновацій також залежить від особистих характеристик користувача, таких як

⁶Rogers E. Diffusion of Innovations / E. Rogers. — New York: The Free Press. — 1995. — Other editions 1983, 1996, 2005.

схильність до прийняття ризику, рівень освіти, космополітизм і вплив комунікацій. Перші користувачі (новатори) схильні до підприємництва, добре освічені і довіряють інформації про інновацію, що надається засобами масової інформації, тоді як користувачі «другої хвилі» більше покладаються на міжособистісні джерела інформації (друзів чи родину) як першоджерело інформації. ТДІ піддавалася критиці через наявність «про-інноваційної помилки», тобто через припущення, що всі інновації мають переваги і, врешті решт, будуть розповсюджені серед усіх споживачів, хоча насправді неефективні інновації можуть швидко зникнути з ринку без адаптації усіма користувачами або бути заміщеними кращими інноваціями до виходу на пікові значення.

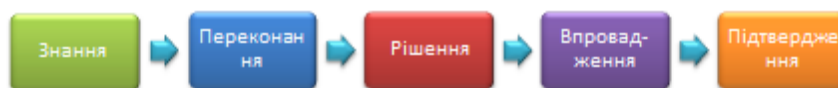


Рис. 4.4. Процес прийняття інновацій

Модель вірогідності розроблення. Запропонована Петті і Кациоппо (1986)⁷ модель вірогідності розроблення (МВР), є теорією подвійного процесу – формування ставлення або зміни. Вона пояснює, як можна впливати на особистостей, щоб змінити їхнє ставлення до певних об’єктів, подій чи процесів, і якою є відносна ефективність таких змін стратегії. МВР вказує, що чиєсь ставлення може бути сформовано двома «маршрутами» впливу – центральним і периферичним, які відрізняються обсягом обробки корисної інформації, або необхідним «розробленням» (рис. 4.5). Центральний маршрут вимагає від людини думати про пов’язані з даним об’єктом аргументи в інформаційному повідомленні й ретельно зважувати переваги та релевантність цих аргументів до того, як вона сформує обґрунтоване судження про досліджуваний об’єкт. На периферичному шляху у формуванні свого ставлення до досліджуваного об’єкту люди покладаються на зовнішні сигнали, такі як кількість попередніх користувачів чи схвалення експертів, а не на якість аргументів. Останній маршрут вимагає менших когнітивних зусиль і, згідно з МВР, маршрути зміни ставлення зазвичай операціоналізуються використанням відповідних конструктів – якістю аргументів і периферичними сигналами відповідно.

Який маршрут – центральний чи периферичний, – оберє людина, залежить від її здібностей і мотивації розробляти центральні переваги аргументу. Така здатність і мотивація до розроблення називається вірогідністю розроблення. Люди з високою вірогідністю розроблення більш схильні до глибокого осмислення представленої інформації, а отже,

⁷Petty R. E., Cacioppo J. T. Communication and Persuasion: Central and Peripheral Routes to Attitude Change/ R. E. Petty, J. T. Cacioppo.—New York: Springer Verlag.—1986.

більше зважають на якість аргументів, тоді як люди з низькою вірогідністю розроблення більш мотивовані периферичними сигналами. Вірогідність розроблення – це ситуаційна характеристика, а не риса особистості. Наприклад, лікар може вдатися до центрального маршруту для діагностики й лікування захворювання, спираючись на свій професійний досвід, але може покластися на периферичні сигнали від автомеханіка, щоб оцінити несправності свого автомобіля. Отже, теорія має широке застосування стосовно того, як попереднє ставлення змінюється по відношенню до нових продуктів, ідей і навіть соціальних процесів.

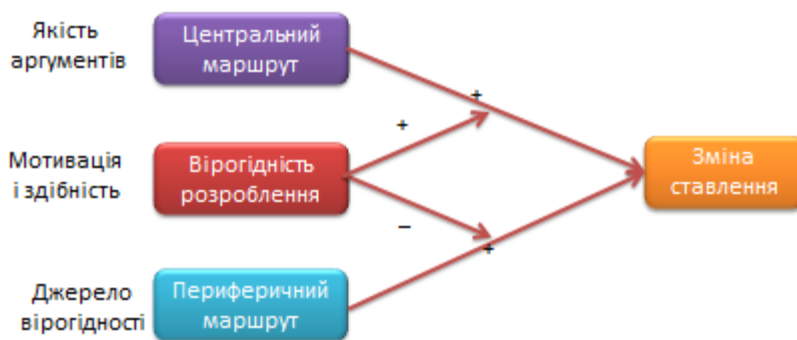


Рис. 4.5. Модель вірогідності розроблення

Загальна теорія покарання. Філософи XVIII ст. Цезаре Бессарія і Джереми Бентам сформулювали загальну теорію покарання (ЗТП) як пояснення злочинності й метод її зниження. Теорія досліджує, чому деякі особи здатні до злочинних асоціальних дій. Теорія передбачає, що люди за своєю природою діють раціонально (у випадку як соціально прийнятної, так і асоціальної поведінки). Оскільки люди зазвичай обирають поведінку, спрямовану на максимальну корисність, асоціальні вибори, спрямовані на отримання особистої вигоди чи задоволення, можуть контролюватися збільшенням ціни такої поведінки у вигляді покарання (контрзаходів) або підвищенням вірогідності покарання. Швидкість, тяжкість і неможливість уникнення покарання є ключовими конструктами в ЗСП.

Тоді як класичні позитивістські дослідження акцентують увагу на соціальних причинах злочинів (таких як бідність, низький рівень освіти, психологічні проблеми) і пропонують відповідні стратегії реабілітації злочинців, ЗТП фокусується на процесі прийняття рішень злочинцями і ситуаційних факторах, що впливають на цей процес. Отже, особисті обставини злочинця (такі як низькі моральні цінності, агресивність чи потреба в грошах) і контекстуальні впливи зовнішнього середовища (захищеність потенційної жертви злочину, ефективність правоохоронних органів і вірогідність того, що злочинець буде звинувачений) відіграють важливу роль у процесі прийняття рішень. Сутність ЗТП не в реабілітації злочинців чи попередженні злочинів у

майбутньому, а в тому, як зробити злочинну діяльність менш привабливою. У цьому відношенні посилення заходів безпеки, таких як встановлення камер відеоспостереження, розвиток навичок самозахисту, робота правоохоронних органів із виявлення мереж розповсюдження наркотиків і посилене патрулювання неблагополучних районів є високоефективним для попередження злочинів. Ця теорія може бути застосована й по відношенню до злочинів «білих комірців», таких як продаж конфіденційної інформації, використання піратського програмного забезпечення чи незаконне розповсюдження відеопродукції.

Розділ 5

ДИЗАЙН ДОСЛІДЖЕННЯ

Дизайн дослідження – це комплексний план для збору даних в емпіричному дослідженні. Це своєрідна «мапа» для проведення емпіричних досліджень спрямована на те, щоб відповісти на конкретні питання дослідження або протестувати конкретні гіпотези, і має визначити щонайменше три процеси: 1) процес збору даних; 2) процес вибору чи розроблення інструментів дослідження; 3) процес формування вибірок. Останні два процеси описуються в наступних розділах, тоді як процес збору даних (який часто неправильно називають «дизайном дослідження»), розглядається в цьому розділі та більш детально – у розділах 9-12.

У загальному сенсі методи збору даних можуть бути згруповані у дві великі категорії: позитивістські та інтерпретативні. Позитивістські методи, такі як лабораторні експерименти й анкетні дослідження, спрямовані на тестування теорій (чи гіпотез), тоді як інтерпретативні методи, такі як дослідження дією й етнографія спрямовані на створення теорій. **Позитивістські методи** спираються на дедуктивний підхід до досліджень, починаючи від теорії та тестування теоретичних постулатів із використанням емпіричних даних. **Інтерпретативні методи**, навпаки, спираються на індуктивний підхід, який бере початок у даних, і намагаються побудувати теорію про досліджувані об'єкти на основі отриманих емпіричних даних. Дуже часто ці методи неправильно прирівнюються до кількісних і якісних досліджень. Кількісні та якісні методи відносяться до типу даних, що збираються (кількісні дані включають числові дані, рейтингові бали і т.п., тоді як якісні дані включають інтерв'ю, спостереження і т.п.), і до того, як ці дані аналізуються (використовуючи кількісні методи, такі як регресійний аналіз, чи якісні техніки, такі як кодування). Позитивістське дослідження користується переважно кількісними даними, однак може залучати і якісні дані. Інтерпретативне дослідження покладається значним чином на якісні дані, хоча використання кількісних даних теж не виключається. Іноді одночасне використання кількісних і якісних даних може призвести до отримання унікальних інтуїтивних здогадок щодо природи складних соціальних феноменів, які не вдалося б отримати, збираючи лише дані одного типу, отже змішаний дизайн, що поєднує кількісні та якісні дані, часто є найбажанішим.

Ключові характеристики дизайну дослідження

Якість дизайну дослідження може бути визначена за чотирима ключовими характеристиками: внутрішня валідність, зовнішня валідність, валідність конструктів і валідність статистичного висновку.

Внутрішня валідність, яка також називається причинністю, оцінює, чи дійсно зміни, що демонструє залежна змінна, викликані відповідними змінами в гіпотетичній незалежній змінній. Причинність вимагає дотримання трьох умов: коваріація причини й ефекту, тобто якщо має місце причина, то має місце й ефект; 2) часова послідовність: у часовому вимірі причина має передувати ефекту; 3) не існує іншого правдоподібного альтернативного пояснення. Певні дизайни дослідження, такі як лабораторні експерименти, мають сильну внутрішню валідність завдяки своїй здатності маніпулювати незалежною змінною (причиною) в окремих спробах і спостерігати ефект (залежну змінну) протягом певного часу після завершення спроби, контролюючи впливи зовнішніх змінних. Інші дизайни, такі як виробничі дослідження, мають низьку внутрішню валідність через неможливість маніпулювати незалежною змінною (причиною) і тому, що причина й ефект вимірюються одночасно, що порушує часову послідовність і уможливорює ситуацію, що навпаки, очікуваний ефект впливає на причину. Порівняно з іншими методами, лабораторні експерименти менш чутливі до помилок внутрішньої валідності, хоча й чутливі до передісторії, тестування, інструментарію, регресії та інших загроз, що будуть обговорюватися надалі в цьому розділі в матеріалах стосовно дизайну експериментів. Тим не менш, різні дизайни експериментів значно варіюють за рівнями внутрішньої валідності.

Зовнішня валідність, або рівень узагальненості, стосується того, наскільки виявлені у вибірці закономірності можуть бути узагальнені до рівня генеральної сукупності (валідність генеральної сукупності), або перенесені на інших людей, організації, контексти і час (екологічна валідність). Наприклад, чи можуть результати, отримані при дослідженні вибірки фінансових компаній, бути узагальнені до рівня країни (екологічна валідність)? Анкетні дослідження, у яких дані отримуються від широкого кола осіб, компаній чи інших об'єктів, мають тенденцію до більшої узагальненості порівняно з лабораторними експериментами зі штучно встановленими умовами та жорстким контролем сторонніх змінних, що знижує можливість узагальнення їх результатів до реальних бізнес-ситуацій, де конкретні умови та сторонні змінні не піддаються контролю. Варіації внутрішньої і зовнішньої валідності для широкого спектру дизайнів дослідження показані на рис. 5.1.

Деякі дослідники вважають, що між внутрішньою і зовнішньою валідністю є компроміс: підвищення зовнішньої валідності можливе лише за рахунок зменшення внутрішньої валідності й навпаки. Однак це не завжди так. Деякі дизайни дослідження, такі як виробничі експерименти, виробниче анкетування в динаміці, багатокейсові дослідження мають високі ступені внутрішньої та зовнішньої валідності. Перевагу потрібно надавати тим дизайнам дослідження, які мають достатньо високий ступінь внутрішньої та зовнішньої валідності, тобто тим, що знаходяться

в конусі валідності, показаній на рис. 5.1. Однак це не означає, що дизайни поза межами цього конусу є менш корисними та цінними. Вибір дизайну дослідниками є питанням їхніх особистих преференцій і компетенції, і вони можуть обрати той рівень внутрішньої та зовнішньої валідності, який бажають.



Рис. 5.1. Внутрішня і зовнішня валідність

Валідність конструкту показує, наскільки добре певна шкала вимірює теоретичний конструкт, для вимірювання якого вона була створена. Багато конструктів, що використовуються в соціальних дослідженнях, такі як опір змінам, креативність чи організаційне навчання, складно визначити і ще складніше виміряти. Наприклад, валідність конструкту має забезпечити, щоб показник креативності вимірював саме креативність, а не інтелектуальні здібності, що може бути достатньо складно, оскільки ці конструкти близькі за значенням. Валідність конструкту оцінюється в позитивістських дослідженнях на основі кореляційного й факторного аналізів даних пілотного проекту, як це буде описано в наступному розділі.

Валідність статистичного висновку оцінює ступінь валідності висновків, зроблених при використанні певної статистичної процедури. Наприклад, вона оцінює, чи правильний статистичний метод був використаний для тестування гіпотез, чи відповідають змінні обмеженням даного статистичного тесту (такі як розмір вибірки чи вимоги до розподілу) і т.п. Оскільки інтерпретативні дизайни дослідження не передбачають використання статистичних тестів, по відношенню до них поняття валідності статистичного висновку не застосовується. Різні види валідності та сфери їх існування на теоретичному й емпіричному рівнях проілюстровані на рис. 5.2.

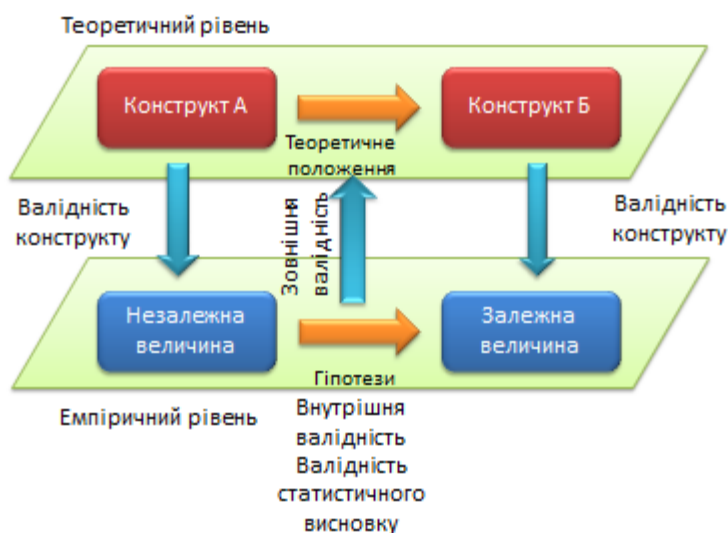


Рис. 5.2. Види валідності в наукових дослідженнях

Підвищення внутрішньої та зовнішньої валідності

Найкращими дизайнами дослідження є ті, що можуть забезпечити високі рівні внутрішньої і зовнішньої валідності. Такі дизайни запобігають хибним кореляціям, гарантують більшу об'єктивність у тестуванні гіпотез і забезпечують те, що результати, отримані на невеликих вибірках, можна узагальнити до рівня популяції. Для забезпечення внутрішньої валідності (причинності) в дизайні дослідження необхідний контроль, який може бути виконаний декількома різними способами: 1) маніпулювання, 2) усунення, 3) включення, 4) статистичний контроль, 5) рандомізація.

Маніпулювання означає, що дослідник змінює незалежні величини (стимули) на одному чи декількох рівнях, що називаються спробами, і порівнює результати їх дії з результатами контрольних груп, які не знаходилися під впливом дії досліджуваних незалежних величин. Стимулом, наприклад, може бути нова методика навчання, яка протягом різного періоду часу застосовувалася в експериментальних групах, тоді як контрольна група навчалася за традиційною методикою. Цей тип контролю досягається в експериментальному чи квазі-експериментальному дизайнах, але не в неекспериментальних дизайнах, якими є, наприклад, анкетні опитування.

Усунення (елімінація) означає видалення сторонніх величин завдяки тому, що вони мають постійні значення при всіх спробах, наприклад, за рахунок включення в досліджувану групу осіб однієї професії чи з однаковим соціально-економічним статусом. За способом **включення** роль сторонніх величин враховується їх включенням в дизайн дослідження і окремим оцінюванням ефектів їх впливу на залежну величину, як, наприклад у факторному аналізі. Такий спосіб дозволяє досягти більшого рівня узагальненості, проте, вимагає суттєво більших

вибірок. У **статистичному контролі** сторонні величини вимірюються та використовуються як коваріантні величини на етапі статистичного тестування.

Спосіб **рандомізації** спрямований на те, щоб усунути впливи сторонніх величин шляхом випадкового формування вибірок, якщо можна гарантувати, що ці впливи мають випадкову, несистематичну природу. Двома типами рандомізації є: 1) випадковий вибір, при якому вибірка з популяції формується випадково і 2) випадковий розподіл, при якому об'єкти добираються не випадково, однак випадково потрапляють до різної експериментальної та контрольної груп.

Рандомізація також забезпечує зовнішню валідність, дозволяючи узагальнення закономірностей, установлених для вибірки, до рівня популяції (генеральної сукупності), із якої вибірка була взята. Випадковий розподіл у групі є обов'язковим, якщо випадковий вибір неможливий через обмеженість ресурсів чи доступу. Однак у цьому випадку складніше забезпечити узагальненість результатів між популяціями, оскільки популяції можуть відрізнятися за багатьма показниками, а в дослідженні можна контролювати лише деякі з цих показників.

Поширені дизайни дослідження

Як відзначалося раніше, залежно від цілей дизайни дослідження можна класифікувати у дві категорії: позитивістські та інтерпретативні. Позитивістські дизайни слугують для перевірки теорій, тоді як інтерпретативні – для побудови теорій. Позитивістські дизайни спрямовані на пошук узагальнених паттернів на основі об'єктивного аналізу реальності, тоді як інтерпретативні дизайни пропонують суб'єктивні інтерпретації соціальних феноменів із точки зору учасників дослідження. Деякі поширені приклади позитивістських дизайнів включають лабораторні експерименти, виробничі експерименти та анкетування, повторний аналіз даних і дослідження ситуацій (кейсів), тоді як прикладами інтерпретативних дизайнів є дослідження ситуацій, феноменологія й етнографія. Зверніть увагу, що дослідження кейсів може використовуватися як для побудови, так і для тестування гіпотез, однак не одночасно. Не всі методики підходять для всіх видів наукових досліджень. Деякі методики, як, наприклад, фокус-групи, краще підходять для пошукових досліджень, інші, як етнографія, - для описових досліджень, а, наприклад, лабораторний експеримент є найкращим для пояснювальних досліджень. Далі будуть описані деякі з цих дизайнів. Додаткова інформація надається в розділах 9-12.

Експериментальні дослідження – це дослідження, спрямовані на тестування причинно-наслідкових зв'язків (гіпотез) у жорстко контрольованих умовах шляхом відділення причини від наслідку в часі.

При цьому одна група досліджуваних (експериментальна група) піддається дії досліджуваного фактора, або причини, тоді як інша (контрольна група) – ні. Дослідником фіксуються усереднені результати, або наслідки в цих групах. Більш складні дизайни можуть включати декілька експериментальних груп. У **справді експериментальному дизайні** досліджувані повинні випадково розподілятися в групи. Якщо випадковий добір не можливий, тоді дизайн перетворюється на **квазі-експериментальний**. Експерименти можуть проводитися в спеціально створених умовах у лабораторії (лабораторні експерименти), або у виробничих умовах, наприклад, на певному підприємстві (виробничий експеримент). Лабораторні умови дозволяють досліднику ізолювати досліджувані показники та контролювати сторонні показники, що не завжди можливо у виробничих експериментах. Отже, висновки, зроблені на основі лабораторних експериментів, мають тенденцію до вищої внутрішньої валідності, а виробничі експерименти мають тенденцію до вищої зовнішньої валідності. Експериментальні дані аналізуються за допомогою кількісних статистичних методів. Важлива перевага експериментального дизайну полягає в його високій внутрішній валідності завдяки здатності ізолювати, контролювати та інтенсивно досліджувати невелику кількість показників, тоді як його головний недолік – обмежена зовнішня валідність через складність узагальнення отриманих результатів, оскільки реальне життя часто набагато складніше (тобто, включає більше сторонніх величин), ніж штучні лабораторні умови. Більш того, якщо дослідження не встановлює не релевантні сторонні величини й не контролює їх, то це може зменшити внутрішню валідність і призвести до хибних кореляцій.

Виробничі опитування не є експериментальними дизайнами, вони не контролюють і не маніпулюють незалежними величинами чи умовами, однак вимірюють ці величини та перевіряють їх вплив, застосовуючи статистичні методи. Виробничі дослідження фіксують «моментальну фотографію» дій, переконань чи ситуацій із випадкової вибірки досліджуваних у виробничих умовах за допомогою анкетних опитувань, або, рідше – структурованих інтерв'ю. У крос-секційних виробничих дослідженнях незалежні та залежні величини вимірюються одночасно (наприклад, в одній анкеті), тоді як у динамічних (longitudinal) виробничих дослідженнях залежні величини вимірюються пізніше, ніж незалежні. Перевагами виробничих досліджень є зовнішня валідність (оскільки дані збираються в реальних умовах), здатність виявляти та контролювати велике число показників, а також досліджувати проблему комплексно, спираючись на численні теорії. Однак завдяки труднощам з доведенням внутрішньої валідності (зв'язок причина-наслідок) й уникненням помилок респондентів (досліджувані можуть, наприклад, надавати не правдиві, а соціально бажані відповіді), внутрішня валідність цих досліджень страждає.

Вторинний аналіз даних – це аналіз даних, що раніше були зібрані й оброблені іншими авторами. Такі дані можуть включати дані державних установ, а також дані інших авторів, які часто використовуються в мета-аналітичних дослідженнях, або дані з відкритих джерел, наприклад, біржових відомостей. Цим вторинні дослідження суттєво відрізняються від інших дизайнів, де збір первинних даних є частиною роботи дослідника. Вторинний аналіз даних може бути ефективним засобом досліджень, якщо первинний збір даних занадто дорогий або неможливий, тоді як вторинні дані є на рівні аналізу, що дозволяє отримати відповіді на питання, поставлені в даному дослідженні. Недоліки цього дизайну полягають у тому, що задіяні дані не обов'язково збиралися в систематичний і науковий спосіб, а отже, можуть не підходити для наукових досліджень і не узгоджуватися з цілями даного дослідження. Проблематичною є інтервальна валідність вторинного аналізу даних, оскільки часові співвідношення між причиною і наслідком неочевидні.

Кейс-дослідження – це поглиблене дослідження проблеми в одній чи декількох реальних ситуаціях протягом певного періоду часу. Дані можуть збиратися як комбінація інтерв'ю, особистих спостережень й аналізу внутрішніх чи зовнішніх документів. За своєю природою кейс-дослідження може бути позитивістським (для тестування гіпотез) чи інтепретативним (для побудови теорій). Сила цього методу полягає у здатності встановлювати широкий спектр соціальних, культурних і політичних факторів, які потенційно пов'язані з досліджуваним феноменом, однак невідомі для дослідника заздалегідь. Аналіз кейсів має тенденцію бути якісним за своєю природою, однак значно залежить від контексту та має багато нюансів. Інтерпретація результатів може залежати від спостережливості дослідника та його здатності до узагальнення. Нестача контролю може ускладнити встановлення причинності, і результати одного кейсу можуть виявитися недостатніми для узагальнення в інших ситуаціях. Рівень узагальненості може зрости завдяки повторенню та порівнянню результатів аналізу інших кейсів у **багатокейсових дизайнах**.

Дослідження фокус-груп є таким типом досліджень, що передбачає створення групи з 6-10 осіб і організацію в ній дискусії щодо досліджуваного феномену протягом 1,5-2 годин. Дискусія спрямовується кваліфікованим ведучим-фасилітатором, який встановлює порядок денний зустрічі, ставить перед учасниками ряд питань, забезпечує можливість кожного представити свої ідеї та поділитися досвідом, намагається побудувати цілісне розуміння проблемної ситуації на основі коментарів і власного досвіду учасників. Внутрішня валідність цього дизайну не може бути встановлена через недостатній контроль, а результати не можуть бути узагальнені до інших ситуацій через невеликий розмір групи учасників. Отже, фокус-групи зазвичай не

використовуються при проведенні пояснювальних і описових досліджень, а більш доречні для досліджень пошукового типу.

Дослідження дією передбачає, що складний соціальний феномен краще розуміється через впливи або активні дії дослідника стосовно феномену і спостереження за результатами цих дій. У цьому методі дослідником зазвичай виступає зовнішній консультант або член організації, який занурений в соціальний контекст організації, ініціює певні дії (такі як нові організаційні процедури чи нові технології) у відповідь на реальну проблему (наприклад, падіння прибутку або труднощі зі збутом). Вибір дослідником конкретних дій має базуватися на теорії, яка пояснює, чому та як ці дії можуть призвести до бажаних змін. Дослідник спостерігає за результатами дій, вносить у разі потреби зміни й одночасно навчається, генеруючи теоретичні уявлення щодо досліджуваної проблеми й ефекту запроваджених дій. Первинна проблема валідизується настільки, наскільки успішно обрані дії вирішують досліджувану проблему. Одночасне вирішення проблеми та генерування інтуїтивних уявлень про неї є ключовою рисою, що відрізняє дослідження дією від інших методів дослідження, а отже, дослідження дією є прекрасним методом, що, як міст, зв'язує наукові дослідження з практикою. Цей метод також доречний для вивчення унікальних соціальних проблем, що не можуть бути відділені від контексту, у якому вони існують. Водночас метод є предметом помилок і суб'єктивізму дослідника. У підсумку узагальненість результатів дослідження дією часто обмежена тим контекстом, у якому проводилося конкретне дослідження.

Етнографія – це інтерпретативний дизайн дослідження, який виник за прикладом антропології й підкреслює, що досліджуваний феномен повинен вивчатися в контексті певної культури. Дослідник глибоко занурюється в певну культуру протягом тривалого періоду часу (від 8 місяців до 2 років) і в цей період фактично живе життям досліджуваних. Спостерігаючи та фіксуючи особливості їх культури, він робить теоретичні висновки стосовно еволюції та соціальних процесів у цій культурі. Дані збираються, головним чином, методом спостереження, фіксацією особливостей формальної та неформальної взаємодії між учасниками спільноти. Аналіз даних включає пояснення, або «визначення змісту». Дослідник повинен до найменших дрібниць документувати свої спостереження і досвід, щоб інші особи могли повною мірою відчути особливості даної культури. Переваги етнографії полягають у високій чутливості до контексту, глибокому й насиченому деталями розумінні досліджуваного феномену та мінімальній похибці відповідей респондентів. Однак це надзвичайно витратний підхід із точки зору часу й ресурсів, а результати, добре відображаючи особливості даної культурної спільноти, не можуть бути узагальнені до інших культур.

Вибір дизайну дослідження

То який же дизайн дослідження повинен обрати науковець, ураховуючи різноманітність дизайнів дослідження? Зазвичай, дослідники намагаються обрати ті дизайни дослідження, із якими вони найбільш обізнані і мають досвід їх проведення, однак у ідеалі вибір повинен робитися, виходячи з сутності досліджуваного феномену. На попередніх етапах дослідження, коли проблема ще не чітко визначена, і дослідник намагається встановити її природу та масштаб, фокус-групи (коли об'єктом дослідження є особистість), або кейс-дослідження (коли об'єктом дослідження є організація) є ідеальною стратегією. Коли дослідник глибше пізнає досліджуваний феномен, розуміє, що не існує теорій для його пояснення, і ставить перед собою мету побудувати таку теорію, що закрие прогалину в даній галузі знань, корисними можуть виявитися такі дизайни, як кейс-дослідження чи етнографія. Якщо існує декілька альтернативних теорій для пояснення досліджуваного феномену, і науковець прагне їх перевірити або інтегрувати в одну, більш загальну теорію, більш доречними будуть такі позитивістські дизайни, як експеримент, опитування чи вторинний аналіз даних.

Незалежно від типу обраного дизайну дослідження науковець повинен прагнути зібрати кількісні та якісні дані, використовуючи комбінацію таких методик, як анкетування, інтерв'ю, спостереження, документи й дані інших джерел. Наприклад, навіть у дуже структурованій анкеті, розробленій для збору кількісних даних, дослідник може залишити декілька відкритих питань. Це дозволить отримати якісні дані, що можуть наштовхнути на неочікувані здогадки, які не можна зробити лише на основі структурованих кількісних даних. Аналогічно, якщо кейс-дослідження передбачає, головним чином, особисте інтерв'ю для збору якісних даних, не можна ігнорувати потенціал і цінність кількісних даних. Скажімо, під час дослідження процесу прийняття організаційних рішень дослідник може зафіксувати такі цифрові дані, як кількість днів для прийняття рішення, кількість осіб, що беруть участь у цьому процесі, кількість альтернативних варіантів, що розглядаються і т.п. Це може призвести до формулювання важливих припущень, які не можна зробити лише на основі аналізу неструктурованих і довільних за формою відповідей учасників інтерв'ю. Незалежно від конкретного дизайну дослідження метою науковця має бути збір якомога чисельніших і різноманітніших даних, що допоможуть генерувати найбільш вірогідні теоретичні пояснення досліджуваного феномену.

Розділ 6

ВИМІРЮВАННЯ КОНСТРУКТІВ

Теоретичні положення представляють собою зв'язки між абстрактними конструктами. Перевірка теорій (а отже, і теоретичних положень) вимагає точного, правильного та науково обґрунтованого вимірювання конструктів до того, як буде оцінюватися сила зв'язків між ними. Вимірювання означає спеціально проведене спостереження дійсності й становить сутність емпіричного дослідження. Якщо такі конструкти в дослідженнях соціальних наук, як форма власності чи розмір організації легко виміряти, інші конструкти, такі як креативність, інвестиційна привабливість чи інноваційний потенціал оцінити кількісно складніше. У цьому розділі будуть розглянуті процеси концептуалізації та операціоналізації, результатом яких є створення вимірюваних показників таких конструктів.

Концептуалізація

Концептуалізація – це процес мислення, завдяки якому розпливчасті та неточні конструкти (концепти) та їхні складові отримують чіткі й однозначні визначення. Наприклад, ми часто використовуємо термін «креативність», і цей термін викликає певний образ у нашому мозку, однак потрібно докласти зусиль, щоб надати точне визначення цього терміну. Якщо одна організація продукує нові ідеї частіше, ніж інші, то чи можна це трактувати як прояв організаційної креативності? Які ідеї слід враховувати під час оцінки організаційної креативності – будь-які нові ідеї, чи тільки ті, що є корисними з точки зору можливості їх упровадження у вигляді інновацій? Якщо є різні види організаційної креативності, то як можна їх визначити? Чи існують різні рівні креативності, наприклад, висока та низька? Відповіді на ці питання є ключовими для правильного виміру креативності. Процес розуміння того, що треба включати чи не виключати в набір концептів даного конструкту й називається процесом концептуалізації.

Процес концептуалізації особливо важливий, ураховуючи неточність, розпливчатість і суб'єктивність багатьох конструктів, що вживаються в соціальних науках. Визначення таких конструктів базується не на об'єктивних критеріях, а скоріше на загальній узгодженості між ментальними образами (концепціями) цих конструктів.

Визначаючи такі конструкти як креативність чи інноваційний потенціал, потрібно розуміти, що іноді вони не існують у реальному житті або не можуть існувати незалежно, оскільки є творіннями нашого розуму. Наприклад, багато людей не можуть дати точного визначення терміну «інноваційний потенціал», однак у наукових дослідженнях ми прагнемо розглядати цей конструкт як такий, що реально існує. Процес розгляду і вживання ментальних конструктів як реально існуючих

називається *матеріалізацією*, і він є центральним для визначення конструктів та ідентифікації їх вимірюваних показників.

Важливе рішення в концептуалізації конструктів полягає у встановленні того, є вони одновимірними або багатовимірними. Передбачається, що одновимірні конструкти можуть бути виміряні однією величиною. Прикладами є такі прості конструкти, як чистий дохід організації, прибуток до оподаткування тощо. Багатовимірні конструкти складаються з двох і більше складових. Наприклад, якщо академічна успішність учня може бути концептуалізована на основі двох складових – математична та лінгвістична успішність, тоді вона є багатовимірним конструктом. Кожна зі складових, що лежить в основі даного конструкту, може бути оцінена окремо, використовуючи різні тести з математики й мови. Для створення конструкту «академічна успішність» потрібно в подальшому об'єднати бали двох тестів, можливо, надаючи їм різної ваги.

Операціоналізація

Після того, як теоретичний конструкт визначений, постає питання про те, як його виміряти. **Операціоналізація** стосується процесу розвитку індикаторів, або показників для виміру конструктів. Наприклад, якщо такий «невидимий» теоретичний конструкт, як соціоекономічний статус родини визначається рівнем її доходу, то він може бути операціоналізованим за допомогою індикатора у вигляді питання «яким є ваш річний дохід?». Ураховуючи високий рівень суб'єктивності й неточність, властиву конструктам соціальних наук, ми намагаємося більшість із них оцінювати за допомогою індикаторів. Цей процес дозволяє дослідити тісноту зв'язку між індикаторами як показник їхньої точності (надійності).

Індикатори діють на емпіричному рівні, на відміну від конструктів, які концептуалізуються на теоретичному рівні. Комбінація індикаторів, яка на емпіричному рівні представляє певний конструкт, називається **змінною**. Як відзначалося в попередньому розділі, змінні можуть бути незалежними, залежними, модераторами чи посередниками, залежно від того, яку роль вони відіграють у дослідженні. Кожний індикатор може мати декілька рівнів (атрибутів), і кожний рівень має певне значення. Наприклад, показник «стать» може мати два рівні: жінки та чоловіки. Аналогічно шкала задоволення споживачів може бути сконструйована таким чином, щоб мати п'ять рівнів: «повністю незадоволений», «скоріше незадоволений», «нейтрально», «скоріше задоволений» і «повністю задоволений». Значення рівнів можуть бути кількісними (числовими) або якісними (нечисловими). Кількісні дані можна аналізувати, використовуючи методики кількісного аналізу даних (наприклад, регресія), тоді як аналіз якісних даних вимагає застосування відповідних методик (наприклад, кодування). Багато показників у дослідженнях соціальних наук є якісними, навіть якщо представлені в числовому

вигляді. Наприклад, можна розробити індикатор задоволеності споживачів із п'ятьма рівнями: «повністю незадоволений», «скоріше незадоволений», «нейтрально», «скоріше задоволений» і «повністю задоволений», а потім присвоїти п'ятьом рівням числа від 1 до 5 відповідно, що дозволяє потім застосувати складні статистичні інструменти для кількісного аналізу даних. Однак важливо розуміти, що цифри є тільки позначками, пов'язаними з персональними оцінками споживачами їхньої задоволеності, і показник, що лежить в основі конструкту, все ще є якісним, хоча й представлений числом.

Індикатори можуть бути рефлексивними або формативними. Рефлексивний індикатор є величиною, що «відображає» конструкт, який лежить в його основі. Наприклад, якщо інтелектуальний потенціал визначити як знання організації, що здатні створювати нову вартість, тоді прибуток, отриманий від застосування знань, може слугувати рефлексивним індикатором інтелектуального капіталу. Формативний індикатор – це величина, що «формує», або робить свій внесок у конструкт, який лежить в його основі. Такі індикатори можуть представляти різні складові досліджуваного конструкта. Наприклад, якщо інтелектуальний капітал визначений як сукупність людського й споживчого капіталу, тоді індикатори, що обираються для виміру кожної із складових інтелектуального капіталу, будуть формативними. Одновимірні конструкти можуть вимірюватися з використанням рефлексивних індикаторів, тоді як багатовимірні конструкти можуть вимірюватися як комбінація багатьох складових, хоча кожна з цих складових може вимірюватися за допомогою одного чи багатьох рефлексивних індикаторів.

Рівні вимірювання

Першочергове рішення, яке потрібно зробити для операціоналізації конструкта, полягає у визначенні рівня вимірювання. **Рівні вимірювання**, які також називають рейтинговими шкалами, відносяться до значень, які може приймати індикатор (хоча нічого не говорить власне про сам індикатор). Наприклад, чоловіки і жінки (Ч і Ж, або 1 і 2) є двома рівнями індикатора «стать». У визначній роботі, названій «Про теорії шкал у вимірюванні», яка була опублікована в журналі «*Science*» у 1946 р., психолог Стенлі Сміт Стівенс визначив чотири загальні типи рейтингових шкал для вимірювання в наукових дослідженнях: номінальну, порядкову, інтервальну та шкалу відношень. Статистичні властивості цих шкал показані в табл. 6.1.

