ШЛЯХ до РЕЗУЛЬТАТІВ

ШЛЯХ до РЕЗУЛЬТАТІВ

планування та проведення ефективних оцінювань розвитку

ЛІНДА ДЖ. МОРРА ІМАС
РЕЙ К. РІСТ

Київ, 2015

Автори
Лінда Дж. Морра Імас
Рей К. Ріст
Оригінальна назва
The Road to Results: Designing and Conducting Effective Development Evaluations
Назва
Шлях до результатів: планування та проведення ефективних оцінювань розвитку
Організація
МБФ «Міжнародний Альянс з ВІЛ/СНІД в Україні»
Переклад
В. Каплун
Верстка
І. Сухомлинова
Обкладинка
Naylor Design, Washington, DC

Шлях до результатів: планування та проведення ефективних оцінювань розвитку. – Морра Імас Л.Дж., Ріст Р.К. – К.: МБФ «Міжнародний Альянс з ВІЛ/СНІД в Україні», 2015. – 580 с.

Вперше видано Світовим банком англійською мовою під назвою «*The Road to Results: Designing and Conducting Effective Development Evaluations*» у 2009 році. Даний переклад українською мовою здійснено на замовлення проекту «Залучення місцевих організацій до розвитку моніторингу та оцінки у сфері ВІЛ/СНІДу в Україні» (МЕТІДА), що впроваджується МБФ «Міжнародний Альянс з ВІЛ/СНІД в Україні». У разі невідповідності перекладу просимо звертатися до оригінальної версії книги.

Morra Imas Linda G., Rist Ray C., 2009. *The Road to Results: Designing and Conducting Effective Development Evaluations*. World Bank. © World Bank. https://openknowledge.worldbank.org/handle/10986/2699 License: CC BY 3.0 IGO.
Морра Імас Лінда Дж., Ріст Рей К., 2009. *Шлях до результатів: планування та проведення ефективних оцінювань розвитку.* Світовий банк. © Світовий банк. https://openknowledge.worldbank.org/handle/10986/2699 Ліцензія: CC BY 3.0 IGO.

Дана книга видана в рамках ліцензії Creative Commons Attribution 3.0 IGO (CC BY 3.0 IGO). Для перегляду екземляру ліцензії відвідайте веб-сторінку http://creativecommons.org/licenses/by/3.0/igo/

Переклад цієї книги не було здійснено Світовим банком і ця публікація не повинна розглядатися як офіційний переклад, здійснений Світовим банком. Світовий банк не несе відповідальність за зміст або невідповідності у даному перекладі. This translation was not created by The World Bank and should not be considered an official World Bank translation. The World Bank shall not be liable for any content or error in this translation.

Переклад та публікація цієї книги стали можливими за підтримки Проекту «Залучення місцевих організацій до розвитку моніторингу та оцінки у сфері ВІЛ/СНІДу в Україні» (МЕТІДА), що впроваджується МБФ «Міжнародний Альянс з ВІЛ/СНІД в Україні», за фінансування Центрів США з контролю та профілактики захворювань (CDC), в рамках Надзвичайного плану Президента США для надання допомоги у зв’язку зі СНІДом (PEPFAR).

Ця публікація підтримана Угодою про співробітництво № U2GGH000840 з Центрами США з контролю та профілактики захворювань (CDC). Відповідальність за зміст публікації лежить виключно на її авторах та не обов’язково відображає офіційну позицію Центрів США з контролю та профілактики захворювань (CDC).

ПРИСВЯТА

Моєму чоловікові Сільвіо, який підтримує та надихає мене в професійній діяльності. Особлива подяка також двом сильним жінкам, які стали для мене зразком для наслідування й наставницями – Еленор Челімські та Лойс-еллін Детта.

ЛІНДА ДЖ. МОРРА ІМАС

Присвячується Моллі, Ітану, Медлін,
Лукасу, Ніні та Анні,
моїм шістьом онукам, які щодня
приносять радість у моє життя.

РЕЙ К. РІСТ

Книжка «Шлях до результатів» є доступною у вигляді інтерактивного підручника на сайті **http://www.worldbank.org/r2r.**

Цей інструмент дозволяє студентам та викладачам обмінюватися замітками та відповідними матеріалами, що сприяє більш ефективному, мультимедійному навчанню.

ЗМІСТ

ПЕРЕДМОВА XV
АВТОРИ XVII
АБРЕВІАТУРИ XIX

ВСТУП 1

ОСНОВИ 5

ГЛАВА 1 Представлення оцінювання розвитку 7
Оцінювання: що це? 8
Витоки та історія оцінювання як дисципліни 19
Контекст оцінювання розвитку 27
Принципи та стандарти оцінювання розвитку 30
Приклади оцінювання розвитку 36
Підсумки 41
Завдання до Глави 1 41
Примітки 42
Посилання та рекомендована література 42
Веб-сайти 47

ГЛАВА 2 Розуміння чинників оцінювання розвитку 49
Огляд оцінювання в розвинених країнах та країнах, що розвиваються 50
Значення виникаючих проблем розвитку 73
Підсумки 97
Посилання і рекомендована література 98
Веб-сайти 106

ПІДГОТОВКА ТА ПРОВЕДЕННЯ ЕФЕКТИВНОГО ОЦІНЮВАННЯ РОЗВИТКУ 109

ГЛАВА 3 Побудова орієнтованої на результат системи моніторингу й оцінювання **111**
Важливість орієнтованої на результат системи моніторингу й оцінювання 112
Що таке орієнтована на результат система моніторингу й оцінювання? 113
Порівняння традиційної та орієнтованої на результат системи моніторингу й оцінювання 114
Десять кроків до побудови орієнтованої на результат системи моніторингу й оцінювання 118
Підсумки 141
Завдання до Глави 3 142
Примітки 145
Посилання та рекомендована література 145
Веб-сайти 146

ГЛАВА 4 Розуміння контексту оцінювання й програмної теорії змін **147**
Аналіз вихідних даних 148
Визначення головного клієнта та основних зацікавлених сторін 150
Розуміння контексту 155
Використання існуючих знань 156
Підсумки 178
Завдання до Глави 4 179
Посилання та рекомендована література 181
Веб-сайти 185

ГЛАВА 5 Розгляд підходу оцінювання **189**
Загальні підходи до оцінювання 190
Підсумки 216
Завдання до Глави 5 221
Примітки 222
Посилання та рекомендована література 222
Веб-сайти 227

ПЛАНУВАННЯ І ПРОВЕДЕННЯ **231**

ГЛАВА 6 Розробка питань оцінювання і початок заповнення матриці плану **233**
Джерела питань 234
Види питань 235
Визначення й вибір питань 242
Як ставити хороші питання 244
Планування оцінювання 246

Підсумки 256
Завдання до Глави 6 257
Посилання та рекомендована література 258
Веб-сайти 258

ГЛАВА 7 Вибір дизайну для причинно-наслідкових, нормативних та описових питань оцінювання 259
Встановлення зв'язку між питаннями та дизайном 260
Планування причинно-наслідкових питань 264
Дизайни для описових питань 286
Дизайни для нормативних питань 288
Потреба у ретельнішому плануванні оцінювання 289
Завдання до Глави 7 295
Посилання та рекомендована література 298
Веб-сайти 301

ГЛАВА 8 Вибір і створення інструментів для збору даних 303
Стратегії збору даних 304
Характеристики хороших вимірів 307
Кількісні й якісні дані 309
Інструменти для збору даних 313
Підсумки 366
Завдання до Глави 8 367
Посилання та рекомендована література 368
Веб-сайти 372

ГЛАВА 9 Вибір стратегії формування вибірки 373
Введення до формування вибірки 374
Види вибірок: випадкові та невипадкові 374
Визначення розміру вибірки 382
Підсумки 385
Завдання до Глави 9 387
Посилання та рекомендована література 389
Веб-сайти 390

ГЛАВА 10 Планування і проведення аналізу даних 391
Стратегія аналізу даних 392
Аналіз якісних даних 392
Аналіз кількісних даних 407
Пов'язання кількісних і якісних даних 420
Підсумки 423
Завдання до Глави 10 424
Посилання та рекомендована література 426
Веб-сайти 429

ПОДОЛАННЯ ВИКЛИКІВ 431

ГЛАВА 11 Оцінювання комплексних інтервенцій 433
Загальна картина через призму оцінювання розвитку 434
Спільні оцінювання 435
Оцінювання національних програм 439
Оцінювання секторальних програм 442
Тематичні оцінювання 445
Оцінювання програм глобального й регіонального партнерства 447
Підсумки 452
Посилання та рекомендована література 453
Веб-сайти 456

УПРАВЛІННЯ 457

ГЛАВА 12 Управління оцінюванням 459
Робота з матрицею плану 460
Проведення оцінювання на підрядній основі 460
Функції й обов'язки різних гравців 464
Управління персоналом, завданнями й бюджетами 469
Підсумки 478
Завдання до Глави 12 479
Посилання та рекомендована література 481
Веб-сайти 484

ГЛАВА 13 Представлення результатів 487
Розробка комунікаційної стратегії 488
Написання звіту з оцінювання 490
Візуалізація інформації 497
Як провести усну презентацію 509
Резюме 512
Глава 13 Вправи 512
Посилання та рекомендована література 513
Веб-сайти 514

ГЛАВА 14 Керівництво для фахівця з оцінювання: етика, політика, стандарти та керівні принципи оцінювання 515
Етична поведінка 516
Корумпованість і омани під час оцінки 516
Політика й оцінювання 520
Стандарти й керівні принципи оцінювання 526

Підсумки 531
Завдання до Глави 14 532
Посилання та рекомендована література 534
Веб-сайти 535

ГЛАВА 15 Погляд у майбутнє 539
Від минулого до сьогодення 540
Майбутнє 542
Завдання до Глави 15 551
Посилання та рекомендована література 552
Веб-сайти 553

Блоки

1.1 Шляхи використання оцінювання 15
1.2 Інститут внутрішніх аудиторів та Міжнародна організація вищих органів фінансового контролю 24
1.3 10 частин Стандартів КСР з оцінювання якості 32
1.4 Актуальність: Оцінка продовольчої допомоги для покращення ситуації та відновлення в Сомалі, проведена Всесвітньою продовольчою програмою 36
1.5 Результативність: Оцінювання DfID підтримки роботи Всесвітньої продовольчої програми в Бангладеш 37
1.6 Ефективність: Оцінювання використання коштів, зібраних для постраждалих від повені в Мозамбіку, проведена Комітетом з ліквідації наслідків стихійних лих 38
1.7 Вплив: Об’єднане оцінювання екстреної допомоги Руанді 39
1.8 Сталість: Оцінювання JICA Програми навчання в сфері інформаційно-комунікаційних технологій для третіх країн 40
2.1 Цілі розвитку тисячоліття 76
3.1 Ефект від вимірювання результатів 113
3.2 Різниця між орієнтованими на результат моніторингом та оцінкюванням 114
3.3 Десять способів використання даних про результати 137
4.1 Як проводити аналіз зацікавлених сторін 154
4.2 Використання фонду знань щодо профілактики злочинності 156
4.3 Огляд літератури щодо програм з поліпшення успішності учнів в південноафриканській країні 162
5.1 Соціальна оцінка проекту розвитку й кредитування сільського господарства Азербайджану 199
5.2 Побудова довіри за допомогою партисипативного оцінювання 205
5.3 Використання синтезованого оцінювання для вимірювання екологічної користі 211

6.1 Оцінювання інтервенції, в якої відсутні стандарти 238
6.2 Оцінювання політик та інтервенцій за допомогою причинно-наслідкових питань 240
6.3 П'ять етапів процесу оцінювання 250
7.1 Чому втекли слони? 260
7.2 Чи ефективно працюють комунальні школи? Оцінка сальвадорської програми EDUCO 275
7.3 Використання крос-секційного дизайну для відповіді на описові питання 281
7.4 Приклад дизайну опису випадків для описових питань 286
7.5 Використання дизайну порівняння стану до й після інтервенції для відповіді на описові питання 287
7.6 Приклад дизайну багаторазового порівняння з перервою для описових питань 288
7.7 Вплив програм з професійного навчання на звільнених робітників 290
7.8 Підтримка Іспанією ретельних дизайнів оцінювання 291
8.1 Правила для збору даних 305
8.2 Застосування структурованого підходу до оцінювання сільськогосподарської інтервенції 306
8.3 Контрольний список з 20 якісних питань Петона 310
8.4 Вимірювання популярності музейних експозицій за допомогою нав'язливих та ненав'язливих методів 313
8.5 Використання програми Google Earth для картування 319
8.6 Питання структурованих та напівструктурованих опитувань 333
8.7 Рекомендації щодо формулювання питань в опитуваннях 338
8.8 Загальні вказівки з проведення опитувань 344
8.9 Рекомендації щодо написання опитувальників 345
8.10 Рекомендації щодо проведення інтерв'ю 348
8.11 Рекомендації щодо проведення інтерв'ю в міжкультурному контексті 349
8.12 Рекомендації щодо формування питань для фокус-груп 354
8.13 Приклад питань для фокус-групового обговорення 356
9.1 Використання серійної випадкової вибірки для відбору людей зі СНІДом для проведення інтерв'ю 379
10.1 Приклад кодування 396
10.2 Використання аналізу змісту для визначення сприйняття позитивних наслідків використання ручних насосів 397
10.3 Стратегії електронного кодування даних 402
10.4 Рекомендації щодо роботи з кількісними даними 407
10.5 Розрахунок стандартного відхилення 414
10.6 Вказівки щодо аналізу даних кількісного дослідження 415
11.1 Приклад методології оцінювання національної програми 440
11.2 Об'єднане зовнішнє оцінювання сектору охорони здоров'я Танзанії, 1999–2006 443

11.3 Приклад тематичного оцінювання, присвяченого дитячій праці на смітниках в Африці, Азії та Європі 445
12.1 Рекомендації щодо написання технічного завдання для оцінювання 463
12.2 Рекомендації щодо розв’язання конфліктів 470
12.3 Рекомендації з покращення командної роботи 471
13.1 Використання інноваційних комунікаційних технологій для підвищення інтересу до результатів оцінювання 491
13.2 Поради щодо створення ефективних рисунків. 508
13.3 Поради щодо підготовки презентацій у PowerPoint та роздаткових матеріалів 511
13.4 Поради щодо використання слайдів 511
15.1 Ключові питання діагностики урядових систем МіО 544
15.2 Основні кваліфікаційні вимоги до фахівця з оцінки програм 548

Рисунки

3.1 Програмна теорія змін (логічна модель) для досягнення довгострокових результатів і впливів 116
3.2 Зразок програмної теорії змін (логічна модель) для скорочення дитячої захворюваності за допомогою пероральної регідратації 117
3.3 Десять кроків до побудови та підтримки орієнтованої на результат системи моніторингу й оцінювання 119
3.4 Спектр методів збору даних 126
3.5 Виявлення очікуваних чи бажаних рівнів поліпшення 128
3.6 Основні види моніторингу 130
3.7 Зв’язки між моніторингом реалізації та моніторингом результатів 131
3.8 Приклад зв’язку моніторингу реалізації та моніторингу результатів 131
3.9 Досягнення результатів за рахунок партнерства 132
4.1 Перехід від внесків до результатів 159
4.2 Потенційні впливи середовища на результати програми 159
4.3 Процес формування теорії змін 161
4.4 Проста теорія змін для покращення прийняття рішень підготовленими фахівцями з оцінювання 164
4.5 Схема простої діаграми теорії змін із зазначеними ключовими припущеннями 166
4.6 Шаблон теорії змін 167
4.7 Теорія змін для програми з поліпшення успішності учнів у читанні шляхом відвідування вчителями домівок учнів 169
4.8 Теорія змін для програми з використання семінарів із залученням зацікавлених сторін для скорочення корупції в уряді 171

4.9 Стандартна блок-схема 173
4.10 Блок-схема з акцентом на припущення 174
4.11 Стандартний ланцюжок результатів 174
A Проста теорія змін програми мікрокредитування 179
B Складніша теорія змін програми мікрокредитування 180
6.1 Використання логічної моделі для формулювання питань оцінювання 235
6.2 Підхід до оцінювання розвитку 252
6.3 Шаблон матриці плану 255
8.1 Приклад інструменту збору даних 323
8.2 Форма для спостережень з прив'язкою до існуючих питань 327
8.3 Зразок форми для структурованого спостереження 328
10.1 Блок-схема дослідження з використанням контент-аналізу 400
10.2 Бланк робочого листа для аналізу якісних даних 402
10.3 Заповнений робочий лист для аналізу якісних даних 404
10.4 Нормальний розподіл 412
10.5 Ненормальний розподіл 413
10.6 Стандартні відхилення за нормального розподілу 413
12.1 Приклад діаграми Ґанта 473
13.1 Двоетапний звіт щодо відстеження управлінських дій НГО MATR 496
13.2 Приклад організаційної діаграми 499
13.3 Приклад діаграми Ґанта 500
13.4 Середнє зниження температури у провінції A, лютий–червень 2008 501
13.5 Відмінності між середніми балами за читання у трьох школах району, 2004/05 502
13.6 Бали за тестом на знання щодо виховання дітей по регіонах 503
13.7 Різниця між відповідями на питання асистентів та директорів 504
13.8 Найвищі тарифи на електроенергію у канадських доларах, 3-й квартал 504
13.9 Середні бали. Порівняння балів за тест по роках навчання 504
13.10 Приклад рисунка з надмірною кількістю елементів 509
13.11 Приклад оптимізованого графіка, на якому більшість площі займає «корисне чорнило» 509

Ілюстрації

13.1 Малюнок дитини, що відображає її враження щодо стану села до та після проекту 498

Таблиці

1.1 Приклади оцінювання політики, програми та проекту 14
1.2 Порівняння моніторингу та оцінювання 17
1.3 Зміни в підході Світового банку до розвитку, 1950–2000 рр. 28
1.4 Критерії та показники для визначення незалежності організацій з оцінювання 33
3.1 Основні складові теорії змін 115
3.2 Матриця для вибору показників 124
3.3 Зразок системи оцінки результативності роботи: довгострокові результати/наслідки та показники 124
3.4 Зразок системи оцінки результативності роботи: довгострокові результати/наслідки, показники й базові значення 127
3.5 Зразок системи оцінки результативності роботи: довгострокові результати/наслідки, показники, базові значення й цілі/цільові значення 129
3.6 Зразок таблиці звітування про довгострокові результати/наслідки 135
4.1 Ролі зацікавлених осіб в оцінюванні 152
4.2 Питання, що ставляться на різних етапах оцінювання 155
4.3 Логічна структура програми зі зниження дитячої смертності 177
5.1 Типи прогнозних питань та шляхи їх вирішення 191
5.2 Риси спільного та традиційного оцінювання 203
5.3 Ціль, характеристики, сильні й проблемні сторони різних підходів оцінювання 216
6.1 Якісні питання, що ставляться на різних етапах політичного циклу 237
6.2 Матриця вибору й ранжування питань оцінювання 244
7.1 Дві ключові характеристики експериментальних, квазі-експериментальних та неекспериментальних дизайнів 263
7.2 Чотири види аудиту результативності роботи 289
7.3 Переваги й недоліки популярних експериментальних дизайнів 293
8.1 Приклад таблиці прийняття рішень щодо методу збору даних 304
8.2 Коли використовувати кількісний та якісний підходи 312
8.3 Переваги й недоліки громадських зборів 316
8.4 Переваги й недоліки використання наявних даних 325
8.5 Причини застосування та приклади спостереження 326
8.6 Використання напівструктурованого спостереження 330
8.7 Переваги й недоліки збору даних шляхом спостереження 332
8.8 Переваги структурованих та напівструктурованих опитувань 334
8.9 Переваги й недоліки опитувань 335
8.10 Етапи розробки змісту анкети 337
8.11 Результати заповнення громадянами карток оцінки задоволення від інфраструктурних послуг 346

8.12 Використання карток оцінки для громадян 347
8.13 Переваги й недоліки використання карток оцінки для громадян для збору даних 347
8.14 Переваги й недоліки інтерв'ю 350
8.15 Чотири етапи фокус-групи та дії модератора 355
8.16 Переваги й недоліки фокус-груп 357
8.17 Рекомендації щодо використання щоденників або журналів 358
8.18 Переваги й недоліки щоденників та опитувальних листів для самостійного заповнення 359
8.19 Переваги й недоліки експертних суджень 361
8.20 Переваги й недоліки методу Дельфі 363
9.1 Мінімальний розмір вибірки для забезпечення довірчого рівня у 95% з 5-відсотоковою похибкою 383
9.2 Розмір вибірки для генеральної сукупності розміром від 1 млн елементів 384
10.1 Рекомендації щодо збору й аналізу якісних даних 393
10.2 Потенційні переваги й недоліки контент-аналізу 401
10.3 Упорядкування й інтерпретація якісних даних 406
10.4 Відсоток міського населення в країнах Латинської та Центральної Америки 411
10.5 Бажані ступені центральної тенденції для різних типів даних 411
10.6 Думки клієнтів щодо медичних послуг у місцевій лікарні (відсоток респондентів) 416
10.7 Думки клієнтів щодо медичних послуг у місцевій лікарні (відсоток респондентів) 416
11.1 Вибір, організація й управління зовнішньою групою оцінювання та звітування про результати 438
12.1 Зразок мапи завдань 473
12.2 20 дій керівника проекту за Ґріром 475
13.1 Зразок контрольного списку для попередньої фази оцінювання та фази планування 489
13.2 Система відстеження рекомендацій 495
13.3 Елементи графіків 501
13.4 Типи та призначення різних видів графіків 502
13.5 Приклад даних в таблиці, яка має забагато ліній. Демографічна інформація про учасників 506
13.6 Приклад даних в таблиці, яка не перевантажена лініями. Демографічна інформація про учасників 506
13.7 Приклад класифікаційної таблиці. Стратегії подолання бідності: країни, де проводилося дослідження 507

Додаток

7.1 Стислий виклад ключових термінів і ознак різних видів дизайнів 296

ПЕРЕДМОВА

Оцінювання розвитку – цікавий компонент загальної мозаїки оцінок. Воно охоплює глибинні, складні та раптово виникаючі сфери розвитку. Оцінка боротьби з бідністю, глобалізації та її впливу на бідних, наслідків глобального потепління для слабких країн, структурної нерівності світових фінансових систем, стратегій допомоги країнам, що пережили конфлікт – ось лише декілька сфер, в яких оцінювання розвитку допомагає краще зрозуміти, як ми реагуємо на нагальні проблеми.

У світі все гучніше лунають вимоги щодо того, що уряди та організації мають відповідально ставитися до наслідків своїх дій, бути більш чутливими до потреб внутрішніх та зовнішніх зацікавлених осіб, а найбільше – щодо необхідності підвищення ефективності розвитку. В цьому контексті оцінювання стає ключовим інструментом для системного пошуку відповідей на запитання: «І що далі?» Тепер тим, хто працює в сфері розвитку, мало показувати, що вони щось роблять. Вони мають доводити, що їхня праця є ефективною.

Оцінювання розвитку також стає об’єктом досліджень. Важко відшукувати способи оцінювання, коли урядові системи слабкі або й зовсім відсутні, інформація часто перекручується з політичних міркувань, інформаційні прогалини великі та істотні, а засобів забезпечення надійності інформації не існує. Стикаючись з цими викликами, оцінювання розвитку демонструє гнучкість, новаторські й винахідливі підходи до пошуку шляхів інформування громадян, урядовців, донорів, громадянського суспільства та ЗМІ про те, чи ефективними є урядові програми.

Цей підручник покликаний посилити роль оцінювання розвитку як інструмента забезпечення інформованості при розробленні та впровадженні як окремих політик і програм, так і систем управління в цілому. Оцінка може бути потужним інструментом державного управління, здатним покращити роботу урядів та організацій, а також досягнення ними результатів.

Та її участь не обмежується лише державним сектором. Оцінка може бути потужним інструментом і для громадянського суспільства, неурядових і донорських організацій, що прагнуть підтримувати покращення становища бідних.

Наведений в цій книзі матеріал було адаптовано з Міжнародної тренінгової програми з оцінювання розвитку (International Program for Development Evaluation Training, IPDET) – програми, яку щоліта реалізують Світовий банк та Карлтонський університет (див. сайт www.IPDET.org). Зібрані на IPDET учасники з усього світу протягом місяця разом вивчають оцінювання розвитку. Матеріал в цьому підручнику є детальніше розгорненим змістом основного курсу IPDET. Ми надаємо його для того, щоб повніше викласти курс IPDET та поділитися ним з усіма, хто цікавиться оцінюванням розвитку.

Ми, автори, вдячні тим, хто давав поради, читав окремі частини рукопису, надавав відгуки та закликав нас продовжувати роботу над цією книгою. Повний список тих, кому ми хочемо подякувати, наведено зі зворотного боку книги. Але частині з них ми маємо подякувати прямо тут – це Майкл Петон, Патрік Грасо, Мартін Абрамс, Нільс Дабелстейн, Грег Джексон, Джин Свімер та Ненсі Портос. Всі вони читали та критикували певні розділи цієї книги. Ми вдячні їм за їхні зусилля. Сантьяго Помбо Бехарано з видавництва Світового банку надавав потужний імпульс нашій праці. Він постійно підтримував нас у роботі над цим рукописом.

Ми також вдячні двом особам, у співпраці з якими втілювали в життя програму IPDET в Карлтонському університеті – це Карен Гінзберг та Барбара Лівайн. Без IPDET книги не було б. Вони чудові партнери й чудові друзі.

Нарешті, ми маємо визнати видатний внесок Дайани Шульц Новак, яка співпрацювала з нами упродовж написання та редагування рукопису. Її відданість, турбота та вміння працювати були настільки важливими для нас, що ми не писали б цю передмову, бо не мали б книги, до якої її зараз пишемо.

Ми ж двоє дружимо та співпрацюємо вже близько 30 років. Познайомилися й стали працювати разом в Рахунковій палаті США в 1981 році. Вже більше 10 років працюємо в Світовому банку. Наші співробітництво та дружба з роками ставали все міцнішими. І це добре, що зараз, підійшовши до вищої точки своєї кар’єри, ми можемо поділитися із зацікавленою в оцінці спільнотою цим результатом нашої праці.

АВТОРИ

Лінда Дж. Морра Імас – радник Незалежної групи оцінювання Світового банку. Широко відома як «мати IPDET» – Міжнародної тренінгової програми з оцінювання розвитку та один із директорів цієї програми. Вона надає консультації та проводить навчання з питань моніторингу й оцінювання по всьому світу. Працювала ад'юнкт-професором в Університеті Джорджа Вашингтона та в Карлтонському університеті. Доктор Морра Аймес приєдналась до Світового банку в 1996 році і з тих пір багато разів керувала програмою оцінювання розвитку. Обіймала посаду старшого директора Рахункової палати США і часто виступала свідком перед комітетами Конгресу з питань освіти та програм зайнятості населення.

Рей К. Ріст – радник Незалежної групи оцінювання Світового банку. Перейшов до Банку в 1997 році. До цього він 15 років працював в уряді США – як у виконавчій, так і законодавчій гілках влади. Доктор Ріст обіймав академічні посади в Корнельському університеті, Університеті Джона Гопкінса, а також Університеті Джорджа Вашингтона. В 1976–77 роках доктор Ріст був старшим членом ради програми Фулбрайта в Університеті Макса Планка в Берліні (Німеччина). Він є автором, редактором або співредактором 25 книжок, написав 135 статей та виступав з лекціями у більш ніж 75 країнах.