Номінальні шкали, які також називають категоріальними, вимірюють категоріальні дані. Ці шкали використовують для показників чи індикаторів, які мають взаємовиключні рівні (атрибути). Приклад включає галузь економіки (промисловість, сільське господарство, фінанси тощо), форма власності підприємства (державне, приватне, публічне

тощо). Якщо ми надамо номери кожному рівню, наприклад, 1 – промисловість, 2 – сільське господарство, 3 – фінанси і т.п., ці номери фактично не означають числових співвідношень між групами і можуть бути замінені на інші позначення: П – промисловість, С – сільське господарство, Ф – фінанси. Номінальні шкали надають назви або позначення різним значенням рівнів. Можливою оцінкою центральної тенденції для номінальних шкал є лише мода. Середнє чи медіана не можуть бути використані. Статистика, що може бути використана – це Хі-квадрат і розподіл частот. Єдина можлива трансформація «1 до 1» (рівність) (наприклад, 1 – чоловіки, 2 – жінки.).

Таблиця 6.1.

Статистичні властивості рейтингових шкал

Шкала	Центральна тенденція	Статистика	Трансформації
Номінальна	мода	Хі-квадрат	Один до одного (рівність)
Порядкова	медіана	Перцентілі, непараметрична статистика	Монотонне збільшення (порядок)
Інтервальна	Середнє арифметичне, ранг, стандартне відхилення	Кореляція, регресія, аналіз варіацій	Позитивні лінійні (уточнення)
Відношень	Геометрична середня, гармонійна середня	Коефіцієнт варіації	Позитивні подібності (мультиплікативні, логарифмічні)
Примітка: усі шкали вищих порядків можуть використовувати будь-яку статистику шкал нижчих порядків			

Порядкові шкали вимірюють упорядковані за рангами дані, як, наприклад, рейтинг студентів групи, що встановлюється за середнім балом диплому або результатами виконання тесту. Однак фактичні або відносні значення рівнів, або різниця у значеннях рівнів не можуть бути оцінені. Наприклад, рангування студентів нічого не говорить про їх фактичний середній бал, або про бали, отримані за виконання тесту, або наскільки одні з них відносно успішніші за інших. Класичний приклад в природничих науках – це шкала твердості мінералів, запропонована Мохом. Вона характеризує твердість різних мінералів за їх здатністю наносити подряпини на інші мінерали. Оскільки діаманти можуть подрятати всі інші природні мінерали, вони вважаються найтвердішими

мінералами. Однак шкала не вказує на фактичну твердість цих мінералів і не надає відносної оцінки їх твердості. Порядкові шкали можуть використовувати такі позначення (якорі), як «поганий», «середній» і «хороший», або «повністю незадоволений», «скоріше незадоволений», «нейтрально», «скоріше задоволений» і «повністю задоволений». В останньому випадку можна стверджувати, що респонденти, які «скоріше задоволені» насправді менш задоволені, ніж ті, хто «повністю задоволені», однак неможливо встановити фактичний рівень задоволеності. Мірою центральної тенденції рангових шкал є медіана або мода, тоді як середнє значення не має сенсу. Статистичний аналіз може включати перцентілі і непараметричний аналіз, однак більш складні методи, як, наприклад, кореляція, регресія чи аналіз варіацій не можуть бути використані. Дозволяється трансформація монотонного збільшення, що зберігає ранги.

Інтервальні шкали – це шкали, у яких вимірювані значення не тільки впорядковані за рангами, але й мають однакову відстань між сусідніми рівнями. Наприклад, у температурних шкалах Фаренгейта і Цельсія різниця між 30 і 40 градусами по Цельсію є такою ж, що і між 80 і 90 градусами по Фаренгейту. Аналогічно, якщо є шкала річного доходу респондентів із такими рангами, як «від \$0 до 10000», «від \$10000 до 20000», «від \$20000 до 30000» і т.д., то це буде інтервальна шкала, оскільки середні точки кожного рангу, тобто \$5000, \$15000, \$25000 і т.д. знаходяться на однакових відстанях одна від одної. Шкала інтелекту (IQ) також є інтервальною, оскільки вона розроблена таким чином, що різниця між балами 100 і 110 є такою ж, що і різниця між балами 110 і 120. Інтервальна шкала дозволяє досліджувати, наскільки більшим є один рівень порівняно з іншим, що неможливо з номінальними й порядковими шкалами. Інтервальна шкала уможливорює вимір таких показників центральної тенденції, як середнє, медіана чи мода, і таких показників варіації (дисперсії), як розмах варіацій і стандартне відхилення. Дозволений статистичний аналіз включає всі види аналізів для номінальних і порядкових шкал, а також кореляцію, регресію, аналіз варіацій і т.п. Можлива трансформація шкали – позитивна лінійна. Якщо бути точними, то шкала задоволення, що розглядалася вище, не є повністю інтервальною, оскільки не можна сказати, чи відстань між «повністю незадоволений» і «скоріше незадоволений» є такою ж, як між «нейтральний» і «скоріше задоволений». Однак у соціальних науках дослідники часто вважають (хоча й неправильно), що ці відстані є однаковими і, відповідно можна застосувати ті методи статистичного аналізу, які застосовуються для інтервальних шкал.

Шкали відношень – це ті, що мають всі риси номінальних, порядкових та інтервальних шкал і, на додаток, мають «істинну нульову точку», де значення «нуль» означає відсутність, або неприйнятність вимірюваного конструкту. Більшість вимірювань показників у природничих науках і в техніці, таких як маса, нахил літака чи

електричний заряд, використовують шкали відношень. Аналогічно деякі показники в соціальних науках, такі як вік працівника, досвід, розмір організації (за числом працюючих чи річним доходом), також є шкалами відношень. Наприклад, якщо розмір компанії дорівнює нулю, це означає, що вона не має працівників чи доходу. Температурна шкала Кельвіна також є шкалою відношень, на відміну від шкал Фаренгейта і Цельсія, оскільки нульова точка на цій шкалі (що дорівнює -273,15 градусів Цельсія) не є довільною, а представляє стан, при якому частини речовини мають нульову кінетичну енергію. Такі шкали називаються шкалами відношень, тому що відношення двох значень на них може інтерпретуватися змістовно. Наприклад, розмір компанії 10 осіб у два рази більше, ніж компанії з 5 особами, це ж можна сказати і при порівнянні компанії з 10000 і 5000 осіб. Всі показники центральної тенденції, включаючи геометричні та гармонійні середні, як і коефіцієнт варіації, є прийнятними для використання. Дозволено використання всіх статистичних методів. Складні трансформації, такі як мультиплікативні чи логарифмічні, також прийнятні.

На основі чотирьох основних шкал, що розглянуті вище, можна створити особливі рейтингові шкали для досліджень у соціальних науках. Загальновизнані шкали включають бінарні, Лікерта, семантичні диференціальні чи Гуттмана. Інші, менш відомі, тут не розглядаються.

Бінарні шкали. Бінарні шкали – це номінальні шкали, що складаються з бінарних показників, які можуть приймати лише одне з двох значень, наприклад, «так» чи «ні», «вірно» чи «невірно». Наприклад, типова бінарна шкала для конструкту «політична активність» може складатися з шести бінарних показників, як показано в табл. 6.2.

Таблиця 6.2.

Бінарна шкала для оцінки політичної активності

Ви коли-небудь писали листа до державних установ?	Так	Ні
Ви коли-небудь підписували політичні петиції?	Так	Ні
Ви коли-небудь надавали гроші з політичних міркувань?	Так	Ні
Ви коли-небудь підтримували матеріально кандидата, що претендував на державну посаду?	Так	Ні
Ви коли-небудь писали політичний лист редактору газети чи журналу?	Так	Ні
Ви коли-небудь переконували інших людей змінити їхню думку на виборах?	Так	Ні

Кожний показник шкали є бінарним, і загальне число відповідей «так», даних респондентом (значення від 0 до 6), може бути використано

як загальна оцінка політичної активності респондента. Щоб зрозуміти, як ці показники були розроблені, зверніться до підрозділу «Створення шкал». Бінарні шкали можуть використовувати інші показники, такі як працюючий/безробітний, комерційна/державна установа і т.п. Якщо модифікувати показник зайнятості, встановивши більш, ніж два значення (наприклад, працюючий повний день/працюючий неповний день/безробітний/пенсіонер), тоді це вже буде не бінарна, а номінальна шкала статусу зайнятості.

Шкала Лікерта. Ця шкала, розроблена Ренсісом Лікертом, є поширеною рейтинговою шкалою для вимірювання порядкових даних у соціальних дослідженнях. Шкала включає показники Лікерта, які є простими твердженнями з декількох слів. Респонденти можуть оцінити, наскільки вони погоджуються чи не погоджуються з кожним із тверджень за п'яти- чи семибальною шкалою, що містить значення від «повністю не погоджуюсь» до «повністю погоджуюсь». Типовим прикладом шкали Лікерта з шести показників (тверджень) є конструкт «робоча самооцінка», що показаний у таблиці 6.3. Шкали Лікерта є такими, що загальний бал може бути сумою значень показників, які обрав респондент.

Таблиця 6.3.

Шкала Лікерта з шести показників для робочої самооцінки

	Повністю не погоджуюсь	Скоріше не погоджуюсь	Нейтрально	Скоріше погоджуюсь	Повністю погоджуюсь
Мені подобається моя робота	1	2	3	4	5
Я маю добрі стосунки з моїми колегами	1	2	3	4	5
Я задоволений своїми стосунками з керівництвом	1	2	3	4	5
Я можу сказати, що інші працівники раді, що я з ними працюю	1	2	3	4	5
Я можу сказати, що мої колеги мене поважають	1	2	3	4	5
Я думаю, що моя робота корисна	1	2	3	4	5

Показники шкали Лікерта дозволяють більшу деталізацію, ніж бінарні показники, оскільки надають більше опцій для вибору, включаючи нейтральне ставлення респондента. Три або дев'ять опцій (які називаються якорями) також можуть використовуватися, однак важливо мати непарне число опцій, щоб було нейтральне значення. Деякі дослідження використовували «підхід примусового вибору», щоб примусити респондентів погодитися чи не погодитися з твердженнями, однак це не є гарною стратегією, оскільки деякі люди дійсно можуть ставитися нейтрально до певних тверджень, і примусовий вибір не дозволяє їм відзначити своє нейтральне відношення. Ключова характеристика шкали Лікерта полягає в тому, що для різних тверджень існують одні і ті ж опції (від «повністю не погоджуюсь» до «повністю погоджуюсь»). Шкали Лікерта є порядковими, оскільки опції не обов'язково є рівновіддаленими, навіть якщо їх вважають інтервальними.

Шкала семантичного диференціалу. Ця шкала складається з декількох показників, і респондентів просять оцінити своє ставлення до кожного з тверджень, використовуючи різні пари прикметників із полярними значеннями. Наприклад, конструкт «ставлення до національної системи охорони здоров'я» може бути оцінений за чотирима показниками, як показано в табл. 6.4. Як і взагалі для шкал Лікерта, загальний бал може бути сумою балів за індивідуальними показниками. У шкалах Лікерта змінюються твердження, але опції залишаються тими ж самими. У шкалах семантичного диференціалу навпаки, твердження є одним і тим самим, тоді як показники (пари прикметників) змінюються. Уважається, що семантичний диференціал є прекрасною методикою для оцінки ставлення чи думки людей щодо певних об'єктів, подій чи процесів.

Таблиця 6.4.

Шкала семантичного диференціалу для оцінки ставлення до національної системи охорони здоров'я

Як Ви оцінюєте національну систему охорони здоров'я?						
	Значною мірою	До певної міри	Складно сказати	До певної міри	Значною мірою	
Хороша						Погана
Корисна						Не корисна
Турботлива						Не турботлива
Професійна						Не професійна

Шкала Гуттмана. Розроблена Луїсом Гуттманом, ця шкала використовує набір показників, що розташовуються у порядку зростання інтенсивності досліджуваного конструкту від найменш інтенсивного до найбільш інтенсивного. Як приклад, конструкт «ставлення до іммігрантів» може вимірюватися за п'ятьма показниками, як показано в таблиці 6.5. Кожний показник у шкалі Гуттмана має вагу (яка не вказана вище), яка змінюється в залежності від інтенсивності шкали, і зважена комбінація кожної відповіді використовується як агрегована оцінка спостереження.

Таблиця 6.5.

Шкала Гуттмана з п'яти показників для оцінки ставлення до іммігрантів

Якою є Ваша думка стосовно іммігрантів?		
Ви не проти того, щоб іммігранти стали громадянами країни?	Так	Ні
Ви не проти того, щоб іммігранти жили з Вами в одному районі?	Так	Ні
Ви не проти того, щоб іммігранти жили поруч з Вами?	Так	Ні
Ви не проти того, щоб іммігрант став Вашим близьким другом?	Так	Ні
Ви не проти того, щоб хтось із Вашої родини узяв шлюб із іммігрантом?	Так	Ні

Розроблення шкал

У попередньому підрозділі йшлося про оцінку відповідей респондентів на розроблені питання чи індикатори, що відносяться до відповідного конструкту. Однак як розробити самі індикатори? Процес створення індикаторів називається **розробкою шкал**. Більш формально, розробка шкал є галуззю вимірювання, яка включає конструювання оцінок шляхом асоціації якісних тверджень щодо конструктів, які не піддаються безпосередньому спостереженню, із кількісними вимірюваними показниками. Стівенс у 1946 р. сказав: «Розробка шкал – це приписування об'єктам чисел за певним правилом». Цей процес виміру абстрактних концептів у конкретних термінах залишається найбільш складним завданням в емпіричних дослідженнях соціальних наук.

Результатом цього процесу є певна шкала, яка є емпіричною структурою для оцінки показників чи індикаторів даного конструкту. Поняття «шкала», яке розглядається в цьому підрозділі, дещо відрізняється від поняття «рейтингова шкала», що використовувалося в попередньому підрозділі. Рейтингова шкала потрібна, щоб відзначити реакцію респондента на певний показник. Наприклад, показник номінальної шкали фіксує реакцію респондентів у термінах «так» чи «ні»,

тоді як показник інтервальної шкали фіксує значення в діапазоні між «повністю не згодний» та «повністю згодний». Приєднання рейтингової шкали до твердження не є розробкою шкали. Скоріше, розробка шкали є формальним процесом створення показників шкал до того, як рейтингова шкала може бути приєднана до цих показників.

Шкали можуть бути одно- чи багатовимірні, залежно від того, чи є конструкт, що лежить в їх основі, одновимірним чи багатовимірним. Одновимірні показники шкали конструюються за однією шкалою і ранжуються від високого до низького значення. Деякі з цих шкал можуть включати численні показники, однак усі вони мають оцінювати один і той же вимір. Це відбувається у випадку багатьох конструктів соціальних наук, таких, наприклад, як самооцінка. Передбачається, що цей конструкт має єдиний вимір від низького до високого значення. Багатовимірні шкали, навпаки, використовують різні показники для оцінки кожного виміру конструкту окремо, і потім бали по кожному виміру об'єднуються для створення оцінки багатовимірного конструкту. Наприклад, академічна успішність може вимірюватися за двома окремими тестами – математичних і вербальних здібностей, бали за якими потім об'єднуються для створення єдиного конструкту «академічна успішність». Оскільки більшість шкал у дослідженнях соціальних наук є одновимірними, далі будуть розглядатися три підходи до створення одновимірних шкал.

Методи розроблення одновимірних шкал були запропоновані протягом першої половини ХХ ст. й отримали назви за іменами своїх авторів. Три найбільш поширені методи одновимірних шкал це: 1) рівномірна шкала Терстоуна; 2) сумарна шкала Лікерта; 3) кумулятивна шкала Гуттмана. Усі три підходи багато в чому схожі, але їхні ключові відмінності полягають у тому, як ранжуються показники шкал експертами, і які статистичні методи використовуються для вибору інтегрального показника. Ці методи описуються нижче.

Метод рівномірної шкали Терстоуна. Луїс Терстоун, один із перших і найбільш відомих теоретиків із розроблення шкал, опублікував метод рівно розподілених інтервалів у 1925 р. Метод починає з чіткого концептуального визначення конструкту, який буде досліджуватися. Спираючись на це визначення, генеруються потенційні показники шкали для виміру цього конструкту. Це роблять експерти, обізнані з сутністю даного конструкту. Первинний набір потенційних показників (в ідеалі від 8 до 100) повинен бути сформульований в однаковому стилі, наприклад, як твердження, із якими респонденти погоджуються чи не погоджуються (а не як, скажімо, питання). Після цього збирається група експертів для відбору окремих показників, які найкращим чином презентують досліджуваний конструкт. Експертами можуть бути науковці, спеціально навчені в процесі конструювання методики, або вибірка досліджуваних респондентів (тобто обізнаних з феноменом людей). Процес відбору відбувається таким чином, що кожний експерт незалежно рангує кожний

показник за шкалою від 1 до 11 на основі того, наскільки даний показник відображає досліджуваний конструкт (1 – погано, 11 – відмінно). Для кожного показника розраховується медіана та інтер-квартільний ранг (різниця між 75-м і 25-м перцентилем, що є мірою дисперсії), які наносяться на графік у вигляді гістограми, як показано на рис. 6.1. Остаточні показники шкали обираються як твердження, що знаходяться на рівній відстані вздовж рангів медіан. Це можна зробити, групуючи показники з загальною медіаною і потім обираючи показник із найменшим інтер-квартільним рангом у межах кожної медіанної групи. Однак, замість того, щоб повністю покладатися на статистичний аналіз при виборі показників, кращою стратегією може бути дослідження потенційних показників на кожному рівні і вибір твердження, що є найбільш точним і змістовним. Значення медіани по кожному показнику шкали представляє вагу, яку потрібно використовувати для агрегації показників в об'єднаний бал шкал, що представляє досліджуваний конструкт. Так отримується шкала, що має вигляд лінійки з одним показником чи твердженням на кожній з 11 точок (з відповідними вагами). Оскільки показники розташовані на рівних відстанях вздовж шкали з 11-ти поділок, ця техніка називається рівномірною шкалою. Терстоун також розробив два додаткові методи для створення одновимірних шкал – метод послідовних інтервалів і метод парних порівнянь, які схожі з методом рівномірних інтервалів, за винятком того, що експертів просять прорангувати дані. Наприклад, метод парних порівнянь вимагає від кожного експерта зробити порівняння кожної пари тверджень замість того, щоб рангувати кожне твердження незалежно за шкалою від 1 до 11. Із великою кількістю тверджень цей підхід вимагає занадто багато часу і програє методу рівномірних інтервалів.

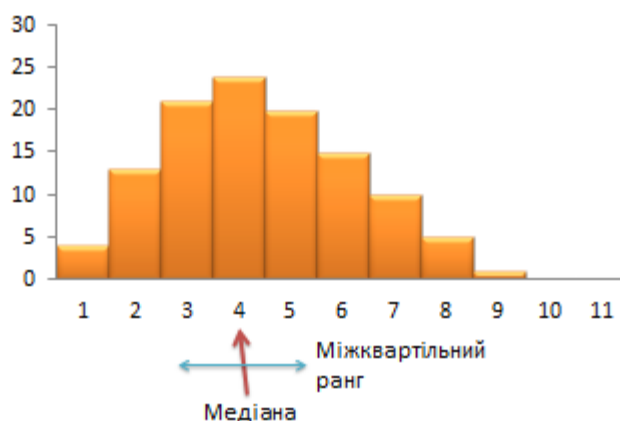


Рис. 6.1. Гістограма рейтингів показників шкали Терстоуна

Метод сумарної шкали Лікерта. Метод Лікерта, або метод розроблення одновимірної шкали, був запропонований Мерфі і Лікертом у 1938 р., і з того часу залишається найбільш поширеним із трьох методів розроблення шкал, що описуються в цьому розділі. Як і у випадку методу Терстоуна, метод Лікерта також починається з чіткого визначення

досліджуваного конструкту й використання групи експертів для генерування 80-100 потенційних показників шкали. Ці показники рангуються експертами за шкалою від 1 до 5 (або від 1 до 7) таким чином: 1 – повністю не погоджуюсь з концептом, 2 – скоріше не погоджуюсь, 3 – складно вирішити, 4 – скоріше погоджуюсь, 5 – повністю погоджуюсь. За цим рейтингом окремі показники можуть бути відібрані до кінцевої шкали наступним чином: 1) розрахунком парної кореляції між рейтингом, наданим експертами за кожним показником, і загальним показником, який створюється додаванням усіх індивідуальних показників за кожним із респондентів і відкиданням тих показників, що мають низьку (менш ніж 0,60) кореляцію з загальним показником; 2) усередненням рейтингу за кожним показником для найвищого і найнижчого квартиля експертів, виконанням *t*-тесту для різниці середніх і подальшого вибору показників, що мають найбільші значення *t*-критерію (тобто найкраще диференціюють між верхнім і нижнім квартилем відповідей). Нарешті, висновки експерта можуть бути використані, щоб отримати відносно невеликий (10-15 показників) набір показників, що мають високі кореляції з сумарним показником і високі значення *t*-критерію. Метод Лікерта передбачає рівні ваги всіх показників, а отже, відповіді респондентів за кожним із показників можна просумувати, щоб створити узагальнений бал респондента. Із цієї причини метод і називається сумарною шкалою. Відзначимо, що кожний показник із зворотним значенням щодо напрямку конструкту потрібно кодувати у зворотному напрямку (тобто, 1 стає 5, 2 – 4 і т.д.) до підрахунку суми.

Метод кумулятивного шкалування Гуттмана. Цей метод, запропонований Гуттманом у 1950 р., базується на техніці соціальної відстані Еморі Богардуса, яка передбачає, що готовність людей брати участь у соціальних зв'язках із іншими людьми змінюється за інтенсивністю, яку можна оцінити за шкалою від «найменш інтенсивна» до «найбільш інтенсивна». Ідея полягає в тому, що люди, які погоджуються з одним показником із набору, також погоджуються з усіма іншими показниками. На практиці рідко можна знайти набір показників, що повністю відповідають цьому кумулятивному паттерну. Щоб дослідити, наскільки близько набір показників відповідає ідеї кумулятивності, використовується аналіз шкалограм.

Аналогічно попереднім методам, метод Гуттмана починається з чіткого визначення досліджуваного конструкту і залучення експертів до створення великого набору потенційних показників. Потім експерти рангують кожний потенційний показник або «так», якщо вважають, що показник відображає конструкт, або «ні», якщо на їх думку показник погано відображає конструкт. Потім створюється матриця, або таблиця, у яку заносять відповіді експертів стосовно всіх можливих показників. Матриця створюється в порядку зменшення – від експертів із найбільшим числом відповідей «так» угорі таблиці до експертів із найменшим числом

відповідей «так» внизу. Для експертів із однаковою кількістю відповідей «так» твердження можуть бути впорядковані зліва направо від найбільшої кількості погоджень до найменшої. Результуюча матриця наведена в табл. 6.6. Зверніть увагу, що шкала зараз є майже кумулятивною, коли читається зліва направо (за показниками). Однак може бути декілька винятків, як показано в табл. 6.6., і тому шкала не є повністю кумулятивною. Щоб визначити набір показників, які найбільше відповідають властивості кумулятивності, може бути використана техніка аналізу даних, що називається аналізом шкалограм (або це можна зробити візуально, якщо число показників не дуже велике). Статистична методика також оцінює бал для кожного показника, який можна використати для розрахунку узагальненого балу респондента за набором показників.

Таблиця 6.6.

**Матриця впорядкованих рейтингів для
шкали Гутмана**

Респон- дент	Показ- ник 12	Показ- ник 5	Показ- ник 3	Показ- ник 22	Показ- ник 8	Показ- ник 7	...
29	Ү	Ү	Ү	Ү	Ү	Ү	
7	Ү	Ү	Ү	-	-	-	
15	Ү	Ү	Ү	Ү	-	-	
3	Ү	Ү	Ү	Ү	-	-	
32	Ү	Ү	Ү	-	-	-	
4	Ү	Ү	-	<u>Ү</u>	-	-	
5	Ү	Ү	-	-	-	-	
23	Ү	Ү	-	-	-	-	
11	Ү	-	-	<u>Ү</u>	-	-	
<u>Ү</u> означає винятки, що не дозволяють матриці стати повністю кумулятивною							

Індекси

Індекс – це об'єднаний бал, що зроблений на основі агрегованих оцінок декількох конструктів (які називаються компонентами) за певними правилами та формулами. Від шкал індекси відрізняються тим, що, хоча шкали також є агрегованими оцінками, вони вимірюють різні складові одного конструкту або один і той же конструкт. Добре відомий приклад індексу – це індекс споживчих цін (ІСЦ). ІСЦ є мірою того, скільки споживачі повинні платити за товари та послуги загалом. Він ділиться на вісім головних категорій (їжа і напої, житло, одяг, транспорт, медичне обслуговування, оздоровлення, освіта і комунікації, інші товари та послуги), які у свою чергу діляться на більш ніж 200 менших показників. Кожний місяць держслужбовці надають інформацію про поточні ціни на тисячі видів товарів і послуг. За складною схемою зважування, що враховує розташування й можливість купівлі кожного

найменування, ці ціни об'єднуються аналітиками, на основі яких за певними правилами та формулами розраховують загальний індекс цін.

Інший приклад індексу – це соціо-економічний статус (СЕС), який також називається соціо-економічним індексом Дункана. Цей індекс є комбінацією трьох конструктів: дохід, освіта та професія. Дохід вимірюється в грошових одиницях, освіта – у роках або отриманих наукових ступенях, а професія класифікується в категорії, або рівні за статусом. Ці дуже різні оцінки об'єднуються для створення загального СЕС із використанням зваженої комбінації «професійна освіта» (відсоток людей у професії, що мають один і більше років навчання в коледжі) і професійний дохід (відсоток людей у професії, які заробляють більш, ніж певний річний дохід). Однак розрахунок індексу СЕС породив критику й неоднозначно сприймається в колах науковців.

Процес створення індексу аналогічний до створення шкали. По-перше, даються визначення індексу та його складових. Хоча це виглядає просто, однак може бути багато розходжень у думках експертів стосовно того, які компоненти (конструкти) потрібно включати в індекс. Наприклад, в індексі СЕС, якщо дохід корелює з освітою і професією, то чи треба включати лише одну складову, або всі три? Аналіз літератури, використання теорій і спілкування з експертами та ключовими стейкхолдерами може допомогти у вирішенні цього питання. Наступним кроком є операціоналізація і вимір кожної складової. Зокрема, як категоризувати професії, особливо якщо деякі з них значно змінилися з часом (наприклад, до виникнення Інтернету не було веб-розробників). Потім потрібно створити правило чи формулу для розрахунку значення індексу. У цьому процесі дуже важко уникнути суб'єктивності. І, нарешті, необхідно валідизувати бали індексу, спираючись на існуючі чи нові дані.

Хоча індекси і шкали мають результатом єдине цифрове значення, що представляє досліджуваний конструкт, між ними є багато відмінностей. По-перше, індекси часто включають складові, що значно відрізняються між собою (як, наприклад, дохід, освіта і професії в індексі СЕС), і оцінюються по-різному. При цьому, шкали зазвичай включають набір подібних показників, що використовують однакову рейтингову шкалу (як, наприклад, 5-бальну шкалу Лікерта). По-друге, індекси часто об'єднують значення, які можуть бути об'єктивно оцінені (ціни, дохід і т.п.), тоді як шкали розробляються для оцінки суб'єктивних чи атрибутивних конструктів (ставлення, упередженість і т.п.) Деякі науковці вважають, що сутність методології розроблення шкал робить їх відмінними від індексів, хоча методологія розроблення індексів може бути не менш складною. У будь-якому разі, індекси та шкали є важливими інструментами в дослідженнях соціологічних наук.

Типології

Шкали й індекси генерують порядкові оцінки одновимірних конструктів. Однак іноді дослідники хочуть об'єднати величини двох чи більше конструктів, щоб створити набір категорій чи типів, який називається **типологією**. На відміну від шкал та індексів, типології є багатовимірними і включають лише номінальні змінні. Прикладом типології є класифікація товарів фірми за матрицею Бостонської групи на такі групи: «важкі діти», «зірки», «дійні корови» та «собаки», інший приклад – рівні навченості людей за класифікацією Блума.

Розділ 7

НАДІЙНІСТЬ І ВАЛІДНІСТЬ ШКАЛ

У попередньому розділі аналізувалися деякі складнощі, пов'язані з вимірюванням конструктів у дослідженнях соціальних наук. Якщо досліджується інвестиційна привабливість, то як можна знати, що оцінюється саме цей конструкт, а не інший, схожий за змістом, наприклад, інвестиційний клімат? Ситуація ускладнюється тим, що часто досліджувані конструкти є віртуальними, тобто не існують у реальності, та багатовимірними (у цьому випадку додається проблема ідентифікації складових конструкту). Отже, недостатньо лише оцінити конструкти за обраними шкалами, необхідно протестувати самі шкали, щоб бути впевненими, що: 1) ці шкали дійсно вимірюють ті віртуальні конструкти, які дослідник хоче виміряти, а отже, шкали є валідними; 2) при повторному вимірюванні бажаних конструктів будуть отримуватися однакові результати, отже, шкали є надійними. Надійність і валідність, які разом називають психометричними властивостями шкал вимірювання, є ключовими критеріями для оцінки адекватності та точності процедури вимірювання, і тому мають оцінюватися при проведенні наукових досліджень.

Показник може бути надійним, але не валідним, якщо він при повторних замірах однаково вимірює один і той же «неправильний» конструкт. Аналогічно, показник може бути валідним і водночас не надійним, якщо він оцінює «правильний» конструкт із недостатньою повторюваністю.

Використовуючи аналогію з пострілом по мішені, як показано на рис. 7.1, оцінка складного конструкту, який є одночасно надійним і валідним, означає точні попадання у вузьку область навколо центру мішені. Оцінка, що є валідною, але ненадійною, ілюструється неточними пострілами, розсіяними в широкій області навколо центру мішені. Оцінка, що є надійною, але не валідною, ілюструється пострілами, які щільно згруповані навколо точки, що знаходиться поза межами центру мішені. Отже, для забезпечення адекватного вимірювання досліджуваного конструкту необхідно, щоб застосовувані методики були одночасно надійними та валідними.



Рис. 7.1. Порівняння надійності і валідності

Надійність

Надійність показує, наскільки щільно співпадають значення оцінки конструкту при повторних замірах цього конструкту за тією ж методикою, якщо досліджуваний феномен не змінився. Прикладом ненадійного показника є судження різних людей стосовно ваги певної людини. Зрозуміло, що люди по-різному оцінять цей показник, і їхні суб'єктивні судження не будуть співпадати, отже техніка «здогадок» є ненадійною. Більш надійний показник дасть застосування вагів, оскільки результати повторних зважувань будуть співпадати між собою (якщо вага між замірами не зміниться).

Зверніть увагу, що надійність означає повторюваність, однак не точність. У попередньому прикладі з вагами, якщо їхня шкала неправильно відкалібрована, тоді вимірювання не дасть істинної величини ваги людини й отримані результати будуть не валідними. Однак неправильно відкалібровані ваги даватимуть надійні результати, оскільки результати замірів будуть повторюватися.

Що є джерелами ненадійних результатів у дослідженнях соціальних наук? Одне з головних джерел – це суб'єктивність дослідника. Якщо, наприклад, соціально-психологічний клімат у колективі оцінюється числом конфліктних ситуацій, тоді різні дослідники один і той же колектив оцінять по-різному: один побачить багато конфліктів, якщо проводив спостереження під час напружених зборів, інший – не побачить конфліктів взагалі, якщо його спостереження припали на період буденної роботи. Окрім того, різні дослідники можуть по-різному трактувати конфлікт – від висловлювання альтернативної думки до відкритої сварки. Взагалі спостереження, що застосовується в даному прикладі досліджень, є якісним методом. Для покращення надійності можна звернутися до кількісних вимірів, наприклад, фіксуючи кількість конфліктів різного ступеня напруженості протягом певного періоду часу.

Інше джерело ненадійних результатів при спостереженні пов'язано з неконкретними чи незрозумілими формулюваннями питань. Наприклад, якщо у респондентів запитати про рівень їхньої зарплати, вони можуть

по-різному інтерпретувати це питання – як про річну зарплату, місячну, з премією чи без і т.п.

Третє джерело ненадійності результатів пов'язано з питаннями, на які респонденти не мають відповіді через необізнаність або відсутність інтересу, наприклад коли директора з персоналу запитують про ефективність логістичної функції компанії.

Отже, як створити надійні показники? Якщо дослідження передбачає збір інформації від респондентів, як це відбувається в багатьох дослідженнях соціальних наук, тоді потрібно перейти від більш суб'єктивних методик (наприклад, спостереження) до таких, де рівень суб'єктивності менше (наприклад, анкетне опитування). При цьому потрібно запитувати лише про ті речі, із якими респондент добре обізнаний, уникати неоднозначних, розпливчастих формулювань, використовувати загальнозживану лексику і уникати складної термінології. Ці стратегії можуть підвищити надійність результатів, навіть якщо вони й не зроблять отримані результати повністю надійними. Ті інструменти вимірювання, що застосовуються в дослідженні, все одно потрібно перевіряти на надійність. Для цього існує багато способів, які й будуть розглянуті далі.

Надійність, пов'язана зі спостерігачами. Надійність, пов'язана із спостерігачами, є мірою узгодженості між двома й більше незалежними спостерігачами в оцінці одного й того ж конструкту. Зазвичай вона оцінюється в пілотному дослідженні одним із двох способів, залежно від рівня оцінки конструкту. Якщо оцінка є категоріальною, тоді визначається набір всіх категорій, і спостерігачі перевіряють, у яку категорію потрапляє кожне спостереження. Відсоток узгодженості між спостерігачами є оцінкою надійності. Наприклад, якщо є двоє спостерігачів, які розподіляли 100 спостережень в одну з трьох категорій і їхні рейтинги співпали у 75% випадків, тоді надійність, пов'язана зі спостерігачами, становитиме 75%. Якщо оцінкою є інтервал чи шкала відношень (наприклад, активність студентів під час групової роботи, яка вимірюється кожні 5 хвилин двома спостерігачами за 7-бальною шкалою), тоді оцінкою надійності, пов'язаною зі спостерігачами, буде слугувати проста кореляція між оцінками двох спостерігачів.

Тест-ретест надійність. Тест-ретест надійність є мірою узгодженості між двома вимірами одного й того ж конструкту в одній і тій же вибірці досліджуваних у різні періоди часу. Якщо оцінка суттєво не змінилася при повторному дослідженні, тоді вона є надійною. Кореляція в спостереженнях між двома тестами є оцінкою тест-ретест надійності. При цьому часовий інтервал між двома тестами є критичним. Зазвичай, що довший цей інтервал, то більша вірогідність розходження в даних двох тестів (через випадкову помилку), і тим меншою буде тест-ретест надійність.

Надійність поділу навпіл. Надійність поділу навпіл є мірою узгодженості між двома половинами оцінки конструкту. Наприклад, якщо є оцінка конструкту з 10-х показників, тоді їх ділять на дві групи по 5 (дозволяються неоднакові за кількістю показників групи, якщо загальне число показників є непарним), і роздають для заповнення вибірці респондентів. Потім розраховують загальні оцінки по кожному респонденту для кожної половини показників. Їх кореляції і вважають мірою надійності поділу навпіл. Чим довшою є анкета, тим більша вірогідність, що обидві половини оцінки конструкту будуть давати близькі значення (оскільки при збільшенні числа показників випадкові помилки мінімізуються), а отже, цей спосіб має тенденцію систематично переоцінювати надійність довгих анкет.

Надійність внутрішньої узгодженості. Надійність внутрішньої узгодженості є мірою узгодженості між різними показниками одного й того ж конструкту. Якщо респонденти для оцінки конструкту заповнюють анкету з численними показниками, ступінь подібності відповідей при оцінці кожного показника є відображенням внутрішньої узгодженості. Ця надійність може бути оцінена як середня кореляція між окремими показниками, середня кореляція між окремими показниками і загальною оцінкою конструкту і, найчастіше, за альфа-коефіцієнтом Кронбаха. Наприклад, для шкали з 6 показників є 15 різних пар показників і 15 коефіцієнтів кореляції. Щоб обчислити середню кореляцію окремих показників із загальним значенням, потрібно спочатку розрахувати загальне значення, додаючи значення шести окремих показників, потім поррахувати кореляції між кожним із них і загальним значенням, і, нарешті, – середнє з шести коефіцієнтів кореляції. Жоден із двох розрахованих показників не бере до уваги число показників у шкалі (шість у даному випадку). Як міра надійності, Лі Кронбахом у 1951 р. був запропонований альфа-коефіцієнт Кронбаха, який враховує розмір шкали. Стандартизований альфа-коефіцієнт Кронбаха може бути розрахований за формулою:

$$\alpha_{\text{standardized}} = \frac{K\bar{r}}{(1 + (K - 1)\bar{r})}$$

де K – число показників у шкалі, \bar{r} – середня кореляція між показниками, тобто середнє значення $K*(K-1)/2$ коефіцієнтів у вищому чи нижчому трикутниках кореляційної матриці.

Валідність

Валідність, яку часто називають валідністю конструкту, стосується того, наскільки показник адекватно представляє конструкт, що лежить в його основі й для виміру якого він був розроблений. Наприклад, наскільки показник групової солідарності дійсно вимірює групову солідарність, а не, скажімо, задоволеність роботою в групі. Валідність

може бути оцінена на основі теоретичних чи емпіричних підходів і в ідеалі повинна вимірюватися на основі них їх обох. Теоретична оцінка валідності фокусується на тому, наскільки добре ідея теоретичного конструкту представлена в операціональній величині. Цей тип валідності називається **трансляційною** (або репрезентативною) валідністю і складається з двох підтипів: очевидна та змістовна валідність. Трансляційна валідність зазвичай оцінюється панеллю експертів, які ранжують кожний показник (індикатор) залежно від того, наскільки добре цей показник відповідає концептуальному визначенню конструкту, за допомогою якісної за природою техніки, що називається Q-сортуванням.

Емпірична оцінка валідності на основі емпіричних спостережень оцінює, наскільки добре дана величина пов'язана з одним чи декількома зовнішніми критеріями. Цей тип валідності називається **валідністю, пов'язаною з критеріями**, і включає чотири підтипи: конвергентний, дискримінантний, конкурентний і прогнозуючий. Якщо трансляційна валідність оцінює, як добре певна величина відображає конструкт, критеріально пов'язана валідність встановлює, чи змінюється цей конструкт так, як має змінюватися, урахувавши його теоретичну сутність. Ця оцінка базується на кількісному аналізі даних дослідження з використанням статистичних методик, таких як кореляційний аналіз, факторний аналіз і т.п. Різниця між теоретичним та емпіричним оцінюванням валідності ілюструється на рис. 7.2. Однак необхідно застосовувати обидва підходи, щоб забезпечити адекватну валідність конструктів у соціальних дослідженнях.

Різні типи валідності, що тут розглядаються, стосуються валідності *процедури вимірювання*, яка відрізняється від валідності процедури тестування гіпотез, такої як внутрішньої валідності (причинності), зовнішньої валідності (узагальненості) чи валідності статистичного висновку. Останні типи валідності будуть розглядатися пізніше.

Очевидна валідність. Очевидна валідність показує, наскільки обґрунтовано індикатор можна вважати відображенням конструкту, що лежить в його основі, за видимими ознаками. Наприклад, кількість упроваджених за певний період часу інновацій можна вважати індикатором інноваційної активності компанії, і тому цей показник має очевидну валідність. Однак сумнівно, що кількість відвідувань бібліотеки персоналом компанії може слугувати валідним показником інноваційної активності. Слід відзначити, що деяким популярним показникам, що використовуються в організаційних дослідженнях, бракує очевидної валідності. Наприклад, абсорбтивна здатність організації (як багато нових знань може поглинути організація для покращення внутрішньоорганізаційних процесів) часто використовується як показник інтенсивності досліджень, тобто витрати на дослідження і розвиток, поділені на чисті прибутки. Якщо дослідження включає конструкти, що є дуже абстрактними або їх важко відокремити концептуально від інших

конструктів, мабуть, слід розглянути можливість залучення експертів, щоб оцінити очевидну валідність показників даного конструкту.



Рис. 7.2. Два підходи до оцінки валідності

Змістовна валідність. Змістовна валідність є оцінкою того, наскільки добре набір показників шкали відповідає релевантному змісту конструкту, який ці показники повинні вимірювати. Наприклад, якщо дослідник хоче виміряти конструкт «задоволеність клієнтів обслуговуванням в ресторані», і сутність цього конструкту визначена як така, що включає якість страв, ввічливість персоналу, тривалість очікування та загальна атмосфера в ресторані (шум, паління і т.п.), тоді для адекватної змістовної валідності цей конструкт повинен вимірюватися через чотири відповідні індикатори, які оцінюють ступінь задоволеності клієнта. Такий підхід вимагає детального опису змісту конструкту як цілого, що може виявитися складною задачею для багатьох конструктів, таких як кадровий потенціал чи креативність. Отже, не завжди можливо адекватно оцінити змістовну валідність. Як і у випадку з очевидною валідністю, рекомендується залучити експертів, щоб з'ясувати змістовну валідність конструкту.

Конвергентна валідність стосується того, наскільки близько певний показник пов'язаний із конструктом, який він має оцінювати. **Дискримінантна валідність** стосується того, наскільки певний показник не вимірює інші конструкти (або відрізняється від інших конструктів), які він не повинен вимірювати. Зазвичай конвергентна і дискримінантна валідність оцінюються разом для набору пов'язаних конструктів. Наприклад, якщо дослідник очікує, що організаційні знання пов'язані з ефективністю діяльності організації, то як він може бути певним, що вимірює саме організаційне знання (для конвергентної валідності), а не ефективність (для дискримінантної валідності)? Конвергентна валідність може бути встановлена порівнянням значень індикатора певного

конструкту зі значеннями інших індикаторів того ж конструкту та виявленням схожості (або високої кореляції) між значеннями цих індикаторів. Дискримінантна валідність встановлюється демонстрацією того, що індикатори одного конструкту не схожі (або мають низьку кореляцію) із іншими конструктами. У наведеному вище прикладі, якщо організаційне знання вимірюється трьома показниками і ефективність діяльності теж вимірюється трьома показниками, тоді можна розрахувати парні коефіцієнти кореляції між кожним із індикаторів знання й ефективності. Якщо ця кореляційна матриця показує високу кореляцію між показниками організаційного знання та високу кореляцію між показниками ефективності і водночас низькі коефіцієнти кореляції між показниками цих конструктів, тоді одночасно спостерігається і конвергентна, і дискримінантна валідність (табл. 7.1.).

Таблиця 7.1.

Парний кореляційний аналіз для оцінки конвергентної і дискримінантної валідності

	SE1	SE2	SE3	SAT1	SAT2	SAT3
SE1	1.00	0.83	0.79	0.23	0.21	0.19
SE2		1.00	0.75	0.11	0.20	0.03
SE3			1.00	0.03	-0.11	0.17
SAT1				1.00	0.84	0.91
SAT2					1.00	0.77
SAT3						1.00

Висока кореляція між показниками одного конструкту (конвергентна валідність)

Низька кореляція між показниками різних конструктів (дискримінантна валідність)

Альтернативний і більш поширений статистичний метод оцінки конвергентної і дискримінантної валідності – *пошуковий факторний аналіз*. Це метод редукції даних, що складаються з певного набору показників, до меншого числа факторів за допомогою кореляційної матриці з використанням статистичної методики, що називається аналізом головних компонентів. Ці фактори повинні ідеально відповідати теоретичному конструкту, який вони мають вимірювати. Загальна норма для виділення факторів полягає в тому, що кожний фактор повинен мати власне значення, що перевищує 1. Виділені фактори можуть потім перевертатися за допомогою різних технік ротації, залежно від того, чи очікується відносна кореляційна незалежність чи, навпаки, залежність відповідних конструктів. Це дозволяє генерувати факторні ваги, які можуть використовуватися, щоб поєднати окремі показники кожного конструкту в єдиний показник. Для адекватної конвергентної валідності очікується, що окремі показники одного й того ж конструкту повинні продемонструвати факторні навантаження від 0,60 і вище (навантаження одного фактора), тоді як для дискримінантної валідності ці показники

повинні мати значення факторних навантажень 0,30 і менше за всіма іншими факторами (міжфакторні навантаження), як показано в матриці обернених факторів (табл. 7.2). Більш складною методикою для оцінки конвергентної та дискримінантної валідності є МТММ (multi-trait-multi-method) підхід. Він вимагає виміру кожного конструкту (характеристики) з використанням двох і більше різних методів (наприклад, анкетування й особисті спостереження або анкетування двох різних груп респондентів – клієнтів і менеджерів для оцінки якості товару. Це складний і тому менш поширений підхід, який не буде розглядатися детально в цій книзі.

Таблиця 7.2.

Пошуковий факторний аналіз для конвергентної та дискримінантної валідності

	Factor1	Factor2
SE1	0.88	0.13
SE2	0.93	0.11
SE3	0.87	0.03
SAT1	0.17	0.93
SAT2	-0.03	0.85
SAT3	0.07	0.78

Високі навантаження одного і того ж фактору (конвергентна валідність)

Низькі міжфакторні навантаження (дискримінантна валідність)

Валідність, пов'язана з вибором критеріїв, може також оцінюватися на основі того, чи добре даний показник пов'язаний із поточними або майбутніми критеріями. Відповідно, розрізняють конкурентну і передбачувальну валідність. **Передбачувальна валідність** – це ступінь того, наскільки успішно показник прогнозує майбутній результат, який очікується отримати на теоретичному рівні. Наприклад, чи можуть бали респондента за анкетною оцінкою інтелектуальних здібностей правильно спрогнозувати академічну успішність? Вимірювання такої валідності вимагає створення номологічної мережі, яка відображає теоретичний зв'язок між конструктами. **Конкурентна валідність** досліджує, наскільки добре певний показник пов'язаний із іншим відомим критерієм, що існує одночасно. Наприклад, оцінки учня з алгебри будуть корелювати з його оцінками з геометрії? Ці оцінки повинні бути пов'язані конкурентно, оскільки вони відносяться до однієї дисципліни – математики. На відміну від конвергентної і дискримінантної валідності, конкурентна і передбачувальна валідність часто ігноруються в емпіричних дослідженнях соціальних наук.

Теорія вимірювання

Після того, як ми ознайомилися з різними видами надійності та валідності, спробуємо узагальнити наше розуміння цих понять за допомогою *класичної теорії тестів*, яка також називається *теорією*

істинних балів. Це психометрична теорія, яка досліджує, як працює певний показник, що він вимірює, а що ні. Теорія постулює, що кожне спостереження має істинне значення (Т), яке можна точно зафіксувати, якщо не було допущено помилок при вимірюванні. Однак присутність помилки вимірювання (Е) призводить до відхилення значення, що спостерігалось в дослідженні (Х), від істинного значення (Т):

$$X = T + E.$$

Дисперсії значень, що спостерігалися в дослідженні, та дисперсії істинних значень пов'язані аналогічним рівнянням:

$$\text{var}(X) = \text{var}(T) + \text{var}(E).$$

Мета психометричного аналізу полягає в оцінці та мінімізації варіації помилки $\text{var}(E)$, щоб значення Х стало хорошою оцінкою істинного показника Т.

Помилки вимірювання можуть бути двох типів: випадкові та систематичні. **Випадкова помилка** – це помилка, що відноситься до набору невідомих та неконтрольованих зовнішніх факторів, що випадково впливають на деякі спостереження. Наприклад, протягом періоду дослідження деякі респонденти мають гарний настрій і тому вище оцінюють задоволеність від праці, ніж ті, що мають поганий настрій. Звичайно, неможливо контролювати всі змінні, що можуть вплинути на досліджуваний показник. Аналогічно, на організаційному рівні, якщо вимірюється ефективність діяльності компанії, зміни в бізнес-середовищі можуть вплинути на діяльність деяких компаній більшою мірою, ніж інших. Отже, випадкова помилка при вимірюванні вважається «шумом» і зазвичай ігнорується.

Систематична помилка – це помилка, що постійно виникає через фактори систематичного впливу на всі спостереження конструкту у всій досліджуваній вибірці. Якщо фінансова криза негативно впливає на всі фінансові установи, тоді можна очікувати зменшення ефективності діяльності досліджуваних фінансових установ і вважати фінансову кризу фактором, що вносить систематичну помилку в діяльність компаній цього сегменту ринку. На відміну від випадкової помилки, яка в досліджуваній вибірці може бути позитивною, негативною чи дорівнювати нулю, систематичні помилки у вибірці мають тенденцію бути або позитивними, або негативними. Отже, систематична помилка іноді розглядається як погрішність вимірювання й може коригуватися.

Оскільки величина показника, що спостерігалась в дослідженні, може включати як випадкові, так і систематичні помилки, рівняння для істинного значення величини показника прийме вигляд:

$$X = T + E_r + E_s ,$$

де E_r та E_s представляють випадкову й систематичну помилки відповідно. Статистичний вплив цих помилок полягає в тому, що випадкова помилка додає варіативності (тобто збільшує стандартне відхилення) розподілу значень вибірки, однак не впливає на його центральну тенденцію (тобто не змінює середню), тоді як систематична помилка впливає на центральну тенденцію, а не на варіативність, як показано на рис. 7.3.

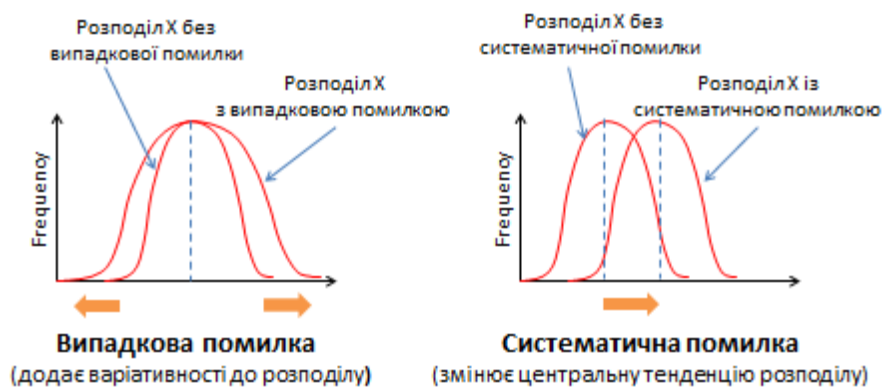


Рис. 7.3. Ефекти випадкової та систематичної помилок

Що означають випадкові та систематичні помилки для проведення процедури вимірювання? При збільшенні варіативності спостережень випадкова помилка зменшує надійність вимірювань. Систематична помилка, навпаки, зміщує центральну тенденцію й зменшує валідність вимірювання. Валідність викликає набагато більш серйозні проблеми при вимірюванні, ніж надійність, оскільки не валідне вимірювання, можливо, оцінює інший конструкт, ніж той, який збирався оцінювати дослідник, отже, через проблему валідності виникають сумніви в результатах статистичного аналізу.

Надійність є відношенням, або фракцією, яка показує, як співвідносяться істинне значення й те значення, що спостерігалось в дослідженні. Отже, надійність може бути виражена як:

$$\text{var}(T) / \text{var}(X) = \text{var}(T) / [\text{var}(T) + \text{var}(E)] .$$

Якщо $\text{var}(T) = \text{var}(X)$, тоді істинне значення має ту ж саму варіативність, як і те значення, що спостерігалось у дослідженні, і надійність дорівнює 1.

Інтегрований підхід до вимірювання валідності

Вичерпна та адекватна оцінка валідності повинна включати як теоретичний, так й емпіричний підходи. Як показано на рис. 7.4, це складний багатокроковий процес, який має враховувати різні типи

надійності та валідності тих шкал, що будуть застосовуватися в дослідженні.

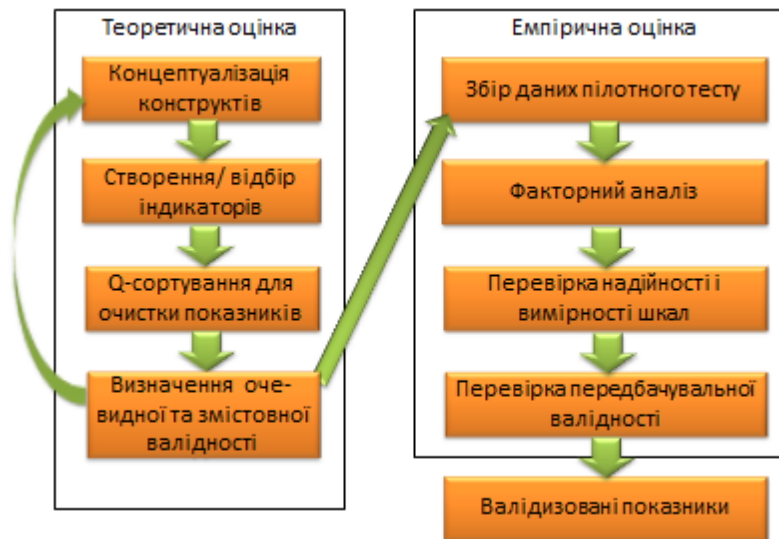


Рис. 7.4. Інтегрований підхід до валідації вимірювання

Інтегрований підхід починається в теоретичній області. Перший крок – це концептуалізація конструктів, що підлягають дослідженню. Мається на увазі визначення кожного конструкту та ідентифікація його складових доменів і/чи напрямів. Далі, базуючись на концептуалізації досліджуваних конструктів, для кожного з них обираються (чи створюються) показники, або індикатори, як представлено в описі процедури розроблення шкал у розділі 5. При виборі індикаторів також може бути корисним огляд літератури. Усі показники формулюються в однаковому стилі і з використанням простої та зрозумілої лексики. Після цього група експертів (дослідників-науковців, що мають досвід у розробленні методів дослідження і/чи представники ключових респондентів) може бути залучена для дослідження кожного індикатора і виконання Q-сортування. У цьому аналізі кожному експерту дається список усіх конструктів із їх концептуальними визначеннями та набір карток із назвами індикаторів кожного конструкта (один індикатор на картку). Експертів просять незалежно читати кожну картку, перевіряти ясність, зрозумілість та семантичне значення показника і сортувати їх за конструктами, до яких за змістом вони мають належати, урахувавши визначення конструктів. Оцінюється надійність, пов'язана з експертами, щоб визначити ступінь узгодженості думок експертів щодо класифікації показників за конструктами. Сумнівні показники, які постійно пропускалися багатьма експертами, можуть бути переглянуті, переформульовані, або відкинуті. Найкращі показники (скажімо 10-15) для кожного конструкту обираються для подальшого аналізу. Кожний показник знову оцінюється експертами щодо очевидної та змістовної валідності. Якщо на цьому етапі не вдається отримати адекватний набір показників, можливе розроблення нових показників, виходячи з

концептуального визначення конструкту, що розробляється. Може знадобитися два чи три раунди Q-сортування, перш ніж експерти дійдуть згоди стосовно набору показників, що найкраще представляють досліджуваний конструкт.

Далі валідація переходить в емпіричну область. Створюється дослідницький апарат, що включає всі «очищені» показники конструкту й розповсюджується в пілотній репрезентативній групі респондентів. Дані, що збираються, заносяться в таблицю й проходять кореляційний аналіз або пошуковий факторний аналіз із використанням таких програмних продуктів як SAS чи SPSS для оцінки конвергентної і дискримінантної валідності. Показники, що не відповідають очікуваним нормам факторного навантаження (навантаження одного фактора вище за 0,60, навантаження між факторами менше за 0,30) на цій стадії вибраковуються. Шкали, що залишаються, оцінюються на надійність, використовуючи міру внутрішньої узгодженості, таку як альфа-коефіцієнт Кронбаха. На цій стадії також має бути перевірена вимірність шкал, залежно від того, чи досліджувані конструкти були концептуалізовані як одновимірні чи багатовимірні. Потім оцінюється здатність кожного конструкту до прогнозування в межах теоретично встановленої номологічної мережі конструкту з використанням регресійного аналізу чи моделювання структурним рівнянням. Якщо показники конструкту задовольняють більшості або всім вимогам надійності та валідності, що були описані в цьому розділі, це буде означати, що операціоналізовані показники є достатньо адекватними й точними.

Інтегрований підхід до оцінки валідності, що тут обговорювався, вимагає від дослідника як зусиль, так і часу. Однак необхідно пройти цей складний багатоетапний процес, щоб ті шкали вимірювання, що будуть застосовуватися в дослідженні, відповідали очікуваним нормам наукових досліджень. Узагалі валідація шкал і вимірювання залишаються одним із найбільш важливих етапів емпіричного дослідження.

Розділ 8

ФОРМУВАННЯ ВИБІРОК

Формування вибірок – це статистичний процес вибору групи (яка називається вибіркою) із генеральної сукупності об'єктів дослідження (популяції) з метою проведення спостережень і статистичних висновків та узагальнень щодо всієї генеральної сукупності. Загалом, дослідження соціальних наук мають на меті встановлення закономірностей поведінки певних генеральних сукупностей. Однак у більшості випадків неможливо дослідити генеральні сукупності через їх обсяг і відповідні ресурсні обмеження на проведення досліджень, тому доводиться формувати репрезентативні вибірки з генеральної сукупності для спостереження та аналізу. Важливо обрати вибірку, яка дійсно представляє генеральну сукупність. Неправильне формування вибірок – це перша причина помилкових суджень щодо поведінки генеральної сукупності.

Процес формування вибірок

Процес формування вибірок об'єднує кілька етапів (рис. 8.1).



Рис. 8.1. Процес формування вибірок

Перший етап – це встановлення генеральної сукупності, що буде досліджуватися. Генеральна сукупність може бути визначена як усі об'єкти, що містять характеристики, які підлягають дослідженню. Об'єктом дослідження може бути окрема людина, група, організація, країна чи будь-який інший об'єкт, стосовно якого дослідник збирається зробити наукові пояснення і висновки. Іноді генеральна сукупність є очевидною. Наприклад, якщо виробник хоче встановити, чи відповідає вироблена ним продукція вимогам якості, тоді генеральна сукупність

буде складатися з повного набору виробленої продукції, що виготовляється на даному підприємстві. У деяких випадках досліджувану генеральну сукупність встановити складніше. Якщо метою дослідження є виявлення рушійних сил, що спонукають до навчання студентів вищих навчальних закладів, тоді альтернативними варіантами генеральної сукупності можуть розглядатися самі студенти, викладачі, адміністрація вищих навчальних закладів чи навіть батьки студентів. Правильним у цьому випадку буде дослідження самих студентів, оскільки, урахувавши мету дослідження, потрібно проаналізувати результати їхнього навчання.

Другим етапом у процесі формування вибірки є визначення **меж вибірки**. Це та частина досліджуваної генеральної сукупності, яка доступна для дослідження, і з якої може сформуватися вибірка. Якщо об'єктом дослідження є фахівці певної професійної сфери, тоді реалістичним буде дослідження працівників місцевих компаній, що дадуть згоду на участь у дослідженні. Вони й будуть представляти межі вибірки.

Межі вибірки не завжди достатньою мірою представляють генеральну сукупність, і в цьому випадку висновки, зроблені щодо вибірки, не можуть бути узагальнені на всю генеральну сукупність. Наприклад, якщо об'єктом дослідження обраний управлінський склад виробничих підприємств, а вибірку складають управлінці декількох вітчизняних підприємств харчової промисловості, то отримані результати не можна поширити на управлінців всіх вітчизняних промислових підприємств.

На останньому етапі процесу формування вибірки за допомогою спеціальних методик у межах вибірки обирається певна група, що, власне, і буде підлягати дослідженню. Способи формування вибірок можна згрупувати у дві великі групи: випадкове й не випадкове формування вибірок. Ідеальним є випадкове формування вибірок, якщо узагальнення результатів є важливим для дослідження, що проводиться, однак є широкий спектр обставин, за яких не випадкове формування вибірок є виправданим. Ці методики висвітлюються в наступних двох підрозділах.

Вірогідне формування вибірки

Вірогідне формування вибірки є таким способом, при якому кожний елемент генеральної сукупності має шанс (не нульову вірогідність) потрапити до вибірки, і цей шанс може бути точно визначеним. Статистичні показники вибірки, такі як середнє та стандартне відхилення, є точними оцінками параметрів генеральної сукупності, якщо елементи вибірки зважені згідно з вірогідністю їх відбору. Усім способам випадкового формування вибірок властиві дві спільні риси: 1) кожний елемент сукупності має відому ненульову

вірогідність потрапити у вибірку; 2) процедура формування вибірки на певному етапі передбачає випадковий добір. Різні типи випадкового формування вибірок є такими.

Просте випадкове формування вибірок. За цим способом усі можливі підгрупи в генеральній сукупності (а точніше в межах вибірки) отримують рівні шанси потрапити до вибірки. Вірогідність відбору будь-якого набору з n елементів із загальної кількості N елементів, що входять до меж вибірки, позначається як ${}^N C_n$. Отже, статистичні показники вибірки є точними оцінками параметрів генеральної сукупності без надання ваги. Просте випадкове формування вибірок включає випадковий добір респондентів із меж вибірки, однак у разі великих меж вибірки рекомендується використовувати таблицю випадкових чисел чи комп'ютеризований генератор випадкових чисел. Наприклад, якщо дослідник хоче обрати 200 компаній для анкетування зі списку, що містить 1000 компаній, можна ввести список в Excel і використати функцію RAND(), щоб згенерувати випадкові числа для кожної з 1000 компаній. Потім компанії потрібно розташувати в зростаючому порядку за отриманими випадковими номерами й у вибірку добрати перші 200 компаній із утвореного таким чином списку. Це найпростіший спосіб випадкового формування вибірок, і простота є його важливою перевагою. Оскільки межі вибірки попередньо не діляться на частини, вибірка добре представляє генеральну сукупність, і отримані в дослідженні висновки будуть найбільш загальними порівняно з іншими способами ймовірнісного формування вибірок.

Систематичне формування вибірок. За цим способом межі вибірки впорядковуються за певним критерієм, і елементи обираються через однакові інтервали в утвореному списку. Систематичне формування вибірок передбачає встановлення початкової точки відліку, від якої добирається кожний із k елементів:

$$k=N/n,$$

де k – відношення розміру меж формування вибірки N і бажаного розміру вибірки n , і називається *відношенням вибірки*. Важливо, що точка відліку не є автоматично першою в списку, а обирається випадково з перших k елементів списку. У попередньому прикладі вибору 200 компаній із 1000 можна відсортувати наявні 1000 компаній у зростаючому порядку за їхнім розміром, випадково обрати одну з перших п'яти компаній створеного списку і потім обирати кожну п'яту компанію. Цей процес забезпечить відсутність завищеного числа великих чи малих компаній у вибірці, отже, вибірка буде репрезентативною вибіркою генеральної сукупності щонайменше за обраним критерієм (розмір компанії).

Стратифіковане формування вибірки. При стратифікованому формуванні вибірки межі вибірки поділяються на однорідні підгрупи, які

не перетинаються між собою (вони називаються стратами), і в межах кожної підгрупи береться проста випадкова вибірка. У попередньому прикладі з вибором 200 компаній із 1000 можна виділити три категорії компаній за їхнім розміром: великі (більш ніж 500 працівників), середні (від 50 до 500 працівників) і малі (до 50 працівників). Потім можна випадково обрати 67 компаній із кожної підгрупи, щоб сформувати вибірку з 200 компаній. Однак, оскільки число малих компаній більше, ніж великих, сформована таким чином вибірка не буде повністю репрезентувати генеральну сукупність (великі компанії становитимуть непропорційно велику частину вибірки). Такий спосіб називається *непропорційним стратифікованим формуванням вибірки*, оскільки пропорції вибірки не відображають пропорції в межах вибірки, і частка невеликої підгрупи (у даному випадку підгрупи підприємств із більш, ніж 500 працівниками) буде завищеною. Альтернативним способом буде добір підгруп вибірки пропорційно їх частці в генеральній сукупності. Наприклад, якщо є 100 великих компаній, 300 середніх і 600 малих, тоді в досліджувану вибірку мають потрапити 20 великих, 60 середніх і 120 малих компаній. У цьому випадку у сформованій вибірці зберігається пропорційне співвідношення між компаніями різного розміру в генеральній сукупності. Такий спосіб називається *пропорційним стратифікованим формуванням вибірки*. Непропорційний підхід важливий, коли потрібно забезпечити репрезентацію малих підгруп і не обов'язково буде менш репрезентативним порівняно з пропорційним підходом, якщо результати, отримані при застосуванні непропорційного підходу зважуються відповідно до пропорції певної підгрупи у генеральній сукупності.

Кластерне формування вибірок. Якщо досліджувана генеральна сукупність розсіяна на великій території, буде проблематично провести просте випадкове формування вибірки з такої сукупності. У цьому випадку може бути доцільним поділити сукупність на «кластери» (зазвичай за географічними кордонами), випадково обрати декілька кластерів і дослідити всі елементи у створених кластерах. Наприклад, якщо потрібно створити вибірку з міських адміністрацій Київської області, можна створити кластери таких адміністрацій за районами області, випадково обрати групу з трьох районів і потім провести інтерв'ю чиновників міської адміністрації цих районів. Однак, зважаючи на відмінності між кластерами, варіативність оцінок вибірки в кластерній вибірці зазвичай буде вищою, ніж у простих випадкових вибірках, отже, і результати будуть менш відображати генеральну сукупність, ніж при формуванні простих випадкових вибірок.

Попарне формування вибірок. Іноді дослідники хочуть порівняти дві підгрупи однієї генеральної сукупності за якимось критерієм. Наприклад, чому одні компанії мають постійно більшу прибутковість, ніж інші? Для проведення такого дослідження потрібно

розбити межі вибірки компаній на дві групи: високоприбуткові та малоприбуткові, базуючись на одному з показників прибутковості. Потім потрібно сформувати просту випадкову вибірку в одній підгрупі й до кожної компанії з такої вибірки дібрати відповідну за розміром, галуззю чи іншими критеріями компанію з іншої підгрупи. Так створюються дві парні вибірки високо прибуткових і низько прибуткових компаній, які потім можуть детально досліджуватися. Спосіб попарного формування вибірок є ідеальним для розуміння біполярних відмінностей між різними підгрупами в певній генеральній сукупності.

Багатоетапне формування вибірок. Описані вище способи ймовірного формування вибірок є прикладами одноетапних способів. Залежно від потреби, можливо скомбінувати одноетапні способи та провести багатоетапне формування вибірок. Наприклад, якщо дослідник стратифікує список компаній, базуючись на їх розмірі, а потім проводить систематичне формування вибірок у кожній страті. Це буде двоетапна комбінація стратифікаційного та систематичного формування вибірок. Аналогічно, можна спочатку сформувати кластери, а потім у них провести просте випадкове формування вибірок. Якщо, наприклад, у такі вибірки потраплять школи декількох районів Києва, а в кожній із них формується проста випадкова вибірка класів, із яких формується вибірка учнів, що братимуть участь у дослідженні, тоді матимемо чотирьохетапний процес формування вибірок, що складається з кластерного й трьох простих випадкових формувань вибірок.

Невірогідне формування вибірок

Невірогідне формування вибірок – це формування вибірок, при якому деякі елементи генеральної сукупності мають нульову вірогідність добору або вірогідність добору не можна точно встановити. При цьому елементи вибірки добираються на основі певних не випадкових критеріїв, таких як квота чи зручність. Оскільки добір є не випадковим, невірогідне формування вибірок не дозволяє оцінити помилки формування вибірок і може стати джерелом похибки. Отже, інформація, отримана при дослідженні вибірки, не може бути узагальнена до рівня генеральної сукупності. Способи невірогідного формування вибірок включають такі.

Формування вибірок за зручністю. Цей спосіб, який також називається формуванням вибірок за можливістю, передбачає, що вибірка формується з доступної для дослідження частини генеральної сукупності або формування вибірки є найбільш зручним і легким із практичної точки зору. Прикладом формуванням вибірок за зручністю є опитування відвідувачів найближчого супермаркету. Це є не вірогідна вибірка, оскільки в дослідження не можуть потрапити відвідувачі інших супермаркетів. Результати анкетування, отримані при дослідженні такої вибірки, можуть виявитися специфічними саме для даного супермаркету й не будуть відображати думку генеральної сукупності, (якщо такою

виступають покупці всіх супермаркетів). Можливість узагальнення таких результатів обмежена. Іншим прикладом формування вибірки за зручністю є дослідження студентів однієї групи. Таке формування вибірки найбільш відповідне для проведення пілотного дослідження, метою якого є тестування інструментів дослідження або валідація вимірювань, а не отримання узагальнених висновків дослідження.

Формування вибірок за квотою. За цим способом, як і у випадку стратифікації, генеральна сукупність сегментується на підгрупи, елементи яких не перетинаються, і потім із кожної з них обирається не випадковий набір спостережень за визначеною квотою. При *формуванні вибірок за пропорційною квотою* пропорція респондентів у кожній підгрупі повинна відповідати її пропорції у генеральній сукупності. *Формування вибірок за непропорційною квотою* менш жорстке в тому сенсі, що досліднику не обов'язково досягати пропорційне представництво, лише потрібно забезпечити мінімальний розмір кожної підгрупи. Жоден із типів формування вибірок за квотою не є репрезентативним відносно генеральної сукупності, особливо непропорційне квотування, однак цей спосіб може бути корисним, щоб отримати дані від невеликих підгруп шляхом збільшення їх частки в загальній кількості досліджуваних.

Експертне формування вибірок. Згідно з цим способом респонденти обираються не випадково за їхніми експертними знаннями щодо досліджуваного феномену. Перевагою цього підходу є те, що залучаються експерти, які краще обізнані з досліджуваним феноменом, ніж інші особи, і тому їхня думка є більш вагомою, ніж думка вибірки, що складається з експертів і неекспертів. При цьому отримані результати не можуть бути узагальнені до рівня генеральної сукупності.

Формування сукупності за способом сніжної лавини. У цьому випадку дослідження починається з ідентифікації декількох респондентів, що відповідають критеріям включення в дослідження. Потім респондентів просять порекомендувати інших осіб, які відповідають встановленим критеріям і можуть взяти участь у дослідженні. Наприклад, якщо досліджуються адміністратори комп'ютерних мереж, можна почати з декількох знайомих, які порекомендують своїх колег із інших компаній. Іноді це може бути єдиним способом доступу до окремих закритих груп генеральної сукупності або таких, для яких немає встановлених меж вибірки.

Статистика формування вибірок

У попередніх підрозділах були введені терміни "параметр генеральної сукупності", "статистичний показник вибірки" і "похибка формування вибірки". У цьому підрозділі ці поняття та зв'язки між ними будуть досліджені більш детально.

Якщо проводиться спостереження певного показника для одного з елементів вибірки, наприклад, фіксується відповідь респондента на питання шкали Лікерта, то таке спостереження називається відповіддю (див. рис. 8.2). Іншими словами, відповідь – це значення показника, отримане від одного елемента вибірки. Кожний респондент надає різні відповіді на різні питання. Відповіді різних респондентів на одне і те ж питання можуть бути представлені графічно як **розподіл частот** на основі частоти їх спостереження. Для великого числа відповідей у вибірці цей розподіл частот прагне утворити так званий **нормальний розподіл**, який можна використати для оцінки загальних характеристик вибірки в цілому, а саме середнього значення та стандартного відхилення. Ці оцінки вибірки називаються статистиками, або **статистичними показниками**, які є оцінками даних спостереження. Генеральні сукупності також характеризуються середніми та стандартними відхиленнями, однак, оскільки дослідити генеральну сукупність неможливо, її характеристики є невідомими й називаються **параметрами генеральної сукупності**, а не статистичними показниками, оскільки вони не оцінюються статистично за даними спостережень. Статистичні показники вибірки можуть відрізнитися від параметрів генеральної сукупності, якщо вибірка недостатньо добре представляє сукупність (не є репрезентативною). Різниця між статистичним показником і параметром називається **похибкою формування вибірки**. Теоретично, із збільшенням розміру вибірки, у міру наближення її розміру до генеральної сукупності, похибка вибірки зменшується й статистичний показник наближається до відповідного параметра генеральної сукупності.