АБРЕВІАТУРИ

ADB	Asian Development Bank		Азійський банк розвитку
AEA	American Evaluation Association		Американська асоціація оцінювання
AfDB	African Development Bank		Африканський банк розвитку
ANOVA	analysis of variance		Дисперсійний аналіз
ANSI	American National Standards Institute		Американський національний інститут стандартів
ASEAN	Association of Southeast Asian Nations	**АСЕАН**	Асоціація країн Південно-Східної Азії
BE	Business Environment		Бізнес-середовище
CAHMI	Child and Adolescent Health Measurement Initiative		Ініціатива щодо вимірювання здоров'я дітей та підлітків
CAQDAS	computer-assisted qualitative data analysis software		Програмне забезпечення для комп'ютеризованого аналізу якісних даних
CATI	computer-assisted telephone interviewing	**КСТО**	Комп'ютеризована система телефонних опитувань
CES	Canadian Evaluation Society		Канадська спілка з оцінювання
CESE	Central, Eastern, and Southern Europe		Центральна, Східна та Південна Європа
CGD	Center for Global Development	**ЦГР**	Центр глобального розвитку
CIDA	Canadian International Development Agency		Канадське агентство міжнародного розвитку
CIS	Commonwealth of Independent States	**СНД**	Співдружність незалежних держав
CODE	Committee on Development Effectiveness	**КЕР**	Комітет з ефективності розвитку

CSF	Community Support Framework		Програма регіональної і соціальної підтримки
CSR	corporate social responsibility	**КСВ**	Корпоративна соціальна відповідальність
DAC	Development Assistance Committee	**КСР**	Комітет сприяння розвитку
DANIDA	Danish International Development Agency		Датське агентство з міжнародного розвитку
DfID	Department for International Development		Міністерство з міжнародного розвитку
DEC	Disasters Emergency Committee		Комітет з ліквідації наслідків стихійних лих
EBRD	European Bank for Reconstruction and Development	**ЄБРР**	Європейський банк реконструкції та розвитку
ECI	Early Communication Indicator	**ПРС**	Показник раннього спілкування
EDUCO	El Salvador's Community-Managed Schools Program		Реалізована в Сальвадорі програма комунальних шкіл
ESEA	Elementary and Secondary Education Act		Закон про початкову та середню освіту
EU	European Union	**ЄС**	Європейський Союз
FAO	Food and Agriculture Organization of the United Nations		Організація з питань продовольства і сільського господарства ООН
FATF	Financial Action Task Force on Money Laundering		Група з розробки фінансових заходів боротьби з відмиванням грошей
FDI	foreign direct investment	**ПІІ**	Прямі іноземні інвестиції
FIT-ED	Foundation for Information Technology and Education Development		Фонд розвитку інформаційних технологій та освіти
GAD	gender and development	**ГіР**	«Гендер і розвиток»
GAO	Government Accountability Office		Рахункова палата
GDP	gross domestic product	**ВВП**	Внутрішній валовий продукт
GICHD	Geneva International Centre for Humanitarian Demining		Женевський міжнародний центр гуманітарного розмінування
GMI	guaranteed minimum income		Гарантований мінімум доходів
GPRA	Government Performance Results Act		Закон про роботу уряду та її результати
GRPPs	Global and Regional Partnership Programs	**ПГРП**	Програми глобального та регіонального партнерства

HIPC	Heavily Indebted Poor Countries	**БКВЗ**	Ініціатива для бідних країн з великою заборгованістю
HSSP2	Health Sector Strategic Plan 2	**СПСОЗ-2**	Стратегічний план для сектору охорони здоров'я №2
IBSTPI	International Board of Standards for Training, Performance, and Instruction		Міжнародна рада зі стандартів фахової підготовки, ефективності роботи та навчання
IDEAS	International Development Evaluation Association		Міжнародна асоціація оцінювання розвитку
IDRC	International Development Research Centre		Міжнародний центр досліджень розвитку
IFAD	International Fund for Agricultural Development		Міжнародний фонд розвитку сільського господарства
IFC	International Finance Corporation	**МФК**	Міжнародна фінансова корпорація
IISD	International Institute for Sustainable Development		Міжнародний інститут сталого розвитку
ILO	International Labour Organization	**МОП**	Міжнародна організація праці
IMF	International Monetary Fund	**МВФ**	Міжнародний валютний фонд
IOB	Policy and Operations Evaluation Department (the Netherlands)		Департамент оцінювання політики й операцій (Нідерланди)
IOM	International Organization for Migration	**МОМ**	Міжнародна організація з міграції
IPDET	International Program for Development Evaluation Training		Міжнародна тренінгова програма з оцінювання розвитку
IRBM	integrated results-based management system	**ІСУОР**	Інтегрована система управління з орієнтацією на результати
INTOSAI	International Organization of Supreme Audit Institutions		Міжнародна організація вищих органів фінансового контролю
IOCE	International Organisation for Cooperation and Evaluation		Міжнародна організація зі співпраці в сфері оцінювання
JEEAR	Joint Evaluation of Emergency Assistance to Rwanda		Об'єднане оцінювання екстреної допомоги Руанді
JICA	Japan International Cooperation Agency		Японське агентство міжнародного співробітництва

LGAB	look good – avoid blame	**ЗОУВ**	Зберігай обличчя – уникай вини
LSMS	Living Standards Measurement Survey Study	**ДКОРЖ**	Дослідження критеріїв оцінки рівня життя
M&E	monitoring and evaluation	**МіО**	Моніторинг і оцінювання
MATR	Management Action Tracking Record		Звіт про відстеження управлінських дій
MBS	Modified Budgeting System	**МБС**	Модифікована бюджетна система
MDBs	Multilateral Development Banks	**МБР**	Міжнародні банки розвитку
MDGs	Millennium Development Goals	**ЦРТ**	Цілі розвитку тисячоліття
MFPED	Ministry of Finance, Planning, and Economic Development	**МФПЕР**	Міністерство фінансів, планування та економічного розвитку
MIS	management information system	**УІС**	Управлінська інформаційна система
MMV	Medicines for Malaria Venture		Підприємство «Ліки від малярії»
MSH	Management Sciences for Health		Організація «Управлінські науки для охорони здоров'я»
NAD	National Audit Department	**НДА**	Національний департамент аудиту
NCSTE	Chinese National Centre for Science and Technology Evaluation		Китайський національний Центр оцінювання науки й технологій
NGO	nongovernmental organization	**НУО**	Неурядова організація
NORAD	Norwegian Agency for Development Cooperation		Норвезьке агентство зі співробітництва в цілях розвитку
NSGPR	National Strategy for Growth and Reduction of Poverty	**НССБР**	Національна стратегія скорочення бідності й розвитку
ODA	official development assistance	**ОДР**	Офіційна допомога розвитку
OECD	Organisation for Economic Co-operation and Development	**ОЕСР**	Організація економічного співробітництва і розвитку
PEAP	Poverty Eradication Action Plan	**ПЗВБ**	План заходів з викорінення бідності
POW	Health Sector Program of Work	**ПРСОЗ**	Програма робіт у секторі охорони здоров'я
PPBS	Planning, Programming, and Budgeting System	**СППФ**	Система планування, програмування і фінансування
PPBS	Program Performance Budgeting System	**СБВП**	Система бюджетування виконання програм

PRSP	Poverty Reduction Strategies Paper	**ДССБ**	Документи стратегії скорочення бідності
RFP	request for proposals	**ЗП**	Запит пропозицій
SDC	Swiss Agency for Development and Cooperation		Швейцарське агентство з розвитку та співробітництва
SHIPDET	Shanghai International Program for Development Evaluation Training		Шанхайська міжнародна тренінгова програма з оцінювання розвитку
SIDA	Swedish International Development Cooperation Agency		Шведське агентство міжнародного розвитку
SIEF	Spanish–World Bank Trust Fund for Impact Evaluation		Програма «Іспанський довірчий фонд Світового банку для оцінювання впливу»
SIR	subjective interpretation of reality	**СІР**	Суб'єктивна інтерпретація реальності
SWAps	sectorwide approaches		Секторальні підходи
TI	Transparency International		«Трансперенсі Інтернешнл»
TOR	terms of reference	**ТЗ**	Технічне завдання
UNDP	United Nations Development Programme	**ПРООН**	Програма розвитку Організації Об'єднаних Націй
UNEG	United Nations Evaluation Group		Група оцінювання ООН
UNEP	United Nations Environment Programme		Програма ООН з навколишнього середовища
UNESCO	United Nations Educational, Scientific and Cultural Organization	**ЮНЕСКО**	Організація Об'єднаних Націй з питань освіти, науки і культури
UNFPA	United Nations Population Fund		Фонд народонаселення ООН
UNHCR	United Nations High Commissioner for Refugees	**УВКБООН**	Управління Верховного комісара ООН у справах біженців
UNICEF	United Nations Children's Fund	**ЮНІСЕФ**	Дитячий фонд ООН
UNSD	United Nations Statistics Division		Відділ статистики ООН
USAID	U.S. Agency for International Development		Агентство США з міжнародного розвитку
WHO	World Health Organization	**ВООЗ**	Всесвітня організація охорони здоров'я
WID	women in development	**ЖвР**	«Жінки в розвитку»
WTO	World Trade Organization	**СОТ**	Світова організація торгівлі

ВСТУП

Аналітичні, концептуальні та політичні основи розвитку зазнають значних змін. Нові виклики вимагають ширшого розуміння секторів, країн, стратегій та політик розвитку. Підкреслюється важливість навчання та постійного зворотного зв'язку на всіх етапах циклу розвитку.

Більше того, оцінювання розвитку можна певною мірою вважати суспільним благом.

Оцінка виходить за межі однієї організації. Якісне оцінювальне дослідження може мати позитивні додаткові ефекти для спільноти розвитку. Оцінювання розвитку має властивості міжнародного суспільного блага (Пічьотто і Ріст, 1995 р., ст. 23).

Масштаби і складність програм розвитку зростають – це стосується і їх оцінки. Спеціалісти, які займаються оцінюванням розвитку, відходять від традиційних моделей, орієнтованих на оцінку активності, і звертаються до моделей, орієнтованих на результати, адже спільнота вимагає результатів і прагне досягнення Цілей розвитку тисячоліття. Спільнота розвитку зміщує увагу з проектів на комплексне вирішення проблем країн, і спеціалісти з оцінювання розвитку шукають методи оцінки результатів на рівні країни, сектору, теми, політики та навіть на глобальному рівні. Визнаючи важливість не лише комплексного, а й координованого підходу до вирішення проблем країн, що розвиваються, а також важливість партнерських відносин, спеціалісти з оцінювання розвитку все частіше почали проводити об'єднані оцінювання. Хоча об'єднані оцінювання мають багато переваг, вони підвищують складність оцінювання розвитку (ОЕСР 2006). Крім того, спеціалісти з оцінювання розвитку все частіше стикаються з проблемою оцінки роботи окремої організації в цьому ширшому контексті та визначення її внеску.

З появою цього більш складного та вимогливого підходу до розвитку розробляти проекти оцінювання стало важче. Доводиться враховувати складні методологічні вимоги та високі стандарти щодо визначення впливу.

Попит на нові підходи оцінки та новий набір навичок виходить за межі чистої економіки. Нагальні питання – такі як зміна клімату – потребують нових підходів до оцінки соціально-екологічної відповідальності. Масштаб екологічних проблем, міжнародні наслідки, складність отримання порівнюваних показників, а також численні випадки неочікуваних наслідків зумовлюють необхідність застосування комплексного, багатометодного підходу до оцінювання.

Цілком імовірно, що жодна дисципліна не може домінувати в ситуації, коли доводиться мати справу з багатьма проблемами, надіями та зусиллями більшої частини людства. За відсутності єдиного об'єднуючого принципу фахівці з оцінювання часто переходять межі дисциплін, все більш еклектично та розкуто підходять до використання інструментів суспільствознавчих наук (Пічьотто і Ріст, 1995 р., ст. 169).

Нарощування потенціалу для проведення оцінки, тобто створення підрозділів оцінювання, навчених практикам і методам оцінювання розвитку, є нагальним завданням, що стоїть перед більшістю країн, що розвиваються. Організація національних асоціацій оцінювання в країнах, що розвиваються, так само, як і відповідних регіональних груп оцінювання – перші важливі кроки щодо підвищення професіоналізму спільноти оцінювання розвитку. В нещодавно створеній Міжнародній асоціації оцінювання розвитку (IDEAS) спеціалісти знаходять для себе міжнародну професійну організацію, діяльність якої присвячена їхнім потребам. Також створенню засобів оцінювання розвитку сприяє зростання кількості університетських курсів та регіональних навчальних центрів.

Ця робота має стати інструментом для підвищення рівня оцінювання розвитку. Вона повинна допомогти спеціалістам розглядати та осмислювати нову архітектуру оцінювання, а особливо розробляти та проводити такі оцінювання, що концентруються на результатах вирішення завдань розвитку.

Міжнародну тренінгову програму з оцінювання розвитку (IPDET) було створено Відділом оцінки операцій Світового банку (нині – Незалежна група оцінювання [IEG]) в 2001 році. IEG спільно з Карлтонським університетом та за підтримки Інституту Світового банку провели першу програму в місті Оттава (Канада). З 2003 року програма пропонувала одно- та двотижневі спеціальні версії IPDET у більш ніж десяти країнах. В 2007 році було засновано Шанхайську міжнародну тренінгову програму з оцінювання розвитку (SHIPDET).

Незважаючи на зміни, що постійно вносяться до програми IPDET задля відображення мінливої природи розвитку, вона усе ще спрямована на всіх, хто працює або ж працюватиме в сфері оцінювання розвитку. Вона надає загальні інструменти для оцінки заходів із забезпечення розвитку (політик, програм та проектів) на місцевому, національному, регіональному та глобальному рівнях. Ця книга призначена для співробітників агенцій двостороннього та багатостороннього розвитку, урядів країн, що розвиваються, громадських організацій, а також парламентарів і приватних консультантів.

Загальною метою IPDET є покращення знань, навичок та вмінь учасників щодо розробки та проведення ефективного оцінювання розвитку з метою прийняття обґрунтованих рішень. В основі програми – 14 навчальних модулів, що окреслюють шлях до ефективної оцінки заходів із забезпечення розвитку.

Наша праця включає та доповнює ці модулі, комплексно розглядаючи проблеми, що постають перед спеціалістами з оцінювання розвитку, а також виступаючи в якості керівництва з проведення оцінювання розвитку. Завдяки цій книзі фахівці отримають загальні інструменти для проведення ефективної оцінки результатів розвитку.

Посилання

OECD (Organisation for Economi Co-operation and Development). 2006. DAC *Guidance for Managing Joint Evaluations.* Paris: OECD. http://www.oecd.org/ dataoecd/28/14/37484787.pdf.

Picciotta, Robert, and Ray C. Rist. 1995. *Evaluating Country Development Policies and Programs: New Approaches for a New Agenda.* San Francisco: Jossey-Bass.

ОСНОВИ

«Справжня геніальність полягає в здатності до оцінювання ненадійної, нестійкої та суперечливої інформації».

– ВІНСТОН ЧЕРЧІЛЛЬ

Глава 1: Представлення оцінювання розвитку

- Оцінювання: що це?
- Витоки та історія оцінювання як дисципліни
- Контекст оцінювання розвитку
- Принципи та стандарти оцінювання розвитку
- Приклади оцінювання розвитку

Глава 2: Розуміння рушійних факторів оцінювання розвитку

- Огляд оцінювання в розвинених країнах та країнах, що розвиваються
- Наслідки виникаючих тенденцій розвитку

ГЛАВА 1

Представлення оцінювання розвитку

В цій главі наводяться визначення та загальні положення оцінювання проектів, програм та політик. Далі розглядається оцінювання заходів із забезпечення розвитку, яку зазвичай називають *оцінюванням розвитку*.

Ця глава складається з п’яти основних частин:

- Оцінювання: що це?
- Витоки та історія оцінювання як дисципліни
- Контекст оцінювання розвитку
- Принципи та стандарти оцінювання розвитку
- Приклади оцінювання розвитку

Оцінювання: що це?

Для початку важливо розібратися, що мається на увазі під оцінюванням, зрозуміти його завдання, а також шляхи використання. В цій частині глави:

- наводиться визначення оцінювання;
- визначається мета оцінювання;
- наводиться перелік вигід від проведення оцінювання;
- вказується, що саме оцінюють фахівці з оцінювання;
- описуються шляхи використання оцінювання;
- розглядається зв'язок між моніторингом та оцінюванням;
- визначаються ролі та функції спеціалістів з оцінювання.

Визначення оцінювання

Оцінювання:
Встановлення цінності проекту, програми або політики

Існує багато визначень ***оцінювання***. Оксфордський словник наводить наступні варіанти:

> 1. Встановлення цінності або вартості (товарів і т.д.); розрахунок або декларація цінності; 2. Оцінювання або встановлення значення (математичного виразу, фізичної кількості і т.д.) або обчислення імовірності, винесення судження про факти тощо.

У контексті оцінювання як дисципліни цей термін отримав кілька значень. Розходження у визначеннях відображають різні погляди на мету оцінювання – забезпечення відповідальності протиставляється вивченню – або на час оцінювання в контексті етапу розвитку програми, проекту чи політики. Насправді загальної згоди стосовно визначення не існує.

> Справді, розглядаючи роль формулювань в оцінюванні, Майкл Скрайвен, один із засновників сучасного оцінювання, зазначив, що існує близько шістдесяти різних термінів, які означають оцінювання і можуть використовуватися в тому чи іншому контексті. Мова йде про такі дієслова, як оцінювати, аналізувати, рецензувати, розглядати, судити, ранжувати, вивчати, перевіряти і т. д. (Fitzpatrick, Sanders and Worthen 2004, p. 5).

Більшість визначень оцінювання включають поняття встановлення цінності або вартості предмета оцінки. Саме це «встановлення цінності» використовується для відокремлення оцінювання від досліджень та моніторингу.

В цій книзі використовується значення, наведене в Глосарії Організації економічного співробітництва і розвитку (OECD)/Комітету сприяння розвитку (КСР) (OECD 2000, p. 21):

> Оцінювання – це процес визначення вартості чи значення заходу, політики або програми. Воно є максимально систематичним та об'єктивним і стосується запланованого, триваючого або завершеного втручання.

Оцінювання бувають формативними, сумативними та перспективними:

> Формативне оцінювання – це оцінювання, спрямоване на підвищення ефективності, [та] яке частіше за все проводиться на етапі реалізації проектів або програм. Формативні оцінювання також можуть проводитися з інших причин – таких як, наприклад, відповідність вимогам, законодавству, або в рамках більш масштабної оціночної ініціативи. Сумативне оцінювання – це, навпаки, дослідження, що проводиться наприкінці заходу (або його етапу) для визначення того, якою мірою було досягнуто запланованих результатів. Сумативне оцінювання має надавати інформацію щодо вартості програми (OECD 2002, pp. 21–22).

Формативне оцінювання:
Оцінювання того, як реалізується програма, політика або проект

Формативне оцінювання розглядає використані шляхи реалізації програми, політики або проекту. Вивчається, чи відповідає прийнята «логіка дій» фактичним діям, встановлюються (безпосередні) наслідки реалізації (або її етапів). Таке оцінювання проводиться на стадії реалізації проекту або програми. Формативне оцінювання іноді називають оцінюванням процесу, тому що воно стосується, передусім, конкретних дій.

Прикладом формативного оцінювання є оцінка, проведена для Міжнародного центру дослідження розвитку (IDRC) стосовно його ініціативи з використання природних ресурсів країн Латинської Америки та Карибського басейну (відомої як «Мінґа») (Adamo 2003). Основною метою ініціативи «Мінґа» було сприяння формуванню спільноти професіоналів з управління природними ресурсами, як чоловіків, так і жінок, в Болівії, Еквадорі та Перу.

Одним з компонентів програми цієї ініціативи, що цікавив IDRC, була стратегія досягнення гендерної рівності. Щоб з'ясувати, як працює цей напрям програми, IDRC замовив проведення формативного оцінювання. Першим етапом формативного оцінювання був аналіз програмних документів стосовно досягнення гендерної рівності та відповідних заходів. Фахівці з оцінювання також перевіряли звіти про відрядження, щоб встановити, якою мірою під час цих відряджень опрацьовувалося питання гендерної рівності. З персоналом програми були проведені інтерв'ю: вивчалося, яких зусиль вони особисто докладали задля досягнення гендерної рівності, які враження при цьому отримали і чому навчилися.

Одним з видів формативного оцінювання є проміжне оцінювання. Як видно з самої назви, проміжне оцінювання проводиться приблизно на середині реалізації проекту, програми або зміни в політиці. Мета проміжного оцінювання – визначити, які функції працюють добре, а які – ні. Проміжне оцінювання може фокусуватися на накопиченому досвіді, а також на акту-

альності, дієвості та ефективності. Накопичений досвід відіграє важливу роль у плануванні майбутніх інтервенцій та покращенні поточних.

Сумативне оцінювання:
Оцінювання, що проводиться наприкінці інтервенції з метою визначити, наскільки вдалося досягти бажаного результату

Сумативне оцінювання, яке нерідко називають *оцінюванням результатів або впливу*, проводять наприкінці інтервенції або стосовно інтервенції, що знаходиться на пізній стадії реалізації, щоб визначити, в якому обсязі було досягнуто очікуваних результатів. Сумативне оцінювання має надавати інформацію щодо вартості і впливу програми. До сумативного оцінювання відносяться оцінки впливу, дослідження економічної ефективності, уявні експерименти, рандомізовані експерименти та вивчення **практичних прикладів**.

Прикладом сумативного оцінювання є оцінка Другої програми фінансового сектору в Монголії, проведена Азійським банком розвитку (ADB 2007). Згадана програма стосувалася реформ фінансового сектору, а саме реструктуризації фінансового сектору та його перетворення з монобанкової системи на дворівневу систему за підтримки АзБР. Сумативне оцінювання було проведено наприкінці другого етапу програми. Сумативні оцінювання використовуються для того, щоб давати відповіді на запитання стосовно актуальності, ефективності, впливу, сталості, зовнішньої функціональності, а також отриманого досвіду.

Різницю між формативним та сумативним оцінюванням можна сформулювати наступним чином:

- Формативне оцінювання фокусується на реалізації та вдосконаленні проекту, програми, політики.
- Сумативне оцінювання фокусується на результатах. Воно дозволяє приймати рішення щодо продовження, відтворення, розширення масштабу або завершення певного проекту, програми, політики.

Зазвичай необхідними є обидва види оцінювання. Вони використовуються організаціями на різних стадіях циклу проекту, програми або політики.

Перспективне оцінювання:
Оцінювання імовірних результатів запропонованого проекту, програми або політики

Перспективне оцінювання стосується ймовірних результатів запропонованих проектів, програм або політик. Воно є дещо схожим з визначенням оцінюваності. Визначення оцінюваності дає відповідь на запитання: «Чи є ця програма або цей проект вартими оцінювання?» та «Чи будуть вигоди вартими витрачених зусиль/ресурсів?» Перспективне оцінювання поєднує висновки оцінок попередніх досліджень, щоб визначити ймовірні результати запропонованих нових проектів, програм або політик. Наприклад, комітети Конгресу США часто звертаються до Рахункової палати (GAO) за допомогою в прогнозуванні ймовірних результатів запропонованих законопроектів.

Давній, але тим не менш цікавий приклад перспективного оцінювання – це проведене в 1986 році GAO дослідження «Підліткова вагітність: 500 000 новонароджених щороку, але мало застосовуваних програм» (U.S. GAO 1986). В ході цього оцінювання використовувалися чотири процедури. Аналі-

зувалися основні характеристики двох законодавчих актів Конгресу, переглядалася статистика стосовно масштабів підліткової вагітності, вивчалися характеристики федеральних та інших програм, аналізувалися оцінювальні дослідження ефективності попередніх програм з допомоги вагітним підліткам та підліткам, що стали батьками, а також програм, спрямованих на зниження ризику вагітності серед підлітків. Фахівці з оцінювання відтворили теорію, на якій засновувалася програма, а також логіку дії обох законодавчих актів Конгресу, аби встановити, чому вважалося, що ці ініціативи працюватимуть таким чином, як запропоновано в законодавстві. Потім спеціалісти порівняли отримані дані з характеристиками запропонованих законодавчих актів.

Цей вид перспективного оцінювання іноді називають *оцінюванням очікувань* (Rossi and Freeman 1993). Оцінювання очікувань, або перспективне оцінювання, часто включає в себе відтворення теорії або прогнозів та дослідження сценаріїв, а також аналіз зведених даних наявних досліджень та оцінку, що проводиться з метою перевірки емпіричного обґрунтування запропонованих ініціатив.

Мета оцінювання

Оцінювання можна застосовувати з різною метою. В рамках цієї дисципліни існують різні погляди щодо мети оцінювання в даному контексті.

Домінуючим поглядом є такий, що виокремлює чотири основні напрями застосування **оцінювання:**

- ***етичний:*** інформувати політичних лідерів та громадян про хід реалізації політики або програми та досягнуті при цьому результати. Ця мета поєднує в собі завдання покращення обліку, обробки інформації та сприяння демократії;
- ***управлінський:*** забезпечити більш раціональний розподіл фінансових та людських ресурсів серед «конкуруючих» програм, підвищити якість управління програмами, а також якість результатів таких програм;
- ***для прийняття рішень:*** підготувати базу для прийняття рішень щодо продовження, припинення або внесення змін до політики чи програми;
- ***освітньо-мотиваційний:*** сприяти освіті та мотивації державних установ та їхніх партнерів, даючи їм можливість розуміти процеси, з якими вони працюють, та співвідносити свою діяльність з відповідними цілями (Scientific and National Councils of Evaluation 1999).

Провідні фахівці з оцінювання-практики описують наступні варіанти мети оцінювання:

- Забезпечити покращення соціального стану.
- Сприяти розвитку дорадчої демократії.
- Забезпечити нагляд та відповідність вимогам.

- Забезпечити відповідальність та прозорість.
- Утворювати знання, ділитися та керувати ними.
- Сприяти організаційному вдосконаленню.
- Сприяти діалогу та співпраці між основними зацікавленими сторонами.
- Визначити актуальність, шлях реалізації, ефективність, дієвість, вплив та сталість проекту, програми або політики.
- Узагальнити отриманий досвід.

Челімські та Шейдіш (1997) використовують глобальну перспективу, поширюючи контекст **оцінювання** на світові проблеми. До викликів вони відносять вплив нових технологій, демографічний дисбаланс між різними країнами, захист оточуючого середовища, сталий розвиток, тероризм, права людини та інші питання, що виходять за межі однієї програми або навіть однієї країни.

Отже, метою будь-якого оцінювання є надання інформації відповідальним особам з тим, щоб вони могли приймати кращі рішення стосовно проектів, програм та політик. Оцінювання покликане допомогти особам, які приймають рішення, краще зрозуміти події, що можуть статися або вже сталися внаслідок певних заходів, а також визначити шляхи отримання більшого обсягу бажаних результатів.

Позитивні наслідки оцінювання

Оцінювання допомагає дати відповіді на запитання стосовно втручання, наприклад:

- Який вплив має інтервенція?
- Чи дає інтервенція очікуваний ефект?
- Чи є розбіжності в дієвості інтервенції в різних місцях її застосування?
- Хто є бенефіціаром даної інтервенції?

Вигоди від проведення оцінювання можна отримувати по-різному. Хтось отримує такі вигоди безпосередньо. Інші отримують їх опосередковано – це люди, що не беруть участі в **інтервенції,** але тим не менш користуються її позитивними результатами. Деякі інтервенції дають короткострокові результати, інші – довгострокові.

Щоб продемонструвати, яким чином люди по-різному можуть отримувати вигоди, наведемо приклад. Міністерство житлового будівництва і міського розвитку США в 1997 р. оцінювало програму опівнічного баскетболу для підлітків у віці 16–20 років, що мешкають в муніципальному житлі. Було проведено дослідження учасників програми до її реалізації та після. Опитування показали: до початку програми 92 відсотки респондентів очікували, що в них можуть виникнути проблеми із законом, а дві третини –

що стануть жертвами насильницьких дій протягом наступних трьох місяців. Після реалізації баскетбольної програми 20 відсотків респондентів заявили, що очікують можливих проблем із законом, і лише 5 відсотків очікували, що стануть жертвою злочину.

Оцінювання дієвості програми опівнічного баскетболу продемонструвало скорочення рівня злочинності серед неповнолітніх у віці від 16 до 20 років, що мешкають в районі, забудованому соціальним житлом, на 78 відсотків. Основною причиною такого скорочення, згідно з твердженням респондентів, було те, що програма опівнічного баскетболу дала їм можливість зайнятися чимось позитивним. В цьому прикладі учасники є безпосередніми бенефіціарами програми, яка, на їх думку, допоможе їм уникнути проблем із законом та не стати жертвою насильницького злочину.

Також було опитано інших членів громади. Респонденти повідомили, що завдяки програмі опівнічного баскетболу в них з'явилося відчуття, що тепер і громада, і діти в більшій безпеці. У цьому випадку жителі району отримують вигоди опосередковано та принаймні у короткостроковій перспективі (залежно від тривалості впливу програми). Хоча вони й не беруть участь у програмі, завдяки їй почуваються безпечніше.

Наведені вище підсумкові результати могли бути використані, щоб продемонструвати жителям та громаді в цілому, що програма успішно запобігає злочинам та загалом знижує рівень насильства. Адміністратори програми також змогли представити висновки міській раді й обґрунтувати запит про подальше фінансування. Адміністратори програми отримують опосередковану вигоду, якщо продовження фінансування дозволяє їм зберегти роботу.

У довгостроковій перспективі суспільство в цілому також виграє, якщо молодь не потраплятиме за грати, адже в цьому випадку не доведеться оплачувати утримання ув'язнених та нести втрати продуктивності. Замість того, щоб сидіти у в'язниці, ці молоді люди мають змогу влаштуватися на роботу, створювати суспільні блага та сплачувати податки.

Оцінювання також може надати відомості щодо процесу реалізації програми. Інші установи, що працюють в сфері державного або муніципального житлового будівництва (нецільові бенефіціари), зможуть користуватися досвідом, отриманим в ході реалізації програми та за результатами оцінки.

Другий приклад – побудова та обслуговування водоочисного комплексу, що забезпечує населення чистою питною водою. Пряма, безпосередня вигода – чиста питна вода, що її отримують люди. В більш довгостроковій перспективі прямою вигодою для громади в цілому та окремих її членів буде зниження рівня захворювань, що передаються через воду. Непрямі середньострокові вигоди можуть включати залучення до громади підприємств завдяки наявності чистої питної води.

Що оцінювати

Оцінювання може розглядати різноманітні аспекти розвитку (див. табл. 1.1). Ось деякі з них:

- ***проекти:*** одна **інтервенція** в одному місці, або ж реалізація єдиного проекту в кількох місцях;
- ***програми:*** інтервенція, що складається з різних видів діяльності, або проекти, що працюють на досягнення спільної мети;
- ***політики:*** стандарти, норми та правила, що встановлюються організацією з метою керування рішеннями щодо розвитку;
- ***організації:*** програми з кількох інтервенцій, що реалізуються організацією;
- ***сектори:*** інтервенції в межах певної сфери (це, наприклад, освіта, лісове господарство, сільське господарство або охорона здоров'я);
- ***теми:*** конкретні питання, нерідко міжгалузеві – такі як гендерна рівність або загальносвітові суспільні блага;
- ***допомога країні:*** виконання країною плану, загальний ефект від допомоги, накопичений досвід.