Якщо вибірка добре представляє генеральну сукупність, тоді *оцінені* статистичні показники вибірки будуть ідентичними відповідним *теоретичним* параметрам генеральної сукупності. Як можна дізнатися, що статистичні показники достатньо наближені до параметрів генеральної сукупності? Для цього потрібно розуміти концепцію розподілу формування вибірки. Уявіть, що було взято три різні випадкові вибірки з генеральної сукупності, як показано на рис. 8.3., і для кожної з них розрахували середнє та стандартне відхилення. Якщо кожна вибірка дійсно є репрезентативною, тоді статистичні показники трьох вибірок мають співпасти (і дорівнювати параметру генеральної сукупності), а варіативність середніх трьох вибірок повинна дорівнювати нулю. Однак така ситуація малоймовірна, оскільки кожна випадкова вибірка міститиме дещо відмінні елементи генеральної сукупності, і середні вибірок будуть відрізнитися. Якщо число таких вибірок збільшити від 3 до 10 чи 100, тоді гістограма частот перетвориться на розподіл формування вибірок. Отже, розподіл формування вибірок є частотним розподілом *статистичного показника* (як середнє вибірки) з набору вибірок, тоді як часто вживаний розподіл частот є *розподілом відповідей* (спостережень) *однієї вибірки*. Як і у випадку розподілу частот, при збільшенні числа вибірок їхні статистичні показники щільніше групуються навколо середньої (яка,

вірогідно, є оцінкою параметра генеральної сукупності). При нескінченно великому числі вибірок цей розподіл наближається до нормального. Варіативність, або щільність значень статистичного показника вибірки в розподілі вибірок називається **стандартною помилкою** (тобто стандартним відхиленням статистичного показника вибірок). Термін стандартне відхилення, навпаки, відноситься до варіативності відповідей спостереження однієї вибірки.

Назви показників	ATT1	ATT2	ATT3	INT1	INT2	INT3	FRSQ	HT3
Індивідуальні відповіді	6	6	6	6	6	6	6	2
Всі відповіді одного респондента	1	1	1	1	1	1	1	10
Набір відповідей всіх респондентів для одного показника становлять розподіл частот	6	5	4	2	3	3	2	4
Середнє чи стандартне відхилення є статистичними показниками вибірки	6	6	6	6	5	5	5	2
	6	6	6	6	5	4	2	2.5
	7	7	7	7	7	6	2	3
	5	4	5	5	4	5	2	1
	6	6	6	6	7	7	3	10
	7	6	6	5	5	4	2	1
	4	4	4	4	4	4	1	0
	5	5	3	4	4	4	2	1
	4	4	4	4	4	4	2	3
	6	6	6	5	5	5	2	4
	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1
	-1	-1	-1	-1	-1	-1	-1	-1
	6	6	6	6	6	6	2	15
	6	6	5	5	5	4	2	2
	1	1	1	1	1	1	1	0
	6	6	6	6	6	6	2	6
	6	6	5	6	6	5	2	3
	5	5	5	2	2	2	2	7
	4	5	6	1	3	2	2	3
	6	7	7	6	7	6	2	3

Рис. 8.2. Статистичні показники вибірки

Середнє значення статистичного показника в розподілі вибірок слід уважати оцінкою невідомого параметра генеральної сукупності. Базуючись на скупченості розташування значень у розподілі вибірок (тобто на стандартній помилці), можливо оцінити довірчі інтервали для прогнозування параметру генеральної сукупності. **Довірчий інтервал** є оцінена вірогідність, що параметр генеральної сукупності лежить у межах певного інтервалу значень статистичного показника вибірки. Для всіх нормальних розподілів діє правило 68-95-99 відсотків (див. рис. 8.4), яке означає, що понад 68% значень у розподілі лежить у межах одного стандартного відхилення від середнього значення ($\mu \pm 1\sigma$), понад 95% значень – у межах двох стандартних відхилень від середнього ($\mu \pm 2\sigma$), і понад 99% значень – у межах трьох стандартних відхилень від середнього ($\mu \pm 3\sigma$). Оскільки розподіл вибірок при нескінченному числі вибірок наближається до нормального, буде діяти те ж правило 68–95–99, і тому можна сказати, що:

- діапазон «середнє статистичного показника \pm одна стандартна помилка» представляє 68% довірчий інтервал для параметра генеральної сукупності;

- діапазон «середнє статистичного показника \pm дві стандартні помилки» представляє 95% довірчий інтервал для параметра генеральної сукупності;
- діапазон «середнє статистичного показника \pm три стандартні помилки» представляє 99% довірчий інтервал для параметра генеральної сукупності.

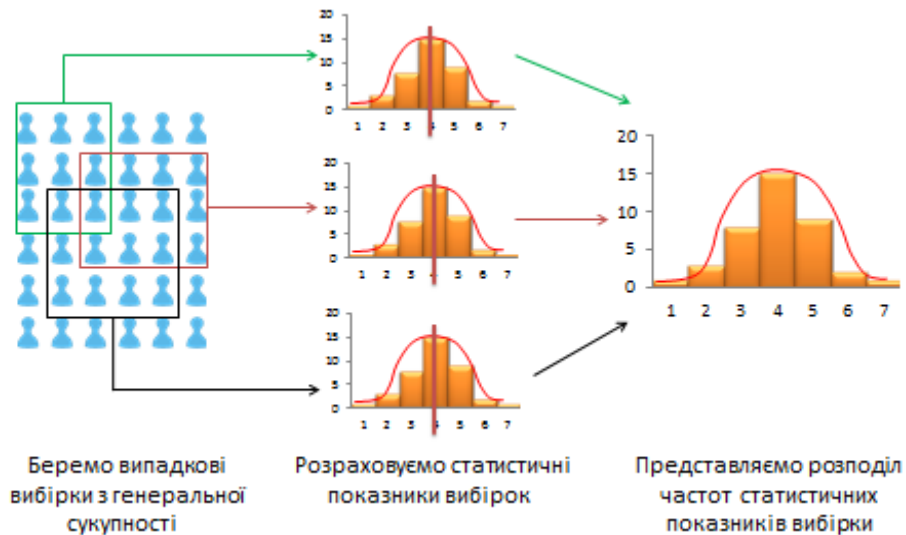


Рис. 8.3. Розподіл вибірок

Вибірка є хибною (нерепрезентативною), якщо розподіл вибірок не може бути оцінений або розподіл вибірок не підпорядковується правилу 68-95-99. Варто відзначити, що в регресійному аналізі, коли оцінюється істотність коефіцієнтів регресії на рівні $p < 0.05$, ми намагаємося встановити, чи прогнозує показник вибірки (коефіцієнт регресії) відповідний параметр (розмір істинного ефекту) із 95% довірчим інтервалом. Цікаво, що стандарт «шість сигм» намагається встановити виробничий дефект поза межами 99% довірчого інтервалу, або шести стандартних відхилень, що означає тестування істотних відмінностей на рівні $p < 0.01$.

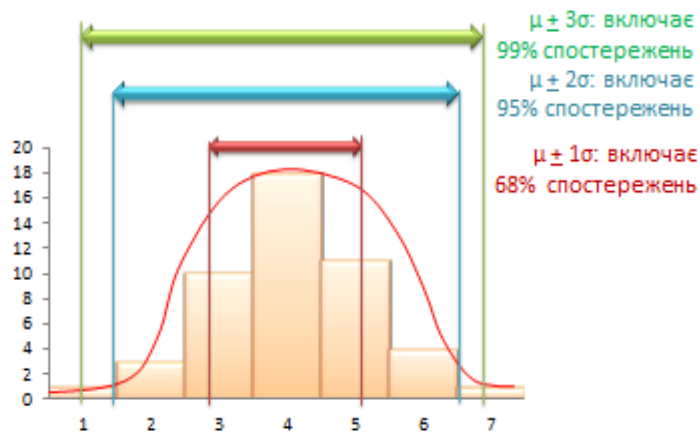


Рис. 8.4. Правило 68-95-99% для довірчого інтервалу

Розділ 9

АНКЕТНЕ ДОСЛІДЖЕННЯ

Анкетне дослідження – це науковий метод, що передбачає використання стандартизованих анкет чи інтерв'ю для збору даних стосовно людей, їх уподобань, думок і поведінки на системній основі. Хоча опитування населення проводилися вже в Давньому Єгипті, анкетування, як формальний метод дослідження, було започатковано в 1930-1940 рр. Паулем Лазарсфельдом для дослідження впливу радіо на формування політичної думки в США. Із того часу метод набув великої популярності для кількісних досліджень у соціальних науках.

Метод опитування може використовуватися для описових, пошукових та пояснювальних досліджень. Він найкраще доречний для досліджень, у яких об'єктом дослідження виступають окремі люди. Хоча інші об'єкти дослідження, такі як групи, організації чи діади (пари організацій, такі як продавці і покупці) також досліджуються за цим методом, у такі дослідження часто залучають окремих осіб як ключові джерела інформації або «проксі». Анкетні дослідження можуть виявитися досить суб'єктивними через вплив похибки респондента, якщо респондент не має необхідних знань, або його думки про досліджуваний феномен хибні. Наприклад, керівники вищого рівня не завжди адекватно оцінюють сподівання своїх підлеглих щодо командної роботи в компанії і тому можуть надати дослідникам хибну інформацію з цього питання.

Анкетне дослідження має декілька важливих переваг порівняно з іншими методами. По-перше, опитування є джерелом отримання широкого кола даних, які неможливо отримати через безпосереднє спостереження, наприклад, щодо преференцій, ставлення, переконань, цінностей, відчуттів чи мотивів. По-друге, опитування також ідеальне для збору даних у великих групах, які через географічну віддаленість складно досягти. У масштабах країни великі території можна охопити за допомогою електронної пошти чи телефонних опитувань, використовуючи спеціально підібрані вибірки, щоб забезпечити їх адекватну репрезентативність. По-третє, завдяки неав'язливій формі анкетування й можливості відповідати в зручному для респондентів форматі, деякі респонденти можуть віддавати перевагу саме анкетному опитуванню. По-четверте, анкети можуть бути єдиним способом отримати дані від певних груп респондентів, для яких не встановлені вибіркові межі. По-п'яте, опитування великих за розміром вибірок дозволяє виявити ефекти, які були б непомітні в маленьких вибірках, навіть якщо мова йде про численні змінні. У залежності від дизайну опитування уможливується порівняльний аналіз підгруп генеральної сукупності (тобто внутрішньогруповий і міжгруповий аналізи). По-шосте,

анкетне дослідження є достатньо практичним стосовно витрат часу, зусиль і ціни порівняно з експериментом чи кейс-дослідженням. Водночас, анкетне дослідження має деякі характерні недоліки. Серед них – пропущені питання, похибки формування вибірок, соціально бажані відповіді, неповернення анкет. Недоліки анкетування будуть обговорюватися в останньому підрозділі цього розділу.

Залежно від того, як збираються дані, анкетні дослідження можуть бути поділені на дві великі категорії: анкетні опитування (які можуть виконуватися по пошті, у групах чи в режимі он-лайн) та інтерв'ю (які можуть бути особистими, телефонними, чи у фокус-групах). Анкети є інструментами, що заповнюються респондентами письмово, тоді як інтерв'ю проводяться інтерв'юерами та базуються на вербальних відповідях респондентів. Кожний тип опитування має свої сильні сторони й обмеження стосовно ціни, покриття необхідного контингенту та гнучкості сценарію проведення дослідження.

Анкетні опитування

Уперше запропоноване Френсісом Гальтоном, **анкетне опитування** є інструментом дослідження, що складається з набору питань (пунктів), які розраховані на отримання відповідей респондентів у стандартній формі. Питання можуть бути неструктуровані та структуровані. Неструктуровані питання передбачають відповіді респондентів у довільній формі, тоді як відповіді респондентів на структуровані питання вимагають вибору з наданих альтернативних варіантів. Відповіді респондентів на індивідуальні питання в структурованій анкеті можуть бути агреговані в одну складну шкалу чи індекс для статистичного аналізу. Питання повинні бути сконструйовані так, щоб респонденти були здатні їх прочитати, зрозуміти та дати усвідомлені відповіді. Отже, анкетне опитування не може застосовуватися по відношенню до дітей і неписьмених.

Більшість анкетних досліджень проводяться **на індивідуальному рівні**, коли анкетні форми надсилаються респондентам поштою. Одна й та ж анкета надходить до великого числа людей, і респонденти, які хочуть взяти участь у дослідженні, можуть заповнити її у зручний час, а потім повернути в проплачених конвертах, що надаються. Перевагою поштових анкет є те, що вони недорогі в розповсюдженні і можуть заповнюватися в зручний для респондента час. Однак рівень їх повернення досить низький, оскільки багато людей їх просто ігнорують. Можливі також довгі затримки із заповненням і поверненням анкет. Отже, дослідник повинен постійно відслідковувати кількість повернених анкет і надсилати нагадування тим, хто не відповів (в ідеалі два-три нагадування протягом 1-1,5 місяців). Анкетні дослідження також не зовсім доречні, якщо респондентам необхідно давати пояснення щодо заповнення анкети або якщо від респондентів очікуються розгорнуті відповіді. Повторне

опитування за однією анкетною одних і тих же респондентів у різні періоди часу можливе, однак із кожним повторним анкетуванням кількість повернених анкет суттєво знижується.

Другий тип анкетування – **дослідження, що проводяться в групах**. Вибірка респондентів збирається разом, і кожного респондента просять самостійно заповнити анкету. Цей формат зручний для дослідника й забезпечує високий рівень повернення заповнених анкет. Зручним є й те, що респондент, у якого виникли питання щодо заповнення анкети, може відразу отримати на них відповіді. У багатьох організаціях, якщо на це є згода керівництва, зібрати групу працівників не складно.

Відносно новий тип опитування – це он-лайн чи веб- опитування. Анкети надсилаються через Інтернет із використанням інтерактивних форм. Проведення таких досліджень не є дорогим, їх результати постійно записуються в он-лайнній базі даних, і за потреби анкета може бути легко модифікована. Однак, веб-анкетування не захищений, тому існує ризик псування анкет чи багаторазове заповнення анкети. Окрім того, таке опитування не зможе охопити тих, хто не має комп'ютера або доступу до Інтернету, що значно збільшує похибку формування вибірки. Вибірка респондентів зміщується в бік молодих людей, які постійно користуються Інтернетом і мають час для заповнення анкети. Розрахунок кількості повернених анкет може стати проблематичним, якщо респонденту надсилається посилання на місцезнаходження анкети, а не сама анкета (на електронну пошту). Із цих причин багато дослідників надсилають анкети як поштою, так і по Інтернету, дозволяючи респондентам самостійно обрати зручний спосіб заповнення анкети.

Конструювання анкети для опитування є мистецтвом. Треба прийняти численні рішення стосовно змісту питань, їх словесної форми, формату та послідовності, які можуть мати важливі наслідки щодо очікуваних відповідей.

Формати відповідей. Питання анкети можуть бути структуровані та неструктуровані. Відповіді на структуровані питання отримуються з використанням таких форматів:

- *Дихотомічна відповідь.* Респондентів просять обрати один із двох можливих варіантів відповідей, таких як правильно/неправильно, так/ні, згоден/не згоден. Приклад такого питання: Ви маєте роботу в даний час?
- *Номінальна відповідь.* Респонденти мають більш, ніж два можливих варіанти вибору з набору категорій. Приклад: У якій галузі Ви працюєте: виробництво, сільське господарство, обслуговування, продаж, освіта, охорона здоров'я, інша.

- *Порядкова відповідь.* Респонденти мають більш, ніж два можливих варіанти вибору з певної градації. Приклад: Яка у Вас освіта: вища освіта, середня спеціальна, середня.

- *Інтервальна відповідь.* Респондентам пропонується 5- чи 7-бальна шкала Лікерта, шкала семантичного диференціалу чи шкала Гуттмана, які розглядалися в попередньому розділі.

- *Числова відповідь.* Респонденти відзначають певне числове значення на шкалі відношень, яка має змістовне нульове значення, наприклад, вік або стаж роботи на даному підприємстві.

Зміст і форма питань. Відповіді, отримані в анкетному дослідженні, дуже чутливі до питань, що ставляться дослідником. Погано сформульовані або незрозумілі питання можуть призвести до сумнівних відповідей, що не мають наукової цінності. Дільман (1978) рекомендує декілька правил для створення коректних питань анкети. Кожне питання потрібно ретельно перевіряти, враховуючи наступні вимоги:

- *Чи є питання чітко сформульованим і доступним для розуміння?* Питання формулюються максимально простою мовою, із уникненням складних слів, жаргону чи професійних термінів, які можуть бути незрозумілими для певних груп респондентів. Винятком є анкетування спеціалістів – лікарів, юристів, IT-спеціалістів, які користуються певною професійною лексикою в повсякденному житті.

- *Чи не сформульоване питання в негативному світлі?* Негативно сформульовані питання, наприклад, стосовно рівня конфліктності в робочому колективі, можуть викликати неоднозначну реакцію респондентів і призвести до нещирих відповідей або взагалі відсутності відповідей. Таких питань потрібно уникати.

- *Чи є питання однозначним?* Питання анкети не повинні містити слова чи вирази, які можуть по-різному інтерпретуватися різними респондентами (наприклад, питання щодо доходу може стосуватися місячного чи річного доходу, особистого доходу чи доходу родини). Різна інтерпретація матиме наслідком не порівнювані відповіді респондентів.

- *Чи не містить питання слів і виразів, що нашкодують респондентів на певну відповідь?* Іноді формулювання питання може схилити респондентів до певної відповіді та внести похибку в дослідження. Наприклад, Кеннет Расінскі (1989) проаналізував декілька досліджень, присвячених ставленню респондентів до соціальних витрат державного бюджету. З'ясувалося, що респонденти більше підтримують «допомогу бідним верствам населення», ніж «грошове утримання бідних», «рішення проблем великих міст», ніж «допомогу великим містам» і т.п., коли насправді йдеться про одні й ті ж державні заходи. Хибне формулювання може внести похибку у відповіді респондентів. Хоча часто складно передбачити наперед, яке саме формулювання може виявитися проблемним, потрібно якомога ретельніше перевіряти питання анкети для уникнення цієї помилки.

- *Чи не є питання подвійним?* Подвійні питання – це ті, що вимагають двох відповідей. Наприклад, питання «Чи задоволені ви рівнем обслуговування і якістю товару?» не має однозначної відповіді, оскільки респондент може бути задоволений якістю товару й не задоволений рівнем обслуговування компанії. Такі питання рекомендується ділити на два, щоб отримати однозначні відповіді.

- *Чи не є питання занадто загальним?* Іноді питання можуть виявитися занадто узагальненими для респондентів, і тому їхні відповіді будуть недостатньо точними й інформативними. Замість того, щоб запитати, чи є певна компанія успішною на думку респондента, краще запитати про кількість нових клієнтів, залучених протягом певного періоду, приріст доходів тощо.

- *Чи не є питання занадто деталізованим?* Необхідно уникати надмірної деталізації, якщо це не сприяє досягненню цілей дослідження. Наприклад, питання стосовно віку дітей, що знаходяться на утриманні члена родини, може бути зайвим і достатньо обмежитися питанням щодо кількості дітей на утриманні. Однак, якщо є сумніви, краще надати перевагу деталізації, ніж узагальненню.

- *Чи не є питання упередженим?* Наприклад, питання щодо переваг зменшення податків передбачає, що респондент бачить певні переваги в зменшенні податків, хоча не всі люди узагалі вітають зменшення податків, оскільки це може призвести до скорочення соціальних програм.

- *Чи не є питання надуманим?* Питання на кшталт «Щоб б ви зробили, якби виграли мільйон?» не можна застосовувати в наукових дослідженнях для отримання обґрунтованих висновків, оскільки більшість респондентів ніколи не стикалися з таким вирашем у реальності, і їхні відповіді не базуються на пережитому досвіді.

- *Чи володіють респонденти необхідною інформацією, щоб правильно відповісти на питання?* Часто дослідник розраховує, що респонденти володіють інформацією про певний феномен, хоча насправді вони її не мають. Навіть якщо відповідь буде надана, вона буде неточною і не матиме наукової цінності. Наприклад, виконавці певного робочого завдання добре знають всі деталі його виконання, тоді як директор компанії володіє лише загальною інформацією про його виконання і тому не є найкращим респондентом для відповідної анкети.

Послідовність питань. Загалом, питання повинні слідувати одне за одним в логічній послідовності. Щоб мати високий рівень повернення анкет, питання мають розташовуватися від менш складних до більш складних, від питань стосовно фактів і поведінки до питань стосовно відношень, від загальних до деталізованих. Існують загальні правила щодо послідовності питань:

- Починайте з простих питань, відповіді на які не несуть загрози для респондента, і на які легко відповісти. Це можуть бути питання щодо

демографічних даних (вік, стать, освіта) для анкет особистісного рівня та питання щодо загальних характеристик компанії (число працівників, галузь, форма власності), для анкет організаційного рівня.

- Ніколи не починайте з відкритих питань.
- Якщо є потреба в дотриманні історичної послідовності подій, дотримуйтеся хронологічного порядку від більш ранніх до більш пізніх подій.
- В одному питанні запитуйте лише про одну річ. Змінюючи тему, використовуйте перехід «Наступний розділ досліджує думки респондентів щодо...».
- Використовуйте фільтри, або інструкції для переходу від однієї групи питань до інших, наприклад, «Якщо Ви відповіли «так» на це питання, переходьте до розділу 2, якщо «ні», переходьте до розділу 3».

Інші золоті правила. Ставтеся до ваших респондентів так, як ви б хотіли, що вони ставилися до вас. Цінуйте їх час, увагу й довіру, дотримуйтеся конфіденційності щодо особистої інформації. При проведенні анкетних досліджень завжди спирайтеся на такі стратегії:

- Поважайте час людей, що беруть участь у дослідженні. Намагайтеся зробити анкету максимально короткою. Респонденти не будуть налаштовані витратити більш, ніж 10-15 хвилин на відповіді, не зважаючи на їх важливість для дослідника. Довгі анкети суттєво зменшують рівень повернення анкет.
- Завжди інформуйте респондентів про рівень конфіденційності дослідження, і те, як їхні відповіді будуть використовуватися (тобто для наукових досліджень), як будуть оприлюднюватися результати досліджень (зазвичай у цілому по групі, в узагальненому вигляді).
- Для організаційних досліджень пообіцяйте надіслати компанії копію звіту й на забудьте виконати обіцянку.
- Подякуйте респондентам за участь у дослідженні.
- Завжди апробуйте анкету в пілотному дослідженні перед тим, як надати її респондентам досліджуваної вибірки. Це допоможе уникнути неточностей, виправити деякі формулювання перед анкетуванням.

Інтерв'ювання

Інтерв'ю є більш персоналізованою формою збору даних, ніж анкетне опитування. Воно виконується підготовленими інтерв'юерами з використанням того ж протоколу дослідження, що й у випадку анкетного опитування (тобто стандартизований набір питань). Однак, на відміну від анкетування, бланк інтерв'ю може містити спеціальні інструкції для інтерв'юера, про які респонденти не будуть знати, а також може включати коментарі і спостереження інтерв'юера. На відміну від анкетування поштою, інтерв'юер має можливість уточнити деякі відповіді або поставити додаткові питання. Водночас, інтерв'ю є ресурсозатратним методом. Для його проведення потрібні спеціальні

навички. При проведенні інтерв'ю інтерв'юер фактично стає частиною інструменту вимірювання та повинен докладати всіх зусиль, щоб не впливати на відповіді респондентів і не вносити похибок у дослідження.

Найбільш типовою формою інтерв'ю є **особисте спілкування**, під час якого інтерв'юер безпосередньо працює з респондентом, задає питання та фіксує відповіді.

Різновидом персонального інтерв'ю є групове, або **фокус-група**. За цим способом невелика група респондентів (зазвичай 6-10) проходять одночасне інтерв'ювання. Інтерв'юер фактично виступає фасилітатором, чия робота полягає в проведенні дискусії і забезпеченні участі в ній усіх респондентів. Фокус-групи дозволяють провести глибше дослідження складних питань, ніж інші форми опитування, оскільки думки одних респондентів можуть активізувати роботу інших і призвести до виникнення нових ідей. Однак у дискусіях фокус-груп може домінувати одна особа, тоді як інші не наважуватимуться вільно висловлювати свої думки в присутності колег і керівництва, особливо коли йдеться про делікатні питання, наприклад, щодо розподілу робочого навантаження чи політики матеріальної винагороди в компанії. Через невелике число учасників фокус-групи зазвичай використовуються на етапі пошукового дослідження, а не описового чи пояснювального.

Третій тип інтерв'ювання – це **телефонні інтерв'ю**. За цим способом інтерв'юери контактують із потенційними респондентами по телефону, спираючись на випадковий відбір респондентів із телефонного довідника, щоб поставити стандартний набір питань. Більш сучасний підхід – це телефонне інтерв'ювання за допомогою комп'ютера, яке все частіше використовується науковцями й комерційними компаніями. Інтерв'юер виступає в ролі телефонного оператора та керується інструкціями комп'ютерної програми, яка висвітлює питання, що потрібно поставити, на екран дисплея. Система обирає респондентів, користуючись технікою випадкового набору номерів, і записує голосові відповіді респондентів. Цей спосіб не дуже підходить для сільської місцевості, де рівень телефонізації нижчий, і не може бути використаний для передачі графічної чи іншої візуальної інформації.

Роль інтерв'юера. Інтерв'юер під час інтерв'ю відіграє складну й багатогранну роль, яка включає виконання таких завдань:

- *Підготуватися до проведення інтерв'ю.* Якість зібраних даних значно залежить від кваліфікації інтерв'юера. Він повинен бути підготовленим до проведення інтерв'ю, обізнаним із цілями, методикою та джерелами помилок дослідження.
- *Визначити респондентів і спонукати їх до співпраці.* Інтерв'юер має проводити інтерв'ю в тому місці і в той час, що найбільше підходить респондентам, навіть якщо це незручно інтерв'юеру.

- *Мотивувати респондентів.* Респонденти часто підживлюються мотивацією інтерв'юера. Якщо ж така відсутня, то спонукати респондентів до конструктивної співпраці дуже складно. Інтерв'юер повинен випромінювати ентузіазм, переконувати у важливості дослідження і бути уважним до потреб респондентів.

- *Відповідати на будь-які питання чи побоювання респондентів стосовно дослідження.* Інтерв'юери повинні бути готові кваліфіковано відповідати на непередбачувані питання респондентів і самі ставити пілотні питання, яких нема в протоколі.

- *Контролювати якість відповідей.* Інтерв'юер контролює якість зібраної інформації і може доповнювати відповіді респондентів, спираючись на особисті спостереження.

Проведення інтерв'ю. До початку інтерв'ю інтерв'юер повинен підготувати все необхідне для його проведення, включаючи лист-дозвіл від керівництва організації, бланки анкет, контактну інформацію респондентів. Він повинен відкрити зустріч, представитися, стисло та переконливо пояснити мету дослідження. Якщо робота фокус-групи буде фільмуватися, інтерв'юер повинен отримати згоду респондентів. Навіть у випадку відеозапису зустрічі інтерв'юер повинен робити нотатки.

Під час інтерв'ю інтерв'юер повинен дотримуватися протоколу дослідження та ставити питання так, як вони написані, не намагаючись їх змінити в будь-який спосіб. Не можна змінювати послідовність питань чи пропускати якісь із них. Усі питання, що виносяться на інтерв'ю, слід попередньо обговорити на тренувальній сесії до початку дослідження. Інтерв'юер не повинен закінчувати речення замість респондентів, навіть якщо їхні відповіді здаються йому очевидними. Якщо респондент дає незрозумілу чи занадто коротку відповідь, інтерв'юер повинен спонукати його пояснити свою думку в більш розгорнутому вигляді. Для цього можна використовувати такі прийоми:

- *Пауза:* просто зробивши паузу після відповіді, яка видається недостатньо розгорнутою, інтерв'юер може показати респонденту, що очікує більшої деталізації.

- *Вербальне заохочення:* заохочуючи респондента до пояснень, інтерв'юер не повинен давати оціночних суджень стосовно того, що говорить респондент.

- *Прохання про пояснення:* інтерв'юер може прямо попросити респондента більш розгорнуто пояснити свою думку.

- *Рефлексія:* Інтерв'юер може спробувати такий психологічний прийом, як повторення останнього речення, сказаного респондентом, щоб спонукати його до продовження думки.

Після завершення інтерв'ю, інтерв'юер повинен подякувати респондентам за участь і повідомити, коли очікується отримати остаточні результати. Інтерв'юеру рекомендується записати ключові

спостереження, які допоможуть в подальшій інтерпретації коментарів респондентів.

Похибки опитування

Не дивлячись на свої сильні сторони та переваги, опитування респондентів як метод дослідження пов'язаний із певними систематичними похибками, які можуть поставити під сумнів висновки, що ґрунтуються на цьому методі. Ці п'ять похибок такі: похибка неповернення анкет, похибка формування вибірки, похибка соціальної бажаності відповідей, мнемонічна похибка та похибка спільного методу.

Похибка неповернення анкет. Загальновідомо, що опитування характеризуються низьким рівнем повернення анкет. Рівень повернення анкет у 15-20% є типовим для опитувань поштою навіть після двох чи трьох нагадувань. Якщо більшість респондентів не взяли участь в опитуванні, тоді виникає правомірне питання, чи не пояснюється цей факт дією якогось системного фактора, а отже, наскільки валідними будуть результати такого дослідження. Наприклад, незадоволені споживачі проявляють більшу активність, ніж задоволені споживачі, і тому з більшою вірогідністю візьмуть участь в опитуванні. У результаті кожна вибірка респондентів матиме вищу пропорцію незадоволених споживачів, ніж у цілому в генеральній сукупності, із якої формувалася вибірка. Отже, результати такого дослідження можна розглядати як артефакт, пов'язаний із формуванням вибірки, і їх важко узагальнити до рівня генеральної сукупності. Щоб збільшити рівень повернення анкет, можна звернутись до таких стратегій:

- *Попереднє сповіщення.* Короткий лист, що пояснює причини проведення дослідження і його значення, налаштовує респондентів на співпрацю і підвищує рівень повернення анкет.
- *Релевантність змісту.* Якщо опитування торкається важливих для респондентів проблем, тоді вірогідність повернення анкет зростає.
- *«Дружній інтерфейс» бланку анкети.* Чим коротше анкета, тим вищий рівень повернення. Аналогічно короткі, зрозумілі та неагресивні за змістом питання підвищують вірогідність повернення анкет.
- *Підтримка керівництва компанії.* Для проведення організаційних опитувань дослідникам важливо заручитися підтримкою керівництва. Це підвищить значущість опитування в очах працівників, а отже, і рівень повернення анкет.
- *Нагадування.* Декілька нагадувань можуть спонукати «повільних» респондентів нарешті заповнити анкету й повернути її досліднику.
- *Кваліфікація інтерв'юерів.* Більш кваліфіковані інтерв'юери мають кращі шанси отримати згоду респондентів на участь в інтерв'юванні.

- *Матеріальна мотивація.* Рівень повернення анкет певними категоріями респондентів зростає, якщо дослідник заохочує їх невеликими подарунками, участю в лотереї тощо.
- *Нематеріальна мотивація.* Компанії більш чутливі до не грошових заохочень. Прикладом заохочення може бути звіт з бенчмаркінгу, у якому участь компанії порівнюється з даними участі інших компаній.
- *Конфіденційність і приватність.* Важливо заздалегідь поінформувати респондентів стосовно захищеності їхніх персональних даних, які вони надають в опитуванні, щоб підвищити рівень повернення анкет.

Похибка формування вибірки. Різні види похибок формування вибірок пов'язані із залученням «неправильних» респондентів, які роблять вибірки нерепрезентативними і знижують можливість узагальнення отриманих результатів.

Похибка соціальної бажаності відповідей. Багато респондентів намагаються уникнути негативних оцінок стосовно себе, колег чи родини. Серед респондентів існує тенденція приховування правди, щоб представити себе у кращому світлі. Пов'язана з цим явищем похибка, яка негативно впливає на валідність результатів опитування, називається похибкою соціальної бажаності. Шляхів подолання цієї похибки при анкетному опитуванні фактично не існує, однак досвідчений інтерв'юер може помітити суперечливість відповідей та поставити додаткові питання, щоб прояснити ситуацію.

Мнемонічна похибка. Відповіді респондентів часто залежать від їх мотивації, пам'яті та здатності відповідати. Це особливо стосується подій, що відбувалися в минулому і можуть неточно оцінюватися респондентами. Щоб зменшити мнемонічну похибку, рекомендується уникати питань, що вимагають від респондента екскурсів у далеке минуле.

Похибка спільного методу. Похибка спільного методу відноситься до хибної коваріації між незалежними та залежними змінними, що одночасно вимірюються. Наприклад, похибка спільного методу може виникнути у міжгалузевих опитуваннях, які проводяться за допомогою одного й того ж методичного інструменту (анкети). У таких випадках досліджуваний феномен не завжди може бути відділений від артефактів вимірювання.

Існують стандартні статистичні методи для тестування похибки спільного методу, такі як однофакторний тест Хармона (Podsakoff et al. 2003), спосіб ринкової змінної Лінделла і Уїтні (2001) та інші. Щоб уникнути цієї похибки, рекомендується вимірювати незалежну та залежну величини в різні моменти часу, використовуючи повторюваний дизайн

опитування, або вимірювати ці змінні різними методами, як, наприклад, комп'ютерна реєстрація залежної величини й анкетне самооцінювання незалежних величин, або об'єктивні показники фінансової звітності підприємства як залежні величини й анкетне опитування як незалежні величини.

Розділ 10

ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ

Експериментальне дослідження, яке часто називають «золотим стандартом» дослідницького дизайну, є найбільш строгим науковим дизайном. При його проведенні дослідник маніпулює однією чи декількома незалежними величинами (факторами) за різних умов (спроб), досліджувані випадково відносяться до груп з різними рівнями факторів і фіксуються результати дії факторів (залежні величини). Унікальною перевагою експерименту є його внутрішня валідність (причинність) через здатність зв'язати причину й результат завдяки зміні умов у різних спробах при одночасному контролі за дією побічних змінних.

Експеримент більш доречний для пояснювальних досліджень, ніж для описових чи пошукових, коли мета дослідника полягає у виявленні причинно-наслідкових ефектів. Експеримент також має переваги, коли дослідження включає відносно невелике і точно визначене число незалежних величин, якими можна маніпулювали і які можна контролювати. Експеримент проводиться в лабораторних чи виробничих умовах. **Лабораторні експерименти** виконуються в лабораторних (штучних) умовах, мають тенденцію до високої внутрішньої валідності, однак є дорогими видами дослідження. **Виробничі експерименти** виконуються в реальних умовах, як то в організаціях, і мають високу внутрішню і зовнішню валідність. Однак виробничі експерименти проводяться доволі рідко, оскільки в реальних умовах складно забезпечити маніпулювання факторами і контроль побічних ефектів.

Експериментальні дослідження можна розділити на дві великі групи: справжні експериментальні дизайни й квазі-експериментальні дизайни. Обидва дизайни вимагають маніпулювання факторами, однак, на відміну від квазі-експериментів, при проведенні справжніх експериментів також необхідно забезпечити випадковий добір досліджуваних у групи. Іноді застосовуються не експериментальні дослідження, які включають усі типи досліджень, що не використовують маніпуляцію факторів у спробах чи випадкове формування груп, такі як анкетне опитування, спостереження чи кореляційні дослідження.

Базові концепції

Експериментальна та контрольна групи. В експериментах досліджувани, що складають **експериментальну групу**, знаходяться під дією досліджуваних факторів, або стимулів, тоді як на інших досліджуваних, що складають **контрольну групу**, ці фактори не діють. Дія експериментальних факторів може вважатися успішною, якщо досліджувані експериментальної групи показують інші значення

результуючих величин порівняно з контрольною групою. Якщо застосовуються різні рівні експериментального фактору, то може бути створена більш, ніж одна експериментальна група. Наприклад, порівнюється ефективність виконання завдання за умов стресу різного рівня (експериментальні групи) і відсутності стресу (контрольна група). При цьому можна порівняти ефективність виконання завдання між усіма експериментальними групами загалом і контрольною групою, щоб з'ясувати вплив стресу, а також порівняти ефективність виконання завдання між різними експериментальними групами, щоб з'ясувати вплив різних рівнів стресу.

Маніпулювання факторами експерименту. Фактори є унікальними рисами експериментального дослідження, які відрізняють певний дизайн від усіх інших. Маніпулювання факторами експерименту допомагає контролювати «причину» в причинно-наслідковому зв'язку. Факторами в експериментальному дослідженні можуть виступати складність завдання, час виконання завдання, рівень винагороди, рівень кваліфікації учасників тощо. Валідність експериментального дослідження залежить від того, наскільки добре можна маніпулювати факторами експерименту. Маніпулювання факторами потрібно перевіряти в пілотних тестах до проведення власне самого експерименту. Значення величин, що вимірюються до початку дії експериментальних факторів, називаються початковими (дотестовими), тоді як значення величин, отриманих після дії факторів експерименту, називають післятестовими.

Випадковий добір. Випадковий добір є процесом випадкового формування вибірки з генеральної сукупності або з доступної для формування вибірки частини генеральної сукупності. Цей підхід зазвичай використовується в анкетних опитуваннях і забезпечує рівні шанси всіх елементів генеральної сукупності потрапити до вибірки. Випадковий добір також передбачає випадкове віднесення того чи іншого досліджуваного в експериментальну чи контрольную групи. Це є стандартною практикою у справжніх експериментах, яка забезпечує рівність усіх досліджуваних груп до початку експериментального дослідження. Випадковий добір пов'язаний із формуванням вибірок, а отже, більше впливає на зовнішню валідність (узагальненість) результатів. Однак випадкове віднесення до тієї чи іншої групи пов'язане з дизайном, а отже, більше впливає на внутрішню валідність. У добре спланованих експериментах можливо забезпечити одночасно і випадковий добір, і випадкове віднесення до груп, тоді як квазіекспериментальне дослідження не передбачає ні першого, ні другого.

Загрози внутрішній валідності. Хоча експериментальний дизайн вважається більш строгим стосовно внутрішньої валідності, ніж інші наукові методи, (завдяки його здатності контролювати причини через маніпулювання факторами), він не дозволяє уникнути всіх загроз

валідності. Деякі з них описуються нижче на прикладі дослідження впливу спеціальної навчальної програми, спрямованої на покращення математичних здібностей школярів.

- *Історична загроза* – можливість того, що ефекти, які спостерігаються (залежні величині), обумовлені дією побічних або історичних подій, а не експериментальним фактором. Наприклад, покращені результати після навчальної програми можуть бути пов'язані не з самою програмою, а з підготовкою учнів до іспиту з математики в школі.

- *Загроза зростання загальної навченості досліджуваних* – можливість того, що отримані ефекти обумовлені природнім процесом зростання загальної навченості школярів, який допомагає краще розуміти складні концепції, а не досліджуваною навчальною програмою.

- *Загроза процедури тестування* – загроза того, що в дизайнах типу «оцінка до/після експерименту» досліджуваним пропонуються однакові завдання до початку і після закінчення програми, і досліджувані можуть запам'ятати правильні відповіді. Цієї загрози можна уникнути, якщо не проводити оцінку до початку програми.

- *Інструментальна загроза* також може виникнути в дизайнах типу «оцінка до/після експерименту» і стосується можливості того, що різниця між результатами оцінки досліджуваних до та після проходження програми пов'язана не з самою програмою, а з відмінностями в рівні складності наданих оціночних тестів.

- *Загроза дострокового виходу з експерименту* – це можливість того, що до закінчення експерименту програму залишать погані учні, які показували найгірші результати до початку програми. Отже, результати оцінки після закінчення програми будуть завищеними через збільшення питомої ваги хороших учнів.

- *Загроза регресії*, яка також називається регресією до середньої, означає статистичну тенденцію щодо загального результату групи після завершення програми. Ці результати наближуються (регресують) до середньої групи. Наприклад, якщо досліджувані показували високі результати до початку програми, вони матимуть тенденцію показувати нижчі результати після закінчення програми (ближчі до середньої), тому що дуже високі результати до початку програми могли бути зумовлені випадковими факторами. Ця проблема зростає в невідповідних вибірках і коли є незначна кореляція двох показників.

Двогрупові експериментальні дизайни

Найпростіший експериментальний дизайн – це той, що передбачає дослідження двох груп – експериментальної та контрольної. Він ідеальний для тестування ефектів однієї незалежної величини, якою можна маніпулювати. Основні різновиди двогрупових дизайнів – це дизайн «оцінка до/після експерименту з контрольною групою», а також

дизайн «оцінка після експерименту з контрольною групою», хоча варіанти можуть включати дизайни з коваріантами. Для позначення цих дизайнів часто використовують стандартні записи. При цьому R означає випадковий відбір досліджуваних у групи, X – фактор, що діяв на експериментальну групу, O – спостереження залежної величини до чи після експерименту (з різними індексами, щоб розрізнити значення контрольної й експериментальної груп «до» і «після»).

Дизайн «оцінка до/після експерименту з контрольною групою». У цьому дизайні досліджувані випадково добираються в експериментальну та контрольну групи. В обох групах проводиться попереднє вимірювання досліджуваних залежних величин. Потім експериментальна група піддається дії досліджуваного фактора (незалежної величини), і в обох групах знову вимірюються залежні величини. Схематичний опис цього дизайну наведений на рис. 10.1.

R	O ₁	X	O ₂	(Експериментальна група)
R	O ₃		O ₄	(Контрольна група)

Рис. 10.1. Дизайн «оцінка до/після експерименту з контрольною групою»

Ефект E дії експериментального фактора в дизайні «оцінка до/після експерименту з контрольною групою» вимірюється як різниця між значеннями «до/після» в експериментальній і контрольній групах:

$$E = (O_2 - O_1) - (O_4 - O_3).$$

Статистичний аналіз цього дизайну включає простий аналіз варіацій (ANOVA) між експериментальною та контрольною групами. Дизайн «оцінка до/після експерименту» усуває декілька загроз внутрішній валідності, таких як зростання загальної навченості, процедури тестування й регресії, оскільки ці загрози однаково діють на експериментальну та контрольну групи. Загроза відбору контролюється завдяки випадковому формуванню груп. Однак, можуть існувати додаткові загрози внутрішній валідності. Наприклад, зростання навченості може стати загрозою, якщо з експерименту достроково вийшло різне число учасників із контрольної й експериментальної груп. Тоді вимірювання до початку експерименту може вносити похибку у вимірювання після нього.

Дизайн «оцінка після експерименту з контрольною групою». Це спрощена версія попереднього дизайну, у якому вимірювання до початку експерименту не проводиться. Схема такого дизайну наведена на рис. 10.2.

R	X	O ₁	(Експериментальна група)
R		O ₂	(Контрольна група)

Рис. 10.2. Дизайн «оцінка після експерименту з контрольною групою»

Ефект впливу досліджуваного фактора вимірюється як різниця значень залежної величини експериментальної та контрольної груп після дії фактора:

$$E = (O_1 - O_2).$$

Коректний статистичний аналіз даних цього дизайну – це двогруповий аналіз варіацій (ANOVA). Простота цього дизайну робить його більш привабливим стосовно внутрішньої валідності, ніж дизайн «до/після». Цей дизайн дозволяє контролювати всі розглянуті вище загрози валідності, окрім дострокового виходу з експерименту.

Коваріантний дизайн. Іноді на значення залежних величин можуть впливати сторонні фактори, які називають *коваріантами*. Коваріанти – це ті змінні, що не є предметом експериментального дослідження і повинні контролюватися, щоб зменшити їх потенційний вплив на залежну величину та провести більш точне вимірювання ефектів досліджуваної незалежної величини. Раніше розглянуті експериментальні дизайни не контролюють коваріанти. Коваріантний дизайн, який також називається дизайном сторонніх змінних, є особливим типом дизайну «оцінки до/після експерименту з контрольною групою», у якому значення величини до експерименту є більше оцінкою досліджуваних коваріантів, ніж залежної величини. Описова схема такого дизайну наведена на рис. 10.3, де С представляє коваріанти:

R	C	X	O ₁	(Експериментальна група)
R	C		O ₂	(Контрольна група)

Рис. 10.3. Коваріантний дизайн

Оскільки значення величини до експерименту не є оцінкою залежної величини, а оцінкою коваріанти, ефект впливу фактора вимірюється як різниця між значеннями величини в експериментальній і контрольній групах після експерименту:

$$E = (O_1 - O_2).$$

Завдяки присутності коваріант коректний статистичний аналіз даних цього дизайну – це двогруповий аналіз коваріант (ANCOVA). Цей дизайн має всі переваги дизайну «оцінка після експерименту», але забезпечує внутрішню валідність, оскільки коваріанти контролюються.

Коваріантний дизайн також може бути розширений до дизайну «оцінка до/після експерименту з контрольними групами».

Факторіальний дизайн

Двогрупові дизайни не є коректними, якщо дослідження вимагає маніпулювання двома та більше незалежними змінними. У цьому випадку необхідні дизайни з чотирма та більше групами. Такі дизайни досить популярні в експериментальних дослідженнях і називаються факторіальними. Кожна незалежна змінна, яка називається *фактором*, має декілька *рівнів*. Факторіальні дизайни дозволяють досліджувати не тільки індивідуальний вплив фактора на залежну змінну (цей ефект називається головним), а й спільний ефект факторів на залежну змінну (цей ефект називається ефектом взаємодії).

Базовим факторіальним дизайном є дизайн 2x2, який включає два фактори, кожний із яких має два рівні (наприклад, високий/низький, або наявний/відсутній). Уявимо, що потрібно порівняти ефективність двох різних навчальних методик – аудиторна робота і навчання он-лайн, які мають тривалість 1,5 або 3 години на тиждень. У цьому випадку маємо два фактори – тип навчання і тривалість навчання. Їх комбінація дає чотири досліджувані групи, як показано на рис. 10.4.



Рис. 10.4. Факторіальний дизайн 2x2

Факторіальні дизайни можуть бути зображені схемою, представленою справа на рис. 10.4. На схемі R означає випадковий добір досліджуваних у групи, X – самі групи (індекси означають рівень кожного фактора), а O – значення залежної змінної. Зрозуміло, що дизайн 2x3 матиме шість досліджуваних груп, а дизайн 2x2x2 – вісім груп. Кожна група у факторіальному дизайні повинна мати щонайменше 20 досліджуваних (це впливає з розрахунку коефіцієнта потужності Кохена, що базується на розмірах середніх ефектів), і дизайн 2x2x2 вимагає щонайменше 160 досліджуваних. Із збільшенням рівнів чи факторів у факторіальному дизайні витрати на збір даних значно зростають. Іноді через обмеженість ресурсів деякі групи можуть виявитися меншими або взагалі порожніми. Це випадок неповного факторіального дизайну, який

суттєво обмежує здатність дослідника робити обґрунтовані висновки через відсутність даних.

У факторіальному дизайні відзначається **головний ефект**, якщо залежна змінна показує істотну відмінність між рівнями одного фактора *на всіх рівнях* іншого фактора. Якщо нема відмінностей у залежній змінній між різними рівнями фактора, то це випадок нульового ефекту (базова лінія), від якого оцінюється головний ефект. Ефект **взаємодії факторів** існує, коли ефект дії одного фактора відрізняється залежно від рівня другого фактора. У прикладі, що розглядався вище, якщо ефект виду навчання на результат навчання більше при 3-годинній тривалості занять, ніж при 1,5-годинній, тоді можна говорити про ефект взаємодії між видом навчання та його тривалістю. У багатьох дослідженнях домінують ефекти взаємодії, що робить іррелевантним інтерпретацію головних ефектів.

Гібридні експериментальні дизайни

У гібридних дизайнах поєднуються риси декількох відомих дизайнів. Серед таких гібридних дизайнів виділимо дизайн випадкових блоків, чотирьохгруповий дизайн Соломона і дизайн перехресного експерименту.

Дизайн випадкових блоків є варіацією дизайнів «після» чи «до/після» з контрольною групою, де досліджувані можуть бути об'єднані у відносно гомогенні підгрупи (їх називають блоками), у кожній із яких і проводиться експеримент. Наприклад, якщо дослідник хоче продублювати дизайн «після» з контрольною групою в групі студентів і працівників бізнесу (дві гомогенні групи), досліджувані в обох блоках будуть випадково розподілені між експериментальними групами (на які діє один і той же фактор) чи контрольними групами (рис. 10.5). Мета такого дизайну полягає в тому, щоб зменшити «шум», або варіацію, пов'язану з відмінностями між блоками, так щоб досліджуваний ефект можна було встановити якомога точніше.

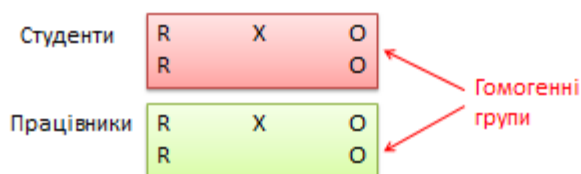


Рис. 10.5. Дизайн випадкових блоків

Чотирьохгруповий дизайн Соломона. У цьому дизайні вибірка ділиться на дві експериментальні та дві контрольні групи. Одна експериментальна й одна контрольна групи мають тестування «до», тоді як інші дві групи не мають. Дизайн являє собою комбінацію дизайну «після» і «до/після» з контрольними групами і має на меті виявити

можливий хибний ефект оцінки «до» на оцінку «після», який може спостерігатися в дизайні «до/після», але не в дизайні «після». Схематичний опис дизайну показаний на рис. 10.6.

R	O	X	O
R	O		O
R		X	O
R			O

Рис. 10.6. Чотирьохгруповий дизайн Соломона

Дизайн перехресного експерименту. Це двогруповий дизайн, що проводиться у дві стадії з трьома хвилями вимірювання. Експериментальна група першої стадії стає контрольною групою на другій стадії, а контрольна група першої стадії стає експериментальною групою на другій стадії, як показано на рис. 10.7. Іншими словами, оригінальний дизайн повторюється, тільки експериментальна та контрольна групи міняються місцями. Таким чином, усі досліджувані підлягають дії експериментального фактора або на першій, або на другій стадії дослідження. Цей дизайн найбільш доречний для організаційних досліджень, де певна організаційна програма (наприклад, навчання персоналу) впроваджується поетапно або повторюється через певні часові інтервали.

R	O	X	O	O
R	O		O	X

Рис. 10.7. Дизайн перехресного повторення

Квазі-експериментальні дизайни

Квазі-експериментальні дизайни майже ідентичні справжнім експериментальним дизайнам, однак їм бракує однієї важливої риси - випадкового віднесення досліджуваних до експериментальної чи контрольної груп. Наприклад, одна організація слугує як експериментальна група, тоді як інша є контрольною. Відсутність випадкового потрапляння досліджуваних у ці групи потенційно призводить до нерівності груп і неповної співставності їх результатів після дії досліджуваного фактора. Квазі-експериментальні дизайни програють справжнім експериментальним в інтервальній валідності завдяки наявності широкого спектру пов'язаних загроз, таких як загроза зростання навченості (експериментальна й контрольна групи можуть навчатися з різною швидкістю), історична загроза (експериментальна й контрольна групи можуть мати різний досвід до початку проведення експерименту), загроза регресії (якщо експериментальна й контрольна групи з різною швидкістю регресують до середніх від значень до початку

експерименту до значень після експерименту), інструментальна загроза (якщо експериментальна й контрольна групи по-різному реагують на процедуру оцінки) і загроза дострокового виходу з експерименту (якщо експериментальну й контрольну групи до закінчення експерименту залишила різна кількість осіб). Ураховуючи ці загрози відбору досліджуваних у групи, рекомендується, якщо це можливо, уникати квазі-експериментальних дизайнів.

Багато справжніх експериментальних дизайнів можуть бути перетворені на квазі-експериментальні, якщо не забезпечується випадковий добір до експериментальної й контрольної груп. Квазі-еквівалентна версія дизайну «оцінка до/після експерименту з контрольною групою» **називається дизайном нееквівалентних груп**, як показано на рис. 10.8, де випадковий добір R замінюється нееквівалентним (невипадковим) добором груп N. Аналогічно, квазі-експериментальна версія дизайну перехресного експерименту називається **дизайном нееквівалентного перехресного експерименту** (див. рис. 10.9).

N	O	X	O
N	O		O

Рис. 10.8. Дизайн нееквівалентних груп

N	O	X	O	X	O
N	O		O	X	O

Рис. 10.9. Дизайн нееквівалентного перехресного експерименту

Варто відзначити, що існує досить багато нееквівалентних дизайнів, які не мають аналогів серед справжніх експериментальних дизайнів. Деякі з найбільш корисних розглядаються нижче.

Дизайн «регресія-розрив». Це нееквівалентний дизайн із оцінкою до/після експерименту, у якому досліджувані відбираються в експериментальну й контрольну групи на основі встановленого розмежувального значення певної величини, що визначається до початку програми експерименту. Наприклад, в експериментальну групу включають працівників зі стажем роботи більш, ніж 10 років, а в контрольну групу – працівників зі стажем до 1 року. Схема такого дизайну представлена на рис. 10.10, де С означає розмежувальне значення.

C	O	X	O
C	O		O

Рис.10.10. Дизайн регресія-розрив

Через використання розмежувального значення можливо, що отримані результати будуть функцією цього розмежувального значення більшою мірою, ніж функцією досліджуваного фактора, що означає нову загрозу для внутрішньої валідності. Однак використання розмежувального значення також означає, що обмежені чи дорогі ресурси, необхідні для проведення дослідження, використовуються більш раціонально, ніж коли вибірки формуються з генеральної сукупності. Бали контрольної групи в дизайні «регресія-розрив» не є точкою розмежування для порівняння балів експериментальної групи, ураховуючи систематичну нееквівалентність між двома групами. Скоріше, якщо нема розриву між балами «до» і «після» в контрольній групі, однак такий розрив зберігається в експериментальній групі, тоді він розглядається як доказ впливу досліджуваного фактора.

Дизайн «заміщення оцінки до експерименту». Цей дизайн, що показаний на рис. 10.11, виглядає схожим на стандартний дизайн «до/після», однак із однією важливою відмінністю: значення оцінки до експерименту збирається після того, як досліджуваний фактор уже діє. Типовим прикладом застосування дизайну є ситуація, коли дослідник вирішує дослідити ефективність програми (наприклад, навчальної) після того, як програма вже розпочалася й зібрати дані до початку дії фактора неможливо. За таких обставин найкращим варіантом буде використання інших відомих показників, наприклад, середній бал навчання студента, як аналог оцінки до експерименту. Варіацією дизайну «заміщення оцінки до експерименту» буде збір даних студентів після експерименту, що може стати причиною мнемонічної помилки, однак надасть оцінку очікуваної зміни залежної величини.

N	O ₁	X	O ₂
N	O ₁		O ₂

Рис. 10.11. Дизайн «заміщення оцінки до експерименту»

Дизайн «різні вибірки для оцінки до/після експерименту». Цей дизайн корисний, якщо з якихось причини неможливо зібрати дані «до/після» для однієї й тієї ж експериментальної групи. Як показано на рис. 10.12, у цьому дизайні є чотири групи, але дві групи створюються з однієї нееквівалентної групи, тоді як дві інші групи – з іншої нееквівалентної групи. Наприклад, дослідник хоче оцінити задоволеність споживачів новим он-лайн сервісом, що запроваджений в одному місті, і не запроваджений в іншому. У цьому випадку споживачі першого міста стають експериментальною групою, а споживачі другого міста – контрольною. Якщо неможливо отримати дані «до/після» від одних і тих же споживачів, можна оцінювати задоволеність споживачів у певний момент часу, потім впровадити нову програму сервісу й оцінити задоволеність споживачів іншої групи після впровадження програми.

Задоволеність споживачів контрольної групи (які не знаходяться під дією експериментального фактору) оцінюється одночасно з експериментальною групою. Дизайн не є ідеальним, оскільки дослідник не може оцінити зміни в значеннях задоволеності конкретних споживачів до та після впровадження експериментальної програми, однак можна оцінити середні значення задоволеності споживачів. Не дивлячись на низьку внутрішню валідність, цей дизайн може бути корисним для збору квазі-експериментальних даних, якщо неможливо зібрати дані «до/після» в одній і тій же групі досліджуваних.

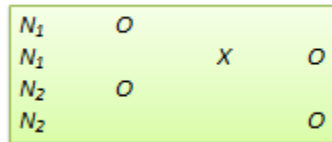


Рис. 10.12. Дизайн «різні вибірки для оцінки до/після»

Небезпеки експериментального дослідження

Експеримент є найбільш складним дослідницьким дизайном і вимагає серйозного ставлення до себе. Він найбільш доречний для вирішення численних методологічних проблем. По-перше, хоча експериментальне дослідження вимагає теоретичного підґрунтя для формулювання гіпотез, більшість експериментів не є теоретичними. Без теорій гіпотези, що тестуються, можуть виявитися випадковими, нелогічними і навіть безглуздими. По-друге, багато методик експериментального дослідження не перевіряються на надійність і валідність та не порівнюються з аналогами. Відповідно, отримані результати експерименту важко порівняти з результатами інших експериментів. По-третє, багато експериментальних досліджень використовують невідповідні дизайни, такі як відсутність контролю, відсутність ефекту взаємодії або неоднакові рівні досліджуваного фактора, що діє на досліджуваних експериментальної групи. Результатам таких експериментів бракує внутрішньої валідності, що ставить під сумнів їх коректність. По-четверте, фактори (наприклад, завдання) що використовуються в експериментах, можуть виявитися неспівставними та навіть невідповідними для певної групи досліджуваних. Використання невідповідних завдань породжує нові загрози для внутрішньої валідності. Зокрема, результати виконання завдань досліджуваними можуть бути артефактом змісту чи складності завдання. Це може призвести до отримання результатів, які складно інтерпретувати, узагальнювати й порівнювати з результатами інших експериментів.

Дизайн коректних експериментальних завдань є надзвичайно важливим, оскільки є сутністю експериментального методу та вимагає особливої уваги дослідника. Щоб розробити адекватне завдання, дослідники повинні використовувати вже перевірені на валідність

завдання, проводити перевірку адекватності завдання (опитуючи досліджуваних після експерименту), проводити пілотні дослідження й надавати перевагу завданням, що є доступними та зрозумілими для досліджуваних.

Цей розділ був присвячений ключовим концепціям експериментального дослідження та опису численних справжніх експериментальних і квазі-експериментальних дизайнів. Хоча розглянуті дизайни значно відрізняються за внутрішньою валідністю, не треба недооцінювати дизайни з меншою внутрішньою валідністю, оскільки за певних умов вони можуть виявитися корисними.

Розділ 11

КЕЙС-ДОСЛІДЖЕННЯ

Кейс-дослідження – це метод інтенсивного дослідження певного феномену в динаміці, що виконується в організаційному середовищі однієї чи декількох компаній. При проведенні кейс-дослідження можуть застосовуватися численні методи збору даних, такі як інтерв'ю, спостереження, архівні документи та вторинні дані, тому висновки щодо досліджуваного феномену відрізняються глибиною, деталізованістю та змістовністю. Метод кейсів може використовуватися в позитивістському плані для тестування теорій, або в інтерпретативному плані для побудови теорій. Цей метод особливо популярний в бізнес-дослідженнях.

Кейс-метод має декілька унікальних переваг порівно з такими методами дослідження, як експеримент й анкетне опитування. По-перше, він може використовуватися для побудови теорій та їх тестування, тоді як позитивістські методи можуть використовуватися лише для тестування теорій. В інтерпретативному кейс-дослідженні конструкти можуть бути відомі заздалегідь, однак можуть і виникати в процесі збору даних дослідження. По-друге, деякі завдання дослідження можуть змінюватися під час його проведення, якщо з'ясується, що раніше заплановані завдання менш релевантні. Це неможливо в будь-якому позитивістському дослідженні. По-третє, кейс-метод може допомогти в наданні більш змістовної, контекстуальної та релевантної інтерпретації досліджуваного феномену, ніж більшість інших методів завдяки його здатності охопити широке коло контекстуальних даних. По-четверте, досліджуваний феномен може вивчатися з позицій численних учасників і з застосуванням численних рівнів аналізу (як індивідуальних, так і організаційних).

Водночас, кейс-метод має декілька суттєвих обмежень. Він не передбачає контрольних груп, тому його валідність невисока. Це загальна проблема всіх методів дослідження, окрім експерименту. Однак, як буде показано нижче, проблема контролю може бути вирішена в кейс-дослідженнях завдяки «природним контрольним групам». По-друге, якість висновків, зроблених з кейс-дослідження, суттєво залежить від аналітичних здібностей дослідника. Досвідчений дослідник може побачити закономірності в тих даних кейс-дослідження, які менш досвідчений дослідник не помітить. Нарешті, оскільки кейс-дослідження є значною мірою контекстуальним, може бути складно проводити узагальнення його результатів до рівня інших організацій.

Важливо розуміти, що кейс-дослідження відрізняються від опису кейсів, що обговорюються в навчальних аудиторіях бізнес-шкіл. Хоча описи кейсів зазвичай представляють організаційну проблему з багатьма

деталіями, щоб стимулювати дискусії та критичне мислення студентів, вони мають мало спільного з кейс-методом, який є формальним дослідницьким методом, що передбачає застосування певного наукового апарату для надання пояснень досліджуваного феномену.

Кейс-метод є складним методом, оскільки вимагає від дослідника значної професійної майстерності, а в разі її відсутності може стати джерелом помилок. Бенбасат і співавт (1987)⁸ описує п'ять проблем, що часто пов'язані з кейс-дослідженнями. По-перше, багато кейс-досліджень починаються без конкретних наукових задач і відповідно закінчуються без будь-яких конкретних відповідей чи знахідок. По-друге, місця проведення кейс-досліджень обираються за наявністю або зручністю, а не за найкращою відповідністю задачі дослідження. По-третє, дослідники часто не проводять оцінку валідності чи узгодження даних, що збираються різними способами. Це може призвести до хибної інтерпретації даних. По-четверте, багато досліджень надають замало деталей стосовно того, як саме дані збиралися (питання інтерв'ю, аналізовані документи, дані учасників інтерв'ю і т.п.), або аналізувалися, що може викликати питання стосовно надійності отриманих результатів. Нарешті, не дивлячись на свої переваги, багато кейс-досліджень не дозволяють відслідковувати феномен у динаміці, а отже, надають лише крос-секційні дані, що не відображають організаційні процеси і феномени довготривалого характеру.

Ключові рішення у кейс-дослідженнях

Дослідник, обираючи кейс-метод для дослідження, повинен прийняти декілька ключових рішень. По-перше, чи є цей метод правильним вибором? Кейс-метод особливо доречний для пояснювальних досліджень, спрямованих на пошук релевантних конструктів у тих наукових галузях, де побудова теорій знаходиться на початкових етапах; для досліджень, де досвід учасників і контекст їхніх дій є критичними; для досліджень, спрямованих на розуміння складних процесів, а не на окремі фактори та причини. Цей метод також доречний для вивчення комплексних організаційних процесів, що включають багатьох учасників і взаємодіючі послідовності подій, такі як організаційні зміни та проект впровадження нових масштабних технологій.

По-друге, що є предметом дослідження при застосуванні кейс-методу? Оскільки кейс-дослідження може одночасно вивчати численні об'єкти, дослідник повинен вирішити, на якому рівні буде проводитись дослідження – на індивідуальному, організаційному чи на обох разом. Наприклад, дослідження щодо прийняття рішень групою може

⁸Benbasat I., Goldstein D. K. Mead M. (1987). The Case Research Strategy in Studies of Information Systems/ I. Benbasat, D. K. Goldstein, M. Mead // MIS Quarterly. – 1987. – Vol. 11:3. – P.369-386.

поєднувати конструкти індивідуального рівня (наприклад, індивідуальна участь в роботі групи) із конструктами групового рівня (наприклад, групова взаємодія та лідерство), щоб отримати більш глибоке розуміння феномену ніж те, що можна отримати при дослідженні конструктів одного рівня.

По-третє, чи повинен дослідник звернутися до однокейсового або до багатокейсового дизайну? Однокейсовий дизайн більш доречний: 1) на початковому етапі створення теорії; 2) якщо ситуація є унікальною, або екстремальною; 3) якщо ситуація раніше була недоступна для наукового дослідження; 4) якщо ситуація представляє критичний кейс, або такий що ставить під сумнів добре відому теорію. Багатокейсовий дизайн більш доречний для: 1) тестування гіпотез; 2) установлення рівня узагальненості висновків; 3) для розвитку більш глибокої та деталізованої інтерпретації феноменів. Йін (1984)⁹ рекомендує застосовувати кейс-метод в різних організаціях за логікою повторення. При цьому кожне місце проведення кейс-дослідження слід уважати одним експериментальним дослідженням, застосовуючи визнані правила проведення наукових досліджень, подібні до тих, що характерні для позитивістських досліджень.

По-четверте, які організації мають бути обрані для кейс дослідження? Ураховуючи контекстуальну природу висновків, що робляться на основі кейс-дослідження, вибір місця його проведення стає особливо важливим, оскільки вибір «неправильного» місця може призвести до «неправильних» висновків. Якщо мета дослідження полягає в тестуванні теорій чи перевірці рівня загальності раніше зроблених висновків, тоді потрібно звернутися до організацій, що відрізняються від попередніх місць дослідження, щоб збільшити варіацію спостережень. Наприклад, якщо мета дослідження полягає в розумінні процесу впровадження технологій компанією, тоді потрібно обрати великі, середні та малі компанії, щоб з'ясувати, чи відрізняється процес упровадження технологій у різних за розміром компаніях. Вибір місця проведення дослідження повинен базуватися на відповідності меті дослідження завдяки процесу, що називається теоретичним формуванням вибірки.

По-п'яте, які способи збору даних треба використовувати в кейс-дослідженнях? Хоча на сьогодні інтерв'ю різного формату залишаються найбільш популярними, дані інтерв'ю можуть бути доповнені або підтверджені іншими методами, такими як безпосереднє спостереження, вивчення документації чи роботи устаткування. Більш того, дослідник повинен об'єднати чи валідизувати отримані дані, порівнюючи відповіді учасників інтерв'ю.

Проведення кейс-дослідження

⁹Yin R. K. Case Study Research: Design and Methods / R. K. Yin. - Thousand Oaks, CA: Sage Publications. – 1984.

Більшість кейс-досліджень є інтерпретативними за своєю природою. Інтерпретативне кейс-дослідження є індуктивним методом, при якому докази, зібрані з різних місць проведення дослідження, систематично аналізуються й синтезуються для виявлення концептів і закономірностей із метою побудови нових теорій або поглиблення існуючих. Ейзенгардт (1989)¹⁰ пропонує «дорожню карту» для побудови теорій на основі кейс-досліджень. Нижче наводиться дещо видозмінена версія цієї карти. Для позитивістського кейс-дослідження необхідно модифікувати деякі з етапів, однак процедура формування вибірок, збір даних та методи їх аналізу зазвичай залишаються незмінними.

Визначення питань дослідження. Як будь-яке інше наукове дослідження, кейс-дослідження має починатися з визначення питань, що представляють теоретичний і практичний інтерес, та ідентифікації певних інтуїтивних очікувань щодо можливих відповідей і попередніх конструктів. У позитивістському кейс-дослідженні попередні конструкти базуються на теорії, тоді як жодних теорій чи гіпотез не висувається до проведення інтерпретативного дослідження. В інтерпретативному кейс-дослідженні питання дослідження та конструкти можуть бути змінені пізніше, тоді як у позитивістському – ні.

Вибір місця дослідження. Дослідник повинен використовувати процес «теоретичного формування вибірок» (а не випадкового), щоб встановити місце проведення кейс дослідження. При цьому підході місця проведення дослідження базуються на теоретичних, а не статистичних міркуваннях, наприклад, повторення попередніх кейсів, поглиблення існуючих теорій або заповнення теоретичних категорій. Важливо, щоб обрані місця кейс-досліджень відповідали природі питань дослідження, мінімізували дію побічних впливів чи «шум» і збільшували варіацію досліджуваних залежних змінних. Наприклад, якщо мета дослідження полягає у вивченні того, чому окремі компанії стають більш інноваційними, ніж інші, дослідник повинен обирати компанії, схожі за розміром і сферою діяльності, але відмінні за рівнем інноваційності, щоб збільшити варіативність досліджуваного фактору. Досліднику потрібно отримати дозвіл керівництва компаній на проведення дослідження.

Розроблення методик і протоколів дослідження. Оскільки в кейс-дослідженнях основним методом збору даних є інтерв'ю, потрібно розробити протокол інтерв'ю, тобто, по суті, список питань, що будуть поставлені учасникам інтерв'ю. Питання можуть бути відкритими (неструктурованими), закритими (структурованими) чи комбінацією обох видів. Важливо чітко дотримуватися протоколу під час інтерв'ю, не змінюючи питання і їх порядок, хоча деякі відхилення все ж можливі для отримання розгорнутих коментарів респондентів. Дослідник під час

¹⁰Eisenhardt K. M. Building Theories from Case Research/ K. M.Eisenhardt // Academy of Management Review. 1989. – Vol. 14:4. – P. 532-550.

проведення інтерв'ю повинен дотримуватися нейтральної позиції, щоб не підштовхувати учасників до певних відповідей. У подальшому техніки проведення інтерв'ю будуть детальніше розглянуті в цьому розділі. Також дослідник повинен спланувати, які внутрішні документи (звіти, інструкції, статті і пресі) і особисті спостереження мають бути використані в даному дослідженні для валідації даних інтерв'ю.

Вибір респондентів. Потрібно сформувати групу учасників інтерв'ю, що включає працівників різних рівнів організації для отримання різносторонніх поглядів щодо досліджуваного феномену. Бажано провести випадкове формування вибірки респондентів, однак, можливим є формування вибірки за принципом сніжної лавини, якщо забезпечується різносторонність думок і поглядів. Учасники інтерв'ю обираються, ураховуючи їхню обізнаність щодо досліджуваного феномену, здатність і бажання адекватно й точно відповідати на питання інтерв'юера.

Початок збору даних. Рекомендується робити аудіо(відео)запис інтерв'ю, щоб зберегти всі деталі. Однак такий запис можна робити лише з дозволу учасників інтерв'ю. У будь-якому випадку інтерв'юер повинен робити нотатки, фіксуючи важливі коментарі, власні спостереження і враження щодо відповідей респондентів. Після завершення інтерв'ю його зміст потрібно зберегти у вигляді текстового документу для подальшого аналізу.

Внутрішньокейсовий аналіз даних. Аналіз даних кейсу може розпочинатися вже після перших інтерв'ю і до закінчення інтерв'ювання всіх респондентів. Це допомагає в разі потреби відкоригувати процес подальшого збору даних і тематику інтерв'ю, ураховуючи аналіз даних перших респондентів. Аналіз даних виконується у два етапи. На першому етапі (внутрішньокейсовий аналіз), інтерв'юер повинен дослідити вихідні концепти окремо для кожної організації і виявити взаємозв'язки між цими концептами, щоб сформулювати первинне теоретичне пояснення проблеми. Дослідник може суб'єктивно осмислити дані кейс-дослідження, спираючись на відповіді респондентів, власні спостереження й досвід. Альтернативний варіант – стратегія кодування, як, наприклад, обґрунтована теорія Глассера і Страусса (1967), із використанням таких методик, як відкрите кодування чи селективне кодування, що може призвести до формування ланцюга доказів і пояснень. Ці методики детально розглядаються в наступному розділі. Також можливе використання графічної презентації даних чи послідовний аналіз (для динамічних рядів). Важливо відзначити, що не існує єдиного алгоритму аналізу різноманітних даних кейс-дослідження і, щоб краще відповідати меті дослідження, можуть бути модифіковані відомі способи аналізу даних.

Міжкейсовий аналіз даних. Якщо кейс-дослідження проводиться в декількох організаціях, другим етапом дослідження має стати

міжкейсовий аналіз даних. При цьому дослідник повинен виявити спільні концепти й закономірності між різними організаціями, ігноруючи контекстуальні відмінності, що можуть вести до поверхневих висновків. Виявлені тренди можуть бути використані для валідації вже існуючої теорії або її уточнення (додаванням чи відніманням концептів і зв'язків), що матиме результатом розвиток більш узагальненої і глибокої теорії. Аналіз може набувати декількох форм. Наприклад, дослідник може обрати категорії (сфера діяльності компанії, її розмір тощо) і шукати внутрішньогрупові подібності та міжгрупові відмінності (наприклад, відмінності між працівниками з високою й низькою продуктивністю, інноваторами й консерваторами тощо). Іншим варіантом може бути попарне порівняння компаній за списком схожих і відмінних рис.