Шляхи використання оцінювання

Існує багато шляхів використання оцінювання. Оцінювання надає клієнтам, державним установам, неурядовим організаціям (НУО), суспільству та багатьом іншим сторонам зворотний зв'язок щодо політик, програм та проектів. Результати містять відомості про те, як використовуються суспільні фонди. Оцінка може надавати керівникам та особам, які формують політику, інформацію про те, що є дієвим, а що – ні з точки зору виконання початкових та переглянутих завдань.

Таблиця 1.1. Приклади оцінювання політики, програми та проекту

Вид оцінювання	Застосування	
	Приватизація систем водопостачання	Переселення
Оцінювання політики	Порівняння зразків підходів до приватизації систем громадського водопостачання	Порівняння стратегій переселення сільських жителів до нової місцевості
Оцінювання програми	Оцінка управління податками урядових систем	Оцінка того, як переселені фермери зберігають попередній рівень життя
Оцінювання проекту	Порівняння покращення показників сплати за водопостачання в двох провінціях	Оцінка агротехнічних підходів переселених фермерів в одній провінції

Джерело: Автори.

Оцінювання допомагає забезпечити відповідальне використання громадських коштів у рамках проектів, програм та політик. Воно дозволяє виявляти, які проекти, програми та політики потребують відтворення, масштабування, покращення або закриття.

Вайс (2004) підкреслює важливість визначення мети проведення оцінювання вже на етапі первісного планування. «Якщо ви не можете визначити основних цільових користувачів та сформулювати шляхи використання оцінювання, таке оцінювання проводити не варто, – відзначає вона. – Оцінювання, результати якого не використовуються – це марнування дорогоцінних людських та фінансових ресурсів» (ст. 1).

Процес оцінювання від початку і до кінця має бути спланований та виконаний відповідно до потреб основного цільового користувача. Такі основні користувачі відповідають за впровадження змін на основі їхньої участі в процесі або ж результатів оцінювання.

Оцінювання може слугувати багатьом цілям і використовуватися по-різному (блок 1.1). Наприклад:

- допомагати з'ясовувати, чому було або не було досягнуто запланованих результатів;
- встановлювати причини отримання незапланованих результатів або наслідків;
- оцінювати шляхи та причини впливу на результати певних дій;
- висвітлювати процес реалізації, невдачі та успіхи на всіх рівнях;
- допомагати робити висновки, висвітлювати сфери досягнень і потенціалу, а також надавати конкретні рекомендації щодо покращення та реформування.

Отже, оцінювання може бути корисним при розгляді:

- питань широкої політичної стратегії та проектування («Чи правильні наші дії?»);
- питань роботи та реалізації («Чи правильно ми робимо те, що робимо?»);
- чи є кращі підходи для вирішення проблеми («Про що ми дізнаємось?»).

Блок 1.1 «Шляхи використання оцінювання»

Оцінювання можна застосовувати з різною метою:

- для прийняття рішень щодо розподілу ресурсів;
- для переосмислення причин проблеми;
- для виявлення виникаючих проблем;
- для підтримки процесу прийняття рішень щодо вибору кращого з кількох альтернативних варіантів;
- для підтримки реформ та інновацій в державному секторі;
- для досягнення консенсусу стосовно причин проблеми та шляхів її вирішення.

Джерело: Kusek and Rist 2004.

Зв'язок між моніторингом та оцінюванням

Моніторинг:
Збір даних, за допомогою яких керівники можуть оцінювати ступінь виконання завдань

Для забезпечення єдності термінології скористаємось визначенням **моніторингу**, наведеним в Глосарії ключових термінів оцінки ОЕСР/КСР:

> Моніторинг – це постійна діяльність з використання систематичного збору даних щодо певних показників розвитку з метою надання керівництву та основним зацікавленим особам відомостей про хід виконання проекту, досягнення цілей та освоєння наданих коштів (OECD 2002, pp. 27–28).

Моніторинг є рутинною, постійною, внутрішньою діяльністю. Його використовують для збору інформації щодо діяльності в рамках програми, її результатів і наслідків з метою оцінки ефективності такої програми.

Прикладом системи моніторингу є використання міністерством охорони здоров'я Малаві 26 показників для відстеження якості послуг з охорони здоров'я, що надаються в Центральній лікарні країни. Показники включають кількість пацієнтів, оглянутих спеціалістом протягом чотирьох тижнів з моменту звернення, кількість смертей пацієнтів в лікарні, кількість смертей безпосередньо під час пологів у медичному закладі, а також кількість днів, проведених хворими в стаціонарі (Government of Malawi 2007).

Регулярно надавані дані за показниками дають міністру охорони здоров'я можливість спостерігати тенденції. Можуть детально вивчатися будь-які різкі коливання. Значне підвищення кількості смертей серед пацієнтів стаціонару, наприклад, може свідчити про високий рівень поширення інфекцій в лікарні, який необхідно негайно знизити. Значне зниження рівня поширення інфекцій може свідчити про ефективність нового дезінфекційного засобу, а отже, про те, що його слід активніше використовувати.

Оцінювання зазвичай проводиться з метою дати відповідь на запитання, яке стоїть за моніторингом даних – наприклад, чому кількість кесаревих розтинів збільшилась в п'яти лікарнях, або чому в трьох із 50 закладів відзначається особливо високий рівень виживання недоношених немовлят. Оцінювання також потрібне для встановлення зв'язку між результатами та конкретною інтервенцією, а не іншими можливими причинами.

І моніторинг, і оцінювання спрямовані на вимірювання та оцінювання ефективності, але вони використовують різні підходи та проводяться в різний час (таблиця 1.2).

- Моніторинг проводиться протягом реалізації програми або проекту.
- Оцінювання є періодичною оцінкою показників виконання програми або проекту. Воно спрямоване на те, щоб дати відповідь на запитання: «чому?».

Моніторинг – це внутрішня діяльність, що її виконує персонал проекту. Зазвичай обов'язки щодо забезпечення проведення моніторингу та використання його результатів покладаються на керівництво проекту. Оцінювання ж може мати як внутрішній, так і зовнішній характер. Обов'язки з оцінки покладаються на фахівця з оцінювання разом з членами персоналу програми (Insideout 2005).

Таблиця 1.2. Порівняння моніторингу та оцінювання

Моніторинг	Оцінювання
Постійний, безперервний	Проводиться періодично, у визначених часових рамках
Внутрішня діяльність	Внутрішня, зовнішня або спільна діяльність
Обов'язок керівництва	Обов'язок фахівця з оцінювання разом з персоналом та керівництвом
Постійний зворотний зв'язок для покращення роботи програми	Періодичний зворотний зв'язок

Джерело: Insideout 2005.

Ролі та діяльність фахівців з оцінювання

Поступова еволюція поняття та завдань оцінювання супроводжувалася такою самою еволюцією ролей та діяльності фахівців з оцінювання. Фахівці з оцінювання виконують багато ролей та займаються різноманітною діяльністю. Їх роль залежить від природи та завдань оцінювання. Як зазначили Фітцпатрік, Сандерс і Уортен (2004 р., ст. 28), «фахівці з оцінювання виконують багато ролей – наукових експертів, координаторів, планувальників, учасників колективної роботи, помічників осіб, що приймають рішення, а також конструктивних критиків». Вони також виступають як арбітри, довірені особи, наставники та агенти соціальних змін.

Хто проводить оцінювання?

Фахівці з оцінювання можуть брати участь в проведенні внутрішньої, зовнішньої або спільної оцінки.

Глосарій ОЕСР/КСР дає наступне визначення ***внутрішнього оцінювання***:

> оцінювання інтервенції з розвитку, що її проводить підрозділ або окремі особи, підзвітні керівництву донора, партнера або організації, що займається реалізацією (2002 р., ст. 26).

Внутрішнє оцінювання:
Оцінювання виконується людьми зсередини організації

Зовнішнє оцінювання:
Оцінювання виконується людьми ззовні організаціїї

Зовнішнє оцінювання визначається наступним чином:

> оцінювання інтервенції з розвитку, що його проводять юридичні та/або фізичні особи, які не входять до складу донора, партнера або організації, що займається реалізацією (2002 р., ст. 23).

Залучення як внутрішніх, так і зовнішніх фахівців з оцінювання має свої переваги та недоліки. Внутрішні фахівці з оцінювання зазвичай більше знають про програму, проект або політику порівняно із зовнішніми. Оцінка інтервенції також може доручатися особі, що розробляє цю інтервенцію та керує нею. Таким особам зазвичай більше відомо про історію, організацію, культуру, задіяних людей, проблеми та успіхи. Завдяки цим знанням внутрішні фахівці з оцінювання можуть задавати більш актуальні та нагальні запитання; вони знають, яким чином можна дізнатися, що насправді відбувається всередині організації.

Однак ця перевага може бути і недоліком. Внутрішні фахівці з оцінювання можуть бути настільки наближеними до програми, проекту або політики, що виявляться нездатними ефективно її розглядати, помічати рішення або зміни, очевидні для інших. Внутрішні фахівці з оцінювання також можуть піддаватися більшому тиску або впливу з боку осіб, які приймають рішення щодо програми та персоналу. Такі фахівці з оцінювання, скоріше за все, розглядатимуть всю організацію лише з перспективи своєї посади в ній. Більше того, зовнішні зацікавлені сторони сприйматимуть їхні висновки як менш достовірні порівняно з оцінками зовнішніх фахівців з оцінювання.

Зовнішні фахівці з оцінювання зазвичай користуються більшою довірою і сприяють тому, щоб оцінка виглядала об’єктивнішою. Окрім того, більшість зовнішніх фахівців з оцінювання володіють більш спеціалізованими навичками, що можуть виявитися необхідними для ефективного проведення оцінки. Вони також не залежать від адміністративно-фінансових рішень щодо програми (Fitzpatrick, Sanders, and Worthen 2004).

Тим не менш, зовнішній характер оцінювання не гарантує незалежність та достовірність її результатів, зокрема у випадку, коли консультанти раніше вже були пов’язані з програмою. Зовнішні консультанти також можуть бути занадто послужливо налаштовані до керівництва програми, сподіваючись отримати замовлення і в майбутньому.

> В якості третьої альтернативи все частіше розглядають спільне оцінювання. У ході спільного оцінювання фахівці з оцінювання співпрацюють з представниками установ та зацікавлених осіб у питаннях планування, виконання та інтерпретації оцінки (OECD 2002). Є декілька фундаментальних відмінностей між спільним оцінюванням та внутрішнім і зовнішним оцінюванням. «Спільне оцінювання являє собою додатковий і радикальніший крок у бік від моделі незалежної оцінки. ... [Це] така форма оцінювання, за якої різниця між експертами

та неспеціалістами, дослідником та піддослідним стирається та переосмислюється. ... Фахівці з оцінювання ... діють переважно в якості координаторів та інструкторів, що допомагають іншим людям проводити оцінювання» (Molund and Schill 2004, p. 19).

Зверніть увагу на різницю між спільнним оцінюванням та методами спільного оцінювання. Методи спільного оцінювання можуть використовуватися і у внутрішній, і в зовнішній оцінці.

Діяльність фахівця з оцінювання

Фахівці з оцінювання здійснюють діяльність у відповідності зі своїми ролями. Внутрішні фахівці з оцінювання можуть працювати над плануванням, реалізацією та стратегіями охоплення проекту або програми. Зовнішні фахівці з оцінювання зазвичай обмежують свою участь в управлінні програмою. Зазвичай всі фахівці з оцінювання:

- консультуються з усіма основними зацікавленими особами;
- керують бюджетом оцінки;
- планують оцінювання;
- проводять оцінювання або залучають до його проведення підрядників;
- визначають стандарти дієвості (на основі керівних документів або інших джерел);
- збирають, аналізують, тлумачать дані та результати оцінки, складають відповідну звітність.

Для досягнення поставлених цілей фахівці з оцінювання потребують різноманітних навичок. В процесі оцінки вони можуть допомагати накопичувати знання та поширювати отриманий досвід.

Витоки та історія оцінювання як дисципліни

Оцінювання як сучасна дисципліна утворилася із соціологічних досліджень, що базуються на застосуванні наукового підходу. Але оцінка застосовувалася ще в стародавні часи. Археологічні знахідки свідчать про те, що стародавні єгиптяни регулярно відстежували кількість зерна і худоби в країні більше 5 тис. років тому. Стародавні китайці та греки також проводили оцінювання.

Свідчення про проведення оцінювання в громадському секторі датуються початком третього тисячоліття до нашої ери, коли китайські чиновники проводили екзамени для державних службовців з метою вивчення професійного рівня претендентів на державні пости. В сфері ж освіти Сократ використовував усну оцінку як частину навчального процесу (Fitzpatrick, Sanders, and Worthen, p. 31).

Деякі експерти пов'язують виникнення сучасних методів оцінювання зі становленням природничих наук та підвищенням уваги до спостереження явищ (емпіричний метод), що мали місце у 17-му сторіччі. В ті часи у Швеції з'явилися спеціальні політичні комісії, що проводили певні оцінки. Система комісій у Швеції використовується і до сьогодні: там існує кілька сотень комісій.

На початку 19-го сторіччя в кількох англосаксонських країнах почали проводитися перші оцінювання освітніх і соціальних програм. У Британії оцінювання програм проводилося урядовими комісіями, покликаними розслідувати й оцінювати невдоволення освітніми та соціальними програмами. Існуючий нині інститут зовнішніх інспекторів шкіл виріс саме з таких комісій.

У Сполучених Штатах на самому початку 19-го сторіччя робились перші спроби оцінити якості шкільної системи за допомогою спеціальних тестів оцінки досягнень. Ця робота ведеться і понині: результати таких тестів є ключовим показником для визначення якості освіти в навчальних закладах. Також у згаданий період бере початок акредитація середніх шкіл і університетів США.

Історія оцінювання в 20-му сторіччі

Офіційне оцінювання та акредитацію американських і канадських навчальних закладів медичного профілю було вперше запроваджено на початку 1900-х років. Інші сфери, в яких у цей період також проводяться дослідження/вимірювання та оцінка – охорона здоров'я, житлове будівництво, продуктивність праці, демократичне та авторитарне врядування, а також стандартизоване освітнє тестування. Це в основному були роботи обмеженого масштабу, які проводилися урядовими установами та соціальними службами.

Россі та Фрімен (1993 р.) пов'язують прагнення до систематичного оцінювання програм в Сполучених Штатах із зусиллями щодо підвищення рівня грамотності, надання професіональної підготовки та скорочення смертності від інфекційних захворювань на зламі століття. У сфері розвитку «спроба впровадити кип'ятіння води в якості міри охорони здоров'я мешканців селищ Близького Сходу є одним із видатних досліджень в міжвоєнній емпіричній соціологічній літературі» (Rossi and Freeman 1993, p. 10).

Прикладні соціальні дослідження в Сполучених Штатах різко активізувалися після введення Президентом Франкліном Д. Рузвельтом політики «Новий курс». Федеральний уряд швидко збільшився у розмірах, створювалися нові установи для керівництва національними програмами та їх реалізації. До таких програм відносилися сільськогосподарські субсидії для фермерів, громадські роботи та схеми для створення робочих місць, сільська електрифікація та соціальне забезпечення. Новизна та експериментальний характер цих масштабних програм зумовили зростання потреби в оцінці

їхньої ефективності з точки зору реанімації економіки, створення робочих місць та впровадження системи соціального захисту населення.

Потреба оцінювання посилилася в ході Другої світової війни та після її закінчення, коли було розроблено та впроваджено ще більше масштабних програм у військовій сфері, міському житловому будівництві, професійній підготовці та в сфері охорони здоров'я. В цей час мала місце активна участь в міжнародних програмах, що охоплювали планування сім'ї, охорону здоров'я і харчування, а також розвиток сільських громад. Відповідні витрати були великими, тож супроводжувалися вимогами щодо відомостей про їх результати.

В 1950-х та 1960-х роках оцінювання стали більш регулярно використовувати в Сполучених Штатах та Європі з метою оцінки програм, пов'язаних з освітою, охороною здоров'я та психічного здоров'я, соціального забезпечення, попередження злочинності та реабілітації злочинців. Програма Президента Ліндона Б. Джонсона «Війна бідності» протягом 1960-х років стимулювала підвищення інтересу до оцінки. Також розширювалася робота в країнах, що розвиваються, по всьому світі. При цьому оцінювалися програми в сферах сільського господарства, розвитку місцевої інфраструктури, планування сім'ї, охорони здоров'я та харчування. Оцінки проводилися переважно з використанням традиційних соціологічних інструментів – таких як опитування та статистичний аналіз.

У 1949 році перша Гуверівська комісія рекомендувала структурувати бюджетні відомості для національного уряду Сполучених Штатів за видами діяльності, а не за окремими статтями витрат. Комісія також рекомендувала супроводжувати звіти про виконання бюджету інформацією про показники ефективності (Беркхед, 1956 р.; Майксел, 1995 р.). Ця бюджетна реформа стала відомою як бюджетування за результатами (Тайлер і Віленд, 1997 р.).

В 1962 році Міністерство оборони США, очолюване Робертом Макнамарою, розробило Систему планування, програмування та фінансування (СППФ). Завданням СППФ було підвищення ефективності та покращення роботи уряду. Система включала:

- встановлення цілей довгострокового планування;
- аналіз витрат і вигід альтернативних програм, здатних забезпечити досягнення поставлених цілей;
- перетворення програм в бюджетно-законодавчі пропозиції та довгострокові прогнози.

СППФ змінила традиційний бюджетний процес, змістивши увагу на цілі та пов'язавши планування і складання бюджету (Офіс міністра оборони, 2007 р.). Перші дії СППФ у кінцевому підсумку привели до руху «моніторингу заради результатів».

Наприкінці 1960-х років багато західноєвропейських країн розпочали оцінювання програм. У ФРН, наприклад, Бундестаг вимагав, щоб федеральний уряд звітував про реалізацію та вплив різних соціально-економічних і податкових програм. Серед відповідних звітів були звіти про Закон про сприяння здійсненню права на працю (1969 р.), Закон про загальні освітні кредити (1969 р.), спільну федеральну програму з покращення регіональної економічної структури (1970 р.), програму інвестування в лікарняні заклади (1971 р.), а також різні звіти щодо субсидій та податків (Derlien 1999). Протягом цього періоду канадський уряд також почав рух в напрямку оцінювання урядових програм та їх ефективності. Канадські міністерства заохочувалися до створення підрозділів планування та оцінювання.

Спочатку це не давало значних результатів. У Канаді, ФРН та Швеції, «незважаючи на інституціоналізацію оцінювання програм в різноманітних сферах, системи залишилися досить фрагментарними, а кількість досліджень, що проводилися, була відносно невеликою» (Derlien 1999, p. 146).

Закон про початкову та середню освіту (ESEA) 1965 року став визначною віхою для розвитку оцінювання в Сполучених Штатах. Цей закон вимагав від уряду оцінки успішності школярів та стандартів якості роботи вчителів. Ним також передбачались ресурси (перші кошти, адресно виділені з бюджету США на оцінку) для проведення відповідної діяльності. Так відбулася інституціоналізація оцінки. Наприкінці 1960-х та на початку 1970-х років, коли федеральний уряд став спрямовувати кошти на оцінку, в Сполучених Штатах та деяких країнах ОЕСР з'являються численні статті та книги, присвячені оцінюванню розвитку. Були розроблені університетські програми з оцінки рівня магістратури, покликані забезпечити підготовку нових спеціалістів з оцінки відповідно до зростаючих потреб щодо прозорості та ефективності соціально-економічних програм, фінансованих урядом – таких як гранти на отримання середньої освіти, а також програми «Велике суспільство», в тому числі програми боротьби з бідністю, дошкільної освіти, захисту прав громадян, професійної освіти для молоді.

Канада, ФРН і Швеція в 1960-х роках проводили програмну оцінку нових урядових програм у сферах освіти, охорони здоров'я та соціального захисту.

> В цьому контексті утворилися офіційні системи планування, що обмежувалися середньостроковим фінансовим плануванням (у ФРН) або навіть намагалися інтегрувати формування бюджету зі складанням програм (в Швеції та Канаді). Так чи інакше, оцінювання розглядалося або як частина цих систем планування, або як певний фактор, зумовлений інформаційними потребами програм інтервенцій. ... Відповідно оцінка переважно використовувалася керівниками програм для реалізації існуючих та нових програм (Derlien 1999, pp. 153–54).

З середини 1970-х до середини 1980-х років оцінювання стало повноцінною сферою професійної діяльності в багатьох країнах ОЕСР. Було утворено професійні об'єднання, введено нові програми підготовки фахівців з оцінювання, засновано профільні періодичні видання, і оцінка почала виходити за межі фінансованих урядом програм, охоплюючи також діяльність корпорацій, фондів та релігійних установ. У Франції, наприклад, систематично розвивалася оцінка державної політики, при цьому багато університетів – в тому числі ***Великі школи*** – пропонували курси та інформацію про оцінювання в рамках своєї навчальної програми.

Багато країн ОЕСР запровадили програми навчання проведенню оцінюваня для службовців – урядові або надавані сторонніми підрядниками. Розглядалися нові методики та моделі, при цьому акцент робився на інформаційних потребах споживачів, вивченні незапланованих результатів та розробці стандартів і цінностей. Кількість присвяченої оцінюванню літератури також зросла – кількісно і якісно (Fontaine and Monnier 2002).

Починаючи з 1985 року комп'ютери та технології значно розширили можливості фахівців з оцінювання в сфері збору даних, їх аналізу та звітування про висновки оцінок, а також можливості обміну даними з іншими сторонами.

Оцінювання та аудит розвитку

Оцінювання розвитку зародилося з традицій аудиту та соціології. Між цими традиціями існують важливі спільні риси, відмінності та зв'язки.

Традиція аудиту

Аудит бере початок в Британії 19-го сторіччя, коли комерційний та промисловий розвиток сприяв зростанню:

> потреби у надійних фінансових даних, точність яких підтверджувалася б документарно... Робота аудиторів слугувала підтримкою зростаючої капіталістичної інфраструктури Заходу. Думки аудиторів були авторитетними завдяки їх компетентності та незаангажованості (Brooks 1996, p. 16).

Традиція аудиту є орієнтованою на дослідження, керування фінансами та бухгалтерський облік. Аудит покликаний визначати, чи досягла програма поставлених цілей, чи відповідали фінансові витрати правилам, нормам та вимогам програми. Використовуються такі концепції, як засоби внутрішнього контролю, ефективне керівництво і управління, а також перевірка фактів. Акцент робиться на прозорості та відповідності вимогам. Глосарій ОЕСР/КСР (OECD 2002) визначає кілька видів аудиту:

Стандартний аудит:
Об'єктивна перевірка, покликана покращити роботу організації

Фінансовий аудит:
Аудит, в центрі уваги якого знаходиться виконання положень відповідних статутів та нормативних документів

Аудит ефективності:
Аудит, що оцінює актуальність, економічність, дієвість та ефективність

- ***стандартний аудит:*** незалежна об'єктивна перевірка, що має створити додану вартість та покращити роботу організації (допомагає організації у досягненні цілей через використання системного дисциплінованого підходу до оцінки та вдосконалення ефективності керування ризиками, контролю та управлінських процесів);
- ***фінансовий аудит:*** аудит, в центрі уваги якого знаходиться виконання положень відповідних статутів та нормативних документів;
- ***аудит ефективності:*** аудит, що оцінює актуальність, економічність, результативність та ефективність.

Аудит може бути внутрішньою або зовнішньою функцією (блок 1.2). Внутрішній аудит допомагає організації у досягненні цілей через використання системного дисциплінованого підходу до оцінки та вдосконалення ефективності керування ризиками, контролю та управлінських процесів. Уряд може користуватися послугами зовнішніх аудиторів у разі необхідності забезпечення незалежності аудиту від керівництва програми.

Блок 1.2 «Інститут внутрішніх аудиторів та Міжнародна організація вищих органів фінансового контролю»

Інститут внутрішніх аудиторів, заснований в 1941 році, є основним міжнародним об'єднанням, що займається підтримкою та розвитком внутрішнього аудиту. Національні інститути розташовані по всьому світу, організація охоплює 165 країн.

У 1947 році Інститут внутрішніх аудиторів видав «Положення про обов'язки внутрішніх аудиторів», яке стало основою для розробки стандартів внутрішнього аудиту. В 1978 році було затверджено «Професійні стандарти внутрішнього аудиту».

Майже всі розвинені країни та країни, що розвиваються – від Афганістану до Зімбабве – представлені в Міжнародній організації вищих органів фінансового контролю (INTOSAI), що виступає в якості головної організації для світової спільноти зовнішнього урядового аудиту. Заснована в 1953 році, INTOSAI сприяє здобуттю та поширенню знань для вдосконалення урядового аудиту в усьому світі. Її «Кодекс етики та стандартів аудиту» та «Практичне керівництво з аудиту ефективності» знаходять широке застосування. Станом на 2008 рік до складу INTOSAI входить 188 країн.

Джерело: http://www.theiia.org, http://www.intosai.org.

Той факт, що традиція аудиту є сильною в країнах, які розвиваються, привів до утворення сильної традиції аудиту виконання нормативних вимог. Національний департамент аудиту (НДА) Малайзії, наприклад, уже протягом ста років займається забезпеченням підзвітності державного сектору. НДА проводить аудиторські перевірки для того, щоб:

- забезпечити виконання вимог законів та нормативних актів;
- виявити небажані фактори, що призводять до недієвості, неефективності та використання економічно не обґрунтованих процедур;
- визначити, чи є підготована фінансова звітність достовірною і об'єктивною, а також чи була відповідна документація належним чином складена згідно із загальноприйнятими стандартами бухгалтерського обліку та аудиту.

Внутрішній аудит:
Аудит, що проводиться всередині організації з метою оцінки фінансової діяльності та заходів

Внутрішній аудит охоплює широкий спектр «видів фінансової діяльності та фінансових операцій, включаючи системи, виробництво, інженерно-проектні роботи, маркетинг та кадрову діяльність» ((http://www.theiia.org). Його ефективність забезпечується існуванням професійної атестації, якої для оцінювання ще не передбачено (в главі 15 розглядаються аргументи «за» і «проти» атестації фахівців з оцінювання).

Оцінювання розвитку запозичила з аудиту таку рису, як висока увага до відповідності законодавчим та процедурним вимогам. Це можна спостерігати в ході оцінювання проектів з орієнтацією на результат у випадках співпраці сторін-донорів, що надають допомогу на двосторонній основі, та банків розвитку. Наприклад, «Стандарти добросовісної практики оцінювання підтримуваних МБР операцій в державному секторі», розроблені Групою співпраці з оцінки Міжнародних банків розвитку (МБР), включають врахування досягнення цілей. Цей підхід ґрунтується на тому, що «оцінювання досягнення цілей сприяє відповідальності» (Evaluation Cooperation Group of the Multilateral Development Banks n.d., p. 9.).

Континуум між аудитом та оцінюванням

Аудит та оцінювання можна розглядати як частину континууму, що надає споріднену, але різнопланову інформацію про виконання нормативних вимог, відповідальність, вплив та результати. Існує «перекривання в певних сферах, таких як ефективність оперативної діяльності та економічна ефективність. … при цьому оцінка зосереджується на аналізі політики та результатів, а аудит – на внутрішніх засобах фінансового контролю та управлінських системах» (Treasury Board of Canada Secretariat 1993, p. 3). Як аудит, так і оцінка покликані допомагати особам, які приймають рішення, «надаючи їм систематичну та достовірну інформацію, що може бути використана в процесі створення, керівництва, нагляду, внесення змін та, в певних випадках, припинення дії програм» (Wisler 1996, p. 1).

Багато було написано про відмінності між аудитом та оцінкою, а також їх спільні сфери. Коріння відмінностей – у походженні цих дисциплін: витоки аудиту криються у фінансовому обліку, а оцінка походить переважно від суспільних наук. Аудит більшу увагу приділяє виконанню вимог, а оцінка більше фокусується на встановленні зв'язку між спостережуваними змінами та політикою, програмою або проектом. Аудитори зазвичай шукають відповіді на нормативні питання (що є порівняно з тим, що має бути?), в той час як фахівці з оцінювання шукають відповіді на описові та причинно-наслідкові запитання (Wisler 1996).

Суспільствознавча традиція

З відходом урядів та організацій від концентрації на перевірці та відповідності вимогам до підвищеної уваги до результатів для проведення оцінювання почали застосовувати техніки з суспільствознавчих наук. В оцінюванні розвитку використовувалися науковий метод та методи соціологічних досліджень.

Науковий метод – системний підхід до отримання інформації, що намагається об'єктивно відокремити істину від переконання. В рамках цього підходу дані збираються шляхом спостереження та експерименту і ґрунтуються на формулюванні та випробуванні гіпотез. Дослідники, які використовують науковий метод:

- виявляють проблему, досліджують її та розглядають попередні пояснення;
- висувають гіпотезу про причинно-наслідкові відносини та викладають її у вигляді, що піддається вимірюванню;
- перевіряють гіпотезу, проводячи експеримент та збираючи дані;
- аналізують дані та роблять висновок;
- документують та розповсюджують результати своєї роботи;
- використовують результати для вдосконалення гіпотези.

Оцінка запозичила інші методи з суспільних наук – таких як соціологія, антропологія, політологія та економіка.

> Застосування методів соціологічного дослідження в оцінці збігається у часі зі зростанням та вдосконаленням самих цих методів, а також з ідеологічними, політичними та демографічними змінами, що мали місце в цьому сторіччі. Особливо важливими були формування і посилення позицій суспільних наук в університетах та підтримки соціальних досліджень. Факультети суспільних наук стали центрами перших робіт з оцінки програм, і надалі продовжували займати важливе місце в цій сфері (Rossi and Freeman 1993, p. 9).

Оцінка також багато запозичує із суспільствознавчих досліджень в таких сферах, як побудова теорії, проектування, підхід, методологія збору даних, аналіз та інтерпретація, статистика, опитування та складання вибірки.

Контекст оцінювання розвитку

Оцінювання розвитку виникло як окремий розділ дисципліни оцінювання. Його початок пов'язаний переважно із зусиллями щодо реконструкції та розвитку після Другої світової війни. В 1944 році було створено Всесвітній банк, а в 1972 році він заснував першу незалежну групу оцінювання. Європейський банк реконструкції та розвитку (ЄБРР) засновано в 1991 році. Інші міжнародні банки розвитку також з'явилися в 1990-х. Створені були й двосторонні організації, такі як британське Міністерство з міжнародного розвитку (DfID) або Канадське агентство міжнародного розвитку (CIDA). Оцінювання розвитку розпочалося із заснування цих організацій у зв'язку з тим, що донорів цікавили витрати за проектами та отримані результати. В свою чергу, міністерствам країн, що розвиваються, потрібно було виконувати вимоги щодо звітності про результати проекту, для цього використовувалися системи оцінки проектів, розроблені донорами з метою отримання інформації та забезпечення відповідальності.

За останні кілька десятиліть поняття розвитку змінилося. Це стосується й оцінювання розвитку. Наприклад, за час свого існування Світовий банк змістив акценти своєї діяльності, що знайшло відображення в ускладненні оцінювання розвитку (табл. 1.3).

ОЕСР відіграла важливу роль в становленні оцінювання розвитку. Місією цієї організації, заснованої в 1961 році, було «допомагати урядам досягати сталого економічного зростання і забезпечення зайнятості населення, а також підвищувати стандарти життя в країнах-членах, при цьому утримуючи фінансову стабільність і сприяючи, таким чином, розвитку світової економіки» (http:// www.oecd.org). Члени ОЕСР проводять зустрічі в спеціалізованих комітетах, до яких належить Комітет сприяння розвитку (КСР), в складі якого вже тривалий час діє робоча група з оцінювання розвитку (нині – Мережа КСР з оцінювання). Метою Мережі КСР з оцінювання є підвищення ефективності міжнародних програм розвитку за рахунок підтримання надійної, інформованої та незалежної оцінки. Мережа поєднує 30 двосторонніх та багатосторонніх агенцій з розвитку.

Широко застосовується дане КСР визначення оцінювання розвитку. Воно дещо відрізняється від загального визначення, наведеного на початку цієї глави. За версією КСР, оцінювання розвитку це:

> систематична та об'єктивна оцінка діючого або завершеного проекту, програми або політики, їх плану, реалізації та результатів. Завдання полягає у визначенні актуальності та міри виконання цілей, дієвості, ефективності, впливу та сталості розвитку. Оцінювання має надавати достовірну та корисну інформацію, що дозволяє включати отримані висновки в процес прийняття рішень як реципієнтами, так і донорами (OECD 1991b, p. 4).

У спільноті, що займається оцінюванням розвитку, використовується широкий спектр методик та практик. Загальновизнано, що поєднання різних теорій, аналітичних стратегій та методик часто дає найкращий результат в оцінці розвитку, особливо зважаючи на зростаючий масштаб та складність проектів, програм і політик, спрямованих на забезпечення розвитку. Комбінування підходів може допомогти покращити якість оцінки.

Цей набір методів, що його називають ***методологічною тріангуляцією***, означає:

> використання кількох теорій, джерел або видів інформації та/або видів аналізу з метою перевірки та підкріплення оцінки. Поєднуючи різні джерела даних, методи, види аналізу або теорії, фахівці з оцінювання прагнуть подолати суб'єктивність, пов'язану з використанням одиничних джерел інформації, одного методу, одиничних спостерігачів або теоретичних досліджень (OECD 2002, p. 37).

Таблиця 1.3. Зміни в підході Світового банку до розвитку, 1950–2000 рр.

Десятиріччя	Об'єкти уваги	Підхід	Дисципліна
1950-ті	Відбудова, реконструкція, технічна допомога, інженерно-проектні роботи	Технічна допомога	Інженерна справа
1960-ті	Економічне зростання, фінансування і створення проектів з надією на те, що посилене економічне зростання витягне більшу кількість населення з бідності	Проекти	Фінанси
1970-ті	Соціальні сектори або базові потреби (освіта, охорона здоров'я та соціальний захист); довгострокове планування та інвестиції в соціальний сектор	Секторне інвестування	Планування
1980-ті	Політики та кредитування, спрямовані на структурні перетворення; кредитування перетворень з конкретними умовами використовувалося для підтримки масштабних політичних реформ та допомоги щодо виведення країн з фінансових та боргових криз	Кредитування перетворень	Неокласична економіка
1990-ті	Переважно комплексні заходи спрямовані на країни в цілому, а не окремі проекти; більше уваги приділялося нарощуванню потенціалу та створенню інституцій в країнах, що розвиваються	Допомога країнам	Багатодисциплінарний підхід
2000-ті	Боротьба з бідністю, партнерства, участь, галузеві підходи та орієнтація на результат	Партнерства	Управління з орієнтацією на результат

Джерело: За Picciotto 2002.

Як зазначають Челімські та Шейдіш (1997 р., ст. 11):

> Оцінювання набуває все більшої методологічної різноманітності. Встановлено, що інструментарій фахівця з оцінювання має містити повний набір методів суспільних наук – мова йде про інструменти психології, статистики, освіти, соціології, політології, антропології та економіки.

Вибір способу та методології (або комбінації способів та методик) визначається тим, які питання ставляться та яка інформація потребується.

Зростання професійних об'єднань з оцінювання

Професійні об'єднання з оцінювання утворюють систему підтримки фахівців з оцінювання та забезпечують їхню професійність. Така підтримка сприяє нарощуванню потенціалу в сфері оцінювання розвитку.

У 1980-х роках існувало лише три регіональних або національних спілки з оцінювання. З тих пір відбулося вибухове зростання кількості національних, регіональних та міжнародних об'єднань з оцінювання в усьому світі; на даний час в розвинених країнах та країнах, що розвиваються, існує понад 75 об'єднань з оцінювання (http:/www.ioce.net/members/eval_associations.shtml). Значна частина цього зростання пов'язана з утворенням об'єднань з оцінки в країнах, що розвиваються. Наприклад, об'єднання на національному рівні існують в Малайзії та Шрі-Ланці з 1999 р., в Уганді – з 2002 р. На регіональному рівні були засновані Австрало-азійська спілка з оцінювання (1991), Європейська спілка з оцінювання (1994) та Африканська асоціація з оцінювання (1999).

Важливою міжнародною організацією в сфері оцінювання є Міжнародна організація зі співпраці в сфері оцінювання (IOCE) – вільне об'єднання регіональних та національних організацій з оцінки (асоціацій, спілок та мереж) з усього світу. Метою IOCE є формування керівного та робочого потенціалу в сфері оцінки у країнах, що розвиваються, а також сприяння взаємному плідному співробітництву щодо теорії та практики оцінки в усьому світі. Задля цього спеціалісти з оцінювання повинні використовувати глобальніший підхід до участі у виявленні та вирішенні світових проблем (http://ioce.net/overview/general.shtml).

Ще однією важливою міжнародною організацією з оцінювання є Міжнародна асоціація оцінювання розвитку (IDEAS). IDEAS було створено в 2001 році з метою нарощування потенціалу в сфері оцінки у країнах, що розвиваються. Її місією є «сприяння та поширення практики оцінювання розвитку шляхом вдосконалення методів, зміцнення потенціалу та розширення реальної зацікавленості» (http://www.ideas-int.org). IDEAS – єдине об'єднання професіоналів-практиків з оцінювання розвитку.

Стратегія IDEAS полягає у тому, щоб:

- сприяти оцінюванню розвитку в контексті результатів, прозорості та відповідальності суспільної політики і витрат;
- надавати пріоритет нарощуванню потенціалу в сфері оцінювання;
- стимулювати найвищі інтелектуальні та професійні стандарти оцінювання розвитку.

Принципи та стандарти оцінювання розвитку

Галузь оцінювання потребує принципів та стандартів, що сприяють відповідальності, забезпечують порівнянність та підвищують надійність і якість надаваних послуг (Picciotto 2005). Більшість пов'язаних з розвитком організацій користуються принципами, стандартами та критеріями ОЕСР/КСР. Основний документ (OECD 1991a) визначає п'ять критеріїв оцінювання сприяння розвитку:

- ***відповідність:*** міра відповідності завдань інтервенції розвитку запитам бенефіціарів, потребам країн, глобальним пріоритетам, а також політиці партнерів та агенцій розвитку;
- ***результативність:*** міра, якою допомога забезпечує досягнення поставлених цілей;
- ***ефективність:*** результати – якісні та кількісні – в порівнянні з використаними ресурсами (цей економічний термін свідчить про використання найдешевших з доступних ресурсів для досягнення бажаних результатів. Вимірювання ефективності зазвичай вимагає порівняння альтернативних підходів до досягнення однакових результатів з тим, щоб дізнатися, чи був застосований найефективніший процес);
- ***вплив:*** позитивні та негативні зміни, спричинені інтервенцією розвитку, прямо або опосередковано, заплановані або незаплановані (вимірювання впливу включає визначення основних наслідків діяльності на місцеві показники соціального, економічного, екологічного та іншого розвитку. В центрі уваги мають бути як заплановані, так і незаплановані результати. Розглядаються позитивні та негативні наслідки впливу зовнішніх факторів, такі як зміни торгово-фінансових умов);
- ***сталість:*** стійкість потоків чистої вигоди до ризику з плином часу (поняття сталості є особливо актуальним стосовно оцінки (не вимірювання) ймовірності продовження надходження вигоди від діяльності чи програми після зупинення донором фінансування. Проекти і програми повинні бути сталими як екологічно, так і фінансово (OECD 1991b)).

КСР розробив принципи сприяння оцінюванню розвитку (OECD 1991b). Ці принципи охоплюють наступні питання:

- мета оцінювання;
- неупередженість та незалежність;
- достовірність;
- корисність;
- участь донорів та реципієнтів;
- співпраця донорів;
- програмування оцінки;
- планування і реалізація оцінок;
- звітність, поширення інформації та зворотний зв'язок;
- застосування цих принципів.

Прийняті КСР «Принципи оцінювання сприяння розвитку» переглядалися в 1998 р. Вони порівнювалися з принципами інших організацій з метою гармонізації та пошуку можливостей розширення. Члени Комітету рекомендували наступні варіанти змін принципів:

- Змінити заяву про цілі.
- Прямо обговорити питання децентралізованого та спільного оцінювання.
- Детальніше викласти принципи та практичні засади участі реципієнтів та співпраці донорів.
- Врахувати останні досягнення в сфері оцінювання – такі як вимірювання показників ефективності, статус та системи оцінки успішності – а також розробку типології оціночної діяльності (OECD 1998).

Цей перегляд заклав підвалини для подальших публікацій КСР.

У 1994 році Американська асоціація оцінювання (AEA) видала «Стандарти оцінювання програм в США». Ці стандарти були затверджені Американським національним інститутом стандартів (ANSI) як Американські національні стандарти оцінювання програм. Їх було оновлено в 1998 році. Ці стандарти були прийняті й іншими об'єднаннями з оцінки, в тому числі з країн, що розвиваються.

У березні 2006 року Мережа оцінки КСР запровадила «Стандарти КСР з якості оцінювання» (OECD 2006) (блок 1.3). Стандарти, що нині знаходяться на стадії пробного використання, визначають основні засади процесу і продукту якісного оцінювання:

- Встановити стандарти процесу (проведення) і продуктів (результатів) оцінювання.
- Сприяти порівнянню оцінювання в різних країнах (мета – оцінка).

- Сприяти утворенню партнерств та співробітництва в рамках об'єднаної оцінки.
- Покращувати можливості країн-членів щодо використання висновків та звітів про оцінювання одне одного (в тому числі передової практики та отриманого досвіду).
- Оптимізувати проведення оцінювання.

Блок 1.3 «10 частин Стандартів КСР з оцінювання якості»

ОЕСР встановила 10 критеріїв визначення якості оцінювання:

- обгрунтування, мета та цілі оцінки
- поле оцінювання
- контекст
- методологія оцінювання
- джерела інформації
- незалежність
- етика оцінювання
- забезпечення якості
- застосовність результатів оцінювання
- повнота.

Джерело: OECD 2006.

На прохання Мережі ОЕСР/КСР з оцінювання, а також інших мереж з оцінки Незалежна група з оцінювання Світового банку розробила орієнтовні узгоджені принципи та стандарти для оцінювання Програми глобального та регіонального партнерства (ПГРП), певні риси якої ускладнюють проведення оцінювання. Ці орієнтовні принципи та стандарти проходять перевірку на практиці та за кілька років будуть переглядатися і затверджуватися (посилання на веб-сайт, де викладено ці принципи та стандарти, наведене в кінці цієї глави. Принципи та стандарти детальніше розглядаються в главі 14).

Незалежне оцінювання:
Оцінювання, яке проводять фахівці з оцінювання, нічим не зобов'язані особам, які займалися плануванням та реалізацією втручання

Важливим компонентом достовірності оцінювання розвитку є незалежність. Глосарій ОЕСР/КСР визначає ***незалежне оцінювання*** як «оцінювання, що проводиться організаціями та особами, які не контролюються відповідальними за планування та реалізацію інтервенції розвитку» (OECD 2002, p. 25). При цьому зазначається:

> Достовірність оцінювання залежить, зокрема, від того, наскільки незалежним було його проведення. Незалежність означає свободу від політичного впливу та адміністративного тиску. Для нього характерні повний доступ до інформації та повна автономія у проведенні досліджень і наданні звітів про отримані висновки.

Незалежність не означає ізоляції: взаємодія між фахівцями з оцінювання, керівниками програми, її персоналом та бенефіціарами здатна покращити оцінювання та його використання. Оцінювання може проводитися – внутрішньо або зовнішньо – фахівцями з оцінювання, підпорядкованими особам, відповідальним за прийняття рішень щодо планування та реалізації програмних інтервенцій (тобто керівництву). Така оцінка не є незалежною. Вона сприяє отриманню інформації, а не підвищенню відповідальності.

Керівники міжнародних банків розвитку з питань оцінювання, що регулярно зустрічаються як члени Групи співпраці з оцінювання, визначили чотири виміри (або критерії) незалежності оцінювання:

- організаційна незалежність;
- поведінкова незалежність;
- захист від зовнішнього впливу;
- уникнення конфлікту інтересів.

У табл. 1.4 наведено критерії та показники для оцінювання незалежності організації оцінки. Ці критерії та показники взято з різноманітних джерел.

Таблиця 1.4. Критерії та показники для визначення незалежності організацій з оцінювання

Критерій	Аспект	Показники
Організаційна незалежність	Структура і роль підрозділу з оцінювання є належними	Чи має підрозділ з оцінювання керівний документ, в якому вказано, що сфера відповідальності підрозділу охоплює всі операції організації та що його функціональне підпорядкування, персонал, бюджет та функції є організаційно незалежними від операційного відділу, а також від відділів, що займаються питаннями політики, стратегії та прийняття відповідних рішень
	Підрозділ підпорядковується голові або заступнику голови організації чи її керівному органу та звітує перед ними про результати оцінювання	Чи встановлено пряме підпорядкування підрозділу керівництву або керівній раді організації
	Підрозділ організаційно відокремлений від персоналу та прямого керівництва оцінюваної програми, діяльності або юридичної особи	Положення підрозділу в організації відносно оцінюваної програми, діяльності або юридичної особи
	Підрозділ регулярно звітує перед аудиторським комітетом або іншим наглядовим органом більшої організації	Підзвітність та частота звітування перед наглядовим органом
	Підрозділ є достатньо захищеним від політичного тиску, щоб мати змогу звітувати про результати оцінювання, не боячись наслідків	Міра незалежності підрозділу оцінювання та його співробітників від політичного керівництва і захищеності від участі в політичній діяльності

(продовження)

Критерій	Аспект	Показники
	Співробітники підрозділу захищені кадровою системою, в якій оплата праці, підготовка, перебування на посаді та підвищення залежать від особистих заслуг	Ступінь наявності та реалізованості меритократичної системи, що регулює питання оплати праці, підготовки, перебування на посаді та підвищення
	Підрозділ має доступ до всієї необхідної інформації та інформаційних джерел	Наскільки вільним є доступ підрозділу оцінки до персоналу, документів, співфінансуючих організацій та інших партнерів, клієнтів організації, програм, заходів або суб'єктів, що їх ця організація фінансує або спонсорує
Поведінкова незалежність	Підрозділ має змогу і бажання надавати жорсткі, безкомпромісні звіти	Рівень надання підрозділом оцінки звітів, що привертають увагу суспільства (з урахуванням обмежень, необхідних для захисту конфіденційної або виробничої інформації, а також для зниження інституційних ризиків) – мова йде про звіти щодо висновків, отриманих від програм та заходів організації; пропонування стандартів роботи, що випереджають використовувані на даний час організацією стандарти; критика результатів програм, заходів та суб'єктів організації
	Підрозділ має змогу надавати чесні звіти	Наскільки організація санкціонує передачу підрозділом оцінки звітів керівництву/раді керівників після перегляду та внесення зауважень відповідними підрозділами, але без обмеження керівництвом тем та коментарів
	Звітування про висновки оцінювання є прозорим	Наскільки прийняті в організації правила публікації звітності дозволяють групі оцінки повідомляти про важливі висновки зацікавленим особам, як внутрішнім, так і зовнішнім (з урахуванням обмежень, необхідних для захисту конфіденційної або виробничої інформації, а також для зниження інституційних ризиків)
Захист від зовнішнього втручання	Оцінювання належним чином планується і проводиться	Наскільки підрозділ оцінювання має змогу визначати план, масштаб, час та проведення оцінок без втручання керівництва
	Належне фінансування оцінки	Наскільки підрозділ оцінки захищений від обмежень у грошових та інших ресурсах, що можуть негативно впливати на його здатність виконувати свої завдання
	Думки фахівців з оцінювання щодо змісту звітів не відкидаються	Наскільки думка фахівця з оцінювання стосовно належного змісту звіту не відкидається та не піддається зовнішньому впливу

Критерій	Аспект	Показники
	Встановлені незалежні кадрові процедури стосовно керівника підрозділу оцінювання	Якою мірою керівний документ забезпечує незалежність керівника підрозділу з оцінювання від операційного керівництва через встановлені процедури прийому на роботу, звільнення, строк повноважень, оцінку роботи та оплату праці керівника підрозділу з оцінювання
	Підрозділ контролює прийом на роботу, підвищення та/або звільнення персоналу	Міра контролю, що його має підрозділ за прийомом на роботу, підвищенням, збільшенням платні та звільненням персоналу в рамках меритократичної системи
	Працевлаштування фахівця з оцінювання не залежить від результатів оцінювання	Якою мірою працевлаштування фахівця з оцінювання залежить від якості його роботи, компетентності та потреби в послугах фахівця з оцінювання
Уникнення конфлікту інтересів	Відсутні офіційні, професійні, особисті або фінансові стосунки, що можуть спонукати фахівця з оцінювання до обмеження в проведенні досліджень, оприлюдненні інформації, або ж призвести до формування пом'якшених чи тенденційних висновків	Наскільки існуючі політики та процедури визначення відносин фахівця з оцінювання, здатних порушувати незалежність оцінювання, донесені до співробітників шляхом проведення тренінгів та іншим чином та як часто вони застосовуються
	Фахівець з оцінювання не має упереджених думок, поглядів або переконань соціального/політичного характеру, здатних вплинути на висновки оцінки	Якою мірою встановлені та застосовуються політики і процедури, що вимагають від фахівців з оцінювання оцінювати свої персональні упередження, здатні вплинути на об'єктивність їхніх оцінок та повідомляти про такі упередження; при цьому про упередження фахівця з оцінювання мають інформуватися зацікавлені особи (з метою компенсації таких упереджень)
	Фахівець з оцінювання не є і не був задіяний у програмі, заході або роботі суб'єкта оцінки на рівні прийняття рішень або фінансового управління, або ж на бухгалтерській посаді, та не намагається працевлаштуватися в такій програмі, заході або роботі суб'єкта протягом проведення оцінки	Якою мірою встановлені та застосовуються правила або кадрові процедури, що не допускають участі персоналу в проведенні оцінки програм, заходів або суб'єктів, де вони обіймали посади, пов'язані з прийняттям рішень, фінансовим управлінням, або куди вони намагаються працевлаштуватися
	Фахівець з оцінювання не має фінансової зацікавленості в програмі, заході або суб'єкті оцінки	Якою мірою встановлені та застосовуються правила або кадрові процедури, що не допускають участі персоналу в проведенні оцінки програм, заходів або суб'єктів, стосовно яких мають фінансовий інтерес
	Близькі родичі або члени родини не задіяні або не мають змоги здійснювати безпосередній та значний вплив на програму, захід або суб'єкт, що оцінюється	Якою мірою встановлені та застосовуються правила або кадрові процедури, що не допускають участі персоналу в проведенні оцінки програм, заходів або суб'єктів, стосовно яких мають вплив члени їх родини

Джерела: Danish Ministry of Foreign Affairs 1999; OECD 1991b; CIDA 2000; Institute of Internal Auditors 2000; European Federation of Accountants 2001; INTOSAI 2001; U.S. GAO 2002.

Приклади оцінювання розвитку

Блоки 1.4–1.8 взято зі звітів про оцінювання (ALNAP 2006). Кожен з них є прикладом одного з описаних вище критеріїв: актуальності, результативності, ефективності, впливу та сталості.

Блок 1.4 «Актуальність: Оцінка продовольчої допомоги для покращення ситуації та відновлення в Сомалі, проведена Всесвітньою продовольчою програмою»

Цю оцінку провели два експатріанти, що перебували в Сомалі з тритижневим візитом в середині липня 2001 р. Оцінка охоплювала три роки підтримки, в рамках якої серед 1,3 млн осіб було розподілено 63 млн тонн продовольчих товарів. Бюджет програми склав 55 млн доларів США. 51% цієї допомоги мав бути спрямований на реабілітацію та відновлення, 30% – на допомогу в надзвичайних ситуаціях, а 19% мало надійти до суспільних установ. Основною метою довгострокової операції з надання надзвичайної допомоги та відновлення було «сприяти формуванню нових можливостей роботи інтегрованих програм відновлення в Сомалі, а також підтримувати гнучкість для того, щоб мати можливість використовувати перспективні напрямки розвитку та реагувати на надзвичайні ситуації» (WFP 2002, p. 4). Таким чином, оцінювалася актуальність такого набору напрямків розподілу допомоги, а також відповідність кожного виду інтервенції.

Загальна актуальність інтервенції розглядалася в контексті політичної економії допомоги в Сомалі. Оцінка стосувалася обґрунтування надання продовольчої допомоги. Серед аргументів за надання продовольчої допомоги зазначалося, що в Сомалі зазвичай має місце дефіцит продовольства, люди в багатьох місцях ізольовані від звичних ринків, а також те, що багато сомалійців втратили основну роботу та статки. Проти продовольчої допомоги говорило те, що більш ефективним могло б бути надання бенефіціарам коштів, за які вони купували б місцеву їжу там, де це можливо, використовуючи формулу «гроші за працю» або «продовольство за працю». Однак такі програми зазвичай є довгостроковими проектами без чіткої стратегії припинення. Розгляд як загальних, так і конкретних питань забезпечує комплексність проведеного аналізу актуальності.

Блок 1.5 «Результативність: Оцінювання DfID підтримки роботи Всесвітньої продовольчої програми в Бангладеш»

У вересні 2000 р. повені в шістьох південно-західних районах Бангладеш завдали серйозної шкоди 2,7 млн осіб. DfID взяло участь у Всесвітній продовольчій програмі щодо надання трьох партій продовольства, що включали повні пайки рису, бобів та олії. Під час розподілення першої партії продовольство отримали 260 000 бенефіціарів; під час розподілення другої та третьої партій – 420 000 бенефіціарів. Оцінка DfID (DfID 2001) містила комплексний аналіз виконання завдань проекту стосовно розміру пайків, набору продуктів та графіка розподілення допомоги.

В рамках оцінювання застосовано як кількісні, так і якісні методи. Кількісні дані було зібрано в 2644 випадково обраних домашніх господарствах в селищах, охоплених проектом. Якісні дані отримано шляхом оцінки засобів до існування в шести репрезентативних селищах, під час якої розглядалися режими отримання засобів до існування, становище та перспективи в громадах, постраждалих від повені. Друга, менша група оцінювання почала роботу через п’ять тижнів після закінчення першої якісної оцінки, щоб дізнатися про думки та поведінку громади стосовно продовольчих пайків. При цьому розглядалися такі питання, як своєчасність розподілення, потреба у наданих видах продовольства та режими використання допомоги. Кількісні та якісні дані аналізувалися в комплексі.

Звіт включає більшість ключових елементів оцінки результативності, в т.ч.:

- розгляд розвитку виконання завдань втручання, включаючи аналіз логічної матриці;
- оцінку критеріїв, використаних для вибору бенефіціарів, включно з поглядами головних зацікавлених осіб щодо цих критеріїв;
- аналіз механізмів реалізації, включаючи рівні участі громади;
- оцінка дієвості цільового підходу з розбивкою за гендерною ознакою та соціально-економічними групами;
- оцінка наданих ресурсів (враховується як розмір пайків, так і набір продуктів), включно з причинами їх надання (це питання також можна розглядати з точки зору критерію актуальності);
- розгляд адекватності графіків розподілення;
- аналіз думок бенефіціарів щодо інтервенції.

Блок 1.6 «Ефективність: Оцінювання використання коштів, зібраних для постраждалих від повені в Мозамбіку, проведене Комітетом з ліквідації наслідків стихійних лих»

Після повеней 2000 р. в Мозамбіку Комітет з ліквідації наслідків стихійних лих (DEC) провів оцінювання гуманітарної допомоги, здійсненої його агенціями (DEC 2001). Метою оцінки було інформувати британську громадськість про те, яким чином та де використано її кошти, а також виокремити корисний досвід для аналогічних операцій в майбутньому. Метод оцінювання включав численні інтерв'ю, попередні дослідження, виїзди на місця, а також детальні опитування бенефіціарів.

Розділ, присвячений ефективності, містить перелік багатьох ключових елементів, необхідних для оцінювання, включаючи аналіз:

- використання Комітетом військових ресурсів, оцінюваного в контексті недостатньої підтримки гелікоптерами проведення оцінки потреб, високої вартості застосування військових сил Заходу порівняно з комерційними засобами гуманітарної допомоги, а також порівняних витрат Королівських військово-повітряних сил, армії США та Південноафриканських сил національної оборони (у звіті відзначається: на витратні військові операції пішла значна частина коштів, що призвело до обмеженого фінансування проектів НУО);
- наслідків малої розвиненості ринку підрядних послуг для ефективності (наприклад: хоча залучення місцевих підрядників дозволило агенціям проводити роботи, що потребують активного застосування обладнання – такі як ремонт доріг – без значних капітальних інвестицій, залучені підрядники часто не виконували роботу вчасно);
- ефективності вибору заходів (пряме втручання за рахунок операційних програм, робота через місцевих партнерів або членів міжнародної мережі); оцінка виявила, що склад персоналу був важливішим фактором ефективності, ніж вибір заходу (це питання також можна було б віднести до оцінки за критерієм актуальності);
- порівняної ефективності роботи агенцій із залученням місцевого або іноземного персоналу;
- порівняної ефективності роботи тих агенцій, що мали партнерів, і тих, що таких партнерів не мали;
- як інвестування в підготовку допомогло забезпечити більш ефективні дії;
- ефективності систем бухгалтерського обліку.

Було зроблено спробу порівняти витрати на ресурси різних агенцій, але це виявилося неможливо зробити через те, що надавалася різна допомога та використовувалися різні канали її надання. Замість цього оцінювання базувалося на загальних фінансових наслідках використаних підходів – зокрема, витратах на складування та транспортування. Оцінка також включала розбивку витрачених коштів за секторами та для кожної агенції Комітету – за витратами на запаси й матеріали, витратами на персонал та не пов'язаними з ним, а також адміністративними витратами агенцій.

Блок 1.7 «Вплив: Об'єднане оцінювання екстреної допомоги Руанді»

Об'єднане оцінювання екстреної допомоги Руанді (JEEAR 1996) – наймасштабніша та найбільш комплексна оцінка гуманітарної допомоги в історії. До проведення оцінювання було залучено 52 консультанти та дослідники. Звіт встановив стандарти об'єднаного оцінювання впливу політичної діяльності (бездіяльності) в складних умовах надзвичайних ситуацій.

Вплив оцінювався переважно з точки зору недостатньої участі міжнародної спільноти у подіях в Руанді, незважаючи на достатні сигнали про те, що певні сили в країні готували настрої та структури для геноциду і політичних вбивств. У ході оцінювання гуманітарна допомога розглядалася як комбінація політичних та соціально-економічних заходів. Такий підхід зумовив проведення аналізу політичних структур, що головним чином визначають характер і вплив гуманітарної допомоги.