Побудова і тестування гіпотез. Базуючись на вихідних концептах і питаннях, узагальнених для всіх досліджуваних організацій, конструюються попередні гіпотези. Ці гіпотези покроково порівнюються з наявними доказами, щоб упевнитися в їх відповідності отриманим даним. У разі потреби конструкти чи взаємозв'язки уточнюються. Також необхідно порівняти вихідні конструкти та гіпотези з тими, що представлені в науковій літературі, щоб установити їх валідність і рівень узагальненості. Не треба відкидати суперечливі дані. Замість цього необхідно розвивати більш узагальнені теоретичні уявлення, які б охоплювали й нові дані. Процес побудови теорії продовжується, доки теорія й дані кейс-дослідження не будуть приведені в повну відповідність.

Підготовка звіту кейс-дослідження. При написанні звіту дослідник повинен детально описати процедуру формування вибірок, аналізу даних і розвитку гіпотез, щоб читачі могли самостійно оцінити обґрунтованість, силу та логічність отриманих висновків. Особиста думка й переконання дослідника не повинні впливати на отримані результати. Щоб досягти цього, необхідний високий рівень прозорості в описі методів дослідження.

Приклад інтерпретативного кейс-дослідження

Найкращий спосіб зрозуміти сутність інтерпретативного кейс дослідження – це звернутися до прикладу. Ейзенхард (1989)¹¹ досліджувала, як керівники приймають рішення в умовах високої невизначеності бізнес-середовища (ВНБС). Якщо одні команди керівників в умовах ВНБС приймають швидкі рішення, то інші команди – ні. При цьому також досліджувалось, як швидкі рішення впливають на ефективність діяльності компаній – покращують чи погіршують. ВНБС було визначене, як середовище, у якому вимоги споживачів, конкуренція й технології змінюються так швидко і різноспрямовано, що наявна

¹¹Eisenhardt K. M. Making Fast Strategic Decisions in High-Velocity Environments / K. M. Eisenhardt. // Academy of Management Journal. – 1989. – Vol. 32:3. – P. 543-576.

інформація про ці середовища часто є неповною, неточною й неоднозначною. Попередні інтуїтивні припущення були такі: 1) складно приймати швидкі рішення при неадекватній інформації в умовах ВНБС; 2) швидкі рішення можуть бути неефективними та призводять до погіршення результатів роботи компанії.

Аналіз літератури щодо прийняття рішень керівництвом компанії дозволив Ейзенхард виявити декілька закономірностей, хоча жодна з них не була характерною саме для ВНБС. Література свідчила, що заради швидкої реакції компанії спираються на інформацію з обмеженої кількості джерел, розглядають менше можливих варіантів рішення, проводять неглибокий аналіз, обмежують участь клієнтів у прийнятті рішень, централізують процедуру прийняття рішень і не допускають внутрішніх конфліктів. Однак Ейзенхард припустила, що ці погляди не обов'язково пояснюють, як керівники приймають рішення у ВНБС, де рішення повинні прийматися швидко, при неповній інформації і одночасному забезпеченні високої якості рішень.

Для дослідження цього феномену Ейзенхард провела індуктивне дослідження восьми компаній, що працюють у галузі інформаційних технологій. Ця галузь у 1980-х роках переживала революційні зміни в технологіях, пов'язані з упровадженням операційної системи UNIX, архітектури RISC та 64 KB випадкового доступу пам'яті, а також із зростаючою конкуренцією через входження IBM в бізнес персональних комп'ютерів і зростаючими вимогами споживачів. Це був багатокейсовий дизайн за логікою повторювання (реплікації), де очікувалося, що кожний кейс підтвердить або спростує висновки, зроблені на основі попередніх кейсів. Місця проведення досліджень були обрані з міркувань зручності і географічної близькості, однак усі обрані компанії працювали в галузі персональних комп'ютерів і були розташовані в Силіконовій долині. Географічна близькість компаній однієї галузі знімала питання щодо можливого «шуму», або варіації залежних величин (швидкість рішень і ефективність компанії), пов'язаних із відмінностями у сфері діяльності й розташуванні.

У дослідженні застосовувалося декілька рівнів аналізу: рішення (порівняння стратегічних рішень у межах компанії), команди керівників (порівняння різних команд, відповідальних за стратегічні рішення) і компанії (ефективність її діяльності). Дані збиралися з п'яти джерел:

- Інтерв'ю з представниками керівного складу компаній: були поставлені питання стосовно конкурентних стратегій, унікальних компетенцій, ключових конкурентів, результатів діяльності й нещодавно прийняті (або тих, що перебувають у процесі прийняття) головних стратегічних рішень. Базуючись на цих інтерв'ю, для подальшого дослідження були відібрані декілька стратегічних рішень. Відбір проводився на основі чотирьох критеріїв: 1) рішення, що включають

стратегічне позиціонування компанії; 2) рішення, що мають великий вплив; 3) рішення, що включають численні функції; 4) рішення, що представляють процес прийняття стратегічних рішень у компанії.

- Інтерв'ю з керівниками управлінь (дивізіонів): кожному з керівників було поставлено 16 відкритих питань, що стосувалися конкурентної стратегії компанії, функціональної стратегії, вищого керівництва компанії, частоти і природи взаємодії з командою, типового процесу прийняття рішень і тривалості прийняття попередніх рішень. Інтерв'ю тривали 1,5-2 години іноді – до 4-х годин. Щоб зосередитися на реальних фактах і подіях, а не на їх інтерпретації респондентами, застосовувався стиль «судового засідання» з акцентом на питаннях «що саме трапилося», «коли», «де» і т.п. Інтерв'ю проводилося двома особами, і дані валідизувалися перехресною перевіркою фактів і вражень інтерв'юерів. Робилися аудіозаписи інтерв'ю, одночасно велося документування інформації, яку інтерв'юери вважали важливою. Детальні нотатки щодо інтерв'ю робилися в термін до 24 годин після його закінчення, щоб уникнути мнемонічних помилок.

- Бланки анкет: керівники кожної фірми заповнювали анкети, щоб надати якісні дані стосовно рівня конфліктності й розподілу влади в компанії.

- Вторинні дані: аналізувалися огляди галузі та внутрішні документи компаній, такі що стосуються автобіографічних даних керівників, фінансової звітності компаній та інші.

- Особисті спостереження: дослідник відвідував одноденну сесію з розвитку стратегії й щотижневу раду керівництва у двох компаніях із досліджуваної вибірки.

Аналіз даних включав комбінацію якісних і кількісних методів. Аналізувалися якісні дані щодо конфліктів і влади в розрізі компаній і рішень. Кількісні дані інтерв'ю були об'єднані в профіль клімату прийняття рішень із використанням тих рис профілю (наприклад, довіра), що були наведені більш, ніж одним респондентом. Для внутрішньокейсового аналізу по кожному стратегічному рішенню були створені їх описи шляхом відтворення хронологічної послідовності ключових подій. Для міжрейсового аналізу порівнювалися пари компаній стосовно їх схожості й відмінностей, які були категоризовані в досліджуваних величинах, таких як швидкість прийняття рішень та ефективність компанії. Базуючись на такому аналізі, індуктивно були запропоновані попередні конструкти й теоретичні припущення для кожного випадку прийняття рішень у межах категорії компаній. Кожний кейс із прийняття рішень був переглянутий, щоб підтвердити запропоновані взаємозв'язки. Пропозиції (теоретичні припущення) порівнювалися з даними літератури, щоб узгодити дані дослідження із вже встановленими результатами, а також для отримання нових ідей. Валідизовані пропозиції були синтезовані в індуктивну теорію

стратегічного прийняття рішень компаніями в умовах невизначеного бізнес-середовища.

Висновки, зроблені з цього дослідження суперечили деяким очікуваням на основі літературного аналізу трендам. По-перше, менеджери, що приймають швидкі рішення, використовують більше, а не менше інформації, як очікувалося з даних літератури. Однак ті, хто приймає швидкі рішення, частіше звертаються до *актуальної* інформації (це раніше не було відомо), що допомагає їм швидше відповідати на виклики, можливості й зміни бізнес-середовища. По-друге, менеджери, що приймають швидкі рішення, досліджують більше (а не менше) альтернатив. Однак, вони розглядають ці альтернативи *одночасно*, тоді як менеджери, що приймають повільні рішення, досліджують альтернативи *попередньо*. По-третє, менеджери, що приймають швидкі рішення, не централізують процес прийняття рішень і не обмежують участі інших, як вважалося з попередньої літератури. Скоріше, ці компанії використовують дворівневий процес прийняття рішень, у якому досвідчені фахівці висловлюють свою думку на першому етапі, тоді як на другому етапі відбувається швидке порівняння та відбір рішень. По-четверте, менеджери, що приймають швидкі рішення, не мають менше конфліктів, як очікувалося за даними літератури, однак застосовують кращі способи вирішення конфліктів, щоб зменшити конфліктність і збільшити швидкість прийняття рішень. Нарешті, менеджери, що приймають швидкі рішення, демонструють високу ефективність діяльності завдяки когнітивним, емоційним та організаційним процесам у компанії.

Приклад позитивістського кейс-дослідження

Кейс-дослідження також може базуватися на позитивістських позиціях і використовуватися для тестування теорій чи гіпотез. Такі дослідження зустрічаються не часто, однак Маркус (1983)¹² надає гарний приклад, який стосується впровадження нової технології в компанії. Мета даного дослідження полягала в тому, щоб пояснити, чому нова інформаційна система фінансової звітності, яка дозволяє підвищити продуктивність бухгалтерів, була підтримана в корпоративному центральному офісі, однак викликала супротив бухгалтерів філій. Ураховуючи унікальність даного випадку, було проведено однокейсове дослідження.

Для пояснення причин супротиву в корпорації Маркус запропонувала три альтернативи: 1) теорія впливу технічної системи: супротив може бути визначений факторами, що пов'язані з неадекватною системою, такими як технічні недоліки, поганий ергономічний дизайн чи недружній інтерфейс; 2) теорія впливу людського фактору: супротив

¹²Markus M. L. Power, Politics, and MIS Implementation / M. L. Markus // Communications of the ACM. –1983.– Vol. 26:6. – P.430-444.

викликаний внутрішніми по відношенню до користувачів факторами, такими як стиль мислення, або особистісні риси бухгалтерів, які погано сумісні з системою; 3) теорія взаємодії: супротив викликаний взаємодією факторів перших двох груп. Зокрема, теорія взаємодії припускала, що впровадження інформаційної системи звітності призвело до перерозподілу внутрішньо організаційної влади, і бухгалтери, що втратили організаційний статус чи владу, будуть сприймати її негативно, на відміну від тих бухгалтерів, чий вплив збільшився.

Для тестування наведених теорій Маркус спрогнозувала три альтернативні результати кожного теоретичного пояснення й проаналізувала ступінь їх узгодженості з власними спостереженнями в компанії. Наприклад, якщо теорія впливу технічної системи передбачає незадоволеність бухгалтерів технічною стороною, тоді рішення технічних проблем зменшить опір змінам. Після того, як версія фінансової звітності була оновлена й були внесені інші важливі зміни в систему, бухгалтери філії пропрацювали з новою системою один рік, однак і після цього противники системи не змінили свого ставлення до неї. Теорія впливу технічної системи була відхилена.

Теорія впливу людського фактору прогнозувала, що заміна окремих супротивників системи або їх співпраця з більш лояльними до системи бухгалтерами зменшить супротив. Відповідно, корпорація започаткувала політику ротації й мобільності кадрів, переміщуючи бухгалтерів із незадоволених відділів, однак ці зусилля виявилися марними: супротив не тільки залишився, а й навіть зріс! В одному випадку бухгалтер, що був одним із розробників системи та її палким прихильником, після переміщення до відділу контролю філії почав оцінювати її негативно. У результаті теорія впливу людського фактору теж була відхилена.

Нарешті, теорія взаємодії прогнозувала, що технічні зміни системи чи людський фактор (тобто рівень освіти чи політика ротації посад) не зменшать супротив, якщо не буде надано належної уваги питанням незбалансованості й перерозподілу порівняно з початковим етапом (до впровадження системи). До початку впровадження нової системи бухгалтери філії відчували себе власниками всіх фінансових даних своєї філії. Вони зберігали ці дані в паперовому форматі й контролювали доступ до них сторонніх осіб, а також могли вносити певні правки у звітність до того, як фінансові звіти надавалися керівництву. Бухгалтери центрального офісу значно залежали від бухгалтерів філії у питанні доступу до даних філії під час створення консолідованого звіту. Коли нова система уможливила автоматичний збір бухгалтерської звітності та її консолідації в єдиній базі даних корпорації, значущість бухгалтерів філії зменшилася і їхня робота стала менш важливою. Бухгалтери центрального офісу отримали прямий доступ до даних філії через корпоративну базу даних, а отже, і можливість виявляти похибки в роботі

бухгалтерів філій. Це мало наслідком руйнацію влади філій і супротив її працівників. Хоча Маркус не протестувала цю теорію емпірично, її спостереження щодо перерозподілу організаційної влади та відхилення двох попередніх теорій призвели до підтвердження теорії взаємодії.

Порівняння з традиційними дослідженнями

Позитивістське кейс-дослідження, спрямоване на тестування гіпотез, часто стає предметом критики дослідників природничих наук за відсутність контрольних груп, відтворюваності й узагальненості даних, тобто за недотримання традиційних принципів позитивістських досліджень. Цю критику можна подолати завдяки застосуванню відповідних дизайнів кейс-дослідження. Наприклад, проблема контрольованого спостереження стосується складності отримання контролю в кейс-дослідженнях. Однак, дослідники, що застосовують кейс-метод, можуть компенсувати недостатній рівень контролю, звернувшись до «природнього контролю». Природнім контролем в дослідженні Маркус (1983) був бухгалтер корпорації, який спочатку підтримував систему, однак змінив свою думку після переходу у відділ контролю. Зміна у ставленні до системи може бути пов'язана саме зі зміною посади. Однак такий природний контроль складно запланувати заздалегідь, і дослідники можуть його легко пропустити, якщо не будуть цілеспрямовано докладати зусиль для його пошуку.

Проблема контрольованої дедукції стосується відсутності адекватних кількісних доказів для підтвердження висновків, виходячи з переважно якісної природи даних кейс-досліджень. Не дивлячись на відсутність кількісних даних для тестування гіпотез, контрольована дедукція в кейс-дослідженнях може бути отримана шляхом генерування прогнозів, що базуються на теоретичних міркуваннях і тестуванні цих прогнозів у часовому вимірі. Маркус використала цю стратегію у своєму дослідженні, запропонувавши три альтернативні теоретичні гіпотези й відхиливши дві з них через невідповідність фактичним даним. У цьому випадку гіпотези тестувалися, спираючись здебільшого на логічні судження, а не на математичні тести, що є таким же валідним, як і статистичні висновки, оскільки математика є підрозділом логіки.

Крім того, проблема відтворюваності стосується складності отримання того ж самого феномену, ураховуючи унікальність і специфіку конкретного місця дослідження. Однак, використовуючи три теорії Маркус як ілюстрацію, інший дослідник може перевірити їх прийнятність, протестувавши три різні припущення в іншій організації. Іншими словами, можливо відтворити висновки кейс-дослідження, хоча місце його проведення і контекст не завжди можуть бути відтворені.

Кейс-дослідження спрямоване на вивчення унікальних і часто не відтворюваних феноменів, які не можуть бути узагальнені до рівня всіх

компаній. Узагальненість у природничих науках установлюється завдяки додатковим дослідженням. Аналогічно проводяться додаткові кейс-дослідження в різних умовах із різними прогнозами, що допомагає встановити рівень узагальненості результатів, якщо такі результати постійно спостерігаються в усіх дослідженнях.

Нарешті, британський філософ Карл Поппер описав чотири вимоги до наукових теорій: 1) можливість їхньої перевірки й спростування; 2) вони повинні бути логічно послідовні; 3) вони повинні мати адекватну прогностичну цінність; 4) вони повинні надавати краще пояснення, ніж конкуруючі теорії. У кейс-дослідженнях перші три вимоги можуть бути збільшені шляхом зростання ступеня свободи спостережень (таких як збільшення кількості задіяних компаній, кількості альтернативних припущень, кількості досліджуваних рівнів аналізу). У дослідженні Маркус це було зроблено завдяки спостереженням за поведінкою численних груп (бухгалтери філії та центрального офісу) та наданням трьох альтернативних пояснень. Четверта умова за Поппером була виконана, оскільки одна із запропонованих теорій надавала краще пояснення феномену, ніж дві інші конкуруючі гіпотези.

Розділ 12

ІНТЕРПРЕТАТИВНЕ ДОСЛІДЖЕННЯ

Попередній розділ був присвячений інтерпретативним кейс-дослідженням. Цей розділ аналізує інші види інтерпретативних досліджень. Згадайте, що позитивістські чи дедуктивні методи, такі як лабораторні експерименти чи опитування, спрямовані на тестування теорій (гіпотез), тоді як інтерпретативні чи індуктивні методи, такі як дослідження дією чи етнографія, спрямовані на побудову теорій. На відміну від позитивістського методу, коли дослідник починає з теорії та перевіряє теоретичні припущення, спираючись на емпіричні дані, при застосуванні інтерпретативних методів дослідник починає з емпіричних даних і намагається розробити на їх основі теорію про досліджуваний феномен.

Термін «інтерпретативне дослідження» часто невірно використовується як синонім «якісного дослідження», хоча ці дві концепції значно відрізняються. Інтерпретативне дослідження – це парадигма дослідження (див. розділ 3), яка базується на припущенні про те, що соціальна дійсність не є однозначною чи об'єктивною, а скоріше формується людським досвідом і соціальним контекстом (онтологія), а отже, краще досліджується в соціально-історичному контексті шляхом узгодження суб'єктивних інтерпретацій різних учасників дослідження (епістемологія). Оскільки дослідники, що використовують методи інтерпретативного дослідження, розглядають соціальну реальність як таку, що вкорінена й не може бути відділеною від свого соціального контексту, вони інтерпретують її як процес «надання змісту», а не як процес тестування гіпотез. Таке трактування контрастує з позитивістською чи заснованою на функціоналізмі парадигмі, яка передбачає, що реальність є відносно незалежною від контексту, може бути відділена від контексту та досліджуватися за структурними елементами у функціональному стилі з використанням об'єктивних методик. Буде дослідник виконувати інтерпретативне чи функціональне дослідження залежить від парадигмальних міркувань щодо сутності досліджуваного феномену і бачення шляхів його найкращого дослідження.

Однак протиставлення кількісних і якісних досліджень стосується емпіричних, орієнтованих на дані, міркувань щодо типу даних, які потрібно зібрати та як їх потрібно потім аналізувати. Якісні дослідження покладаються головним чином на нечислові дані, такі як інтерв'ю чи спостереження. Кількісні дані, навпаки, використовують числові дані. Отже, якісні дослідження не підлягають статистичним процедурам аналізу, таким як регресійний аналіз, але кодуються з використанням

контент-аналізу. Іноді кодовані якісні дані представляються кількісно як частоти кодів, але такі дані статистично не аналізуються. Багато дослідників, що дотримуються інтерпретативної парадигми, відхиляють кодування як марну спробу знайти консенсус чи об'єктивність у соціальних феноменах, які по суті є суб'єктивними.

Хоча інтерпретативне дослідження значною мірою покладається на якісні дані, кількісні дані можуть додати більше точності та чіткості в розуміння феномену, ніж виключно якісні дані. Наприклад, Ейзенхард (1989) у своєму інтерпретативному дослідженні процесу прийняття рішень (опис якого наводився в попередньому розділі) збирала числові дані стосовно того, скільки часу потрібно компанії для прийняття певних стратегічних рішень (значення знаходились у межах від 1,5 до 18 місяців), скільки альтернативних варіантів розглядається при прийнятті кожного рішення, і проводила анкетування респондентів стосовно оцінок організаційних конфліктів. Такі кількісні дані допомогли їй чітко відділити компанії, які швидко приймають рішення, від компаній, де процес прийняття рішень триває довше, не покладаючись при цьому на суб'єктивні враження респондентів. Потім це дозволило їй дослідити кількість альтернатив, що розглядаються компаніями, і рівень конфліктності в компаніях із різної швидкістю прийняття рішень. Інтерпретативне дослідження має прагнути до збору як якісних, так і кількісних даних, що відносяться до досліджуваного феномену, як і позитивістські дослідження. Спільне використання якісних і кількісних даних, яке часто називається змішаним дизайном, може призвести до унікальних ідей і високо цінується науковою спільнотою.

Інтерпретативне дослідження своїм походженням зобов'язано антропології, соціології, психології, лінгвістиці та семіотиці. Воно використовувалося ще на початку XIX ст., задовго до виникнення позитивістського підходу. Із огляду на суб'єктивну природу якісних даних та процесу інтерпретації, який при цьому застосовується, багато дослідників-позитивістів вважають інтерпретативне дослідження неточним і навіть хибним. Однак нездатність багатьох позитивістських досліджень отримати цікаві наукові ідеї чи нове знання призвела до відродження інтересу до інтерпретативних досліджень починаючи з 70-х років XX ст., хоча і в поєднанні з більш точними методами й строгими критеріями для забезпечення надійності та валідності висновків цих досліджень.

Відмінності від позитивістських досліджень

На додаток до фундаментальних відмінностей на рівні парадигми в онтологічних й епістемологічних припущеннях, наведених вище, інтерпретативні й позитивістські дослідження відрізняються в деяких інших аспектах. По-перше, інтерпретативне дослідження застосовує стратегію теоретичного формування вибірок, при якій місце дослідження,

респонденти та ситуації добираються, виходячи з теоретичних міркувань, таких як відповідність досліджуваного феномену меті дослідження (наприклад, практики сталого розвитку можуть досліджуватися лише в компаніях, де ці практики впроваджені) чи наявність певних характеристик, які роблять їх найбільш доречними для проведення дослідження. Позитивістські дослідження, навпаки, звертаються до випадкового формування вибірок із генеральної сукупності, щоб потім досягти можливості узагальнення даних. Отже, на відміну від позитивістських досліджень, формування вибірок за зручністю та невеликі за розміром вибірки вважаються прийнятними в інтерпретативних дослідженнях, якщо вони відповідають сутності й меті дослідження.

По-друге, в інтерпретативних дослідженнях роль дослідника значно зростає. При застосуванні етнографії, дослідження дією чи спостереження дослідник стає фактично частиною соціального феномену, і його специфічна роль та участь у дослідницькому процесі повинні бути чітко визначені під час аналізу даних. При застосуванні інших методів, таких як кейс-дослідження, дослідник повинен займати нейтральну позицію, щоб зберегти неупереджений погляд на ситуацію під час збору та аналізу даних. У позитивістських дослідженнях дослідник не втручається в перебіг досліджуваного феномену, є незалежним щодо контексту дослідження, що повинно не допустити похибки при зборі й аналізі, пов'язаної з особистістю дослідника.

По-третє, інтерпретативний аналіз є більше цілісним і контекстуальним, ніж редукційним й ізоляціоністським. Інтерпретативні пояснення більше фокусуються на мові, знаках і значенні з точки зору учасників соціального феномену, на відміну від статистичного аналізу, який інтенсивно використовується в позитивістських дослідженнях. Точність в інтерпретативних дослідженнях розглядається в термінах систематичного й прозорого підходу до збору й аналізу даних, а не в термінах статистичних тестів для оцінки валідності конструктів чи оцінки істотності.

Нарешті, в інтерпретативних дослідженнях збір й аналіз даних можуть виконуватися одночасно й за необхідності повторюватися. Наприклад, дослідник може проводити інтерв'ю та кодування даних до початку нового інтерв'ю. Одночасний аналіз допомагає досліднику коригувати перебіг інтерв'ю, уносити зміни в протокол дослідження, щоб краще виявити характерні риси досліджуваного феномену. Дослідник може навіть переформулювати мету дослідження, якщо в ході проведення інтерв'ю стає зрозумілим, що такі зміни допоможуть отримати більш цікаві з наукової точки зору результати. Це дуже цінна, але часто недооцінювана перевага інтерпретативних досліджень, якої немає в позитивістських дослідженнях, де дослідницький проект не може змінюватися чи удосконалюватися після початку збору даних.

Переваги й обмеження інтерпретативних досліджень

Інтерпретативні дослідження мають декілька унікальних переваг. По-перше, вони найбільш доречні для виявлення прихованих причин складних, суперечливих і багатозначних соціальних процесів, таких як відносини між компаніями або внутрішня політика компанії, де кількісні докази можуть бути хибними, неточними або недоступними. По-друге, вони часто допомагають у побудові теорії в тих галузях, де теоретичні напрацювання відсутні чи знаходяться на початкових етапах формування. По-третє, вони доречні для дослідження контекстуально обумовлених, унікальних подій та процесів. По-четверте, інтерпретативні дослідження можуть допомогти у виявленні цікавих наукових проблем для подальших наукових досліджень.

Водночас, інтерпретативні дослідження мають свої обмеження. По-перше, вони є більш ресурсозатратними порівняно з позитивістськими дослідженнями на етапі збору й аналізу даних. Невелика кількість даних можуть призвести до поспішних чи хибних припущень, тоді як велика кількість даних не завжди може бути ефективно оброблена дослідником. По-друге, інтерпретативні дослідження вимагають висококваліфікованих дослідників, які будуть здатні виявити й пояснити складні соціальні феномени з позиції учасників, узгодити їх різноманітні погляди й міркування, уникнувши перенесення власних оцінок і припущень на висновки дослідження. По-третє, усі учасники та джерела даних можуть бути не завжди надійними й добре обізнаними щодо досліджуваного феномену або навіть переслідувати власні неочевидні «політичні» цілі, що може призвести до неточних і навіть хибних висновків. Неадекватна довіра між учасниками й дослідником може завадити повному та відвертому висловлюванню думок, а побудова такої довіри вимагає часу. Важливою частиною роботи дослідника є вміння «читати між рядками» приховані мотиви та враження й зрозуміти істинну природу проблеми. По-четверте, ураховуючи висококонтекстуальну природу висновків, що робляться в результаті проведення інтерпретативного дослідження, такі висновки часто складно підтвердити чи узагальнити. Нарешті, інтерпретативні дослідження можуть іноді не надати відповіді на поставленні в дослідженні питання чи передбачити подальшу динаміку досліджуваного феномену.

Характеристики інтерпретативного дослідження

Усі інтерпретативні дослідження повинні відповідати певному набору принципів, що наводяться нижче.

Природне середовище. Соціальний феномен повинен вивчатися у своєму природному середовищі, оскільки інтерпретативне дослідження передбачає, що соціальні феномени не можуть бути виокремлені із свого соціального контексту. Інтерпретації феноменів повинні також

проводитися в соціально-історичному контексті. Це означає необхідність урахування контекстуальних змінних при пошуку пояснень досліджуваного феномену, навіть якщо контекст може обмежувати узагальненість отриманих висновків.

Дослідник як інструмент. Дослідники часто занурені в соціальний контекст, який вони вивчають, і є частиною інструменту збору даних в тому відношенні, що вони повинні використовувати свої здібності до спостереження і довірчі стосунки з учасниками, щоб знаходити надійну інформацію. Їхні особисті ідеї, знання та досвід є критичними для адекватного трактування й пояснення досліджуваного феномену. Водночас, дослідники повинні повністю усвідомлювати свої особисті переконання та упередження, щоб вони не деформували уявлення про досліджуваний феномен.

Інтерпретативний аналіз. Спостереження мають інтерпретуватися з точки зору учасників, що діють у певному соціальному контексті. Інтерпретація проводиться на двох рівнях. Перший рівень включає бачення соціального феномену очима учасників. Другий рівень передбачає розуміння досвіду учасників, щоб надати глибокий змістовний опис досліджуваного феномену та з'ясувати, чому учасники діють саме таким чином.

Використання виразної мови. Документування вербальної та невербальної мови учасників та подальший аналіз цієї мови є інтегральною складовою інтерпретативного аналізу. Дослідження повинно забезпечити, щоб ситуація бачилася очима живих людей, а не машин, а отже, має фіксувати емоції та досвід учасників.

Динамічна природа. Інтерпретативне дослідження часто не ставить на меті пошук конкретних відповідей, а радше спрямоване на розуміння чи знаходження змісту динамічних соціальних процесів у хронологічній послідовності. Отже, таке дослідження вимагає повного «вживання» дослідника в ситуацію протягом тривалого періоду часу, щоб повністю відслідкувати еволюцію досліджуваного феномену.

Цикл герменевтики. Інтерпретативне трактування є повторюваним процесом руху вперед і назад від певних фрагментів спостереження (текст) до цілісної картини соціального феномену (контекст) із метою подолання їх видимих суперечностей і конструювання теорії, яка узгоджується з різноманітними суб'єктивними поглядами й досвідом учасників дослідження. Такі ітерації між розумінням феномену і спостереженнями повинні продовжуватися, доки не буде досягнуто «теоретичного очищення», і подальші ітерації не додаватимуть нових знань стосовно досліджуваного феномену.

Збір даних інтерпретативного дослідження

Дані інтерпретативного дослідження збираються різноманітними способами. Найбільш поширеною є техніка **інтерв'ю** (віч-на-віч, телефонне чи у фокус-групах). Типи та стратегії інтерв'ю детально розглядалися в попередньому розділі. Друга техніка – це **спостереження**. Воно може бути *прямим*, коли дослідник є нейтральним чи пасивним зовнішнім спостерігачем і не втручається в досліджувану ситуацію (як у випадку кейс-дослідження), чи *спостереженням учасника*, коли дослідник стає активним учасником ситуації, і його особистий внесок та присутність впливають на досліджуваний феномен (як у дослідженні дією). Третя техніка – це **документування**, коли внутрішні та зовнішні документи можуть бути використані для встановлення важливої інформації щодо сутності феномену або як документальне підтвердження доказів дослідника.

Дизайни інтерпретативного дослідження

Кейс-дослідження. Як говорилося в попередньому розділі, кейс-дослідження є інтенсивним довготривалим дослідженням певного феномену в одній чи декількох організаціях (або в інших місцях проведення дослідження) із метою виявлення детальних контекстуальних закономірностей і розуміння динамічних процесів, що лежать в основі певного феномену. Кейс-дослідження – це унікальний дизайн, оскільки він може використовуватися в інтерпретативній формі для побудови теорій чи в позитивістській формі для тестування теорій. Дослідник при використанні цього дизайну виступає нейтральним спостерігачем (пряме спостереження) у певному соціальному середовищі, а не активним учасником подій (спостереження учасника). Як і в разі будь-яких інтерпретативних досліджень, формулювання змістовних висновків кейс-дослідження значною мірою залежить від здібностей до спостереження та синтезуючих здібностей дослідника.

Дослідження дією. Дослідження дією – це якісний позитивістський дизайн, спрямований скоріше на тестування гіпотез, ніж на їх побудову. Це інтерпретативний дизайн, який передбачає, що складні соціальні феномени можна найкраще зрозуміти в процесі впровадження змін, активних втручань, тобто «дій», одночасно спостерігаючи за результатами впливу цих дій на досліджуваний феномен. При застосуванні цього методу дослідник зазвичай виступає консультантом чи працівником організації, який ініціює дії у відповідь на соціальну проблему та вивчає наслідки впливу цих дій на досліджуваний феномен, водночас навчаючись і генеруючи нові уявлення щодо взаємозв'язків між дією та феноменом. Прикладами дій можуть бути програми організаційних змін, таких як упровадження нових організаційних процесів, процедур, кадрових чи технологічних змін, які ініціюються з метою покращення показників діяльності чи прибутковості компанії в

конкретних умовах бізнес-середовища. Вибір дій дослідника повинен базуватися на теорії, яка потенційно пояснює, чому і як такі дії можуть принести бажаний результат. Теорія валідизується настільки, наскільки обрані дії (заходи) виявляться успішними у вирішенні конкретної проблеми. Одночасне вирішення проблеми та генерування пояснень є ключовою рисою, яка відрізняє дослідження дією від інших методів дослідження (які можуть не включати рішення певної проблеми) і від консультування (яке не обов'язково вимагає формулювання теоретичних пояснень). Отже, дослідження дією є прекрасним методом для встановлення зв'язку теорії та практики.

Існує декілька варіантів цього методу. Найбільш популярним є метод участі в дослідженні, запропонований Сусманом й Евередом (1978)¹³. Цей метод слідує циклу дослідження дією, який включає п'ять етапів: 1) діагностика; 2) планування дій; 3) реалізація дій; 4) оцінка; 5) навчання (див. рис. 10.1). *Діагностика* включає ідентифікацію та визначення проблеми в соціальному контексті. *Планування дій* включає ідентифікацію та оцінку альтернативних рішень проблеми, а також прийняття рішення стосовно подальших дій (базуючись на теоретичних міркуваннях). *Реалізація дій* – це впровадження запланованих заходів. На етапі *оцінки* встановлюється, якою мірою впроваджені заходи виявилися успішними в рішенні поставленої проблеми, тобто чи підтвердилися теоретичні постулати практикою. На етапі *навчання* використовується досвід і зворотний зв'язок від оцінки, щоб знайти теоретичне пояснення проблеми і запропонувати подальші модифікації дій. Спираючись на оцінку та навчання, дії можуть бути удосконалені, щоб краще вирішувати аналогічні проблеми, і цикл досліджень дією повторюється знову, включаючи модифіковані заходи. Уважається, що повний цикл дослідження дією має відбутися не менше двох разів, щоб навчання в результаті першого циклу могло бути впроваджене в другому циклі. Головний спосіб збору даних полягає в спостереженні учасника, хоча інші способи, такі як інтерв'ю чи вивчення документів, можуть також застосовуватися, щоб підтвердити спостереження дослідника.

¹³Susman G.I., Evered R.D. An Assessment of the Scientific Merits of Action Research / G.I.Susman, R.D.Evered // Administrative Science Quarterly. – 1978. – Vol. 23. – P. 582-603.



Рис. 10.1. Цикл дослідження дією

Етнографія. Метод етнографічного дослідження походить значною мірою з антропології, роблячи акцент на дослідженні феномену в його культурному контексті. Дослідник має бути глибоко зануреним у соціальну культуру протягом тривалого періоду часу (від 8 місяців до 2 років) і повинен бути залученим до неї, щоб спостерігати та фіксувати щоденне життя досліджуваної культури та її учасників у природньому середовищі. Головним способом збору даних є спостереження учасника, тоді як аналіз даних включає підхід «знаходження змісту». На додаток, дослідник повинен вести детальний польовий щоденник, фіксуючи свій досвід у всіх подробицях, щоб читачі щоденника змогли повною мірою відчувати особливості досліджуваної культури. За цим методом дослідник виконує дві ролі: 1) висуває нові наукові ідеї, покладаючись на свої унікальні знання і поглиненість в феномен; 2) переконує наукову спільноту, що виявлені закономірності досліджуваного феномену виходять за межі окремого кейсу.

Феноменологія. Феноменологія – це метод дослідження, який підкреслює необхідність вивчення свідомого досвіду як спосіб дослідження навколишньої дійсності. Метод базується на ідеях німецького філософа початку ХХ ст. Едмунда Хассерла, який уважав, що людський досвід є джерелом усіх знань. Феноменологія спрямована на систематичну рефлексію (відображення) та аналіз феноменів, пов'язаних зі свідомим досвідом, таким як людські судження, сприйняття та дії з метою: 1) оцінити та описати соціальну реальність із різних суб'єктивних точок зору, які мають учасники дослідження; 2) зрозуміти символічне значення («глибоку структуру»), що лежить в основі цього досвіду. Феноменологічне дослідження вимагає, щоб дослідники відмовилися від будь-яких попередніх припущень і персональних упереджень, «вжилися» в ситуацію кожного учасника й налаштувалися на екзистенціальні виміри ситуації, щоб повністю зрозуміти глибокі структури, що є драйверами свідомого мислення, почуттів і поведінки учасників дослідження.

Деякі дослідники розглядають феноменологію більше як філософію, ніж як метод дослідження. У відповідь на критику Джорджі та Джорджі (2003)¹⁴ розробили *метод екзистенціальних феноменологічних досліджень* як керівництво для досліджень у цій галузі. Метод, що ілюструється на рис. 10.2., складається зі збору даних та їх аналізу. На етапі *збору даних* учасники, занурені в соціальний феномен, проходять інтерв'ювання, щоб виявити їх суб'єктивний досвід і погляди стосовно досліджуваного феномену. Прикладами питань, що ставляться перед учасниками, можуть бути такі: «Як ви опишете типовий день у вашому відділі?» або «Ви можете описати конкретно цей випадок якомога детальніше?». Ці інтерв'ю записуються для подальшого аналізу. Під час аналізу даних дослідник читає записи, щоб: 1) сформуванати цілісну картину ситуації; 2) встановити «одиниці значимості», які будуть вірно представляти суб'єктивний досвід учасників. Прикладами одиниць значимості є такі концепти, як «сприйнятий простір» чи «сприйнятий час», які потім використовуються, щоб задокументувати психологічний досвід учасників. Наприклад, чи відчувають учасники себе в безпеці, вільними, спокійними, коли стикаються з певним феноменом («сприйнятий простір»). Феноменологічний аналіз має брати до уваги часовий вимір учасників (їх відчуття минулого, теперішнього та майбутнього), і дослідник повинен помістити себе в уявні почуття ситуації, у якій знаходилися учасники (тобто, ужитися в їхнє життя). Життєвий досвід учасників описується у формі розповіді або за певними темами. Подальший аналіз зосереджується на цих темах, щоб ідентифікувати численні рівні змісту, максимально зберігаючи крихкість і неоднозначність досвіду учасників.

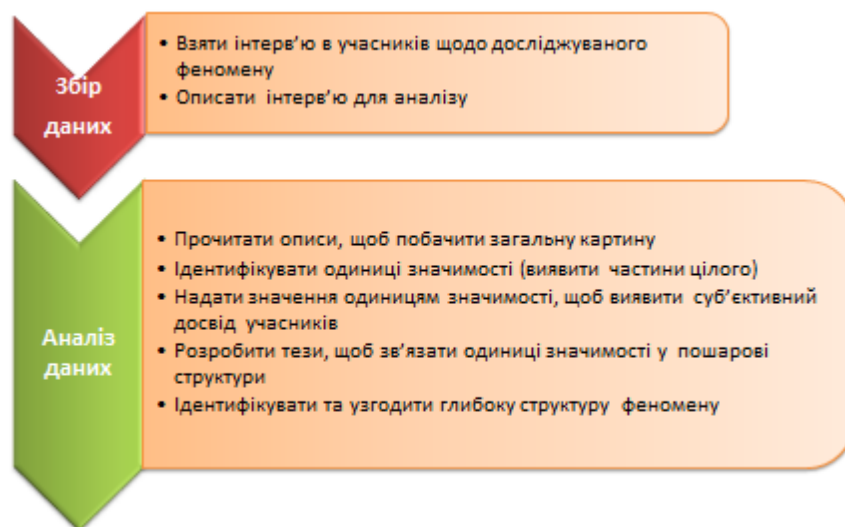


Рис. 10.2. Метод екзистенціальних феноменологічних досліджень

¹⁴Giorgi A., Giorgi B. Phenomenology/ A. Giorgi, B. Giorgi. //Qualitative Psychology: A Practical Guide to Research Methods, – A. Smith (ed.). – London: Sage Publications. – 2003.

Строгість в інтерпретативних дослідженнях

Якщо позитивістське дослідження звертається до редукціонізму шляхом спрощення дійсності у вигляді загальних теорій і законів, інтерпретативні дослідження намагаються інтерпретувати соціальну реальність через суб'єктивні погляди учасників у тому контексті, в якому досліджуваний феномен і відбувається. Ці інтерпретації дуже контекстуальні і, звичайно, складно встановити, наскільки вони будуть прийнятними відносно інших контекстів. Однак, оскільки інтерпретативний аналіз є суб'єктивним і залежним від індивідуального досвіду та уявлень дослідника, дослідники-позитивісти часто вважають цей аналіз менш точним. Оскільки інтерпретативне дослідження базується на різних наборах онтологічних та епістемологічних припущень щодо соціального феномену порівняно з позитивістським дослідженням, позитивістські поняття строгості, такі як надійність, внутрішня валідність й узагальненість, не можуть застосовуватися так, як вони застосовуються в позитивістському аналізі. Однак, Лінкольн і Губа (1985)¹⁵ надають альтернативний набір критеріїв для оцінки строгості в інтерпретативних дослідженнях.

Залежність. Інтерпретативне дослідження може вважатися надійним, якщо два дослідники, оцінюючи один і той же феномен, спираються на один і той же набір доказів і при цьому приходять до однакових висновків, або якщо один і той же дослідник, спостерігаючи за тим же або подібним феноменом, робить аналогічні висновки. Ця концепція близька до концепції надійності в позитивістських дослідженнях, якщо два незалежних дослідники притримуватися схожих понять внутрішньої надійності. Схожість двох спостережень одного й того ж феномену одним і тим же дослідником аналогічна за трактуванням до тест-ретест надійності. Для забезпечення залежності дослідники-інтерпретативісти повинні надавати достатньо детальну інформацію стосовно досліджуваного феномену та соціального контексту, у якому він відбувався, щоб читачі могли самостійно оцінити істинність висновків інтерпретації.

Довіра. Інтерпретативне дослідження може вважатися таким, що викликає довіру, якщо читачі переконуються в правдоподібності його висновків. Ця концепція аналогічна концепції внутрішньої валідності в дослідженні на основі функціональної парадигми. Довіра до інтерпретативного дослідження може зрости, якщо надаються докази інтенсивного залучення дослідника в спостереження феномену і демонструється узгодженість серед досліджуваних чи методик збору даних, а також забезпечується активне управління даними й процедурами аналізу, такими як словесний опис інтерв'ю, точна фіксація контактів й

¹⁵ Lincoln Y. S., Guba E. G. *Naturalistic Inquiry* / Y. S. Lincoln, E. G. Guba.. – Beverly Hills, CA: Sage Publications. – 1985.

інтерв'ю, зрозумілі пояснення щодо теоретичних і методологічних рішень, які в сукупності дозволяють у разі потреби провести незалежний аудит збору й аналізу даних.

Можливість підтвердження. Можливість підтвердження стосується ступеня того, як результати інтерпретативного дослідження можуть бути незалежно підтвержені іншими (зазвичай учасниками дослідження). Це схоже до визначення об'єктивності в позиціях на основі функціональної парадигми. Оскільки інтерпретативне дослідження заперечує поняття об'єктивної реальності, можливість підтвердження демонструється в термінах «інтерсуб'єктивності», тобто підтвердження учасниками висновків, зроблених дослідником. Наприклад, якщо учасники дослідження, вивчивши звіт або наукову статтю про дослідження, у цілому погоджуються зі зробленими висновками, то останні можуть уважатися такими, що підтверджуються.

Можливість перенесення. Можливість перенесення в інтерпретативному дослідженні стосується ступеня того, наскільки результати можуть бути узагальнені для інших випадків. Ця ідея аналогічна зовнішній валідності в дослідженнях на основі функціональної парадигми. Дослідник має надати детальні та глибокі описи контексту, структури, припущень та процесів, виявлених при аналізі даних, щоб читач міг незалежно оцінити, наскільки загальними є отримані висновки, тобто наскільки вони можуть бути застосовані для пояснення інших аналогічних феноменів.

Розділ 13

ЯКІСНИЙ АНАЛІЗ

Якісний аналіз – це аналіз якісних даних, таких як тексти інтерв'ю учасників. На відміну від кількісного аналізу, при якому статистичні показники значно не залежать від дослідника, якісний аналіз значно залежить від аналітичних та синтетичних здібностей дослідника, його особистого знання соціального контексту, у якому збираються дані. Наголос у якісних дослідженнях робиться на розуміння і «знаходження змісту» феномену, а не на прогнозування чи пояснення. Для якісного аналізу потрібен креативний і дослідницький тип мислення, що базується на етично виваженому ставленні до учасників, бажанні зрозуміти їх позицію і виробити правильну стратегію дослідження. Цей розділ надає короткий огляд деяких стратегій проведення якісного аналізу.

Обґрунтована теорія

Як проаналізувати великий набір якісних даних, отриманих в результаті спостереження за учасниками, проведення тривалих інтерв'ю, фокус-груп, аудіо- та відеозаписів чи вторинних документів? Одним із таких способів аналізу текстових даних є **обґрунтована теорія** – індуктивна техніка інтерпретації записаних даних про соціальний феномен, спрямована на побудову теорії щодо сутності феномену. Цей спосіб був запропонований Гласером і Штрауссом (1967)¹⁶ у формі методу постійного порівняльного аналізу й у подальшому доопрацьований Штрауссом і Корбіним (1990)¹⁷, щоб проілюструвати специфічні техніки кодування – процес класифікації та категоризації сегментів текстових даних у набір кодів (концептів), категорій (конструктів) і взаємозв'язків. Інтерпретації базуються (ґрунтуються) на емпіричних даних спостереження. Щоб забезпечити, що теорія базується виключно на доказах спостереження, підхід обґрунтованої теорії вимагає від дослідника утримуватися від будь-яких попередніх теоретичних очікувань і покладатися при формулюванні теоретичних положень виключно на отримані дані.

Штраусс і Корбін (1998) описують три техніки кодування для аналізу текстових даних: відкрите, осьове та селективне. **Відкрите кодування** є процесом, спрямованим на ідентифікацію **концептів** чи ключових ідей, прихованих у текстових даних, які потенційно пов'язані з досліджуваним феноменом. Дослідник вивчає первинні текстові дані рядок за рядком, щоб виявити окремі події, випадки, ідеї та значимі

¹⁶Glaser B., Strauss A. The Discovery of Grounded Theory: Strategies for Qualitative Research / B. Glaser, A. Strauss. – Chicago: Aldine. – 1967.

¹⁷ Strauss A., Corbin J. Basics of Qualitative Research: Grounded Theory Procedures and Techniques/ A. Strauss, J. Corbin. – Beverly Hills, CA: Sage Publications. – 1990.

взаємодії, які кодуються як концепти й називаються кодами *invivo*. Кожний концепт пов'язаний зі специфічною порцією тексту (одиницею кодування) для подальшої валідації. Деякі концепти можуть бути простими, ясними й однозначними, тоді як інші досить складними й можуть по-різному трактуватися різними дослідниками. Одиниці кодування можуть варіюватися залежно від виділених концептів. Прості концепти можуть включати лише декілька слів, тоді як складні (наприклад, місія організації) можуть описуватися на декількох сторінках. Концепти можуть називатися, спираючись на власне бачення дослідника або стандартизовані позначення, що вже існують у науковій літературі. Після того, як базовий набір концептів ідентифікований, вони можуть бути використані для кодування інших даних, водночас упроваджуючи нові концепти та уточнюючи старі. Під час кодування важливо ідентифікувати характерні риси кожного концепту, такі як розмір або рівень, щоб потім схожі концепти могли бути згруповані. Ця техніка кодування називається відкритою, тому що дослідник залишається відкритим до активного пошуку нових концептів, релевантних для досліджуваного феномену.

Далі схожі концепти групуються в **категорії** вищого порядку. Хоча концепти можуть бути специфічними за змістом, категорії часто є досить широкими й узагальненими та в кінцевому рахунку формують конструкти в обґрунтованій теорії. Категорії потрібні для зменшення кількості концептів, із якими працює дослідник, і для побудови «широкої картини» питань, важливих для розуміння соціального феномену. Категоризація може виконуватися поетапно, об'єднанням концептів у підкатегорії, які потім створюють категорії вищого порядку. Конструкти, установлені на основі існуючої літератури, можуть використовуватися для назв цих категорій, особливо якщо мета дослідження полягає в розвитку існуючих теорій. Однак потрібно бути обачними, використовуючи вже існуючі конструкти, оскільки вони можуть бути результатом упереджень у поглядах і похибок. Для кожної категорії повинні бути ідентифіковані її характеристики (чи властивості), а також виміри кожної характеристики. Вимір представляє значення характеристики в певному континуумі. Наприклад, категорії «засоби комунікації» може мати характеристику, що називається «швидкість» і вимірюється як висока, середня та низька. Така категорія допомагає диференціювати між різними видами засобів комунікації й дозволяє дослідникам виявити певні паттерни даних, наприклад, які комунікаційні засоби використовуються для кожного виду завдань.

Другий етап обґрунтованої теорії – це **осьове кодування**, коли категорії та підкатегорії об'єднуються в причинні зв'язки чи гіпотези, які можуть орієнтовно пояснити досліджуваний феномен. Осьове кодування, хоча й відрізняється від відкритого, може виконуватися одночасно з ним. Зв'язки між категоріями можуть бути очевидними, або більш

прихованими та неявними. В останньому випадку дослідники повинні використовувати схему кодування (вона часто називається парадигмою кодування, однак відрізняється від парадигм, що були представлені в розділі 3) для розуміння того, які категорії представляють *умови* (обставини, в яких спостерігається феномен), *дій/взаємодій* (відповіді окремих осіб на події, що відбуваються при зазначених умовах) та *наслідки* (результати дій/взаємодій). Після того як виявлені умови, дії/взаємодії і наслідки, починають вимальовуватися теоретичні положення, і дослідник може перейти до пояснення того, чому даний феномен спостерігається за певних умов і які наслідки він має.

Третій і останній етап обґрунтованої теорії – це **селективне кодування**, яке включає ідентифікацію центральної категорії, або ключового показника й систематичне, логічне пов'язування цієї центральної категорії з іншими категоріями. Центральна категорія може розвинути на основі існуючих категорій або бути категорією вищого порядку, що вбирає в себе раніше закодовані категорії. Нові дані ретельно добираються, щоб валідизувати центральну категорію та її зв'язки з іншими категоріями (тобто сформувані попередню теорію). Селективне кодування обмежує діапазон аналізу та пришвидшує його. Водночас, при такому кодуванні потрібно слідкувати за іншими категоріями, які можуть виникнути на основі нових даних і бути пов'язані з досліджуваним феноменом (відкрите кодування), що може призвести до подальшого «очищення» первинної теорії. Отже, всі види кодувань можуть виконуватися одночасно. Кодування нових даних і уточнення теорії триває, доки не буде досягнуте **теоретичне насичення**, при якому додаткові дані не призводять до якихось суттєвих змін категорій або їх зв'язків.

Процес постійного порівняння означає постійне переформатування, об'єднання та уточнення категорій і покрокову взаємодію чотирьох етапів діяльності: 1) порівняння випадків/текстів, що відносяться до кожної категорії; 2) інтеграція категорій та їх властивостей; 3) розширення меж застосування теорії (фокусування на ключових концептах та ігнорування менш важливих концептів); 4) опис теорії з використанням таких технік, як створення коротких нотаток і розповідей, а також побудова діаграм, що буде обговорюватися в наступному розділі. Наявність центральної категорії не обов'язково означає, що всі інші категорії можуть органічно навколо неї згрупуватися. Щоб ідентифікувати ключові категорії, такі як умови, дії/взаємодії, наслідки центральної категорії, Штраус і Корбін (1990) рекомендують декілька інтегративних технік, таких як опис історії, використання нотаток чи картування концептів. При **описі історії** категорії та зв'язки використовуються для того, щоб пояснити історію досліджуваного феномену. Нотатки є теоретичними записами ідей щодо суттєвих концептів та їх теоретично закодованих зв'язків, як вони виявляються в

процесі аналізу теорії. Вони є важливими інструментами для відслідковування ідей, що розвиваються в процесі аналізу, та їх кристалізації. **Створення коротких нотаток** дозволяє виявити тренди та взаємозв'язки між категоріями за допомогою таблиць 2x2, діаграм, рисунків, або інших ілюстративних матеріалів. **Картування концепту** – це графічна презентація концептів і зв'язків між категоріями з використанням квадратів і стрілок. Головні концепти зазвичай представлені на декількох аркушах паперу, дошках чи за допомогою графічних програм і пов'язані між собою стрілками, щоб найкращим чином відповідати даним спостереження.

Після створення обґрунтованої теорії вона має бути «очищена» з точки зору внутрішньої узгодженості та логіки. Дослідник повинен забезпечити, щоб центральний конструкт мав заявлені характеристики та виміри, у іншому випадку аналіз даних має бути продовжений. Також потрібно зазначити, що характеристики та виміри всіх категорій варіюють. Наприклад, якщо категорією є частота прояву певної поведінки, тоді дані повинні забезпечити докази того, що є особи, які демонструють високу частоту такої поведінки, на відміну від інших осіб, для яких частота такої поведінки є низькою. Нарешті, теорія повинна бути валідизованою шляхом порівняння її з первинними даними. Якщо теорія суперечить наявним доказам, процес кодування потрібно повторювати, доки всі суперечності чи не пояснені раніше варіації будуть усунені.

Контент-аналіз – це систематизований аналіз змісту тексту (наприклад, хто що кому сказав, і які наслідки це мало) в кількісній та якісній формах. Контент-аналіз зазвичай виконується таким чином. По-перше, коли є багато різних текстів, які будуть аналізуватися (газети, звіти, огляди тощо), дослідник починає з формування вибірки текстів із наявної сукупності текстів. Це не є випадковим процесом, навпаки, спеціально обираються тексти, що найбільше відповідають меті дослідження. По-друге, дослідник обирає та застосовує правила поділу кожного тексту на сегменти, які можуть розглядатися як окремі одиниці аналізу. Цей процес називається *виділенням одиниць аналізу*. Наприклад, аналізуючи указані в тексті припущення, наслідки, бар'єри, учасники можуть створювати окремі одиниці аналізу. По-третє, дослідник розробляє і застосовує один чи декілька концептів до кожного сегменту тексту в результаті процесу, що називається *кодуванням*. Для кодування використовується схема кодування, що базується на темі дослідження й додаткових питаннях, які виявляються в процесі класифікації тексту. По-четверте, закодовані дані аналізуються зазвичай кількісно та якісно, щоб установити, які теми висвітлюються найчастіше, у якому контексті та як вони пов'язані між собою.

Простим прикладом контент-аналізу є *аналіз почуттів* – техніка, що використовується для виявлення думок чи ставлення людей до різних

об'єктів, осіб чи феноменів. Прикладом такого аналізу є ознайомлення з он-лайнними повідомленнями щодо політичного кандидата та класифікація цих повідомлень на групи позитивних, нейтральних та негативних. У цьому випадку кожне повідомлення представляє одиницю аналізу. Такий аналіз допомагає встановити, чи загалом позитивно, нейтрально чи негативно аудиторія ставиться до певного кандидата. Вивчення змісту он-лайн оглядів – це інший приклад контент-аналізу. Хоча він може бути виконаний вручну, для аналізу великих масивів даних (мільйони текстових записів) удаються до відповідних комп'ютерних програм. Це допомагає автоматизувати процедуру контент-аналізу і відслідкувати динаміку зміни ставлення до досліджуваного феномену в часі.

Часто контент-аналіз критикують за відсутність систематичних процедур, які б дозволили зробити його доступним для відтворення іншими дослідниками. Відповідаючи на критику, Шиллінг (2006)¹⁸ об'єднав різні процедури контент-аналізу в спіральну модель. Вона складається з п'яти рівнів, або фаз інтерпретації тексту: 1) конвертування записів тексту в первинні текстові дані або стенограми для контент аналізу; 2) конвертування первинних даних у стислі протоколи; 3) конвертування протоколів у попередню систему категорій; 4) використання системи категорій для генерування кодованих протоколів; 5) аналіз кодованих протоколів для створення інтерпретацій про досліджуваний феномен.

Контент-аналіз має декілька обмежень. По-перше, процес кодування обмежений інформацією, що міститься в текстовій формі. У разі відсутності текстових даних про досліджуваний феномен, аналіз не може бути проведений. По-друге, потрібно ретельно проводити добір текстів для аналізу, щоб уникнути помилок формування вибірки. Наприклад, якщо дослідник вивчає лише опубліковані літературні джерела за темою дослідження, тоді поза його увагою залишаться неопубліковані дані, зокрема й результати нових досліджень, які ще не встигли опублікувати.

Герменевтика

Герменевтика – це спеціальний тип контент-аналізу, при якому дослідник прагне зробити інтерпретацію суб'єктивного значення певного тексту в соціально-історичному контексті. На відміну від обґрунтованої теорії чи контент аналізу, який ігнорує зміст і значення текстового документу в процесі кодування, цей вид аналізу якісних даних є виключно інтерпретативним. Метод передбачає, що письмові тексти дають інформацію про досвід автора в соціально-історичному контексті й

¹⁸Schilling J. On the Pragmatics of Qualitative Assessment: Designing the Process for Content Analysis/ J. Schilling // European Journal of Psychological Assessment. – 2006. – Vol. 22:1. – P. 28-37.

повинні інтерпретуватися в цьому контексті. Отже, дослідник постійно переходить від певної інтерпретації тексту (частини) до цілісного розуміння контексту (цілого) для розвитку більш повного розуміння феномену в його контексті. Цей процес німецький філософ Мартін Хайдеггер назвав *циклом герменевтики*. Слово єдиний (герменевтичний) стосується одного конкретного методу чи аспекту інтерпретації.

У більш загальному сенсі герменевтика – це вивчення інтерпретації, її теорії та практики. Історично метод має походження з релігії та лінгвістики. У ХХ ст. Хайдеггер припустив, що більш прямим та неопосередкованим шляхом розуміння соціальної реальності є практичне випробування, а не стороннє спостереження. Він запропонував філософську герменевтику, у якій фокус зміщується від інтерпретації до екзистенціального розуміння. Хайдеггер уважав, що тексти є засобами, за допомогою яких читачі можуть не тільки дізнатися про досвід автора, але й відтворити цей досвід. Сучасна, або модерна герменевтика, розвинута учнями Хайдеггера, зокрема Хансом Георгом Гадамером, досліджує межі письмових текстів для комунікації соціального досвіду й рухається далі, пропонуючи рамки інтерпретативного процесу, охоплюючи всі форми комунікацій (письмові, вербальні, невербальні тощо), та виявляючи фактори, що обмежують комунікаційні здатності письмових текстів, таких як передумови, структура мови (граматика, синтаксис тощо) і семіотика (дослідження письмових знаків – символів, метафор, аналогій тощо).

Розділ 14

КІЛЬКІСНИЙ АНАЛІЗ: ОПИСОВА СТАТИСТИКА

Числові дані наукових досліджень можуть аналізуватися кількісно за допомогою статистичних методів. Це відбувається двома шляхами. Описовий аналіз стосується статистичного опису, поєднання та презентації досліджуваних конструктів або зв'язків між цими конструктами. Пояснювальна статистика стосується статистичного тестування гіпотез (тестування теорій). У цьому розділі будуть досліджуватися статистичні методи описового аналізу, у наступному розділі – пояснювального. У сучасній науці більшість видів кількісного аналізу виконується за допомогою програмного забезпечення, наприклад, SPSS або SAS.

Підготовка даних

У дослідницьких проектах дані можуть збиратися з різноманітних джерел: анкетних опитувань, інтерв'ю, експериментальних даних, спостережень і т.п. Ці дані потрібно конвертувати таким чином, щоб вони могли читатися, а потім аналізуватися комп'ютерною програмою – у вигляді таблиці чи текстового файлу. Підготовка даних відбувається у декілька етапів.

Кодування даних. Кодування – це процес конвертації даних у числовий формат. Бажано створити книгу кодів для адміністрування процесу кодування. Книга кодів – це комплексний документ, що містить детальні описи кожної змінної дослідження, час заміру та її показники, формат (цифровий чи текстовий і т.п.), шкали відповідей для кожного показника (номінальна, порядкова, інтервальна чи шкала відношень, число градацій шкали і т.п.), а також як кожне значення буде переводитися (кодуватися) в числовий формат. Наприклад, номінальні дані, такі як галузь підприємства, можуть кодуватися за такою схемою: 1 – промисловість, 2- торгівля, 3 – фінанси і т.п. Зрозуміло, що номінальні дані не можуть аналізуватися статистично. Такі дані як вік, стаж, дохід кодуються так, як були вказані респондентом. Іноді дані можуть змінювати первинний формат. Наприклад, вимірюючи такий конструкт, як переваги використання комп'ютера, в анкетному бланку можна представити заданий набір переваг, із якого респонденти обиратимуть одну чи декілька опцій. При цьому загальне число опцій може бути використано як узагальнена величина переваг. Багато видів даних (наприклад, стенограма інтерв'ю) не можуть бути конвертовані в числовий формат для статистичного аналізу. Кодування є особливо важливим для великих складних досліджень, що включають багато змінних і показників, а процес кодування виконується різними людьми.

Кодування допомагає дотриматися узгодженості, а також уможливить розуміння та інтерпретацію даних іншими людьми.

Введення даних. Закодовані дані можуть бути введені в таблиці, текстовий файл чи безпосередньо в статистичну програму на зразок SPSS. Більшість статистичних програм мають редактор даних для їх введення. Однак дані зберігаються у форматі, властивому саме для даної програми, що не дозволяє їх безпосереднє використання в інших статистичних програмах. Отже, краще вводити дані в таблиці чи бази даних, де вони можуть бути за потреби реорганізовані, використані в різних програмах, а їхні фрагменти можуть бути взяті для окремого аналізу. Менші масиви даних (до 65 тис. спостережень і до 256 показників) можуть зберігатися в Microsoft Excel, тоді як більші масиви вимагають створення баз даних. Зазвичай окремі спостереження вводяться в рядки, а показники – в стовпчики. Уведені дані треба перевіряти, щоб не допустити випадкових помилок при наборі.

Пропущені дані. У емпіричних дослідженнях фактично неможливо уникнути пропущених даних. Респонденти опускають деякі питання, якщо не впевнені, що правильно їх зрозуміли, або якщо питання є дуже делікатними. Такі проблеми повинні бути виявлені та виправлені до початку основного дослідження. Під час введення даних деякі статистичні програми автоматично вважають пропущеними даними незаповнені клітинки, тоді як інші вимагають введення певних чисел (наприклад, -1 або 999) замість пропущених даних. Під час аналізу даних у більшості програм ігнорується все спостереження, яке містить пропущену клітинку. Таке видалення може значно зменшити розмір досліджуваної вибірки, та унеможливити виявлення другорядних ефектів. Деякі програми передбачають опцію заміни пропущених даних певними числами завдяки процесу заміни значень. Наприклад, якщо пропущеним є одне спостереження в шкалі з багатьма показниками, тоді замінене значення може бути середнім із відповідей респондента на інші питання. Якщо пропущеним є значення одновимірної шкали, тоді для його заміни вводиться середнє значень по групі респондентів. Така заміна може бути хибною, якщо пропущене значення має систематичну, а не випадкову природу. Два методи, що можуть дати відносно правильні оцінки при заміні – це процедура максимальної вірогідності та методи множинної заміни, які підтримуються в таких програмах як SPSS і SAS.

Трансформація даних. Іноді виникає потреба трансформувати значення даних до того, як вони будуть змістовно інтерпретуватися. Наприклад, застосовується зворотнє кодування, коли трансформуються показники, що мають протилежне по відношенню до певного конструкту значення (при застосуванні інтервальної шкали від 1 до 7 потрібно від 8 відняти значення, що спостерігалось), щоб трансформовані значення можна було об'єднати зі значеннями, що вимірюються за прямими шкалами. Іншими видами трансформацій може бути створення

показників додаванням індивідуальних значень, отриманих за певними шкалами, або створення зваженого індексу з набору величин, що спостерігалися, і об'єднання численних показників у декілька категорій (наприклад, віднесення значень індивідуальних доходів до декількох виділених груп).

Одноваріантний аналіз

Одноваріантний аналіз, або аналіз однієї величини, стосується набору статистичних методів, які можуть описати загальні властивості цієї величини. Одноваріантний аналіз включає такі методи, як: 1) розподіл частот; 2) центральна тенденція; 3) дисперсія. **Розподіл частот** змінної – це сума частот (чи відсотки) індивідуальних значень змінної. Наприклад, розглянемо гіпотетичні дані щодо числа конфліктів в організації, у яких респонденти брали участь (як показник їх конфліктності), використовуючи категоріальну шкалу з 7 опціями (рис. 14.1.). Якщо порахувати число (чи відсоток) спостережень у кожній категорії (за винятком опції «не відповіли», яка є скоріше пропущеними даними, ніж категорією) і представити їх у формі таблиці, то отримаємо частотний розподіл. Цей розподіл також може бути проілюстрований у формі стовпчикової діаграми (див. рис. 14.1.), на якій на горизонтальній осі представлені опції категорії, а на вертикальній – їх частоти.

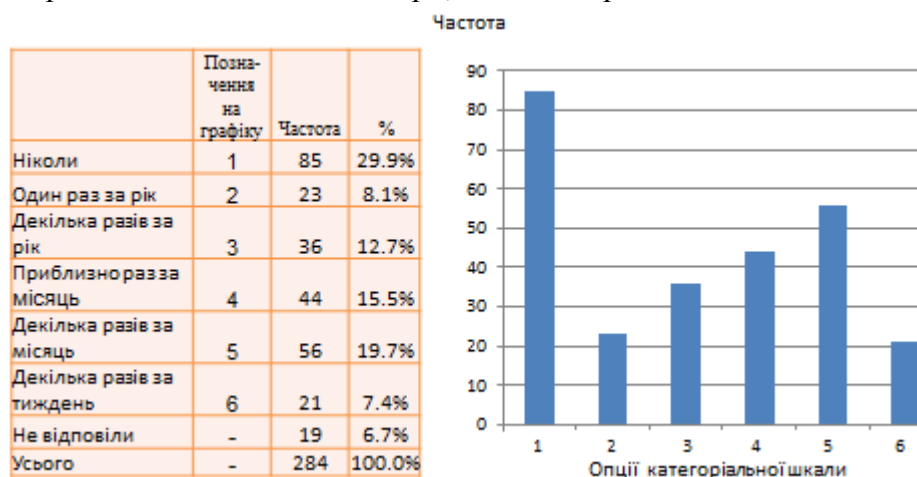


Рис. 14.1. Гіпотетичний частотний розподіл конфліктності працівників

У випадку дуже великих вибірок, де спостереження є незалежними та випадковими, розподіл частот наближається до кривої, що має форму перевернутого дзвіночка, схожого до того, що представлений на рис. 14.2. При цьому більшість спостережень розташовані навколо центру розподілу, і з віддаленням від центру кількість спостережень стає все меншою. Така крива називається *нормальним розподілом*.

Центральна тенденція – це оцінка центру розподілу змінної. Існує три оцінки центральної тенденції: середнє, медіана та мода. Середнє арифметичне – це середнє всіх значень у даному розподілі. Інші типи

середніх включають середнє геометричне та середнє гармонійне, однак, ці середні не дуже поширені в статистичному аналізі даних соціального дослідження. Іншою мірою центральної тенденції є **медіана**, або значення, що знаходиться в середині розподілу. Вона знаходиться, якщо всі дані розподілу розмістити в зростаючому порядку й обрати те значення, що розташоване посередині. При парному числі значень у розподілі медіана розраховується як середнє двох значень, що розташовані в середині розподілу. **Мода** – це значення, що найчастіше зустрічається в даному розподілі.

Усі оціночні показники, що розраховуються на основі даних вибірки (середнє, медіана, мода чи інші) називаються **статистичними показниками**.

Дисперсія характеризує, як індивідуальні значення розташовані навколо центральної тенденції. Два поширених показника дисперсії – це розмах і стандартне відхилення. **Розмах** – це різниця між максимальним і мінімальним значеннями в розподілі. Діапазон є дуже чутливим до випадкових значень (артефактів), на відміну від стандартного відхилення, другого показника дисперсії. Стандартне відхилення бере до уваги, наскільки близько чи далеко кожне індивідуальне значення розташоване від середньої:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}},$$

де σ – стандартне відхилення, x_i – i -те спостереження (значення), μ – середнє арифметичне, n – загальне число спостережень. Корінь квадратний стандартного відхилення називається варіацією розподілу. У нормальному розподілі частот 68% спостережень знаходяться в межах одного стандартного відхилення від середнього ($\mu \pm 1 \sigma$), 95% значень – у межах двох стандартних відхилень ($\mu \pm 2 \sigma$) і 99,7% значень – у межах трьох стандартних відхилень ($\mu \pm 3 \sigma$), як показано на рис 14.2.

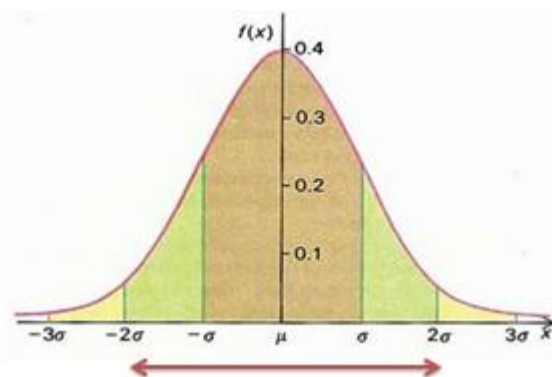


Рис. 14.2. Нормальний розподіл

Двоваріантний аналіз

Двоваріантний аналіз показує, як дві величини пов'язані між собою. Найбільш поширеним показником у двоваріантному аналізі є коефіцієнт кореляції, який являє собою число від -1 до $+1$ й показує силу зв'язку між двома змінними. Уявимо, що потрібно встановити, як вік пов'язаний із самооцінкою людини. Якщо з віком самооцінка збільшується, то це приклад позитивної кореляції, зменшується – негативної, залишається незмінною – дорівнює нулю. Щоб розрахувати значення кореляції, розглянемо гіпотетичний набір даних, представлений в таблиці 14.1.

№ спост	Вік (x)	Самооцінка (y)	xy	x ²	y ²
1	39	4.1	159.9	1521	16.81
2	45	4.6	207	2025	21.16
3	29	3.8	110.2	841	14.44
4	42	4.4	184.8	1764	19.36
5	19	3.2	60.8	361	10.24
6	22	3.1	68.2	484	9.61
7	39	3.8	148.2	1521	14.44
8	30	4.1	123	900	16.81
9	33	4.3	141.9	1089	18.49
10	23	3.7	85.1	529	13.69
11	20	3.5	70	400	12.25
12	18	3.2	57.6	324	10.24
13	24	3.7	88.8	576	13.69
14	22	3.3	72.6	484	10.89
15	29	3.4	98.6	841	11.56
16	35	4.0	140	1225	16.00
17	36	4.1	147.6	1296	16.81
18	37	3.8	140.6	1369	14.44
19	35	3.4	119	1225	11.56
20	32	3.6	115.2	1024	12.96
Сума (Σ)	609	75.1	2339.1	19799.0	285.45

Таблиця 14.1. Гіпотетичні дані «вік» та «самооцінка»

Дві змінні в наборі даних – це вік (x) та самооцінка (y). Вік оцінюється за шкалою відношень, тоді як самооцінка – це середній бал, розрахований за анкетною, у якому набір показників оцінювався за 7-бальною шкалою Лікерта з опціями від «повністю не згодний» до «повністю згодний». Гістограма кожного показника показана з лівого боку рис. 14.3. Формула для розрахунку коефіцієнта кореляції між двома змінними має вигляд:

$$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}},$$

де r_{xy} – коефіцієнт кореляції, \bar{x} і \bar{y} – середні, s_x і s_y – стандартні відхилення змінних x і y . Розрахований вручну коефіцієнт кореляції становить 0,79. Рисунок показує, що вік має значну позитивну кореляцію з самооцінкою. Такий тренд можна побачити і візуально, порівнявши гістограми цих показників на рис. 14.3. Зверніть увагу, що на ординаті представлені значення змінних, що спостерігалися, а не частоти розподілів (як на рис. 14.1), тому це є скоріше гістограми. Графік з правого боку рис. 14.3., на якому представлені значення віку (по абсцисі) та відповідні їм значення

самооцінки (по ординаті), називається скатерограмою. Оскільки кореляція між досліджуваними змінними позитивна, при збільшенні віку самооцінка збільшується (позитивний тренд). Якби при збільшенні віку самооцінка зменшувалася, то скатерограма мала б негативний тренд. При відсутності кореляції скатерограма була б горизонтальною.