Розгляд недостатньої участі світової спільноти проводився в двох розрізах: аналіз історичних факторів, що пояснювали причини геноциду, та детальний опис подій, що безпосередньо призвели до геноциду. Цінність спільної оцінки полягає в тому, що замість обмеженого окремими секторами вивчення інтервенцій було проведено аналіз політичної економії. Політично-економічний підхід пов'язаний з оцінкою дієвості гуманітарної допомоги.

Цей підхід можна протиставити підходу, використаному під час оцінювання інших криз – таких як конфлікт у Косово і його наслідки, наслідки урагану Мітч, інтервенції в Афганістані. В кожному з цих випадків приймалися рішення щодо проведення оцінювання окремими агенціями в окремих секторах, через що переважно не вдавалося оцінити політичну природу події та реакцію на неї. У випадках з Косовом та Афганістаном це призвело до нестачі уваги фахівців з оцінювання до питань захисту та прав людини (ALNAP 2001, 2004). У випадку з ураганом Мітч – недостатня увага до того, наскільки гуманітарні заходи відповідають принципам політики трансформацій, запропонованої в рамках Стокгольмської декларації (ALNAP 2002).

Об'єднана оцінка в Руанді є незвичайною з точки зору оцінювання впливу, тому що особливу увагу приділено питанню з'ясування причин відсутності значної зацікавленості у проведенні інтервенції в Руанді (перш за все, через низьке геополітичне значення країни), не зупиняючись лише на переліченні подій та їх наслідків. Одним з отриманих фахівцями з оцінювання уроків стало те, що оцінка впливу повинна стосуватись не лише проведених інтервенцій, а й того, що могло б статися за інших обставин та інших інтервенцій.

Блок 1.8 «Сталість: Оцінювання JICA Програми навчання в сфері інформаційно-комунікаційних технологій для третіх країн»

Японське агентство міжнародного співробітництва (JICA) провело оцінювання проекту на Філіппінах. Цей проект був спрямований на те, щоб надати учасникам з Камбоджі, Лаосу, М'янми та В'єтнаму можливість вдосконалити свої знання та навички в сфері інформаційно-комунікаційних технологій для підприємництва.

За результатами оцінювання (JICA 2005b) сталість було характеризовано як високу, з урахуванням зобов'язання Фонду розвитку інформаційних технологій та освіти (FIT-ED) проводити навчальні програми і в майбутньому для досягнення цілей проекту. FIT-ED організував електронну групу для забезпечення можливості спільної роботи учасників в мережі та обміну інформацією і розширення можливостей FIT-ED. FIT-ED – організація, що займається підвищенням рівня обізнаності з інформаційними технологіями в урядовому та комерційному секторах країн АСЕАН (Асоціації країн Південно-Східної Азії), вона і надалі працюватиме на передньому краї в сфері інформаційно-комунікаційних технологій в регіоні АСЕАН.

Адекватне та своєчасне розподілення Фондом ресурсів для проведення трьох курсів навчання стало підтвердженням його прагнення до збереження програми. Учасники також висловили готовність підтримувати ініціативу. Вони розуміють важливість інформаційно-комунікаційних технологій для свого бізнесу: 84 відсотки опитаних вже застосовували знання та навички, отримані в ході навчальної програми в своїй роботі (зокрема, при розробці веб-сайтів, в комунікаціях, в текстильній та швейній галузях, в імпорті та експорті ремісничих виробів, будівництві, вирощуванні кави, державному управлінні і т.д.). Респонденти повідомляли, що отримали велику користь від курсу, який вони розглядали як початок навчальної програми. Окрім використання стратегічного електронного бізнес-плану, створеного в ході навчальної програми в якості зразка, учасники також використовували мережу Інтернет для застосування отриманих знань у згаданих вище галузях.

Підсумки

Оцінювання проводяться вже сотні років, однак лише нещодавно було спрямовано на наслідки інтервенцій для розвитку.

Існує три форми оцінювання (формативне, сумативне та перспективне) та чотири напрями його використання (етичний, управлінський, для прийняття рішень, освіта та мотивація). Оцінювання може надавати інформацію щодо стратегії (чи правильні дії виконуються?), операцій (чи правильно робиться те, що робиться?) та навчання (чи існують кращі шляхи?). За способом проведення воно є внутрішнім, зовнішнім або спільним.

Оцінювання розвитку утворилося із соціальних досліджень, наукового методу та аудиту. З плином часу роль фахівця з оцінювання змінювалася: спочатку він був аудитором, бухгалтером та контролером, а потім більшою мірою ставав дослідником та координатором спільних оцінок.

Оцінювання розвитку базується на критеріях ОЕСР/КСР – це актуальність, результативність, ефективність, вплив та сталість. ОЕСР/КСР також розробив конкретні принципи оцінювання сприяння розвитку та стандарти якості оцінювання.

Важливою складовою достовірності є незалежність. Керівники міжнародних банків розвитку визначили чотири виміри, або критерії, незалежності оцінювання: організаційна незалежність, поведінкова незалежність, уникнення конфлікту інтересів та захист від стороннього втручання.

Завдання до Глави 1

Практичні вправи 1.1

1. Вас попросили обґрунтувати фінансування оцінювання розвитку за рахунок бюджету нової національної програми. Програма покликана покращити обізнаність родин щодо ефективних заходів з охорони здоров'я. Що б ви сказали на користь оцінювання розвитку?
2. Проведіть інтерв'ю з фахівцем з оцінювання у вашій галузі, щоб з'ясувати, як враховуються стандарти та керівні принципи в оцінках, що їх він або вона спостерігали. (Якщо у вас немає можливості провести таке інтерв'ю, перегляньте останні звіти про оцінки, що проводилися в вашій галузі.) Які ви помічаєте сильні сторони? У чому полягають слабкості? Поділіться своїми думками з колегами-фахівцями з оцінювання, вислухайте їхні зауваження та історії з досвіду. Чи помічаєте ви якісь закономірності?

Примітки

1. Головне контрольно-фінансове управління (General Accounting Office) США в липні 2004 р. було перейменоване в Рахункову палату (Government Accountability Office). При цьому англомовна абревіатура назви (GAO) залишилась незмінною.

Посилання та рекомендована література

Adamo, Abra. 2003. Mainstreaming Gender in IDRC's MINGA Program Initiative: A Formative Evaluation.
https://idl-bnc.idrc.ca/dspace/bitstream/123456789/30972/1/121759.pdf.

ADB (Asian Development Bank). 2007. Mongolia: Second Financial Sector Program. http://www.oecd.org/dataoecd/59/35/39926954.pdf.

ALNAP (Active Learning Network for Accountability and Performance in Humanitarian Action). 2001. Humanitarian Action: Learning from Evaluation. ALNAP Annual Review 2001. ALNAP/Overseas Development Institute, London.

———. 2002. Humanitarian Action: Improved Performance through Improved Learn ing. ALNAP Annual Review 2002. ALNAP/Overseas Development Institute, London.

———. 2006. Evaluating Humanitarian Action Using the OECD/DAC Criteria.
Overseas Development Institute, London.
http://www.odi.org.uk/alnap/publications/eha_dac/pdfs/eha_2006.pdf.

Bhy, Y. Tan Sri Data' Setia Ambrin bin Huang. 2006. The Role of the National Audit Department of Malaysia in Promoting Government Accountability. Paper presented at the Third Symposium of the Asian Organisation of Supreme Audit Institutions (ASOSAI), Shanghai, September 13.
http://apps.emoe.gov.my/bad/NADRole.htm.

Brooks, R. A. 1996. "Blending Two Cultures: State Legislative Auditing and Evaluation." In Evaluation and Auditing: Prospects for Convergence, ed. Carl Wisler, 15–28. New Directions for Evaluation 71 (Fall). San Francisco: Jossey-Bass.

Burkhead, J. 1956. Government Budgeting. New York: John Wiley & Sons.
Callow-Heusser, Catherine. 2002. Digital Resources for Evaluators.
http://www.resources4evaluators.info/CommunitiesOfEvaluators.html.

Chelimsky, Eleanor. 1995. "Preamble: New Dimensions in Evaluation." In Evaluating Country Development Policies and Programs, ed. Robert Picciotto and Ray C. Rist, 3-8. New Approaches for a New Agenda 67 (Fall). Publication of the Ameri can Evaluation Association. San Francisco: Jossey-Bass Publishers.

———. 1997. “The Coming Transformations in Evaluation.” In Evaluation for the 21st Century: A Handbook, eds. E. Chelimsky and W. R. Shadish, pp. 1–26. Thousand Oaks, CA: Sage Publications.

Chelimsky, Eleanor, and William R. Shadish. 1997. Evaluation for the 21st Century: A Handbook. Thousand Oaks, CA: Sage Publications

CIDA (Canadian International Development Agency). 2000. CIDA Evaluation Guide. Ottawa.

Danida, Ministry of Foreign Affairs. 1998, 1999. Guidelines for an Output and Out come Indicator System. Copenhagen: Danida.

DEC (Disasters Emergency Committee). 2001. Independent Evaluation of DEC Mozambique Floods Appeal Funds: March 2000–December 2000. London.

Derlien, Hans-Ulrich. 1999. “Program Evaluation in the Federal Republic of Germany.” In Program Evaluation and the Management of Government: Patterns and Prospects across Eight Nations, ed. Ray C. Rist, 37–52. New Brunswick, NJ: Transaction Publishers.

DfID (Department for International Development). Sector Wide Approaches (SWAps). London. http://www.dfid.gov.uk/mdg/aid-effectiveness/swaps.asp.

———. 2001. Emergency Food Aid to Flood-Affected People in South-Western Bangla desh: Evaluation report. London.

European Federation of Accountants. 2001. The Conceptual Approach to Protecting Auditor Independence. Bruxelles.

Evaluation Cooperation Group of the Multilateral Development Banks. n.d. Good Practice Standards for Evaluation of MDB–Supported Public Sector Operations. Working Group on Evaluation Criteria and Ratings for Public Sector Evaluation. https://wpqp1.adb.org/QuickPlace/ecg/Main.nsf/h_B084A3976FF5F808482571 D90027AD16/1E8F8A36703318324825 7463002F0726/

Feuerstein, M. T. 1986. Partners in Evaluation: Evaluating Development and Community Programs with Participants. London: MacMillan, in association with Teaching Aids at Low Cost.

Fitzpatrick, Jody L., James, R. Sanders, and Blaine R. Worthen. 2004. Program Evaluation: Alternative Approaches and Practical Guidelines. 3rd ed. New York: Pearson Education, Inc.

Fontaine, C., and E. Monnier. 2002. “Evaluation in France.” In International Atlas of Evaluation, eds. Jan-Eric Furubo, Ray C. Rist, and Rolf Sandahl, 63–76. New Brunswick, NJ: Transaction Publishers.

Furubo, Jan-Eric, Ray C. Rist, and Rolf Sandahl, eds. 2002. International Atlas of Evaluation. New Brunswick, NJ: Transaction Publishers.

Furubo, Jan-Eric, and R. Sandahl. 2002. “Coordinated Pluralism.” In International Atlas of Evaluation, eds. Jan-Eric Furubo, Ray C. Rist, and Rolf Sandahl, 115–28. New Brunswick, NJ: Transaction Publishers.

Government of Malawi. 2007. “National Indicators for Routine Monitoring of Qual ity of Health Services at Central Hospital.” http://www.malawi.gov.mw/MoHP/ Information/Central%20Hospital%20Indicators.htm.

Human Rights Education Associated. 1997. Evaluation in the Human Rights Education Field: Getting Started. Netherlands Helsinki Committee. http://www.hrea.org/pubs/EvaluationGuide/.

Insideout. 2005. “M&E In’s and Out’s.” Insideout 3 (October/November) p. 1. http://www.insideoutresearch.co.za/news_l/Newsletter_issue%203.pdf.

Institute of Internal Auditors. 2000. Professional Practices Framework. Altamonte Springs, Florida.

Inter-American Development Bank. 2004. Proposal for Sector-wide Approaches. http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=509733.

INTOSAI (International Organization of Supreme Audit Institutions). n.d. Draft Strategic Plan 2004 to 2009. Vienna, Austria. http://www.gao.gov/cghome/parwi/img4.html.

———. 2001. Code of Ethics and Auditing Standards. Stockholm.

JEEAR (Joint Evaluation of Emergency Assistance to Rwanda). 1996. The International Response to Conflict and Genocide: Lessons from the Rwanda Experience, 5 volumes. JEEAR Steering Committee, Copenhagen.

JICA (Japan International Cooperation Agency). 2005a. JICA Evaluation: Information and Communication. Tokyo. http://www.jica.go.jp/english/operations/evaluation/.

———. 2005b. Results of Evaluation, Achievement of the Project Joint Evaluation of Emergency Assistance to Rwanda. Tokyo.

KRA Corporation. 1997. A Guide to Evaluating Crime Control of Programs in Public Housing. Report prepared for the U.S. Department of Housing and Urban Development. http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/guide_to_ evaluating_crime.html.

Kusek, Jody Zall, and Ray C. Rist. 2004. Ten Steps to a Results-Based Monitoring and Evaluation System. World Bank, Washington, DC. http://www.oecd.org/dataoecd/23/27/35281194.pdf.

Lawrence, J. 1989. “Engaging Recipients in Development Evaluation: The ‘Stakeholder’ Approach.” Evaluation Review 13 (3): 243–56.

MEASURE Evaluation. n.d. Monitoring and Evaluation of Population and Health Programs. University of North Carolina, Chapel Hill. http://www.cpc.unc.edu/measure.

Mikesell, J. L. 1995. Fiscal Administration: Analysis and Applications for the Public Sector, 4th ed. Belmont, CA: Wadsworth Publishing Company.

Molund, Stefan, and Göran Schill 2004. Looking Back, Moving Forward: SIDA Evaluation Manual. Swedish International Development Agency, Stockholm.

OECD (Organisation for Economic Cooperation and Development). 1991a. DAC Criteria for Evaluating Development Assistance. Development Assistance Committee. http://www.oecd.org/document/22/0,2340, en_2649_34435_2086550_1_1_1_1, 00.html.
———. 1991b. Principles for Evaluation of Development Assistance. Development Assistance Committee. http://siteresources.worldbank.org/ EXTGLOREGPARPRO/Resources/ DACPrinciples 1991.pdf.
———. 1998. Review of the DAC Principles for Evaluation of Development Assistance.
Development Assistance Committee. http://www.oecd.org/ dataoecd/31/12/2755284.pdf.
———. 2002. OECD Glossary of Key Terms in Evaluation and Results-Based Management. Development Assistance Committee, Paris.
———. 2006. Evaluation Quality Standards for Test Phase Application. Development Assistance Committee.
http://www.oecd.org/dataoecd/30/62/36596604.pdf.
———. 2007a. General information. http://www.oecd.org/document/48/0,3343, en_2649_201185_1876912_1_1_1_1, 00.html.
———. 2007b. Development Cooperation Directorate DCD-DAC. http://www.oecd.org/department/0,2688,en_2649_33721_1_1_1_1_1,00.html.
Office of the Secretary of Defense Comptroller Center. 2007. The Historical Context. http://www.defenselink.mil/comptroller/icenter/budget/histcontext.htm.
Picciotto, Robert. 2002. “Development Evaluation as a Discipline.” International Program for Development Evaluation Training (IPDET) presentation, Ottawa, July.
———. 2005. “The Value of Evaluation Standards: A Comparative Assessment.” Jour nal of Multidisciplinary Evaluation 3: 30–59. http://evaluation.wmich.edu/jmde/content/JMDE003content/PDFspercent 20JMDE percent20003/4_percent20The_Value_of_Evaluation_ Standards_A_ Comparative_Assessment.pdf.
Quesnel, Jean Serge. 2006. “The Importance of Evaluation Associations and Networks.” In New Trends in Development Evaluation 5, UNICEF Regional Office for Central and Eastern Europe, Commonwealth of Independent States, and International Program Evaluation Network. http://www.unicef.org/ceecis/New_trends_Dev_EValuation.pdf.
Quinn, Michael 1997. Utilization-Focused Evaluation. 3rd ed. Thousand Oaks, CA: Sage Publications.
Rossi, Peter, and Howard Freeman. 1993. Evaluation: A Systematic Approach. Thou sand Oaks, CA: Sage Publications.

Scientific and National Councils for Evaluation. 1999. A Practical Guide to Program and Policy Evaluation. Paris: National Council of Evaluation.
Sonnichsen, R. C. 2000. High-Impact Internal Evaluation. Thousand Oaks, CA: Sage Publications.
Treasury Board of Canada Secretariat. 1993. "Linkages between Audit and Evaluation in Canadian Federal Developments."
http://www.tbs-sct.gc.ca/pubs_pol/dcgpubs/TB_h4/evaluation03_e.asp.
Tyler, C., and J. Willand. 1997. Public Budgeting in America: A Twentieth Century Retrospective. Journal of Public Budgeting, Accounting and Financial Management 9 (2): 189–219.
http://www.ipspr.sc.edu/publication/Budgeting_in_America.htm.
U.S. Department of Housing and Urban Development. 1997. "A Guide to Evaluating Crime Control of Programs in Public Housing." Paper prepared by the KRA Corporation, Washington, DC.
http://www.ojp.usdoj.gov/BJA/evaluation/guide/documents/benefits_of_evaluation.htm.
21.S. GAO (General Accounting Office). 1986. Teenage Pregnancy: 500,000 Births a Year but Few Tested Programs. Washington, DC: U.S. GAO.
———. 2002. "Government Auditing Standards, Amendment 3," GAO, Washington, DC.
Weiss, Carol. 2004. "Identifying the Intended Use(s) of an Evaluation." Evaluation Guideline 6.
http://www.idrc.ca/ev_en.php?ID=58213_201&ID2=DO_TOPIC p 1.
WFP (World Food Programme). 2002. Full Report of the Evaluation of PRRO Somalia 6073.00, Food Aid for Relief and Recovery in Somalia. Rome.
Wisler, Carl, ed. 1996. "Evaluation and Auditing: Prospects for Convergences." In New Directions for Evaluation 71 (Fall), 1–71. San Francisco: Jossey-Bass.
World Bank. 1996. World Bank Participation Sourcebook.
http://www.worldbank.org/wbi/sourcebook/sbhome.htm.

Веб-сайти

Організації з оцінювання

African Evaluation Association. http://www.afrea.org/.
American Evaluation Association. http://www.eval.org.
Australasian Evaluation Society. http://www.aes.asn.au.
Brazilian Evaluation Association. http://www.avaliabrasil.org.br.
Canadian Evaluation Society. http://www.evaluationcanada.ca.
Danish Evaluation Society. http://www.danskevalueringsselskab.dk.
Dutch Evaluation Society. http://www.videnet.nl/.
European Evaluation Society. http://www.europeanevaluation.org.
Finnish Evaluation Society. http://www.finnishevaluationsociety.net/.
French Evaluation Society. http://www.sfe.asso.fr/.
German Society for Evaluation Standards. http://www.degeval.de/.
Institute of Internal Auditors. http://www.theiia.org.
International Development Evaluation Association (IDEAS). http://www.ideas-int.org/. (IDEAS Web page with links to many organizations: http://www.ideasint.org/Links.aspx.)
International Organisation for Cooperation in Evaluation. http://ioce.net/overview/general.shtml.
International Organization of Supreme Audit Institutions (INTOSAI). http://www.intosai.org/.
International Program Evaluation Network (Russia and Newly Independent States). http://www.eval-net.org/.
Israeli Association for Program Evaluation. http://www.iape.org.il.
Italian Evaluation Association. http://www.valutazioneitaliana.it/
Japan Evaluation Society. http://ioce.net/members/eval_associations.shtml
Latin American and Caribbean Programme for Strengthening the Regional Capacity for Evaluation of Rural Poverty Alleviation Projects (PREVAL). http://www.preval.org/.
Malaysian Evaluation Society. http://www.mes.org.my.
Nigerian Network of Monitoring and Evaluation. http://www.pnud.ne/rense/.
Polish Evaluation Society. http://www.pte.org.pl/x.php/1,71/Strona-glowna.html.
Quebecois Society for Program Evaluation. http://www.sqep.ca.
Red de Evaluacion de America Latina y el Caribe. http://www.relacweb.org.
South African Evaluation Network. http://www.afrea.org/webs/southafrica/.
South African Monitoring and Evaluation Association. http://www.samea.org.za.
Spanish Evaluation Society. http://www.sociedadevaluacion.org.
Sri Lankan Evaluation Association. http://www.nsf.ac.lk/sleva/.
Swedish Evaluation Society. http://www.svuf.nu.

Swiss Evaluation Society. http://www.seval.ch/de/index.cfm.
Ugandan Evaluation Association. http://www.ueas.org.
United Kingdom Evaluation Society. http://www.evaluation.org.uk.
Wallonian Society for Evaluation. http://www.prospeval.org.

Стандарти оцінювання

African Evaluation Association. Evaluation Standards and Guidelines. http://www.afrea.org/.
American Evaluation Association. Guiding Principles. http://www.eval.org/Publications/GuidingPrinciples.asp.
Australasian Evaluation Society. Ethical Guidelines for Evaluators. http://www.aes.asn.au/content/ethics_guidelines.pdf.
ECGnet [The Evaluation Cooperation Group] https://wpqp1.adb.org/Quick-Place/ ecg/Main.nsf/h_B084A3976FF5F808482571D90027AD16/1E8F 8A36703318324 8257463002F0726/.
German Society for Evaluation. Standards. http://www.degeval.de/standards/standards.htm.
Italian Evaluation Association. Guidelines. http://www.valutazioneitaliana.it/statuto.htm#Linee.
OECD (Organisation for Economic Cooperation and Development). Evaluation Quality Standards (for test phase application). Development Assistance Committee. http://www.oecd.org/dataoecd/30/62/36596604.pdf.
———. Evaluating Development Cooperation: Summary of Key Norms and Standards. Development Assistance Committee, Network on Development Evaluation. http://www.oecd.org/dac/evaluationnetwork.
Program Evaluation Standards (updated 1998). http://www.eval.org/Evaluation Documents/progeval.html.
Swiss Evaluation Society. Standards. seval.ch/.
UNEG (United Nations Evaluation Group). Norms for Evaluation in the UN System. http://www.uneval.org/docs/ACFFC9F.pdf.
———. Standards for Evaluation in the UN System. http://www.uneval.org/docs/ ACFFCA1.pdf.
UNFPA (United Nations Population Fund). n.d. List of Evaluation Reports and Findings. http://www.unfpa.org/publications/index.cfm.
UNICEF Regional Office for Central and Eastern Europe, Commonwealth of Independent States, and International Program Evaluation Network. 2006. New Trends in Development Evaluation 5. http://www.unicef.org/ceecis/New_trends_Dev_EValuation.pdf.
United Nations Development Project Evaluation Office. http://www.undp.org/eo/.

ГЛАВА 2

Розуміння чинників оцінювання розвитку

Сфера оцінювання розвитку є відносно новою. Вона змінюється відповідно до проблем, які виникають в розвинених країнах та країнах, що розвиваються. В цій главі ми розглянемо деякі з нагальних питань, що зачіпають обидві згадані категорії країн.

Ця глава складається з двох основних частин:

- Огляд оцінювання в розвинених країнах та країнах, що розвиваються
- Наслідки виникаючих проблем розвитку

Огляд оціннювання в розвинених країнах та країнах, що розвиваються

Завдяки оцінюванню країни можуть дізнаватися про те, як добре та якою мірою вони досягають своїх цілей у сфері розвитку, в тому числі Цілей розвитку тисячоліття (ЦРТ). Різні особи (зокрема ті, що формують політику) можуть використовувати висновки та рекомендації за результатами оцінки для ініціювання змін. Оцінювання дозволяє країнам використовувати отриманий досвід для кращого планування та реалізації діючих проектів, програм і політик та зміни напрямів діяльності.

З метою сприяння розвитку багато розвинених країн та країн, що розвиваються, запровадили системи моніторингу та оцінювання (МіО). Ці системи можуть організовуватися по-різному, залежно від потреб та наявних ресурсів.

Оцінювання в розвинених країнах

Більшість із 30 країн Організації економічного співробітництва і розвитку (ОЕСР) нині мають досконалі системи МіО. Процес їх становлення був складним та нелінійним. Існують розбіжності – нерідко серйозні – щодо шляхів, підходів, стилів та рівнів розвитку в цих країнах.

Фурубо, Ріст та Сандаль (2002) описали культури оцінки в країнах ОЕСР для пояснення наявних патернів. Вони розглянули 23 країни за дев'ятьма критеріями, оцінивши кожну з них за шкалою від 0 (низький рівень) до 2 (високий рівень). Ось ці критерії:

1. Оцінка проводиться в багатьох політичних царинах.
2. Існує певна кількість фахівців з оцінювання, які є фахівцями з різних дисциплін, володіють різними методами оцінки та власне її проводять.
3. Існує національний дискурс з оцінки, в рамках якого загальні обговорення адаптуються до національної специфіки.
4. Існує професійна спільнота з власними об'єднаннями або часто відвідуються зустрічі міжнародних об'єднань, або ж принаймні ведеться обговорення професійно-етичних норм.
5. Уряд організував інституціональні механізми проведення оцінок та поширення їх результатів серед осіб, що приймають рішення.
6. В парламенті або інших законодавчих органах діють інституціональні механізми для проведення оцінок та поширення результатів серед осіб, що приймають рішення.

7. Наявний елемент плюралізму, тобто в кожній політичній сфері працюють різні люди або установи, що замовляють та проводять оцінки.
8. Вищий орган фінансового контролю також займається оцінкою.
9. Оцінки не повинні фокусуватися лише на технічних аспектах або на зв'язку між залученими ресурсами та результатами. Необхідно оцінювати, зокрема, результати програм та політик в державному секторі (Furubo, Rist, and Sandahl 2002).

Відповідно до цих критеріїв, Австралія, Канада, Нідерланди, Швеція та США станом на 2002 рік мали найвищі «показники культури оцінювання» серед країн ОЕСР.

Країни ОЕСР створили культуру оцінювання та системи МіО у відповідь на різні фактори внутрішнього та зовнішнього тиску. Франція, Німеччина та Нідерланди, наприклад, створювали культуру оцінювання у відповідь на сильні фактори як внутрішнього, так і зовнішнього тиску (переважно, пов'язані з Європейським Союзом). Водночас такі країни, як Австралія, Канада, Республіка Корея та Сполучені Штати, діяли переважно під впливом сильного внутрішнього тиску (Furubo, Rist, and Sandahl 2002).

Країни ОЕСР першої хвилі прагнули до впровадження культури оцінювання, передусім через сильний внутрішній тиск – мова йде про першочергові потреби у господарчому плануванні, складанні програм і формуванні бюджетів нових соціально-економічних програм, а також про законодавчий нагляд. На формування культури оцінювання в країнах-першопрохідцях вплинули кілька факторів. Багато країн-піонерів у цій сфері до відповідних дій спонукала наявність демократичних політичних систем, міцних емпіричних традицій, держслужбовців з гуманітарною освітою (а не суворо юридичною), а також ефективних адміністративних систем і установ.

Країни, що вирізнялися високим рівнем витрат на освіту, охорону здоров'я і соціальний захист, запровадили механізми оцінки в цих сферах, а потім поширили їх в інших галузях державної політики. Перші країни ОЕСР, що впровадили в себе культуру оцінювання, також активно сприяли її розповсюдженню, поширюючи ідеї та інформацію щодо оцінки і засновуючи організації з оцінювання, навчальні заклади, мережі та консалтингові фірми.

І навпаки, багато країн, які приєдналися до ОЕСР (це, зокрема, Ірландія, Італія та Іспанія), реагували на пов'язані з оцінкою питання переважно через сильний зовнішній тиск – це перш за все вимоги, що їх потрібно виконати для вступу до Європейського Союзу та отримання доступу до фондів структурного розвитку. Ці країни зазнали сильного впливу культури оцінювання країн першої хвилі, а також тих міжнародних організацій, з якими вони взаємодіяли.

Тавістокський інститут (2003) дає опис моделі (або мапи) руху зміцнення потенціалу в сфері оцінки, що складається з чотирьох етапів та проміжних пунктів призначення:

- **Етап 1:** Санкціонування оцінювання
- **Етап 2:** Координація оцінювання
- **Етап 3:** Інституціоналізація оцінювання
- **Етап 4:** Побудова системи оцінювання

Етап 1 зазвичай починається із зовнішнього тиску, що вимагає запровадження оцінювання під впливом норм, правил або цільових завдань. Навіть коли рушійний фактор діє зсередини, імовірним є певний вплив критики ззовні.

Етап 2 включає два види дій у відповідь на формальну та засновану на правилах політику оцінювання з першого етапу. Перший вид – створення керівної документації та базових інструментів, другий – підвищення фахового рівня кадрів з метою покращення якості оцінювання.

Етап 3 зазвичай починається з повноцінної роботи центральної організації. Він складається з двох кроків, що звичайно реалізуються одночасно: створення децентралізованих підрозділів та зміцнення джерел підвищення кваліфікації спеціалістів.

Етап 4 включає побудову повнофункціональної системи оцінювання, де оцінка інтегрується в процеси створення політики, керування програмами та врядування. На цьому етапі також встановлюються міцніші внутрішні зв'язки всередині системи та відкриваються мережі для зовнішніх зацікавлених осіб.

Країни-першопрохідці та наступні члени ОЕСР по-різному підходили до створення систем моніторингу і оцінювання. Використаних підходів було три:

- загальноурядовий підхід;
- анклавний підхід;
- змішаний підхід.