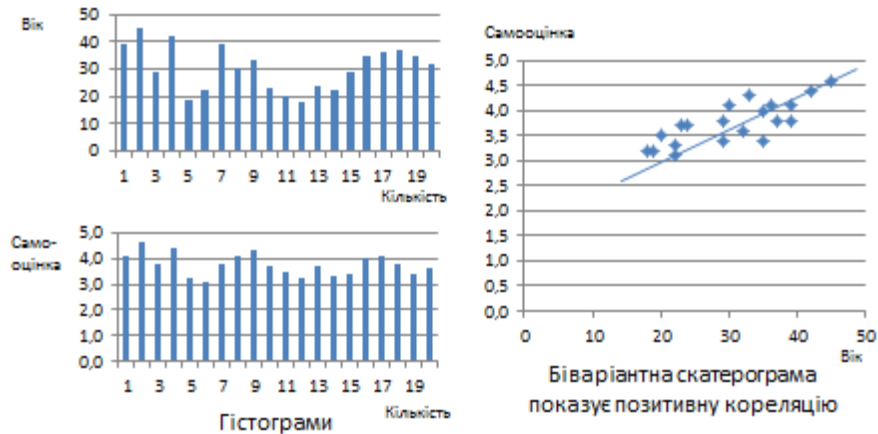


Рис. 14.3. Гістограми та скатерограма віку та самооцінки

Після розрахунку парної кореляції дослідники зазвичай хочуть знати, чи є вона істотною, тобто закономірною, чи зумовлена випадковими причинами. Щоб відповісти на це питання, необхідно протестувати такі гіпотези:

$$H_0: r = 0;$$

$$H_1: r \neq 0.$$

H_0 називається нульовою гіпотезою, H_1 – робочою, або альтернативною гіпотезою. Хоча ці гіпотези виглядають як різні, фактично вони є крайніми випадками однієї гіпотези. Також варто відзначити, що H_1 не є спрямованою гіпотезою, оскільки вона не вказує більшим чи меншим 0 є коефіцієнт кореляції. Спрямовані гіпотези будуть мати вигляд $H_0: r \leq 0$; $H_1: r > 0$ (якщо тестується наявність позитивної кореляції). Істотність тестування спрямованої гіпотези проводиться за допомогою однохвостового t-тесту, тоді як при неспрямованих гіпотезах використовується двоххвостовий t-тест.

У статистичному аналізі робоча гіпотеза не може тестуватися безпосередньо, скоріше вона тестується відхиленням нульової гіпотези з певним рівнем вірогідності. Статистичне тестування є завжди вірогідним, оскільки ніколи не можна з упевненістю сказати, чи отримані для певної вибірки висновки будуть правильними для генеральної сукупності. Вірогідність того, що статистичний висновок обумовлений випадковістю, називається **p-величиною**. Її значення порівнюється зі **статистичним рівнем (α)**, який представляє максимальний рівень ризику неправильності висновку, який дослідник готовий прийняти. Для більшості видів статистичного аналізу α встановлюється на рівні 0,05. Значення p , менше за $\alpha=0.05$, означає, що існує достатньо статистичних доказів для

відхилення нульової гіпотези, а отже, прийняття робочої гіпотези. Якщо ж $p > 0.05$, у дослідника немає достатньо статистичних доказів для відхилення нульової гіпотези та прийняття робочої.

Найпростіший спосіб перевірки вказаної гіпотези полягає в тому, щоб узяти критичні значення r із статистичних таблиць, які наводяться в підручниках зі статистики або в Інтернеті. Більшість статистичних програм також виконують аналіз істотності. Критичне значення r залежить від обраного рівня істотності (α), ступеня свободи (df) і того, чи є обраний тест одно- чи двоххвостовим. **Ступінь свободи** – це число значень, що можуть вільно змінюватися при будь-яких статистичних розрахунках. У випадку кореляції df дорівнює $n - 2$. Для даних таблиці 14.1. $df = 20 - 2 = 18$. Існують різні статистичні таблиці для однохвостового та двоххвостового тесту. У таблицях двоххвостових тестів критичне значення r для $\alpha = 0.05$ і $df = 18$ дорівнює 0.44. Оскільки розраховане значення $r = 0,79$ перевищує табличне (0,44), кореляція між досліджуваними показниками є істотною. Вірогідність того, що кореляція обумовлена випадковим фактором, становить менше 5%. Отже, можна відхилити нульову гіпотезу, що опосередковано дозволяє зробити висновок про правильність робочої гіпотези ($r > 0$).

У багатьох дослідженнях аналізується більш ніж, дві змінні. Якщо є n змінних, тоді число можливих кореляцій між ними становить:

$$n*(n-1)/2.$$

Ці кореляції легко розраховуються статистичними програмами і представляються в кореляційній матриці, як показано в таблиці 14.2. Оскільки кореляції не є спрямованими, кореляція між змінними V_1 і V_2 є такою ж, як і кореляція між V_2 і V_1 . Значення, розташовані на головній діагоналі, завжди дорівнюють одиниці. Далі, в кореляційній матриці нижній трикутник (значення, розташовані нижче головної діагоналі) є дзеркальним відображенням верхнього трикутника (значення, розташовані вище головної діагоналі), отже для спрощення достатньо представити лише один трикутник матриці. Якщо кореляції включають змінні, що вимірюються інтервальними шкалами, цей окремий вид кореляції називається **кореляцією Пірсона**.

Інший корисний шлях представити двоваріантні дані полягає в побудові крос-таблиць. **Крос-таблиця** – це таблиця, що описує частоту (чи відсоток) усіх комбінацій двох чи більше номінальних чи категоріальних даних. Припустимо, що потрібно встановити, які з трьох видів печива (А, В і С) більше подобаються чоловікам, а які – жінкам. Проста крос-таблиця даних показує спільний розподіл статі та виду печива, тобто скільки осіб кожної статі надали перевагу певному виду печива, і має 2x3 клітинок. Таблиця дозволяє побачити, наскільки рівномірно розподілилися вподобання досліджуваних щодо вибору виду

печива А, В і С між особами чоловічої та жіночої статі. Дані крос-таблиці (таблиця 14.3.) показують, що розподіл уподобань за видом печива А зміщений у бік жінок. У вибірці з 10 чоловіків та 10 жінок 5 жінок віддали перевагу печиву А порівняно з лише одним чоловіком. Навпаки, розподіл по виду печива С зміщений у бік осіб чоловічої статі: його вподобали три чоловіки і лише одна жінка. Розподіл по печиву В більш рівномірний. Останній рядок та останній стовпчик таблиці вказують суми значень за кожною категорією і називаються **маргінальними** сумами.

Таблиця 14.3.

Приклад крос-таблиці

Слост	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Стать	F	M	F	M	F	M	M	M	F	F	M	M	M	F	F	M	F	F	F	M
Бал	A	B	B	B	C	A	C	B	B	A	B	C	C	B	B	B	A	A	B	B

Гіпотетичний набір даних

		Бал атестату			Усього	
			A	B	C	
Стать	Чоловіки	Кількість	1	6	3	10
		Очікувана кількість	2.5	5.5	2.0	
	Жінки	Кількість	4	5	1	10
		Очікувана кількість	2.5	5.5	2.0	
Усього			5	11	4	20

Маргінальні суми

Хоча існує помітний тренд у розподілі вподобань за статтю досліджуваних осіб, його істотність вимагає перевірки. Чи можна з упевненістю стверджувати, що наявні частоти не є випадковими? Для відповіді на це питання застосовується статистичний тест Хі-квадрат. Критерій Хі-квадрат розраховується як середнє різниць між тими частотами, що спостерігалися в дослідженні, і тими частотами, що відповідають випадковому розподілу спостережень. Останні розраховуються для кожної клітини як добуток маргінальної суми стовпчика на маргінальну суму рядка, поділений на загальне число спостережень. Порівняння критерію Хі-квадрат із критичним значенням, узятим зі стандартної таблиці для певного рівня істотності ($p < 0,05$) і ступеня свободи df , яке розраховується за формулою: $df = (m - 1) * (n - 1)$, де m і n – кількість рядків та стовпчиків, дозволяє зробити висновок щодо істотності результатів. У даному випадку $df = (2 - 1) * (3 - 1) = 2$ і критичний рівень становить 5,99. Розраховане значення Хі-квадрат дорівнює 1, що менше критичного рівня 5,99. Отже, відмінності в розподілі осіб чоловічої та жіночої статі за вподобаннями трьох видів печива не є статистично істотними.

Розділ 15

КІЛЬКІСНИЙ АНАЛІЗ: ПОЯСНЮВАЛЬНА СТАТИСТИКА

Пояснювальна статистика – це статистичні процедури, які використовуються, щоб зробити висновки стосовно зв'язків між змінними. Ця статистика відрізняється від описової статистики тим, що її процедури спеціально розроблені для тестування гіпотез. У цю категорію потрапляють численні статистичні методи, які підтримуються сучасним програмним забезпеченням, таким як SPSS і SAS. Цей розділ надає короткий огляд лише найбільш поширених і відносно нескладних процедур.

Базові поняття

Британський філософ Карл Поппер сказав, що теорії ніколи не можуть бути доведені, однак вони можуть бути спростовані. Як приклад, чи можна довести, що сонце завтра зійде на небі? Поппер каже, якщо сонце сходить кожний день, це ще не означає, що вони зійде завтра, оскільки теорії, що створилися індуктивно, є тільки здогадками, які можуть чи не можуть прогнозувати майбутній феномен. Замість цього можна припустити правильність того, що сонце сходить кожний день, без доказу, але якщо сонце не зійде хоча б один день, теорія має бути відхилена. Аналогічно, ми можемо відхилити гіпотезу, спираючись на альтернативний доказ, але ми ніколи не можемо повністю її прийняти, оскільки наявність доказів не означає, що в майбутньому не з'являться аргументи, що її спростовують. Оскільки ми не можемо повністю прийняти робочу (альтернативну) гіпотезу, формулюється нульова гіпотеза, яка суперечить робочій, і потім використовуються емпіричні докази для відхилення нульової гіпотези, щоб продемонструвати опосередковану вірогіднісну підтримку робочої гіпотези.

Інша проблема з тестуванням гіпотетичних зв'язків у дослідженнях соціальних наук полягає в тому, що залежна змінна може знаходитися під впливом численних побічних факторів, які неможливо виміряти та контролювати. Отже, навіть якщо дві змінні начебто мають зв'язок за даними досліджуваної вибірки, насправді вони не обов'язково будуть пов'язані між собою в генеральній сукупності. Із цих міркувань пояснювальна статистика не є детерміністською, а завжди носить ймовірнісний характер.

Як можна встановити істотність залежності між двома змінними? Рональд Фішер, один із найбільш видатних статистиків в історії людства, визначив базові принципи тестування істотності. Він установив, що статистичний результат можна вважати істотним, якщо буде доведено, що вірогідність відхилення цього результату як випадкового становить 5% чи

менше. У пояснювальній статистиці вірогідність називається p , і позначається як $p \leq 0,05$. **Рівень істотності** – це максимальний рівень ризику, який дослідник готовий прийняти, базуючи свої висновки на вибірці з генеральної сукупності. Якщо $p \leq 0,05$, тобто менше 5%, це означає, що існує 5% вірогідність помилки відхилення нульової гіпотези, коли насправді вона вірна. Така помилка називається помилкою типу 1. Якщо $p > 0,05$, це означає, що у дослідника недостатньо аргументів для відхилення нульової гіпотези чи прийняття робочої гіпотези.

Існує три пов'язаних статистичних поняття: вибірковий розподіл, стандартна помилка та довірчий інтервал. **Вибірковий розподіл** – це теоретичний розподіл нескінченної кількості вибірок із певної генеральної сукупності. Однак, оскільки вибірка ніколи не співпадає з генеральною сукупністю, кожна вибірка має певний рівень помилки, який називається **стандартною помилкою**. Якщо стандартна помилка невелика, тоді статистичні оцінки, розраховані для вибірки (наприклад, середнє), є достатньо добрими оцінками генеральної сукупності. Точність оцінок вибірки визначається в термінах **довірчого інтервалу**. Наприклад, 95% довірчий інтервал визначається як діапазон від мінус до плюс два стандартних відхилення від середньої вибірки. Отже, якщо дослідник установлює вибіркові оцінки з довірчим інтервалом в 95%, то це означає, що параметр генеральної сукупності в 95% випадках знаходиться в межах двох стандартних відхилень від статистичного показника вибірки. Разом рівень p та довірчий інтервал дають добре уявлення щодо вірогідності результату дослідження й того, наскільки цей результат відображає параметр генеральної сукупності.

Загальна лінійна модель

у дослідженнях соціальних наук більшість процедур пояснювальної статистики мають своїм джерелом загальну лінійну модель (ЗЛМ). Модель – це оціночне математичне рівняння, яке можна використовувати для презентації набору даних, а слово «лінійна» стосується прямої лінії. Отже, ЗЛМ – це система рівнянь для відтворення лінійних трендів зв'язку в досліджуваних даних.

Найпростіший тим ЗЛМ – це двоваріантна лінійна модель, яка досліджує зв'язок між однією незалежною змінною (причина чи фактор) і однією залежною змінною (ефект чи результат). Припустимо, що ці дві змінні – вік та самооцінка відповідно. Біваріантна скатерограма їх зв'язку наведена на рис. 15.1. Як видно, окремі спостереження представляють комбінації віку та самооцінки, які в цілому розташовані навколо уявної прямої під гострим кутом до осі X .

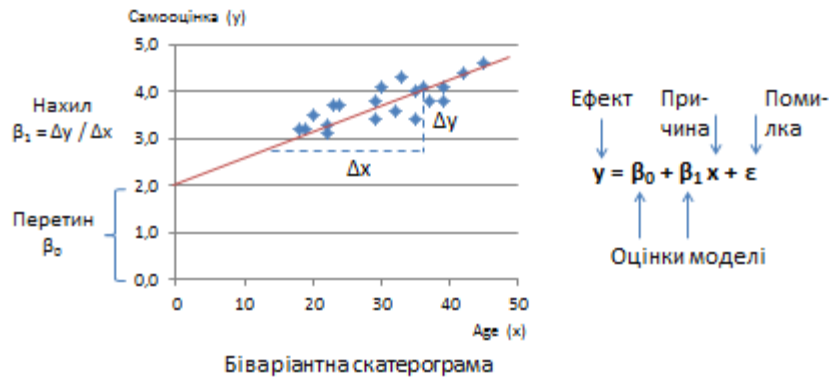


Рис. 15.1. Двоваріантна лінійна модель

Можна оцінити параметри цієї лінії, наприклад, кут нахилу і перетин з віссю Y та представити рівняння цієї прямої у вигляді:

$$y = mx + c ,$$

де m – нахил прямої, який показує наскільки змінюється y , якщо x змінюється на одиницю. У ЗЛМ це рівняння формально записується так:

$$y = \beta_0 + \beta_1x + \varepsilon ,$$

де β_0 – нахил, β_1 – перетин, ε – похибка, яка показує відхилення фактичних спостережень від їхніх оцінюваних значень, оскільки більшість спостережень, хоча й знаходяться близько до прямої, однак не знаходяться на ній безпосередньо (тобто ЗЛМ не є ідеальною). Лінійна модель може мати більш, ніж два фактори. Для візуалізації такої моделі уявімо куб, де результат (y) нанесений уздовж вертикальної осі, а два параметри, наприклад, x_1 та x_2 – уздовж двох горизонтальних осей, що утворюють основу кубу. Така лінія, що описує зв'язок між більш, ніж двома змінними, називається лінією регресії, β_0 , β_1 та інші бета-коефіцієнти називають коефіцієнтами регресії, а процес оцінки коефіцієнтів регресії називається регресійним аналізом. ЗЛМ для регресійного аналізу з n факторами має вигляд:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \dots + \beta_nx_n + \varepsilon .$$

У цьому рівнянні змінні x можуть представляти незалежні змінні чи коваріанти (контрольні змінні). Коваріанти – це змінні, які не становлять теоретичного інтересу, однак можуть впливати на залежну змінну y , і тому мають контролюватися, щоб залишкові ефекти незалежних змінних, що є предметом дослідження, могли бути визначені точніше. Коваріанти дають систематичні помилки в регресійному рівнянні, тоді як помилка (ε) відображає випадкові впливи. Хоча більшість змінних в ЗЛМ є інтервальними чи такими, що вимірюються шкалами відношень, насправді це не зовсім так. Деякі фактори можуть бути номінальними й прийматимуть лише одне з двох значень (0 чи 1). Їх називають *dummy*-змінними. Набір таких змінних представляється $n-1$

dummy-змінними. Наприклад, набір номінальних змінних галузей економіки, що складається з промисловості, сільського господарства і послуг, може бути представлений комбінацією двох dummy-змінних x_1 та x_2 : (0,0) – промисловість, (0,1) – сільське господарство, (1,1) – послуги. При цьому не має значення, який рівень номінальної шкали кодується як 0, чи 1, оскільки 0 та 1 розглядаються як дві різні номінальні групи.

ЗЛМ є потужним статистичним інструментом, оскільки це не один метод, а скоріше група методів, які можуть використовуватися для проведення аналізу з різними типами й кількістю факторів та результуючих змінних. При наявності dummy-змінної як фактора й необхідності порівняння ефектів двох рівнів (0 чи 1) цих змінних на результуючу змінну, виконується *аналіз варіацій* (ANOVA). Якщо виконується ANOVA при контролі ефектів однієї чи декількох змінних-коваріант, робиться *аналіз коваріант* (ANCOVA). У дослідженні може бути декілька результуючих змінних (тобто y_1, y_2, \dots, y_n), які представляються системою рівнянь, що складається з різних рівнянь для кожної змінної (кожне зі своїм унікальним набором факторів). Якщо декілька результуючих змінних моделюються на основі одного й того ж набору факторів, такий аналіз називається *мультиваріантною регресією*. Якщо виконується ANOVA чи ANCOVA з багатьма результуючими змінними, то аналіз називається *мультиваріантним ANOVA* (MANOVA) чи *мультиваріантним ANCOVA* чи MANCOVA відповідно. Якщо результат, що моделюється в одному рівнянні регресії, є фактором в іншому рівнянні у взаємопов'язаній системі рівнянь регресії, тоді це випадок аналізу, що називається *моделюванням структурними рівняннями*. Найбільш важлива проблема в ЗЛМ полягає в *специфікації моделі*, тобто в тому, як визначити рівняння регресії (чи систему рівнянь) для кращої презентації досліджуваного феномену. Специфікація моделі має базуватися на теоретичних аргументах щодо феномену, а не на найкращій відповідності даним. Роль даних полягає у валідизації моделі, а не у визначенні її типу.

Двогрупове порівняння

Один з найпростіших аналізів пояснювальної статистики – це порівняння даних досліджуваної та контрольної груп у дизайні «контрольна група після тесту», наприклад при порівнянні успіхів у математиці студентів, які пройшли спеціальну навчальну програму, і які навчалися за традиційною програмою. В цьому випадку фактор – це dummy-змінна (1 – досліджувана група, 0 – контрольна група), результуюча змінна, успішність, вимірюється за шкалою відношень. Аналітичний метод для цього простого дизайну – це *одноваріантний ANOVA* (оскільки є лише один фактор) і статистичний тест, що використовується для оцінки відмінностей між групами, – це *t-тест Стьюдента*. Цей тест, запропонований Вільямом Госсетом, включає розрахунок значення *t-критерію* і статистично оцінює різницю між

середніми двох груп (непрямий, або двоххвостовий тест), або чи є середнє однієї групи статистично меншим (більшим) порівняно з середнім іншої групи (прямий, або однохвостовий тест). У нашому прикладі застосовується прямий однохвостовий t-тест. Гіпотеза формулюється таким чином:

$$H_0: \mu_1 \leq \mu_2;$$

$$H_1: \mu_1 > \mu_2,$$

де μ_1 – середнє академічних оцінок із математики студентів, що пройшли спеціальну програму підготовки (досліджувана група), μ_2 – середнє академічних оцінок із математики студентів, що навчалися за традиційною програмою (контрольна група). Зверніть увагу, нульова гіпотеза завжди містить знак рівності, а мета всіх статистичних тестів полягає у відхиленні нульової гіпотези.

Як можна пояснити різницю в середніх генеральної сукупності на підставі даних вибірок? Як видно з гіпотетичних розподілів частот досліджуваної та контрольної груп, наведених на рис. 15.2., контрольна група має розподіл, близький до нормального з середнім значенням 45 (за 100 бальною шкалою), тоді як досліджувана група має середнє значення 65. Середні виглядають як відмінні, але це середні вибірок, які відрізняються від середніх генеральних сукупностей через помилку вибірки. Середня вибірка є ймовірнісною оцінкою середньої генеральної сукупності в межах певного довірчого інтервалу.

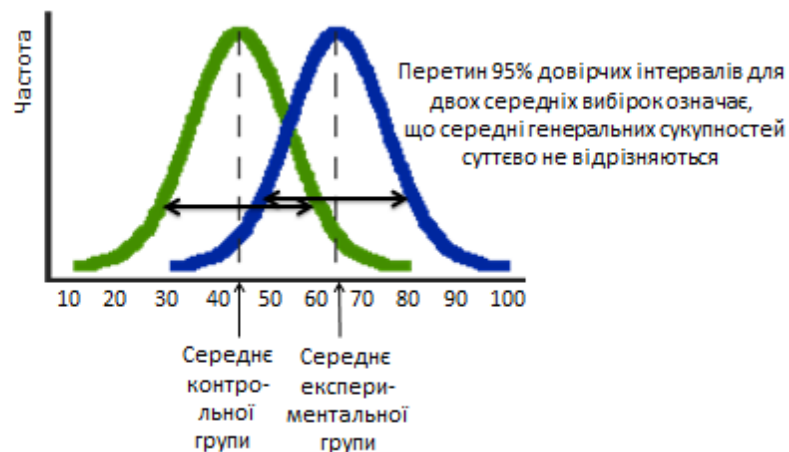


Рис. 15.2. Ілюстрація t-тесту Стьюдента

Статистична істотність середньої генеральної сукупності залежить не тільки від середніх порівнюваних вибірок, але й від стандартної помилки, або ступеня концентрації значень частот розподілу навколо середньої. Якщо концентрація велика, тоді криві розподілів будуть менше перетинатися між собою, ніж у випадку «розлогих» розподілів. При цьому довірчі інтервали розподілів теж менше перетинаються, а отже, можна сподіватися, що середні двох генеральних сукупностей мають

відмінності на рівні $p < 0,05$. Якщо концентрація індивідуальних значень розподілу навколо середньої невелика, тоді 95% довірчі інтервали середніх перетинаються, і не можна з впевненістю стверджувати, що середні генеральних сукупностей відрізняються між собою ($p > 0,05$).

Для виконання t-тесту потрібно розрахувати t-критерій, який показує відмінності середніх досліджуваних вибірок. Цей критерій є відношенням різниці середніх до різниці в їх варіативності (стандартної помилки):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}},$$

де \bar{X}_1 та \bar{X}_2 – середні досліджуваної та контрольної груп, а $s_{\bar{X}_1 - \bar{X}_2}$ – стандартна помилка різниць між двома групами, яка може бути оцінена за формулою:

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

де S_1 і S_2 – варіації груп, n – розмір кожної групи. Далі потрібно встановити, чи достатньо великим є значення t-критерію для того, щоб уважати відмінності не випадковими. Для цього потрібно за статистичною таблицею (або в Інтернеті) подивитися на вірогідність p , яка пов'язана з розрахованим t-критерієм. Значення p є функцією t-критерію для одно- чи двоххвостового тесту та ступеня свободи (df) – числа спостережень у вибірці, які можуть вільно варіювати при розрахунку критерію, і дорівнює в даному випадку $(n_1 + n_2 - 2)$. Якщо значення p менше бажаного рівня істотності (як правило, обирається значення $\alpha = 0,05$), тоді відхиляється нульова гіпотеза і приймається робоча. Це означає, що відмінності між середніми вибірок є істотними.

Наступним кроком є встановлення *розміру ефекту*, який мала дія досліджуваного фактору (у даному випадку – навчання за спеціальною програмою) відносно контрольної групи. Розмір ефекту можна визначити регресійним аналізом, у якому бали успішності є результуючими змінними (y), а *dummy*-змінна наявності чи відсутності програми є фактором (x) у двоваріантній ЗЛМ. Коефіцієнт регресії при x (β_1), який є кутом нахилу регресійної прямої ($\beta_1 = \Delta y / \Delta x$), є оцінкою розміру ефекту. У даному прикладі, оскільки x є *dummy*-змінною з двома значеннями (0 і 1), $\Delta x = 1 - 0 = 1$, а розмір ефекту є різницею між середніми досліджуваної та контрольної груп ($\Delta y = y_1 - y_2$).

Факторіальні дизайни

Продовжуючи розглянутий приклад, уявимо, що спеціальна і традиційна програми мали по два різновиди – 3 чи 6 годин занять на

тиждень. Отже, маємо факторіальний дизайн 2x2 з двома факторами – тип програми й тривалість заняття. Такий дизайн допоможе оцінити незалежний ефект кожного фактора, який називається *головним ефектом*, і також оцінити спільний ефект двох факторів, який називається *ефектом взаємодії*. ЗЛМ для даного дизайну має вигляд:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon,$$

де y – бали успішності учнів після навчання, x_1 – вид програми (спеціальна чи традиційна), x_2 – тривалість заняття (3 чи 6 год./тижд.). При цьому x_1 та x_2 є *dummy*-змінними, і x_2 має лише дві категорії (3 чи 6 год./тижд.), хоча і виглядає як змінна, що вимірюється за шкалою відношень. Регресійні коефіцієнти β_1 та β_2 дають оцінки розміру ефекту для головних ефектів, а β_3 – для ефекту взаємодії. Інший варіант аналізу – двоваріантний ANOVA. Регресійний аналіз, що включає декілька факторів, іноді називається множинною регресією, яка відрізняється від багатоваріантної регресії, що використовує декілька результуючих змінних.

Якщо коефіцієнт β_3 є істотним, це означає, що вплив виду програми залежить від тривалості занять. У цьому випадку не можна змістовно інтерпретувати головні ефекти програми (β_1), або тривалості занять (β_2), оскільки ці ефекти не можна відділити один від одного. Головні ефекти можна інтерпретувати, тільки якщо ефект взаємодії не є статистично істотним.

Коваріанти можуть бути включені у факторіальний дизайн як нові змінні з новими коефіцієнтами регресії (β_4 і т.п.). Вони можуть вимірюватися за інтервальною шкалою чи за шкалою відношень, навіть якщо досліджувані фактори є *dummy*-змінними. Інтерпретація коваріант проводиться за тими ж правилами, що й факторів.

Інші види кількісного аналізу

Існує багато корисних статистичних методів аналізу даних, що базуються на ЗЛМ. Відзначимо декілька з них:

Факторний аналіз – це техніка стискання даних для статистичної агрегації великої кількості досліджуваних показників у невелику групу змінних (факторів), які безпосередньо не спостерігалися в дослідженні, базуючись на їх кореляційних зв'язках. Цей метод широко використовується для оцінки конвергентної та дискримінантної валідності в шкалах з багатьма показниками.

Дискримінантний аналіз – це метод класифікації, який має на меті помістити певне спостереження в одну з декількох номінальних категорій, базуючись на лінійній комбінації незалежних змінних. Метод схожий до методу множинної регресії, тільки залежна величина є

номінальною. Метод поширений у дослідженнях з маркетингу, наприклад, при класифікації споживачів чи товарів за категоріями на основі характерних ознак, виявлених в опитуваннях із використанням багатьох шкал.

Логістична регресія – це ЗЛМ, у якій результуюча змінна є бінарною величиною (0 чи 1) і, вірогідно, має логістичний розподіл. Мета регресії полягає у прогнозі вірогідності успішного результату за відповідністю даних логістичній кривій.

Пробіт-регресія – це ЗЛМ, у якій результуюча змінна може варіювати від 0 до 1 (або можна припустити дискретні значення від 0 до 1) і відповідає стандартному нормальному розподілу. Мета регресії полягає в прогнозуванні вірогідності кожного результату. Цей метод поширений у фінансовій сфері, страхуванні, кредитуванні (наприклад, кредитна оцінка клієнта на основі його кредитної історії, доходів, розміру кредиту і т.п.). Пробіт- та логістична регресії мають схожі коефіцієнти регресії у порівнюваних ситуаціях (бінарні результати), однак логістичну регресію легше побудувати та інтерпретувати.

Аналіз шляху – це методика багатоваріантної ЗЛМ для аналізу спрямованих зв'язків у наборі змінних. Метод дозволяє вивчення складних номологічних моделей, де залежна змінна в одному рівнянні є незалежною змінною в іншому рівнянні. Метод широко використовується в сучасних дослідженнях соціальних наук.

Аналіз часових рядів – це аналіз даних часових рядів чи змінних, що постійно змінюються з часом. Прикладом є прогноз флуктуацій на біржі. Метод популярний в економетриці, фінансуванні тощо.

Розділ 16

ЕТИКА НАУКОВИХ ДОСЛІДЖЕНЬ

Словник Вебстера визначає етику як дотримання стандартів діяльності певною професійною чи іншою групою. Такі стандарти часто визначаються на організаційному рівні як професійний кодекс етики. Навіть у разі відсутності формальних документів науковці повинні усвідомлювати та дотримуватися загальних положень щодо професійної етики виконання наукових досліджень. Наприклад, науковці не повинні маніпулювати даними досліджень, результатами аналізу чи інтерпретації в такий спосіб, що суперечить принципам проведення досліджень чи науковій методології. Чому етика дослідження є такою важливою? Тому що наукові результати досить часто використовувалися людьми чи організаціями для досягнення власних цілей усупереч загальноприйнятим нормам і стандартам проведення наукових досліджень. Класичним прикладом є справа фармацевтичного гіганту «Меркс», який приховав фатальні побічні ефекти нового медичного препарату від наукової спільноти, що призвело до смерті 3468 пацієнтів клініки «Віокс» від зупинки серця. У 2010 р. компанія погодилася на виплату 4,85 млрд. доларів компенсації і призначила два незалежних комітети для моніторингу безпеки нових препаратів. Поведінка «Меркс» була визнана неетичною і такою, що порушує наукові принципи збору, аналізу та інтерпретації даних.

Етика – це моральна категорія, що диференціює правильне від неправильного. Неетична поведінка не обов'язково означає порушення закону чи прийнятих у суспільстві юридично-правових норм. Однак неетична поведінка може зруйнувати професійну репутацію науковця, стати предметом осуду наукової спільноти та навіть призвести до втрати роботи. Етичні норми дещо відрізняються в різних країнах. У цій книзі розглядаються стандарти, що застосовуються в галузі проведення наукових досліджень у країнах Заходу.

Етичні принципи в наукових дослідженнях

Відзначимо найбільш важливі принципи етичної поведінки, прийняті науковою спільнотою Заходу.

Добровільна участь і безпечність дослідження. Учасники дослідження повинні знати, що їх участь у дослідженні є добровільною, і вони можуть вийти з дослідження на будь-якому етапі без будь-яких наслідків. Участь чи відмова від участі в дослідженні не має нести небезпеки чи завдавати шкоди учасникам досліджень.

Усі учасники повинні письмово підтвердити згоду на участь, у якій чітко описуються їхні права, включаючи право відмови від участі на будь-якому етапі, а також право відкликати результати власного опитування.

Анонімність і конфіденційність. Для захисту прав учасників їхня ідентичність має бути захищена. Це досягається дотриманням принципу анонімності й конфіденційності. **Анонімність** означає, що дослідник і всі, хто може ознайомитися з результатами досліджень, не зможуть ідентифікувати відповіді конкретного учасника. Іноді складно отримати правдиві відповіді від учасників, якщо вони не впевнені в повній анонімності.

У деяких випадках, наприклад, в особистому інтерв'ю, неможливо забезпечити анонімність. У випадку довготривалих досліджень, коли потрібно співставити дані одного і того є учасника в різні періоди часу, анонімність неможлива. У таких випадках учасникам повинна гарантуватися **конфіденційність**. Це означає, що хоча дослідник може ідентифікувати учасника, він обіцяє не розкривати ідентичність учасника в будь-якій формі.

Надання інформації. Зазвичай дослідники мають зобов'язання надати потенційним учасникам загальну інформацію до початку дослідження, щоб останні могли прийняти зважене рішення стосовно своєї участі в дослідженні. Зокрема, потрібно повідомити, хто виконує дослідження, із якою метою, якими є очікувані результати і які наслідки це дослідження може мати. Іноді мета дослідження не може бути повідомлена без ризику впливу на відповіді учасників. У цьому випадку учасників інформують щодо мети дослідження одразу після його завершення.

Аналіз та звітування. Дослідники також мають етичні зобов'язання перед науковою спільнотою стосовно того, як дані аналізуються й оприлюднюються. Негативні, або неочікувані результати повинні надаватися повністю, навіть якщо вони можуть поставити під сумнів дизайн дослідження чи висновки. Аналогічно, після завершення дослідження можуть виявитися неочікувані тенденції та зв'язки. Є неетичним представляти ці результати як заплановані та очікувані. Іншими словами, у позитивістських дослідженнях гіпотези не можуть розроблятися після того, як дослідження завершилося, оскільки роль таких досліджень полягає в тестуванні гіпотез, а не в їх розробленні. Також не є етичним дробити дані на окремі фрагменти, щоб підтвердити чи відхилити досліджувану гіпотезу, або писати статті, що стосуються лише окремої частини даних. Сумнівні та погано обґрунтовані висновки, що базуються на фрагментарних даних або неповному їх аналізі, не можна представляти як валідні. Наука розвивається завдяки відкритості та чесності, і служіння науці вимагає від дослідника повного висвітлення

проблем, що виникли в процесі дослідження, щоб всі науковці могли побачити реальну картину і прийняти заходи щодо їх уникнення.

Професійні коди етики

Більшість професійних асоціацій дослідників мають формальні коди проведення досліджень, які регламентують, що вважати прийнятною чи непринятною професійною поведінкою для членів даної асоціації. Нижче, як приклад, наводяться основні положення кодексу Асоціації інформаційних систем (АІС), що об'єднує дослідників у галузі інформаційних систем. Повна версія кодексу знаходиться за посиланням: <http://home.aisnet.org/displaycommon.cfm?an=1&subarticlenbr=15>). Інші галузі наук у своїй практиці керуються схожими кодексами.

Кодекс АІС поділяє порушення етики на дві категорії. До першої включають серйозні порушення, такі як плагіат чи фальшування даних, процедур, аналізу, які можуть мати наслідком виключення з АІС, втрату роботи, судові позови та втрату професійної репутації. Друга категорія включає менш серйозні порушення, такі як неповага до прав учасників дослідження, неправильна інформація щодо проекту чи відсутність посилань на літературні джерела, що можуть завдати шкоди професійній репутації, санкцій щодо журналу і т.п. Кодекс також надає орієнтири стосовно того, що вважати правильною практикою для дослідника і що роботи, коли були виявлені етичні порушення (по відношенню до порушника та постраждалого), а також процесу розслідування випадків порушення кодексу АІС. Хоча подібні кодекси не можуть повністю попередити неетичну поведінку, вони дозволяють установити межі етичної поведінки в науковому середовищі та зменшити число випадків порушень етики.

Кодекс етики Асоціації інформаційних систем

Категорія 1. Положення цієї категорії повинні ЗАВЖДИ виконуватися, і їх недотримання є серйозним етичним порушенням.

1. Не допускайте плагіату.
2. Не фабрикуйте та не фальсифікуйте дані, процедури досліджень чи результати аналізу даних.

Категорія 2. Положення цієї категорії рекомендуються як норми етичної поведінки. Явне нехтування цими та іншими нормами професійної етики, хоча й менш серйозні, ніж віднесені до категорії 1, можуть зашкодити професійній репутації, призвести до санкцій наукових видань та юридичних наслідків.

1. Поважайте права учасників дослідження, особливо їх право на інформаційну приватність та інформацію про дослідження і конкретні завдання, які їм треба виконувати в межах дослідження.

2. Не надавайте неправдиву інформацію редакторам наукових журналів та комітетам конференцій стосовно походження статей, які ви їм надаєте.
3. Не зловживайте повноваженнями, які вам надали як редактору, рецензенту чи керівникові наукового проекту, не дозволяйте особистим відношенням упливати на ваші висновки та судження.
4. Повідомляйте про будь-який конфлікт інтересів, що може вплинути на вашу здатність об'єктивно й неупереджено оцінювати наукові матеріали, заявки на гранти, програмне забезпечення або добирати фахівців зовнішніх організацій.
5. Не використовуйте опубліковані дані інших авторів без посилання чи неопубліковані дані без дозволу і посилання.
6. Відзначайте внесок усіх учасників наукових досліджень, як колег, так і студентів, відповідно до їх частки в розробленні інтелектуального продукту.
7. Не використовуйте неопубліковані документи, інформацію, ідеї, концепції чи дані, які ви можете бачити в результаті рецензування, без дозволу автора.
8. Використовуйте архівні дані лише відповідно до правил джерела архіву.

Порада. Як захистити себе від конфліктів, суперечок, помилок і навіть юридичних позовів щодо авторства.

1. Зберігайте документацію і дані, необхідні для валідації вашого авторства в будь-якому науковому проекті, у якому ви брали участь.
2. Не публікуйте повторно ваші старі ідеї, видаючи їх за новий інтелектуальний продукт.
3. Погоджуйте питання авторства даних до початку їх збору.
4. Звертайтеся за порадою до колег, коли у вас виникають сумніви як правильно вчинити.