Загальноурядовий підхід:
Встановлення всеосяжної, комплексної системи моніторингу і оцінювання на урядовому рівні

Загальноурядовий підхід

Загальноурядовий підхід було застосовано в деяких з перших країн, що запровадили МіО – наприклад в Австралії. Цей підхід включає встановлення всеосяжної, комплексної системи МіО на урядовому рівні. Інфраструктуру для загальноурядового підходу неможливо створити миттєво; для забезпечення її стійкої інтеграції потрібно щонайменше десять років (World Bank 1999).

Успішність такого підходу вимагає постійної підтримки уряду, вироблення необхідних навичок, а також формування структур і систем державної служби для повноцінного використання результатів МіО. Країни, що розвиваються, повинні також забезпечити постійну підтримку агенцій сприяння розвитку.

Після запровадження ЦРТ багато країн, що розвиваються, прагнуть розробити та впровадити комплексні загальноурядові системи оцінювання. Нині, коли все більше уваги приділяється саме результатам міжнародної допомоги, донори, уряди та інші установи допомагають країнам, що розвиваються, створювати кадровий потенціал та системи оцінювання.

Австралійський підхід до оцінювання. Австралія у 1987 році стала першопрохідцем в розробці систем моніторингу та оцінювання. Країна змогла використати наступні переваги для побудови міцної культури і структури оцінювання:

- високий кадровий, інституціональний та управлінський потенціал в державному секторі;
- державна служба, що відзначається сумлінністю, чесністю та професіоналізмом;
- розвинені фінансові, бюджетні та облікові системи;
- традиція відповідальності та прозорості;
- авторитетні, законні політичні лідери.

Побудові сильних систем оцінки в Австралії сприяли два головних чинники. По-перше, бюджетні обмеження змушували уряд шукати шляхів забезпечення більшої економічної ефективності. По-друге, існували дві важливі організації, що відстоювали необхідність оцінювання – Міністерство фінансів та Національне аудиторське управління Австралії. Підтримка також надходила від членів уряду та ключових міністрів, які надавали велике значення використанню результатів оцінки для прийняття більш інформованих рішень (Mackay 2002).

Початок першого покоління оцінювання (1987–97) припав на часи складних бюджетних умов. Багато реформ державного сектору включали надання автономії департаментам прямої дії та відповідним агенціям, однак моніторинг і оцінювання не проводилися. Через це уряд змусив департаменти та агенції займатися оцінкою. Завдання цього першого покоління систем МіО передбачали допомогу в прийнятті рішень щодо бюджету, зміцнення відповідальності всередині уряду, а також сприяння роботі керівників у міністерствах і відомствах.

Розробку системи МіО першого покоління та управління нею виконувало Міністерство фінансів. Оцінювання кожної програми проводилося один

раз на 3–5 років. Профільні міністри мали підготувати постійні трирічні плани основних оцінок.

Застосовувався широкий спектр видів оцінювання. Станом на середину 1990-х років одночасно проводилося близько 160 оцінок. Формальних вимог щодо збору даних та звітності за показниками ефективності було мало.

Результати всіх оцінок поширювалися серед міністерств і розглядалися дуже серйозно. Фактично близько 80 відсотків нових політичних пропозицій та дві третини проектів забезпечення економії мали вплив на бюджетні рішення уряду. До інших сильних сторін цієї системи належать активне використання результатів оцінки профільними відомствами та агенціями, а також спільна робота щодо заходів з оцінювання.

Проте система першого покоління мала і недоліки. Якість оцінювання була нерівномірною. Підтримка центром додаткової освіти в сфері оцінювання була недостатньою, так само як і формальні вимоги стосовно збору даних та звітності щодо показників ефективності. Міністри заявляли про те, що система стала тягарем для адміністрації.

Початком другого покоління оцінювання в Австралії стали вибори, на яких переміг новий консервативний уряд. Зміни включали значне скорочення штатів державної служби, усунення системи політичних консультантів бюджетного процесу, послаблення централізованого нагляду та «бюрократичних» правил, а також значне скорочення штатів міністерства фінансів, що привело до зменшення його ролі в наданні консультацій під час бюджетного процесу.

Ці зміни потребували відповідних новацій і в системі МіО. Попередню стратегію оцінювання було зруйновано, а сам процес оцінювання «дерегульовано» – тепер його застосування лише заохочувалося, але не було обов'язковим. Особлива увага приділялася моніторингу ефективності результатів. Відповідні звіти надавалися Парламенту як заздалегідь, так і ретроспективно.

Національне аудиторське управління Австралії повідомляло про те, що ефективність системи другого покоління є недостатньою. Збір даних проводився погано через використання слабких стандартів. Недостатньо застосовувалися цільові показники та порівняння. Збиралося багато інформації про наявні результати роботи уряду, але мало – про довгострокові. Не вистачало дійсного аналізу інформації про ефективність. Ці недоліки викликали невдоволення парламентських комітетів. Тим не менш, декілька міністерств (зі справ сім'ї та побутового обслуговування, освіти, охорони здоров'я) все ж змогли скористатися належно проведеними оцінками.

Створення третього покоління (починаючи з 2006 року) було зумовлено занепокоєнням стосовно проблем впровадження комплексних урядових програм та забезпечення «зв'язності» (координації дій міністерств і відомств як на федеральному, так і загальнонаціональному рівнях). Окрім того, Міністерство оборони прагнуло змінити свою роль в процесі формування бюджету і політики (Mackay 2007).

Було запроваджено два види контролю з метою забезпечення ефективності витрат, їх результативності та відповідності пріоритетам уряду. Стратегічні перевірки (сім разів на рік) зосереджувалися на цілях, планах, управлінні, результатах та необхідних майбутніх покращеннях в основних сферах політики та витрат. Програмні перевірки (також сім разів на рік) зосереджувалися на окремих програмах та їх відповідності завданням уряду, дієвості, дублюванні, перекриванні та економії.

Функцію управління перевірками покладено на Міністерство оборони. Було також прийнято рішення щодо оптимізації системи. Уряд виділив 17 млн дол. США на проведення перевірок протягом чотирирічного терміну. Очолити дві пілотні програми запросили відставних державних службовців. Також було збережено вимоги другого покоління щодо системи моніторингу ефективності.

Які висновки можна зробити з еволюції австралійської програми оцінювання?

- Питання координації, реалізації та ефективності (результатів) програми завжди є викликами для будь-якого уряду.
- Методи прийняття рішень на рівні уряду визначають ступінь потреби в інформації МіО (та її перевірці).
- Побудова системи контролю МіО потребує часу.
- Важко збалансувати спрямовані зверху вниз/централізовані та спрямовані знизу вгору/децентралізовані потреби в інформації.
- Більшість міністерств не прагнуть проводити оцінювання, вважаючи їх витратними та небезпечними.

Закон США про роботу уряду та її результати, 1993 р.

Ключовим етапом розвитку оцінювання на рівні уряду в Сполучених Штатах за останні 20 років стало прийняття в 1993 році Закону про роботу уряду та її результати (GPRA), яким було затверджено впровадження оцінювання з орієнтацією на результат в усіх урядових установах. Закон прямо торкається способу проведення оцінювання в органах американського уряду. GPRA передбачає загальноурядовий підхід. Впровадженню змін передували пробні проекти.

Оцінка ефективності в США розпочалася в 1970-х роках на рівні місцевого врядування. Згодом вона поширилася на уряди штатів, а потім і федеральний рівень – з прийняттям у 1993 році GPRA. Федеральний уряд США почав застосовувати оцінку ефективності пізніше, ніж органи врядування інших рівнів.

Цілі GPRA:

1. Зміцнити віру американського народу в дієздатність федерального уряду, систематично перевіряючи досягнення федеральними установами програмних цілей.
2. Ініціювати реформу ефективності програм за рахунок серії пілотних проектів з постановки програмних цілей, вимірювання результатів програм у порівнянні з цілями та публічного звітування про їх результати.
3. Підвищити ефективність та прозорість федеральних програм, змістивши увагу на їх результати, якість послуг та рівень задоволення клієнтів.
4. Допомогти федеральним управлінцям покращити процес надання послуг, вимагаючи від них планування, спрямованого на виконання завдань програми, а також надаючи їм інформацію про результати програми та якість послуг.
5. Покращити якість рішень Конгресу шляхом надання більш об'єктивної інформації про досягнення цілей програм та відносну дієвість і ефективність федеральних програм і витрат.
6. Покращити внутрішнє управління федерального уряду. GPRA спонукав федеральні установи до зосередження уваги на їхніх місіях та цілях, шляхах досягнення цілей та вдосконалення структурних організацій і бізнес-процесів. Згідно із законом, установи повинні подати п'ятирічні стратегічні плани своїх програм та корегувати їх кожні три роки. Вони також повинні виявляти будь-які «ключові зовнішні фактори», що можуть мати значний ефект на здатність установи досягати своїх цілей та виконувати завдання. Установи зобов'язані публікувати щорічні звіти про виконання програм (U.S. Department of Labor, 1993, p. 1).

Установи також повинні вимірювати свою ефективність, щоб розуміти, чи досягають вони поставлених цілей та чи є їх рішення інформованими. Вимірювання ефективності повинно базуватися на пов'язаних з програмами параметрах та бути повним, точним і послідовним. Дані, що збираються, мають використовуватися для покращення організаційних

процесів, виявлення недоліків та встановлення контрольних цілей (U.S. GAO 2003).

Під час проведеного в 2003 році вивчення 16 програм дванадцяти урядових агенцій США було встановлено, що для вдосконалення багатьох федеральних програм вже використовуються регулярно отримувані дані про довгострокові результати. Наприклад, для ініціювання корегуючих заходів, виявлення та заохочування «передових практик», мотивації та визнання заслуг персоналу, а також планування і складання бюджету.

В той же час дослідження виявило певні стійкі перешкоди, що заважають використовувати дані про довгострокові результати, а саме:

- брак зацікавленості або повноважень для внесення змін;
- обмежене розуміння та використання даних про довгострокові результати;
- проблеми з даними про довгострокові результати (застарілі дані, недеталізовані дані, неконкретність даних, потреба в проміжних даних тощо);
- небажання «розгойдувати човен» (Hatry and others 2003).

Дію GPRA в 2003 році було розширено для того, щоб пов'язати ефективність і бюджетування. Уряд також намагався тісніше зводити в часі стратегічне та щорічне планування і звітність за GPRA. Шанна Сора підсумовує, що GPRA це:

> чудова річ. Його вимоги надали міністерствам інструменти для ведення справ у найпростіший раціональний спосіб: встановлювати цілі роботи та вимірювати як довгострокові, так і короткострокові результати. Будь-яка організація, що прагне забезпечити підвищення рівня життя, збільшити кількість послуг та покращити загальну якість їх надання, повинна мати бачення і місію, встановлювати цілі й завдання, а також вимірювати результати (2003, pp. 5–6).

В ході проведеного в 2003 році дослідження Головне контрольно-фінансове управління (GAO) виявило, що багато установ США стикаються зі значними проблемами в ході встановлення цільових результатів організації. Опитані федеральні менеджери повідомили, що керівники установ не завжди демонстрували прагнення до досягнення результатів. Більше того, вони стверджували, що установи не завжди відзначають працівників, які допомагають досягати стратегічних цілей. GAO також повідомило про те, що ефективні організації намагаються змістити головну увагу та відповідальність керівництва з дій та процесів на внески працівників та досягнення результатів. Хоча багато опитаних федераль-

них менеджерів заявили, що несуть відповідальність за результати своїх програм, лише декілька зазначили, що мають достатні повноваження щодо прийняття рішень, аби допомагати установам досягати стратегічних цілей. Нарешті, було встановлено, що хоча менеджери все частіше повідомляли про проведення вимірювання ефективності своїх програм з орієнтацією на результати, вони стали менше використовувати отримувану інформацію в основній управлінській діяльності. Проведене дослідження дозволило виявити потребу трансформування культури організацій таким чином, щоб зробити ці організації більш орієнтованими на результат, зосередженими на потребах клієнтів та співробітництві.

Провідні державні організації в Сполучених Штатах та інших країнах виявили, що стратегічне управління людським капіталом повинно займати центральне місце в будь-якій серйозній ініціативі з управління змінами або роботі щодо перетворення культури державних установ. Системи управління ефективністю є невід'ємною складовою стратегічного управління людським капіталом. Такі системи можуть виступати ключовими інструментами максимізації ефективності: показники ефективності організацій пов'язуються з показниками окремих працівників, встановлюється зв'язок між індивідуальними цілями та цілями організації. Провідні організації роблять свої системи управління ефективністю ключовим інструментом для узгодження ефективності на рівні організації, підрозділу та окремого працівника, досягнення результату, прискорення змін, повсякденного управління організацією, сприяння спілкуванню протягом року таким чином, щоб обговорення індивідуальної та організаційної ефективності було інтегрованим і безперервним.

Анклавний підхід

Можливість запровадження системи оцінювання в різних міністерствах часто є різною. Через це реалізувати загальноурядову стратегію одночасно у всіх міністерствах може виявитися неможливим; в такому випадку потрібно послідовно запроваджувати такі системи в різних міністерствах. Інновації на одному рівні часто проникають горизонтально й вертикально на інші рівні уряду.

Анклавний підхід:
Одночасне запровадження систем моніторингу і оцінювання лише в одній частині чи в одному секторі уряду, наприклад в одному з міністерств

Анклавний підхід передбачає концентрацію на одній частині або секторі уряду, наприклад на одному міністерстві. Приміром, Мексика спрямувала зусилля на соціальний розвиток, Йордан – на планування, Киргизстан – на галузь охорони здоров'я. Робота з одним міністерством, де працює активний прибічник оцінювання, може виявитись оптимальним підходом для країн, яким бракує потенціалу для застосування загальноурядового підходу.

Змішаний підхід

Такі країни, як Ірландія, використовують ***змішаний підхід*** до оцінювання. В той час як певні сфери (зокрема, проекти, що фінансуються структурними фондами ЄС) оцінюються комплексно, іншим сферам приділяється менше уваги. Уряд Ірландії почав створення власної системи оцінювання з використанням анклавного підходу, але зробив вибір на користь більш комплексного підходу стосовно програм державних витрат (Lee 1999). Змішаний підхід також може бути ефективною альтернативою для деяких країн, що розвиваються.

Змішаний підхід:
Комплексне запровадження систем моніторингу і оцінювання в певних частинах уряду, при цьому інші його частини отримують менше уваги

Нарощування потенціалу оцінювання розвитку в Ірландії. Наприкінці 1960-х років Ірландія, як і багато інших країн, була зацікавлена в раціональному аналізі та його застосуванні в плануванні та бюджетному процесі. Керівні кола в уряді виявили потребу в об'єктивних дослідженнях соціальних програм та кадрах, здатних отримати необхідну кваліфікацію для запровадження таких досліджень (Friis 1965; Public Services Organisation Review Group 1969). З метою розвитку навичок в сфері оцінки було реалізовано кілька ініціатив.

Незважаючи на ці зусилля, обсяги оцінювання були обмежені і до кінця 1980-х років мало впливали на прийняття рішень. Це було зумовлено частково відсутністю в Ірландії міцної традиції в сфері оцінки політик і програм, а частково – тим фактом, що оцінювання проводилося в період економічної кризи, коли оцінка як інструмент якісного управління не вважалася настільки важливою, як контроль державних витрат.

Існувало виключення – оцінка витрат ЄС в Ірландії. Кошти ЄС витрачаються через низку операційних програм, що проводяться в рамках спільного плану співпраці Ірландії та ЄС – Програми регіональної і соціальної підтримки (CSF). ЄС – значне джерело фінансової допомоги – вимагав проведення послідовної та систематичної оцінки. Оцінювання фінансованих ЄС програм значно вплинули на дві головні сфери політики: (а) виробниче навчання та схеми створення робочих місць і (б) програми боротьби з бідністю та розвитку громади. У сфері ринку праці в центрі уваги опинилися кількісні вимірювання довгострокових результатів, а в сфері ініціатив розвитку громади – якісні описові методи (Boyle 2005).

Між 1989 та 1993 роками було організовано два незалежних підрозділи з оцінювання – один заснований Європейським соціальним фондом, інший – галузевим підрозділом з оцінювання. З 1989 року оцінювання структурних фондів ЄС було обов'язковою вимогою для отримувачів допомоги. Це привело до подальшого розвитку оцінювання. В період з 1994 по 1999 рік засновано центральний підрозділ з оцінювання, підпорядко-

ваний Міністерству фінансів. Третій підрозділ з оцінювання заснований з метою проведення оцінок в сфері сільського господарства і розвитку сільської місцевості. Були призначені зовнішні фахівці з оцінювання операційних програмних витрат та плану CSF. З 1999 по 2006 рік в Ірландії спостерігалося відновлення інтересу до національної оцінки державних витрат. Можливості центрального підрозділу оцінювання розширилися, як і його обов'язки, а діяльність незалежних підрозділів з оцінювання було припинено. Для проведення проміжного оцінювання операційних програм і плану національного розвитку були залучені зовнішні фахівці з оцінювання (Boyle 2005).

Прийняття нового підходу до оцінювання у Франції. До 2001 року Франція входила до числа країн ОЕСР, що найповільніше просувалися в напрямку створення зрілої системи оцінювання. Більше того, Франція відставала в цій галузі від багатьох перехідних економік та країн, що розвиваються. Наприкінці 1980-х та протягом 1990-х років здійснювалися спроби поступового реформування.

У 2001 році уряд провів значні законодавчі зміни, замінивши принципи організації фінансово-кредитної системи 1959 року, скасувавши постатейний підхід до складання бюджету та запровадивши новий програмний підхід. Нові конституційні регламенти, що поступово вводилися протягом п'ятирічного терміну (з 2001 по 2006 рік), були спрямовані на досягнення двох цілей: (а) реформувати систему державного управління, переорієнтувавши її на результати і ефективність; (б) зміцнити законодавчий нагляд. Тодішній прем'єр-міністр Ліонель Жоспен зазначав: «Представлення бюджету в формі програм, де витрати групуються за основними напрямами державної політики, має представити як членам Парламенту, так і громадянам чітку картину пріоритетів уряду, витрат і результатів дій» (Republic of France 2001). Уряд обрав близько ста програм, на які було виділено фінансові ресурси. Бюджетний документ кожної програми, що подавався на розгляд законодавців, повинен був супроводжуватися переліком конкретних цілей та показників ефективності. Державні менеджери отримали більше свободи та автономії стосовно розподілу ресурсів, але стали нести більшу відповідальність за результати. Таким чином, новий бюджетний процес був орієнтований на результат.

Запити на виділення додаткових коштів з бюджету повинні були включати річні плани з порівнянням очікуваних результатів з фактичними – для кожної програми. Бюджетні огляди також включали річні звіти про виконання фінансового плану. Ці кроки мали покращити здатність законодавців до оцінювання ефективності урядових програм.

Ця ініціатива змінила певні політичні та інституціональні відносини всередині французького уряду, надавши Парламенту більше повноважень щодо формування бюджету. «Стаття 40 Конституції раніше забороняла депутатам пропонувати поправки, що збільшують видатки та скорочують доходи. Вони можуть змінювати розподіл асигнувань серед програм в певній місії». Законодавці можуть ставити на голосування кошториси, асигнування на конкретні місії, обмеження кількості створюваних робочих місць в державному апараті, а також спеціальні рахунки та окремі бюджети. На додачу парламентські фінансові комітети мають стосовно бюджету повноваження в частині моніторингу та нагляду.

Державні службовці негайно відреагували на зміни. Був сформований новий контролюючий апарат, залучені нові обліковці, стали проводитися нові аудити, з'явилися нові анкети стосовно аудиторських контор та інспекторів-аудиторів, стало надходити більше запитів щодо звітування і т. ін. Керівникам було важко звикнути до обмежень, що накладалися необхідністю досягати кількісно вимірюваних короткострокових результатів. При цьому ігнорувалася якість послуг, адже вона в цілях не згадувалася.

Що стосується якості послуг, то «жоден механізм конкуренції або тиск з боку споживачів не гарантує її» (Trosa 2008, p. 8). Суспільства є дуже складними. Деякі люди потребують фінансової допомоги, інші – довіри і можливості брати на себе відповідальність; але є й ті, що всім задоволені. Питання полягає в тому, як збалансувати всі ці чинники в одній формулі (Trosa 2008).

Комбінація попередньої моделі оцінювання з новою не уможливила свободи управління, творчості та інновацій. Троза вказує, що потрібна альтернативна модель. Ще один висновок з французького досвіду полягає в тому, що «неможливо покращити внутрішнє керування, не пов'язавши його з внутрішнім управлінням державними секторами» (Trosa 2008, p. 2). Троза зазначає, що нову систему не треба ламати, її треба розширювати, чітко обговорюючи цілі та заохочуючи логіку дій, а не просто використання інструментів.

Оцінювання в країнах, що розвиваються

Країни, що розвиваються, при побудові власних систем оцінювання стикаються з тими самими проблемами, що й розвинені країни. Проте є й відмінності. Для встановлення та зміцнення системи оцінювання в будь-якій країні необхідні зацікавлені сторони, прагнення до прозорості та належне управління. В країнах, що розвиваються, вимагати та встановлювати систему оцінювання важче.

Слабка політична воля та недостатній інституційний потенціал можуть сповільнювати розвиток. Складнощі співпраці та координації міністерств можуть заважати руху до стратегічного планування. Дійсно, брак достатньої співпраці й координації в уряді зустрічається як в розвинених країнах, так і в країнах, що розвиваються.

Для виникнення та становлення систем оцінювання потрібні політична воля уряду та високопоставлені прибічники ідеї, готові йти на політичні ризики у відстоюванні оцінки. Наявність прибічника або прибічників на загальнодержавному рівні значно покращує перспективи країни з точки зору розробки та підтримання системи оцінювання. І навпаки, нам не відомо про випадки, коли б система МіО виникла в державному секторі країни, що розвивається, не маючи активного прибічника.

Багато країн, що розвиваються, все ще мають труднощі щодо встановлення міцних, дієвих інститутів. У певних випадках може бути необхідною реформа державної служби або нормативно-правової системи. Задля цього міжнародна спільнота фахівців з оцінювання намагається вдосконалювати базові структурні елементи, з яких складається система оцінювання. Проблема полягає в побудові інститутів, проведенні адміністративних реформ та реформ державної служби та/або переробці нормативно-правових кодексів паралельно зі встановленням систем оцінювання. Запровадження систем оцінювання могло б допомогти уряду отримувати більше інформації для проведення необхідних реформ у відповідних сферах.

Країни, що розвиваються, повинні мати або створити фундамент для оцінювання. Багато з них зараз рухаються у цьому напрямку. Створення фундаменту означає наявність базових статистичних систем та даних, а також ключових бюджетних систем. Необхідно забезпечити належні якість і кількість даних та інформації. Як і розвинені країни, країни, що розвиваються, повинні усвідомлювати своє актуальне становище в контексті конкретної програми або політики. Для розвитку, підтримки та зміцнення систем потрібен кадровий потенціал. Необхідно навчити посадовців сучасним методам збору даних, моніторингу й аналізу – в багатьох країнах, що розвиваються, це може бути складно (Schacter 2002).

У відповідь на згадані виклики велика кількість організацій, що займаються наданням допомоги, активізувалися з метою побудови інституціонального потенціалу. До методів відносяться технічна та фінансова допомога в побудові статистичних систем, навчання в сфері моніторингу і оцінювання, діагностичну оцінку готовності та результатів, а також бюджетні системи, орієнтовані на ефективність. Зростаюча поширеність орієнтованих на результат стратегій допомоги країнам може допомогти сформувати необхідні процеси. Допомога країнам, що розвиваються, у виробленні власних стратегій боротьби з бідністю також може сприяти формуванню необхідного потенціалу.

В рамках роботи з підтримки місцевого потенціалу в країнах, що розвиваються, організації з розвитку також переходять до створення мереж розвитку, таких як інтерактивні комп'ютерні мережі та колегіальні спільноти, що обмінюються досвідом та інформацією. Прикладами є «Портал розвитку» (Development Gateway) та Азійська спільнота практиків. Можна заперечити, що обставини в певній країні є унікальними, а досвід однієї країни зовсім не обов'язково застосовувати в іншій. Але якщо прийняти за даність, що загальних знань про розвиток дуже мало, що всі знання потрібно збирати, а потім аналізувати, модифікувати, розбирати та адаптувати під місцеві потреби, джерело цих знань стає несуттєвим. Новий девіз – «Шукайте в усьому світі й заново винаходьте у себе» (Fukuda-Parr, Lopes, and Malik, 2002, p. 18).

Міжнародна навчальна програма щодо оцінювання розвитку

У 1999 році Незалежна група оцінювання Світового банку провела дослідження, щоб дізнатися про наявні можливості навчання в сфері оцінювання розвитку. Було виявлено лише кілька разових програм для організацій з розвитку. Це зумовило створення в 2001 році Міжнародної навчальної програми щодо оцінювання розвитку (IPDET). Ця програма – частина масштабних дій, спрямованих на нарощування потенціалу організацій з розвитку та країн, що розвиваються, в сфері оцінювання.

IPDET готує фахівців, які працюють або збираються розпочати роботу в сфері планування та проведення оцінювання розвитку. Щороку в столиці Канади, місті Оттава, в кампусі Карлтонського університету та за підтримки інших організацій проводиться чотиритижнева навчальна програма, що є продуктом співпраці Незалежної групи оцінювання Світового банку, суспільно-політичного факультету Карлтонського університету та інших організацій, серед яких Канадське агентство міжнародного розвитку (CIDA), британське Міністерство з міжнародного розви-

тку (DfID), Швейцарське агентство з розвитку та співробітництва (SDC), Норвезьке агентство зі співробітництва в цілях розвитку (NORAD), Міжнародний дослідницький центр з питань розвитку (IDRC), Женевський міжнародний центр гуманітарного розмінування (GICHD), Шведське агентство міжнародного розвитку (SIDA), Секретаріат Співдружності, Міністерство закордонних справ Нідерландів, Африканський банк розвитку та Датське агентство з міжнародного розвитку (DANIDA).

Курс IPDET починається з двотижневої основної програми, в якій особлива увага приділяється моніторингу й оцінці реалізації стратегій боротьби з бідністю та робить акцент на МіО з орієнтацією на результат і участь зацікавлених осіб. Основна програма включає більше вісімдесяти годин, протягом яких розглядаються інструменти, вивчаються практичні приклади, проводяться обговорення в групах та читання. Майже третина часу присвячується структурованим заняттям в групах, що дають можливість учасникам зі спільними інтересами можливість разом працювати над реальними питаннями оцінювання розвитку та готувати попередні плани оцінювання для програми, яку один з учасників групи повинен оцінити після повернення до роботи.

Після двотижневої основної програми йдуть два тижні індивідуалізованого навчання за допомогою 30 майстер-класів від поважних та відомих інструкторів. Майстер-класи присвячуються, наприклад, наступним темам:

- планування оцінювання впливу за наявності обмежень;
- планування та побудова систем МіО, орієнтованих на результат;
- розуміння підходів Світового банку до оцінювання країн, секторів та проектів;
- використання змішаних методів оцінювання розвитку;
- використання спільного проведення МіО;
- оцінка постконфліктних ситуацій та проведення міжнародного об’єднаного оцінювання;
- оцінювання програм ВІЛ/СНІД;
- оцінювання прихованих популяцій та найбідніших прошарків населення.

IPDET є однією з небагатьох щорічних програм з оцінювання розвитку. Її пройшли більше 2000 фахівців з міністерств, організацій двосторонньої та багатосторонньої допомоги, неурядових організацій (НУО) та інших суб’єктів, що працюють в сфері розвитку.

Основна програма IPDET пропонується на регіональному та національному рівнях. Індивідуалізовані версії IPDET проводилися, напри-

клад, в Австралії (для австралійських НУО), Ботсвані, Канаді, Китаї, Індії, ПАР, Тунісі, Таїланді, Тринідаді й Тобаго, Уганді. Ці скорочені регіональні програми – відомі як «міні-IPDET» – є дуже інтерактивними, в них застосовується суміш з презентацій, дискусійних груп, групових вправ та вивчення практичних прикладів. Програми покликані забезпечити найширші можливості обміну досвідом з особливою увагою до обговорення проблем та їх рішень з реальної практики.

Завдяки успіху, що його демонструє IPDET через щорічні оцінки та оцінювання впливу, Незалежна група оцінювання співпрацює з міжнародним департаментом Міністерства фінансів Китаю, Азійсько-тихоокеанським центром фінансових департаментів та Азійським банком розвитку над першою регіональною програмою. Ця програма, яку називають «Шанхайська IPDET» (SHIPDET), проводиться двічі на рік – один раз на національному, і один – на регіональному рівні. Її створення є результатом зусиль, спрямованих на зміцнення потенціалу в сфері оцінювання розвитку в країнах та регіонах, що розвиваються.

Нові системи оцінювання

Спроби розробити систему оцінювання та висвітлити розподіл ресурсів і фактичні результати можуть стикатися з політичним спротивом, ворожістю та протидією. З урахуванням природи урядів багатьох країн, що розвиваються, побудова системи оцінювання може також призводити до зміни формату політичних відносин. Створення зрілої системи оцінки вимагає взаємного зв'язку, узгодженості та координації на багатьох щаблях державного управління. Досягнення таких умов може бути проблематичним. У багатьох країнах, що розвиваються, уряди мають слабкі внутрішні зв'язки та все ще працюють в напрямку побудови міцної адміністративної культури та прозорих фінансових систем. В результаті деякі уряди не мають чіткої інформації про кількість та розподіл наявних ресурсів. Вони можуть потребувати детальніших відомостей про те, чи використовуються такі ресурси за призначенням. Вимірювання ефективності роботи уряду в такому середовищі буде давати дуже приблизні результати.

Багато розвинених країн та країн, що розвиваються, все ще працюють над встановленням зв'язків між ефективністю та системою чи стратегією державних видатків. Якщо такі зв'язки не встановлено, неможливо визначити ефективність використання бюджетних коштів у рамках програм, для яких вони виділені.

Деякі країни, що розвиваються, починають досягати в цій сфері певних успіхів. Наприклад, в 1990-х роках Індонезія почала використову-

вати оцінювання як частину розподілу річного бюджету. «Оцінювання розглядається як інструмент коригування політики та програм державних видатків через тіснішу прив'язку до Національного плану розвитку та процесу розподілу ресурсів» (Guerrero 1999, p. 5). Деякі країни з середнім рівнем доходів – серед них Бразилія, Чилі та Туреччина – досягли прогресу в напрямку прив'язки видатків до короткострокових та довгострокових цілей. Уряд Бразилії оприлюднює окремі урядові звіти про довгострокові цілі (OECD and PUMA 2002).

Багато країн, що розвиваються, усе ще працюють з двома бюджетними системами – одна для періодичних витрат, а інша – для капітальних витрат/інвестицій. До недавнього часу в Єгипті Міністерство фінансів займалося бюджетом періодичних витрат, а Міністерство планування – капітальним бюджетом. Об'єднання цих бюджетів під крилом одного міністерства полегшило використання урядом всеосяжної системи оцінювання з метою забезпечення досягнення цілей та виконання завдань, що стоять перед країною.

Враховуючи складнощі запровадження систем оцінювання в країнах, що розвиваються, використання анклавного, або часткового, підходу (за якого кілька міністерств або департаментів випробовують та запроваджують такі системи) може бути раціональнішим за використання загальноурядового підходу. Спроби встановити загальноурядовий підхід до оцінювання можуть бути занадто амбітними. За результатами оцінки готовності, проведеної в 2002 році Світовим банком для Киргизької Республіки, наприклад, було рекомендовано, щоб Міністерство охорони здоров'я (де вже існували певні ресурси в сфері оцінювання) підтримувалося в якості потенційної моделі для подальшого загальноурядового впровадження системи оцінювання.

Китай, Малайзія та Уганда звернулися до анклавного підходу. Їхні дії викладено нижче.

Обумовлене економічним зростанням оцінювання в Китаї. Оцінювання є відносно новим явищем для Китаю. До початку 1980-х років воно було майже невідомим. Ця необізнаність з оцінкою відображала тогочасну орієнтацію суспільствознавчих наук, практично повну відсутність спеціальної літератури китайською, а також брак систематичних контактів між китайцями та фахівцями-практиками в сфері оцінювання з інших країн.

Китайці все ж проводили певні заходи, пов'язані з оцінюванням, в тому числі аналіз політики, економічні та управлінські дослідження, опитування, розгляд результатів проектів та узагальнення досвіду. Інституційний та технічний/аналітичний потенціал існував в деяких політекономічних та науково-дослідних інститутах.

Однак лише 1992 року ключові центральні установи, в тому числі Державне аудиторське управління, Міністерство фінансів та Державна планова комісія, приступили до розробки та висування конкретних пропозицій в Державній Раді щодо нарощування потенціалу МіО. Центр оцінювання науки й технологій разом з Департаментом оцінювання політики та операцій Нідерландів провели першу об'єднану оцінку наукових та технологічних програм (NCSTE and IOB 2004).

Завдяки притоку капіталу й допомогзі у розвитку протягом останніх 20 років Китай досяг підвищення рівня компетентності та розуміння інженерно-технічного аналізу, фінансового аналізу, економічного аналізу й моделювання, аналізу соціального та екологічного впливів, аналізу сталості та досліджень впровадження. Можливості Китаю зросли майже раптово.

Рушійною силою оцінки в Китаї є потужні й стабільні національний розвиток та економічне зростання. Річний внутрішній валовий продукт (ВВП) за 9 років, що передували 2007 року, збільшився більш ніж на 7,8 відсотка. Зацікавленість в оцінюванні пов'язана із занепокоєнням Китаю щодо розвитку. Деякі центральні установи, в тому числі Китайська міжнародна інжинірингово-консалтингова компанія (державна консалтингова фірма), Міністерство будівництва та Державний банк розвитку вже запровадили системи оцінювання на найвищих щаблях своїх організацій.

Хоча більшість оцінювань проводяться після завершення проектів, шириться розуміння того, що оцінку слід проводити на всіх етапах циклу проекту розвитку. В Китаї все більше усвідомлюють, що функція оцінювання може застосовуватися на всіх етапах проекту. Існує зацікавленість в зв'язуванні оцінювання з формулюванням та реалізацією проекту або програми. В деяких випадках оцінювання вже проводиться, хоча зазвичай мова йде не про комплексне оцінювання.

Один з прикладів такого оцінювання мав місце в 2006 році, коли Китай вперше включив компонент системного МіО в п'ятирічний план. Відповідна система, включена в план 11-ї п'ятирічки, базується на 10 кроках, визначених К'юзеком, Рістом та Уайтом (2004).

В квітні 2006 року Китай започаткував програму SHIPDET, що проводиться двічі на рік, з метою підготовки національних та регіональних фахівців з оцінювання. Партнерами в тренінговій програмі виступили Міністерство фінансів, Світовий банк, Азійський банк розвитку та Азійсько-тихоокеанський центр фінансів і розвитку. Китай також закладає підвалини оцінювання, хоча поки що зроблено мало.

Для подальшого розвитку оцінювання в структурі уряду та адміністративній ієрархії країни необхідно виконати кілька ключових завдань:

- запровадження сильної центральної організації для загального керування та координації оцінювання;
- організація формальних підрозділів оцінювання, введення потрібних політик та настанов у всіх відповідних міністерствах і банках;
- усвідомлення того, що настав час для початку проведення власного оцінювання органами місцевого самоуправління;
- запровадження в Державному аудиторському управлінні процесу аудиту функції оцінювання для проведення поточного нагляду та аудиту оцінювання в межах міністерств, а також реалізації положень політик і настанов щодо оцінювання, виданих центральними органами оцінки, відповідними міністерствами, провінціями та банками;
- розробка новітніх методів оцінювання для підрозділів та організаційних одиниць;
- посилення функцій моніторингу та нагляду в інвестиційних агентствах;
- розвиток кадрового забезпечення численних міністерств, провінцій, банків, що починають роботу в сфері оцінювання, кваліфікованими фахівцями з оцінювання (Houqi and Rist 2002).

Китай підняв важливе питання – зростання попиту на результати оцінювання. Це питання є проблемним для багатьох країн, де немає традицій прозорого врядування.

Бюджетування з орієнтацією на довгостроковий результат, національне будівництво та глобальна конкурентоспроможність в Малайзії. Малайзія займає одне із провідних місць серед країн, що розвиваються, з точки зору реформування державного управління, особливо в сферах бюджетування та фінансів. Відповідні реформи було започатковано в 1960-х роках в рамках урядових заходів зі стратегічного розвитку країни. Оскільки державний сектор розглядався як головний рушій розвитку, особливий акцент було зроблено на потребі посилити державну службу за допомогою адміністративної реформи.

У 1969 році в Малайзії було впроваджено Систему бюджетування виконання програм (СБВП), яка діяла до 1990-х років. Ця система замінила постатейний підхід до формування бюджету системою, орієнтованою на довгостроковий результат. Хоча установи використовували програмно-цільову структуру бюджету, її практичне втілення все ще нагадувало постатейне формування бюджету з поетапним підходом.

Бюджетна реформа була спрямована передусім на підвищення рівня відповідальності та фінансової дисципліни різних урядових установ, яким було доручено реалізацію планів соціально-економічного розвитку

Малайзії. Уряд також провів низку додаткових реформ, спрямованих, зокрема, на покращення відповідності вимогам фінансового законодавства, управління якістю, продуктивності та управління національними заходами з розвитку.

Бюджетна реформа Малайзії була тісно пов'язана із зусиллями щодо національного будівництва та підвищення глобальної конкурентоспроможності. Однією з рушійних сил цих заходів стала програма «Vision 2020», що передбачає набуття Малайзією статусу розвиненої країни до 2020 року.

У 1990 році уряд замінив СБВП Модифікованою бюджетною системою (МБС). СБВП передбачала мінімальний зв'язок між задіяними ресурсами та короткостроковими результатами; різні політики продовжували фінансуватися навіть за відсутності систематичного відстеження результатів. МБС більше уваги приділила результатам та впливу урядових програм і заходів.

Наприкінці 1990-х в Малайзії було створено власну інтегровану систему управління з орієнтацією на результати (ІСУОР). Компонентами цієї системи стали орієнтовані на результати бюджетна система та система оцінки роботи персоналу. Також було створено дві інших системи з орієнтацією на результат. Це управлінська інформаційна система (УІС) та система МіО, що мали доповнити державну систему управління (Thomas 2007).

Система ІСУОР забезпечує структуру для планування, реалізації, моніторингу та формування звітності щодо організаційної ефективності. Вона встановлює зв'язок між організаційною ефективністю та показниками роботи персоналу. Ціла низка країн (серед яких Афганістан, Ботсвана, Індія, Маврикій та Намібія) поетапно впроваджують ІСУОР, першочергову увагу приділяючи орієнтованому на результат формуванню бюджету та МіО. У Малайзії головним рушієм системи ІСУОР є бюджетна система (Thomas 2007).

Системи УІС та МіО, що доповнюють систему стратегічного планування елементом оцінки ефективності, роблять ІСУОР динамічнішою. УІС та МіО тісно пов'язані. Мета – забезпечити надання системою необхідної інформації потрібним особам у потрібний час. Показники мають бути операційними та орієнтованими на результат. У Малайзії створили та почали застосовувати електронну версію інтегрованої системи управління ефективністю (Thomas 2007).

Малайзійський уряд представив декілька основних висновків зі свого досвіду:

- Необхідно забезпечити функціонування програми з нарощування кадрового потенціалу для ключових щаблів уряду.
- Моніторинг та звітність є трудомісткими.
- Процес планування роботи потребує зміцнення: він повинен бути більше комплексним, ніж поетапним.
- Винагороди та рівні покарання на всіх рівнях різні.
- Інтеграція з іншими ініціативами є обмеженою (Rasappan 2007).

Також було запропоновано кілька рекомендацій:

- Працювати в напрямку встановлення повних вертикальних та горизонтальних зв'язків.
- Уникати розчарувань як на рівні політик, так і на операційному рівні.
- Переглядати та зміцнювати всі допоміжні політики та системи.
- Працювати над тим, щоб інтегрована система управління з орієнтацією на результати зосереджувалася на загальноурядовій ефективності (Rasappan 2007).

Хоча Малайзія є однією з провідних країн з точки зору реформування державного управління, реформи в країні проводилися складно і не завжди послідовно. Тим не менш, МБС була важливою ініціативою уряду, що продемонструвала далекоглядність, інноваційний підхід, динамічність та прагнення до ефективного використання коштів під час реалізації проектів і політик (World Bank 2001).

Боротьба з бідністю як поштовх до розвитку оцінювання в Уганді. Уряд Уганди прагнув забезпечити ефективність надання державою послуг для боротьби з бідністю. Визнання ефективності надання послуг ключовим фактором управління національним розвитком є переконливим свідченням прагнення до результатів. Це прагнення також виявилося в кількох поточних пріоритетах і заходах державного управління.

За останнє десятиліття в Уганді було проведено комплексну економічну реформу та досягнуто макроекономічної стабілізації. Згідно з Концепцією комплексного розвитку в країні розроблено План заходів з викорінення бідності (ПЗВБ), включений до Стратегії боротьби з бідністю. ПЗБВ вимагає скорочення рівня абсолютної бідності з 44% станом на кінець 1990-х років до 10% до 2017 року. Основні цілі ПЗБВ загалом збігаються з ЦРТ; і ПЗБВ, і ЦРТ спрямовані на запровадження відповідальності уряду та його партнерів за хід розвитку.

Уганда стала першою країною, яка отримала право скористатися Ініціативою для бідних країн з великою заборгованістю (БКВЗ) та власне скористалася нею. В 2000 році вона отримала право на збільшену допо-

могу для БКВЗ – це стало визнанням ефективності угандійської стратегії боротьби з бідністю, в рамках якої громадянське суспільство залучалося до консультативного процесу, а також сталого прагнення до забезпечення макроекономічної стабільності.

Уганда запровадила нові заходи, покликані зробити бюджетний процес відкритішим та прозорішим для внутрішніх та зовнішніх зацікавлених сторін. Уряд модернізує фіскальні системи і запроваджує програму децентралізації планування, управління ресурсами та надання послуг на місцях. Міністерство фінансів, планування та економічного розвитку (МФПЕР) також вводить в дію орієнтований на результат підхід до формування бюджету. На додачу передбачається зміцнення урядових установ та посилення їх відповідальності перед суспільством.

Країна все ще стикається з проблемами координації та гармонізації оцінки і ПЗВБ. «Найочевидніша характеристика режиму МіО ПЗВБ – розмежування моніторингу бідності та моніторингу ресурсів, хоча обидва ці напрямки координуються МФПЕР. В цих двох галузях МіО діють різні суб'єкти, використовуються різні звіти та різні критерії оцінки. Моніторинг фінансових ресурсів пов'язаний із залученими засобами, діями та все більшою мірою результатами, в той час як моніторинг бідності базується на аналізі загальних результатів боротьби з бідністю» (Hauge 2001). Інші проблеми координації оцінки пов'язані з утворенням нового національного органу з планування та галузевих робочих груп.

Наприкінці 2007 року Канцелярія прем'єр-міністра (КПМ) представила робочу записку для обговорення МіО національного плану розвитку. Цей документ мав на меті аналіз сильних і слабких сторін ПЗВБ та пошук шляхів розвитку МіО нового національного плану (Uganda OPM 2007a).

У робочій записці зазначалося, що всередині системи існує кілька проблем:

- «Довгострокові та короткострокові результати галузевих міністерств, вимірні показники з відповідними базисами та цілями, ефективні системи моніторингу та стратегічне використання оцінювання для вимірювання ефективності й встановлення причинно-наслідкових зв'язків» (Uganda OPM 2007b, p. 4) чітко не визначені.
- Ведеться контроль і облік витрат, а не істотних показників роботи.
- Обсяг даних, що збираються, не збалансований потребою в даних та рівнем здатності використовувати дані.
- Через те, що моніторинг проводився неузгоджено та з дублюванням, інспекційна діяльність, збір даних та складні формати звітності стали обтяжливим тягарем. Результатом цього став великий обсяг даних щодо відповідності нормам і правилам, які не дають чіткої бази для

оцінки економічної ефективності в державному секторі. В дослідженні, проведеному угандійським КПМ (Uganda OPM 2007a, p. 6), повідомляється, що «причини поганої координації та дублювання роботи, можливо, більше пов'язані з діючою в державній службі системою оплати праці, коли ведення моніторингової діяльності частково зумовлювалося бажанням працівників отримувати добові – як важливий додаток до заробітної плати».

- Брак заохочень та випадки перекриття повноважень в питаннях планування та інших заходів у сфері МіО стали на заваді створенню загальнодержавної робочої групи для вирішення проблем МіО.
- Попри численність оцінювань, вони зазвичай обмежувалися секторами та міністерствами і не були орієнтовані на загальні стандарти.
- Люди на місцях все ще не відчувають, що їх залучають до прийняття рішень.

У робочій записці також було наведено кілька рекомендацій:

- Встановити зв'язок між розподілом бюджетних коштів та досягненням результатів.
- Розглянути можливість запровадження договорів на надання послуг в державному секторі або договорів про результати роботи.
- Своєчасно та в практичній формі надавати інформацію урядовцям.
- Забезпечити відповідність надаваних даних потребі в інформації.
- Запровадити механізми контролю і забезпечення якості.
- Забезпечити корисність аналізу для урядовців.
- Розмежувати функції моніторингу й оцінювання.
- Роз'яснити ролі та сфери відповідальності в уряді стосовно планування, моніторингу, оцінювання та інших функцій із забезпечення якості.

Якщо ж говорити про оцінювання в майбутньому, то Уганда має відстежувати свій рух в напрямку скорочення бідності в рамках ПЗВБ та Національної стратегії боротьби з бідністю і робити належні висновки. Оцінювання неможливо відокремити від практики та стимулів у сфері прийняття рішень, притаманних національним системам і процесам розвитку (Hauge 2001).

Значення виникаючих проблем розвитку

Проблеми розвитку ускладнюють оцінювання. В цьому розділі буде подано короткий огляд згаданих проблем з описом їх наслідків для оцінювання.

Петон (2006) починає розповідь про актуальні тенденції в оцінюванні з визначення її як глобального суспільного блага. Він описує зростання фахових організацій, об'єднань та спілок з оцінки по всьому світі, а також встановлення цими організаціями стандартів та керівних вказівок. Він також вказує на розробку більш ніж ста нових моделей оцінювання, вбачаючи в цьому нову тенденцію.

Ілюструючи зростаючу складність оцінювання, Петон використовує аналогію. В минулому фахівці для проведення оцінювання часто могли скористатися певним рецептом. Петон вказує на вигоди від використання рецептів:

- рецепти перевірені часом, вони забезпечують відтворюваність;
- хоча про якусь особливу кваліфікацію не йдеться, вміння готувати підвищує успішність;
- рецепти на виході дають стандартні продукти.

Рецепти є ефективними в приготуванні їжі, однак вони не забезпечують стандартних результатів в оцінюванні, де необхідною є модель, що передбачає певну реакцію фахівця з оцінювання на складні питання. Аналогія, що її Петон проводить для нової тенденції в оцінюванні – виховання дитини. Рецепт описує процес крок за кроком. Виховання дитини – це, навпаки, дуже складний процес, в ході якого вихователі використовують свої знання для прийняття рішень та реагування на нові обставини.

Петон описує ще одну тенденцію в розвитку оцінювання: перехід до більш формативних ситуацій. До таких формативних ситуацій в оцінюванні він відносить ситуації, коли:

- вказано цільові та очікувані довгострокові результати, але техніки їх вимірювання лише випробовуються;
- модель для досягнення довгострокових результатів викладається у вигляді гіпотези, випробовується та доробляється;
- реалізація інтервенції не стандартизується, але вивчається та вдосконалюється в міру того, як вирішуються проблеми інтервенції (це свого роду ітеративний підхід);
- атрибуція формулюється з можливістю перевірки причинно-наслідкового зв'язку.

Ще одна нова тенденція – перехід від індивідуальних оцінювальних досліджень до серійних (Rist and Stame 2006). Ріст і Стаме описують, як спільнота фахівців з оцінювання почала використовувати системний підхід замість того, щоб покладатися на окремих фахівців з оцінювання чи окремі оцінки для формування оціночних знань. Основною їхньою тезою є те, що просте накопичування звітів про оцінювання мало або й зовсім ніяк не впливає на підсумкові знання, які можна було б отримати, синтезуючи результати тих самих досліджень.

Зміни в розвитку впливають на оцінювання. Напрямки розвитку й надалі будуть змінюватися у відповідь на поточні та нові питання, серед яких глобалізація, почастішання конфліктів у світі, тероризм та відмивання грошей, зростаючий розрив між бідними та багатими країнами, зростання кількості гравців у галузі розвитку, рух в напрямку скорочення боргів, а також нове підвищення уваги до вдосконалення управлінських підходів. Вирішення цих питань висуває до фахівців з оцінювання нові вимоги.

Глобальний рух в напрямку комплексного, координованого, спільного розвитку та демонстрації осяжних результатів також кидає спільноті фахівців з оцінювання нові виклики. Відбулися значні зрушення – перехід від часткового до комплексного розвитку, від індивідуального до координованого підходу (партнерства), від сприяння зростанню до скорочення бідності, від фокусування на реалізації до фокусування на результатах. Що стосується комплексного розвитку. Наприклад, донори, що надають допомогу як на двосторонній, так і багатосторонній основі, «тепер повинні розглядати кожен проект в ширшому контексті та оцінювати його життєздатність і потенціальний вплив на суспільство, політику й економіку в цілому» (Takamasa and Masanori 2003, p. 6). Зазначається наступне:

> Теоретики розвитку тепер вважають найважливішим фактором економічного розвитку не капітал, а належні стратегії та інститути. Цю зміну поглядів спричинив потужний вплив таких економістів, як Норс..., Стігліц... і Сен на економічну науку, в тому числі на економіку розвитку. Ці зміни привели до утворення нинішньої ситуації, коли центральною темою міжнародного сприяння розвитку є боротьба з бідністю (у широкому сенсі), що включає забезпечення людської гідності та політичних і економічних свобод народів країн, що розвиваються.

ЦРТ – це конкретний прояв нового мислення в галузі розвитку. *Звіт про світовий розвиток* 2005 р. (World Bank 2005b) концентрується на тому, що можуть робити уряди для покращення інвестиційного клімату в своїх країнах, на вимірюванні прогресу за допомогою комплексів показників, покликаних оцінювати різні елементи ділового клімату. В звіті

рекомендується впроваджувати інституційні та поведінкові покращення: створювати кращі регулятивні та податкові умови, усувати перепони конкуренції, активніше стимулювати бізнес, боротися з корупцією, сприяти встановленню довіри й законності в суспільстві, а також забезпечувати належну реалізацію нормативних актів і законів.

Вирішення багатьох з цих нових питань сприяння розвитку потребує участі партнерів, в тому числі міжнародних, та означає покладання на країни, що розвиваються, необхідність проведення численних оцінок. Це є підґрунтям для проведення об'єднаних міжнародних оцінок. Такі оцінювання можуть проводитися на рівні проекту, країни, сектору або певної теми. Вони можуть приносити організаціям з розвитку ефективність витрат і економію від масштабу, а також гармонізувати методи оцінки, що забезпечує порівнянність результатів.

Далі надається короткий огляд деяких з основних рушійних сил міжнародного розвитку та їх наслідків для оцінювання. До таких рушійних сил відносяться:

- ЦРТ;
- Монтерейський консенсус;
- Паризька декларація щодо підвищення ефективності зовнішньої допомоги;
- Ініціатива щодо бідних країн з великою заборгованістю;
- роль фондів;
- запобігання конфліктам та відновлення після завершення конфліктів;
- управління;
- боротьба з відмиванням грошей та фінансуванням тероризму;
- перекази коштів працівників;
- гендерні питання;
- розвиток приватного сектору економіки та інвестиційного клімату;
- екологічна та соціальна сталість;
- глобальні суспільні блага.

Автори зазначають, що перелік рушійних сил постійно змінюється. На момент виходу в друк цієї книги ключовим рушійним фактором, який міг би тепер очолювати наведений вище список, є наслідки світової фінансової кризи.

Цілі розвитку тисячоліття

ЦРТ:
Набір цілей розвитку, що їх має досягти міжнародна спільнота до 2015 року

У вересні 2000 року 189 країн-членів ООН та численні міжнародні організації затвердили Декларацію тисячоліття Організації об'єднаних націй, на основі якої частково й було сформовано ЦРТ. **ЦРТ** складаються з набору цілей розвитку, що їх має досягти міжнародна спільнота до 2015 року в результаті активної діяльності як розвинених країн, так і країн, що розвиваються (блок 2.1). Ці амбітні цілі спрямовані на подолання бідності, гуманітарний розвиток та створення глобальних партнерств для виконання обох цих завдань. Це відхід від попередньої спрямованості розвитку, перш за все на економічне зростання, яке, як очікувалося, мало допомогти людям подолати бідність. ЦРТ передбачають комплекс заходів, спрямованих на подолання бідності та покращення умов життя бідного населення у світі.

Блок 2.1 «Цілі розвитку тисячоліття»

1. Подолання крайньої бідності і голоду.
2. Забезпечення загального доступу до освіти.
3. Забезпечення гендерної рівності та прав і можливостей жінок.
4. Зменшення дитячої смертності.
5. Поліпшення системи охорони здоров'я матерів.
6. Обмеження поширення ВІЛ-інфекції/СНІДу, малярії та інших хвороб і започаткування тенденції до скорочення їх масштабів.
7. Забезпечення сталого розвитку довкілля.
8. Розвиток глобальної співпраці задля розвитку.

Джерело: http://www.unorg/millenniumgoals.

Вісім ЦРТ включають в себе 18 конкретних цілей та 48 показників для вимірювання ходу їх виконання. (Країни, що розвиваються, використовують різні комбінації 18 конкретних цілей та різні граничні терміни їх досягнення залежно від ситуації, що склалася в цих країнах.) ЦРТ – це прив'язані до результатів цілі, що мають належним чином вимірюватися, відстежуватися та оцінюватися. Вони створюють серйозні виклики для систем оцінки в усіх країнах.

Багатьом країнам, що розвиваються, бракує потенціалу для проведення МіО. Для вирішення цієї проблеми організації з розвитку сприяють нарощуванню потенціалу в сфері статистики та МіО, надають технічну допомогу та підтримку.

ЦРТ стимулюють країни, що розвиваються, до нарощування потенціалу та створення систем МіО. Організації з розвитку покликані надавати технічну допомогу та фінансування. Багато країн, що розвиваються, лише нещодавно приступили до побудови власних систем МіО і повільно просуваються до створення систем з орієнтацією на результати, що допоможуть визначити, якою мірою виконуються завдання в рамках ЦРТ. Оцінка успішності вимагатиме розвитку та ефективного використання систем оцінки. Система оцінювання, в свою чергу, повинна бути інтегрована в галузь політики ЦРТ таким чином, щоб «всім було ясно, чому важливо збирати дані, та яким чином інформацію використовуватимуть для забезпечення інформованості уряду та громадянського суспільства в їх діяльності щодо досягнення ЦРТ» (Kusek, Rist, and White 2004, pp. 17–18).

Світовий банк та Міжнародний валютний фонд (МВФ) щороку публікують *Глобальний звіт моніторингу* ЦРТ. Цей звіт слугує основою звітності щодо глобальної політики розвитку.

Глобальний звіт моніторингу 2004 року був присвячений тому, як проходить у світі реалізація політик та заходів задля досягнення ЦРТ, і відповідним довгостроковим результатам розвитку. Звіт звертає увагу на декілька пріоритетних напрямів посилення моніторингу:

- зміцнення основної статистики розвитку, в тому числі за рахунок своєчасної реалізації плану дій, погодженому з міжнародними статистичними органами;
- проведення досліджень щодо вирішальних чинників ЦРТ, критичних питань (таких, як дієвість допомоги), розробки надійніших систем показників для ключових політичних галузей, таких як управління та вплив політики багатих країн на країни, що розвиваються;
- поглиблення співробітництва з партнерськими організаціями в цій сфері, створення порівняльних переваг, забезпечення логічного зв'язку у підході до моніторингу й оцінки різних установ (World Bank 2004b).

Глобальний звіт моніторингу 2005 року вказав на можливості, створені нещодавнім покращенням економічних показників у країнах, що розвиваються. В ньому було наведено перелік з п'яти заходів, покликаних прискорити прогрес:

- *Забезпечити відповідальність країни за діяльність з розвитку*. Активізувати розвиток за рахунок стратегій боротьби з бідністю, за які має відповідати країна та які вона повинна проводити.

- *Поліпшити середовище для економічного зростання, локомотивом якого має виступити приватний сектор*. Зміцнити фіскальне управління, зробити вільнішим бізнес-середовище та інвестувати кошти в інфраструктуру.
- *Активізувати надання основних послуг для населення*. Різко збільшити штат працівників сфери охорони здоров'я та освіти, забезпечити більше та гнучкіше й передбачуваніше фінансування таких постійних сервісів, що потребують значних витрат, а також зміцнити інституціональний потенціал.
- *Знищити торгові бар'єри*. Використати надбання Доського раунду, що включають серйозну реформу сільськогосподарської торгівлі, знищити торгові бар'єри та активізувати роботу за програмою «Сприяння торгівлі».
- *Вдвічі збільшити допомогу для розвитку протягом наступних п'яти років.* Також підвищити якість допомоги, прискоривши розвиток координації та гармонізації такої допомоги (World Bank 2005a).

У центрі уваги Глобального звіту про моніторинг 2006 року – економічне зростання, покращення якості допомоги, торгівельні реформи та управління як ключові чинники досягнення ЦРТ (World Bank 2006a). У звіті 2007 року йшлося про дві основні проблемні сфери: забезпечення гендерної рівності й прав та можливостей жінок (третя ЦРТ) та специфічні проблеми нестабільних держав, які все більше стають місцем зосередження крайньої бідності.

Монтерейський консенсус

В березні 2002 року представники урядів більш як 170 країн, в тому числі більше 50 глав держав, провели зустріч, яка мала на меті обговорення проекту Монтерейського консенсусу з фінансування розвитку. Проект відображав спробу виділити більше коштів на користь найбіднішого населення світу, представникам якого доводиться жити менш ніж на 1 долар США на день.

Найбільш важливим з точки зору оцінювання розвитку є те, що Монтерейський консенсус поставив акцент на взаємній відповідальності щодо досягнення ЦРТ. Він закликав країни, що розвиваються, до вдосконалення політик та методик управління, а розвинені країни – до посилення підтримки, зокрема через відкриття своїх ринків та надання кількісно більшої та якіснішої допомоги. Цей документ визнав потребу в посиленні фінансової підтримки задля підвищення рівня життя в найбідніших країнах, але не встановив чітких цілей щодо нарощування допомоги, полегшення тягаря заборгованості більшості країн або усунення торгових бар'єрів (Qureshi 2004).

В середині строку – між роком прийняття ЦРТ та їхньою цільовою датою (2015 р.) – Економічна комісія для Африки опублікувала звіт, в якому оцінюється рух африканських країн в напрямку виконання зобов'язань за Монтерейським консенсусом. У звіті надано висновок про те, що в царині полегшення зовнішніх боргових тягарів було досягнуто значних успіхів, але в інших основних напрямах, встановлених консенсусом, прогрес виявився дуже обмеженим. Для фахівців з оцінювання цікавими є наступні зауваження, наведені у звіті:

> Існує розуміння стосовно того, що моніторинг зобов'язань, взятих на себе як африканськими країнами, так і їхніми партнерами з розвитку, є ключовим для реалізації цілей, визначених Монтерейським консенсусом. Африканські лідери визнали це і запровадили механізм моніторингу ходу виконання їхніх зобов'язань та зобов'язань партнерів з розвитку. Інституціоналізація Конференції міністрів африканських країн з фінансування розвитку є сміливим кроком африканських лідерів у цій сфері. Міжнародна спільнота також запровадила механізм моніторингу роботи донорів. Наприклад, було засновано Форум африканського партнерства та Африканський комітет з питань прогресу, які здійснюватимуть моніторинг ходу реалізації ключових зобов'язань щодо фінансування розвитку. Нарешті, дієвість цих механізмів моніторингу повинна буде оцінюватися з точки зору їх здатності втілювати обіцянки партнерів з розвитку в життя. Тому що тільки реалізуючи ці зобов'язання африканські країни та міжнародна спільнота зможуть подолати бідність в регіоні та закласти підвалини кращого майбутнього для його населення (Katjomulse and others 2007).

Паризька декларація щодо підвищення ефективності зовнішньої допомоги

Паризька декларація щодо підвищення ефективності зовнішньої допомоги – міжнародна угода про продовження активізації зусиль з управління допомогою країнам, що розвиваються. Більше ста міністрів, керівників установ та інших високих посадовців підписали цю угоду 2 березня 2005 року.

Однією з важливих, з точки зору оцінки декларації, стала згода використовувати в рамках реалізації угоди дії та показники, що піддаються моніторингу. Для відстеження та заохочування просування в напрямку підвищення ефективності допомоги було розроблено дванадцять показників. Для одинадцяти з них – встановлено цільові значення на 2010 рік (OECD 2005b).

Прийняті показники та їх цільові значення групуються навколо п'яти основних принципів:

- *Відповідальність*. Партнерські країни здійснюють керівництво власними політиками та стратегіями розвитку, а також координують відповідні заходи.
- *Узгодження.* Організації з розвитку надають підтримку, виходячи з національних стратегій розвитку країн-партнерів, беручи до уваги їх установи та процедури.
- *Гармонізація.* Заходи організацій з розвитку більш гармонізовані, прозорі та сукупно дієвіші.
- *Управління з орієнтацією на результат.* Уряди рухаються в напрямку збільшення уваги до управління ресурсами та вдосконалення прийняття рішень задля забезпечення кращих результатів.
- *Взаємна відповідальність.* Організації з розвитку та їх партнери відповідають за результати розвитку (Joint Progress toward Enhanced Aid Effectiveness 2005).

У 2007 році ОЕСР оприлюднила етапний звіт, що підсумовує результати базисного дослідження стану речей у 2005 році. Звіт (OECD 2007b) оцінює дієвість допомоги як в глобальному розрізі, так і стосовно організацій з розвитку.

ОЕСР провела дослідження з метою моніторингу ходу підвищення дієвості допомоги відповідно до положень Монтерейського консенсусу з урахуванням уточнень, внесених Паризькою декларацією. Було встановлено наступне:

- Паризька декларація допомогла підвищити обізнаність та активізувати на рівні країн діалог щодо потреби вдосконалення надання допомоги та керування нею.
- Швидкість змін поглядів та підходів донорів до керування допомогою була невисокою, а операційні витрати, пов'язані з наданням допомоги та керуванням нею, залишаються дуже високими.
- Необхідно зміцнити національні стратегії розвитку, покращити узгодженість донорської підтримки та місцевих пріоритетів, підвищити надійність бюджету як інструменту для управління та розподілу ресурсів, а також добитися більшої точності бюджетних оцінок потоків допомоги.
- Зміна способів надання допомоги та керування нею вимагає додаткових витрат – донори та партнери повинні мати це на увазі.
- Країни та донори повинні використовувати системи оцінки ефективності та застосовувати більш економну звітність з орієнтацією на результати. Донори повинні брати участь в нарощуванні потенціалу та активніше використовувати системи звітності країн.
- Для забезпечення взаємної відповідальності необхідно створити надійніші системи моніторингу (Katjomulse and others 2007).

Ініціатива щодо бідних країн з великою заборгованістю

У 1996 році Світовий банк та МВФ запропонували Ініціативу щодо бідних країн з великою заборгованістю – перший комплексний захід, спрямований на зменшення зовнішнього боргу найбідніших країн світу, а також країн з найбільшим борговим тягарем. Ініціативу підтримали 180 країн.

Вона спрямована на зниження заборгованості бідних країн, що проводять економічні та соціально-політичні реформи, до сталого рівня. Ініціатива застосовується у випадках, коли традиційні механізми зменшення заборгованості не здатні вирвати країни з процесу постійної реструктуризації боргів. Ініціатива щодо бідних країн з великою заборгованістю знижує розмір витрат на обслуговування боргів та збільшує обсяг соціальних виплат.

Ініціатива охоплює як двостороннє, так і багатостороннє списання заборгованості. Очікується, що обслуговування зовнішнього боргу країн, охоплених ініціативою, буде зменшене приблизно на 50 млрд доларів США. Станом на січень 2009 року затверджено пакетні заходи зі скорочення заборгованості для 34 країн, 28 з яких – африканські; також визначено, що ще сім країн також можуть розраховувати на допомогу (IMF 2009).

Ініціатива щодо бідних країн з великою заборгованістю пов'язана з комплексними національними стратегіями скорочення бідності. В 1999 році міжнародна спільнота з розвитку погодилася з тим, що національні документи стратегії скорочення бідності (ДССБ) повинні стати основою для пільгового кредитування та списання заборгованості. Ці стратегії встановлюють погоджені цілі розвитку на трирічний період. Вони також включають таблицю заходів економічної політики, набір відповідних вимірних показників, а також описують систему МіО для відстеження прогресу. Якщо країна досягає поставлених цілей, борги списуються, стимулюючи її до прискорення реформ та підвищення відповідальності уряду за відповідні заходи. Умовою списання боргів є здатність уряду-отримувача проводити моніторинг і оцінку та звітувати про хід реформ та рух у напрямку подолання бідності. Ця умова спричинила зростання попиту на нарощування потенціалу МіО та отримання відповідної допомоги.

Деякі країни, що розвиваються – такі як Уганда – домоглися успіху в розвитку оцінки і отримали право на додаткове списання боргів в рамках Ініціативи. Брак потенціалу в галузі оцінки став проблемою для інших країн-бенефіціарів Ініціативи, зокрема Албанії, Мадагаскару і Танзанії. Ці країни потребують додаткової допомоги для розвитку свого потенціалу в галузі оцінки.

В результаті надання країнам з великою заборгованістю пільгових кредитів або грантів з метою зниження ризику майбутніх боргових криз постало нове питання в сфері оцінки: враховуючи відмінність між грантами та кредитами, як оцінювати їхню дієвість та за якими критеріями? Це питання створює нові виклики для фахівців з оцінювання розвитку.

Вересень 2006 року позначив десяту річницю Ініціативи щодо бідних країн з великою заборгованістю. З 1999 року витрати бідних країн з великою заборгованістю на скорочення бідності зросли, в той час як витрати на обслуговування боргів зменшилися (World Bank 2007e). Це свідчить про те, що країни рухаються в потрібному напрямку.

Роль фондів

Дослідження ОЕСР дозволило оцінити суму коштів, виділену благодійними фондами країнам, що розвиваються. Дослідники

> спробували серйозно оцінити суму коштів, розподілених 15 найбільшими благодійними фондами (частково – в рамках міжнародної допомоги) станом на 2002 рік. Загальна сума склала 4 млрд доларів США, а загальна сума міжнародної допомоги – 2 млрд доларів США. Це близько 4 відсотків від загальної допомоги розвитку та половина всіх внесків, віднесених Комітетом сприяння розвитку до… НУО в цілому (група, що включає фонди) (Oxford Analytica 2004a).

Американська Рада у справах фондів нарахувала в США 56 000 приватних та суспільних фондів, які щороку розподіляють 27,5 млрд доларів. Європейський центр фондів нарахував 25 000 фондів у дев'яти країнах ЄС з річним обсягом витрат понад 50 млрд доларів США.

На світовій арені домінують кілька великих фондів. Серед них – Фонд Білла і Мелінди Гейтс, Фонд Форда, Фонд С'юзан Томпсон Баффет, а також Фонд Сороса/Відкритого суспільства.

Мережа Фонд Сороса/Інститут «Відкрите суспільство» – впливовий гравець на міжнародній арені розвитку, що здійснює програми у більш ніж 50 країнах. Програми забезпечують підтримку освіти, ЗМІ, охорони здоров'я, жінок, прав людини, культури й мистецтв, а також соціальних, економічних та правових реформ (SOROS Foundations Network 2007).

Запобігання конфліктам та відновлення після завершення конфліктів

За період, що настав після закінчення Холодної війни (1989–2001), сталося 56 серйозних збройних конфліктів у 44 регіонах. За оцінками станом на 2003 рік конфлікти впливали на життя більш як 1 млрд осіб. Більшість конфліктів у цей період тривали сім або більше років. Глобальні витрати, пов'язані з громадянськими війнами, є надзвичайно великими. «Створюючи територію, що не контролюється жодним визнаним урядом, збройні конфлікти сприяють нелегальному обігу наркотиків, тероризму та поширенню хвороб» (Collier and others 2003).

Бідність одночасно є і причиною, і наслідком конфлікту. Шістнадцять з двадцяти найбідніших країн світу перенесли масштабну громадянську війну. За статистикою, існує 44-відсоткова ймовірність того, що в повоєнній країні протягом п'яти років після настання миру знову спалахне війна.

Відновлення після конфлікту потребує координації роботи великої кількості двосторонніх та багатосторонніх організацій з розвитку. Наприклад, в Боснії та Герцеговині працювало 60 організацій з розвитку, на Західному березі ріки Йордан та в Секторі Газа – 50, в Афганістані – 82. Відбудова після конфлікту ускладнює роботу механізмів координації допомоги, покликаних забезпечувати задоволення потреб, уникнення дублювання та прогалин у допомозі.

Відновлення після конфлікту – це дещо більше, ніж просто відбудова інфраструктури. Відновлення часто означає підтримку інституціонального будівництва, демократії та виборів, НУО та громадянського суспільства, громадської поліції, забезпечення разових та регулярних бюджетних виплат, списання боргів, підтримку платіжного балансу, вирішення гендерних питань, розмінування територій, переселення біженців та людей, що втратили житло, а також демобілізацію та реінтеграцію учасників бойових дій в цивільне життя.

Через занепокоєність стосовно корупції та потребу використання офіційної допомоги з розвитку, відновлення після конфліктів нерідко потребувало створення нових кредитних інструментів та механізмів. Наприклад, на Західному березі ріки Йордан та в Секторі Газа було утворено трастовий фонд міжнародних організацій з розвитку, метою якого була підтримка разових та регулярних бюджетних витрат на користь нової палестинської адміністрації. Такі інструменти та механізми нині надзвичайно поширені в пост-конфліктних регіонах інших частин світу.

Програми відновлення після конфлікту – багатосекторні програми вартістю в мільярди доларів – становлять новий рівень складності та масштабності оцінки (Kreimer and others 1998). Фахівцям з оцінювання доводиться

оцінювати вплив, що його мають такі потужні підходи до розвитку й відновлення та врегулювання після конфліктів. Оцінювання координаційного процесу, що поєднує великі й різноманітні групи двосторонніх, багатосторонніх та інших підтримуючих сторін, створює нові виклики.

Фахівці з оцінювання повинні розглядати нові проекти та програми в нетрадиційних галузях допомоги розвитку, таких як демобілізація, реінтеграція учасників бойових дій та розмінування територій. Їм також доводиться оцінювати нові види механізмів та кредитних інструментів організацій з розвитку – такі як трастові фонди багатосторонніх організацій з розвитку.

Двосторонні й багатосторонні організації з розвитку все частіше намагаються встановлювати економічні причини та наслідки конфліктів, вони шукають шляхи запобігання конфлікту. Зростає увага до соціальних, етнічних та релігійних спільнот і відносин, управлінських та політичних установ, прав людини, безпеки, економічних структур та показників, довкілля та природних ресурсів, а також зовнішніх факторів. Це означає, що фахівці з оцінювання повинні також розглядати, що робиться та що можна було б зробити в контексті розвитку для запобігання виникненню конфліктів.

Управління

Хоча проблема управління й корупції добре відома зацікавленим сторонам, предметом публічного розгляду в контексті розвитку вона стала лише в середині 1990-х років. З тих пір у світі підписано низку міжнародних угод, спрямованих на вирішення проблеми корупції. ООН та ОЕСР прийняли антикорупційні угоди, які включають положення щодо попередження та криміналізацію корупції, положення щодо міжнародної співпраці з повернення активів, а також протихабарницьких заходів.

Міжнародні банки розвитку також запровадили антикорупційні програми. Надаються кредити для допомоги країнам у побудові ефективних та відповідальних державних установ. Управлінські та антикорупційні заходи включаються в стратегії допомоги країнам. Управлінські програми покликані сприяти:

- протидії корупції;
- управлінню державними витратами;
- реформуванню державної служби;
- судовій реформі;
- адмініструванню, децентралізації, електронному врядуванню та наданню послуг державної служби.

На початку 1990-х років було засновано «Трансперенсі Інтернешнл» (ТІ) – НУО, метою якої є привернення уваги до корупції як глобальної проблеми. На даний час представництва організації працюють у 88 країнах. Вона бореться з корупцією, взаємодіючи з місцевими, національними, регіональними та міжнародними партнерами (урядовими та неурядовими) (http://www.transparency.org/).

Щороку ТІ складає Індекс сприйняття корупції для 140 країн, що базується на сприйнятті рівня корупції державними посадовцями. А Індекс хабародавців, що також складається кожного року, охоплює країни-експортери, оцінюючи їх в контексті поширеності хабарництва.

За деякими оцінками, щороку через корупцію втрачається більше трильйона доларів США. Вимірювання корупції та пов'язаних з нею втрат стало викликом для міжнародної спільноти, але

> зростання доступності досліджень та опитувань, що проводяться багатьма установами та містять дані про різні виміри управління, дозволило створити всесвітній управлінський банк даних. Використовуючи десятки різних джерел та змінних, а також новітню методику узагальнення, цей банк даних охоплює 200 країн з різних частин світу та містить основні узагальнені показники стосовно таких сфер, як верховенство права, корупція, якість нормативної бази, ефективність, прозорість та відповідальність уряду, а також політична нестабільність (World Bank 2007c, p. 1).

Організації з розвитку та фахівці з оцінювання можуть користуватися цими даними для вимірювання дієвості допомоги. Дані вказують на те, що в місцях, де рівень корупції вище, ймовірність розтрати допомоги, відповідно, також буде вищою.

Для визначення та моніторингу найбільш уразливих чинників і установ в управлінській структурі країни використовується управління з орієнтацією на результати. Дані допомагають пролити світло на певні неясні питання в галузі управління та забезпечити більшу об'єктивність в роботі з ними. Зібрані дані також допоможуть спеціалістам з оцінки допомоги зробити більше кількісних висновків стосовно набутого досвіду. В той же час оцінка інвестиційного клімату та бізнес-середовища вимагатиме застосування складних і проблемних концепцій (див. наступний розділ стосовно розвитку приватного сектору та інвестиційного клімату).

Ця нова галузь швидко еволюціонує. Фахівцям з оцінювання доведеться оперативно реагувати на нові події та дані.

Боротьба з відмиванням грошей та фінансуванням тероризму

Відмивання грошей:
Практика проведення фінансових операцій, спрямованих на приховування дійових осіб, джерел та призначення певних коштів

Відмивання грошей та фінансування тероризму – це частина ширшого антикорупційного ландшафту.

> Відмивання грошей – це практика проведення фінансових операцій, спрямованих на приховування дійових осіб, джерел та призначення певних коштів. У минулому термін «відмивання грошей» застосовувався лише до фінансових операцій, пов'язаних з кримінальною діяльністю. Зараз державні регулятори (такі як американська Комісія з цінних паперів і бірж) нерідко розширюють це визначення, включаючи до нього всі фінансові операції, що не є прозорими відповідно до законодавчих вимог. В результаті протизаконна діяльність з відмивання грошей зазвичай ведеться пересічними особами, компаніями малого та великого бізнесу, корумпованими посадовцями та членами організованих злочинних угруповань – наприклад наркоторгівцями або мафіозі (Investor Dictionary. com 2006).

За оцінками МВФ, щороку відмивається 1 трлн доларів США (2–5 % світового ВВП). Це дозволяє вважати відмивання грошей серйозною міжнародною проблемою, що негативно впливає як на країни, що розвиваються, так і на розвинені країни (Camdessus 1998, p.1). Глобалізація, відкриття кордонів, послаблення митного режиму виявилися сприятливими для міжнародної злочинної діяльності та відповідних протизаконних фінансових потоків. З поширенням тероризму глобальні ініціативи з протидії відмиванню грошей набули ще більшої ваги.

Відмивання грошей здатне створювати особливо серйозні проблеми для країн, що розвиваються.

> Нові фінансові ринки та країни, що розвиваються... є привабливими цілями та легкими жертвами для осіб, які займаються відмиванням грошей і постійно шукають нові місця та шляхи, щоб сховатися від пильного ока закону. Наслідки операцій з відмивання грошей можуть бути особливо згубними для країн, що розвиваються. Діючи вільно, злочинці можуть маніпулювати фінансовими системами країни з метою ведення та розширення своєї протиправної діяльності... та швидко підірвати стабільність і розвиток встановлених інститутів (IFAC 2004, p. 5).

У 1989 році Велика сімка створила Групу з розробки фінансових заходів боротьби з відмиванням грошей (FATF) в рамках ОЕСР. Зараз до складу Групи входить 31 учасник з числа країн та територій і дві регіональні організації. Цей міжурядовий політичний орган має на меті розробку і про-

сування національної та міжнародної політики боротьби з відмиванням грошей та фінансуванням тероризму.

До повноважень FATF входять моніторинг та оцінювання реалізації. Вони виконуються на багатосторонній основі, методом партнерської перевірки та взаємної оцінки. Процес МіО означає наступне:

> Кожна країна-учасниця по черзі перевіряється FATF шляхом проведення польових візитів групами з трьох або чотирьох обраних експертів в юридичній, фінансовій та правоохоронній галузях з урядів інших країн-учасниць. Мета візиту – підготувати звіт, що оцінює ступінь просування оцінюваної країни в запровадженні дієвої системи боротьби з відмиванням грошей та виділення сфер, де необхідна додаткова робота (FATF 2007, p. 1).

FATF встановила набір заходів, що вживаються в разі невідповідності вимогам.

Перекази коштів працівників

Сума коштів, що їх щороку переказують на батьківщину працівники-мігранти, перевищує суму щорічної офіційної допомоги розвитку. Якщо в 1998 році мігранти переказували 60 млрд доларів США, то в 2002 – вже 80 млрд, а в 2003 – 100 млрд. Ці цифри є співставними з 50–60 млрд доларів щорічної офіційної допомоги розвитку та 143 млрд доларів потоків приватного капіталу в 2002 році. Перекази зазвичай мають більшу стабільність порівняно з потоками приватного капіталу (World Bank 2003; Oxford Analytica 2004b).

Глобальні перекази мають значний вплив на скорочення бідності. «У середньому 10-відсоткове збільшення частки міжнародних мігрантів у населенні країни призводить до зменшення на 1,9 відсотка частки людей, що живуть в бідності (1,00 долар США на людину на день)» (Adams and Page 2003, p. 1). Міжнародні перекази допомагають фінансувати місцеве споживання в сфері забезпечення житлом, сільському господарстві, промисловості та створенні нових малих та середніх підприємств у країні-реципієнті.

Розвинені країни та країни, що розвиваються, а також організації обізнані про ці тенденції та шукають шляхів до спрямування відповідних потоків на потрібні їм інвестиції. План нещодавнього саміту Великої вісімки закликає учасників та уряди країн, що розвиваються,

> сприяти переказам коштів з інших країн, щоб допомагати родинам та малим підприємствам, в тому числі шляхом зменшення вартості переказів, та щоб створювати певний обсяг коштів у країні для продуктивних інвестицій, покращити доступ отримувачів переказів до фінансових послуг та координацію (G-8 2004a, p. 7).

Перекази через банківську систему, ймовірно, збільшаться, в той час як обмеження щодо неофіційної передачі коштів стануть суворішими завдяки запровадженню ефективніших моніторингових правил, покликаних скоротити фінансування терористичних організацій, що проводиться через неофіційні канали (див. нижче) та зменшити банківські збори за рахунок посилення конкуренції на ринку міжнародних переказів.

Вплив переказів на країни, що розвиваються, все ще до кінця не оцінений та не перевірений. Відстеження міжнародних переказів і спрямування їх на нові види інвестицій та в нові фонди створюватиме нові виклики для фахівців з оцінювання. Оскільки спеціалісти з розвитку ще не розробили способів консолідації переказів та спрямування їх на скорочення бідності, фахівці з оцінювання з великим інтересом спостерігатимуть за цією сферою.

Гендер: від жінок в галузі розвитку до гендеру й розвитку та досягнення гендерної рівності

Гендер – це соціальні ролі, що відводяться жінкам та чоловікам.

Гендерний аналіз:
Аналіз доступу чоловіків та жінок до ресурсів та їх контролю над цими ресурсами

Гендерний аналіз розглядає доступ чоловіків і жінок до ресурсів та їх контроль над цими ресурсами. Також це системний спосіб оцінки потреб та уподобань чоловіків і жінок у сфері розвитку (причому такі потреби та уподобання часто можуть різнитися) та різний вплив розвитку на жінок і чоловіків. Гендерний аналіз враховує те, яким чином клас, раса, етнічна приналежність та інші чинники, взаємодіючи з гендером, дають дискримінаційні результати. Об'єктом гендерного аналізу традиційно були жінки – це зумовлено різницею умов для чоловіків та жінок в галузі освіти, працевлаштування, послуг і т. ін.

Жінки складають трохи більше половини населення світу та відіграють ключову роль в економічному розвитку. Тим не менш, повністю їх потенціал щодо участі в соціально-економічному розвитку не реалізується. Більше того, саме жінки та діти усе ще становлять більшість серед бідного населення.

> Жінки виробляють половину їжі в деяких частинах світу, що розвивається, несуть основну відповідальність за харчування домашніх господарств, вони є чвертю робочої сили в промисловості та третиною – в сфері послуг... Тим не менш, через більш обмежений доступ до навчання та інших можливостей продуктивність праці жінок порівняно з продуктивністю праці чоловіків залишається низькою. Підвищення продуктивності праці жінок може посприяти зростанню, ефективності та скороченню бідності – головним цілям розвитку в усьому світі (World Bank 1994, p. 9).

Останні тенденції стосовно ролі жінок в розвитку супроводжувалися переходом від традиційного підходу «жінки в розвитку» (ЖвР) до підходу «гендер і розвиток» (ГіР), а потім – і до більш комплексного підходу «забезпечення гендерної рівності». В центрі уваги стратегії ЖвР були жінки як особлива цільова або зацікавлена група бенефіціарів проектів, програм та політик. «ЖвР визнає, що жінки є активними, хоча нерідко й не врахованими учасниками процесу розвитку, що відіграють критично важливу роль в економічному зростанні... як не задіяні ресурси; необхідно залучати жінок до процесу розвитку» (Moser 1995, p. 107).

Підхід ГіР концентрується на соціальних, економічних, політичних та культурних силах, що визначають, яким чином чоловіки та жінки долучаються до проектних ресурсів та заходів, отримують від них вигоди та здійснюють над ними контроль. Підкреслюється, що потреби та уподобання жінок і чоловіків часто різняться. Цей підхід зміщує увагу з жінок як групи на соціально визначені відносини між чоловіками та жінками.

ЦРТ включають розвиток гендерної рівності та надання жінкам прав і можливостей, а також визначають конкретні цілі та показники для оцінки їх виконання. Комітет сприяння розвитку (КСР) ОЕСР також склав перелік запитань, що мають допомагати керівникам оцінювати заходи з розвитку. Це, зокрема, наступні питання:

- Чи вдалося проекту добитися результатів щодо забезпечення рівних можливостей та привілеїв для чоловіків і жінок?
- Чи став проект корисним для жінок та чоловіків, чи зашкодив він їм?
- Чи виявився проект дієвим з точки зору інтеграції гендеру в діяльність з розвитку? (Woroniuk and Schalkwyk 1998).

Гендерне бюджетування – це спосіб реалізації та оцінки користі національного бюджету для чоловіків і жінок. Інший спосіб виміряти та оцінити допомогу – встановити, якою мірою допомога з розвитку справді допомагає секторам, «що залучають жінок, допомагають жінкам, надають жінкам права та можливості, а також створюють результати на користь жінок» (Jalan 2000, p. 75). З урахуванням існуючого акценту на комплексних підходах та партнерствах, оцінка політик забезпечення гендерної рівності також повинна проводитися, інтегруватися та координуватися серед взаємодіючих в сфері розвитку країн, організацій та агенцій. За кожної оцінки важливо дивитися на відмінності в тому, як проект, програма або політика впливають на чоловіків та жінок.

Розвиток приватного сектору економіки та інвестиційного клімату

Ціла низка питань постає у зв'язку з розвитком приватного сектору економіки та інвестиційного клімату. До таких питань відносяться роль приватного сектору та прямих іноземних інвестицій в скороченні бідності, приватизація, приватна участь в підприємствах інфраструктури та відповідних партнерствах, створення та підтримка мікропідприємств і підприємств малого та середнього бізнесу за допомогою фіскальних засобів.

Інвестиції в приватний сектор

Інвестиції в приватний сектор все частіше визнаються як критичні для скорочення бідності в країнах, що розвиваються. У 1990 році інвестиції в приватний сектор у таких країнах склали близько 30 млрд доларів США на рік, в той час як допомога розвитку складала близько 60 млрд доларів. У 1997 році рівень інвестицій в приватний сектор у країнах, що розвиваються, досяг 300 млрд доларів США, а допомога розвитку скоротилася до 50 млрд доларів. Таким чином, менш ніж за 10 років інвестиції в розвиток приватного сектору, що складали половину допомоги розвитку, збільшились у шість разів.

Офіційна допомога розвитку

Офіційна допомога розвитку:
Потоки офіційного фінансування, що розподіляються з метою сприяння економічному розвитку та добробуту країн, що розвиваються, і надаються на пільгових умовах

Одним з показників інвестиційного клімату є **офіційна допомога розвитку** (ОДР), яка в Глосарії статистичних термінів ОЕСР (2002а) визначається наступним чином:

> Потоки офіційного фінансування, що розподіляються з метою сприяння економічному розвитку та добробуту країн, що розвиваються, і надаються на пільгових умовах – при цьому на грант припадає не менше 25 % (з використанням фіксованої дисконтної ставки в 10%). Зазвичай потоки ОДР складаються з внесків агенцій урядів-донорів (всіх рівнів), які спрямовуються на користь країн, що розвиваються («двостороння ОДР»), та на користь багатосторонніх установ. Загальні кошти ОДР складаються з виплат двосторонніх донорів та міжнародних установ. [З веб-сторінки офіційної допомоги розвитку].

У 1997 році допомога збільшилася і трималася приблизно на одному рівні до 2001 року. Загальна ОДР від членів КСР в 2001 році зросла на 7%, а в 2003 – на 5%. В 2005 році офіційна допомога розвитку від членів КСР зросла на 32%, переважно в результаті нарощування допомоги за наслід-

ками цунамі 2004 року та списання заборгованості Іраку та Нігерії (OECD 2005a). У 2006 році ОДР зменшилася на 4,6% – це пов'язано з надзвичайно високим рівнем гуманітарної допомоги та списання боргів у 2004 році (OECD 2006).

ОДР стабільно збільшувалася протягом останнього десятиліття і, як очікується, продовжуватиме збільшуватися, адже донори зобов'язалися значно підвищити рівень допомоги задля досягнення ЦРТ. Щоб забезпечити ефективне використання такої збільшеної допомоги на рівні країн, донорам та реципієнтам необхідно вирішити низку проблем реалізації, зокрема:

- досягнення взаємодоповнюваності в контексті національних, регіональних та глобальних пріоритетів та програм розвитку;
- зміцнення потенціалу країн-реципієнтів у сфері ефективного використання потенційно збільшених, оперативно надаваних коштів ОДР – наприклад, в рамках підтримки бюджету (World Bank 2007a).

Прямі іноземні інвестиції

Іншим виміром інвестиційного клімату є **прямі іноземні інвестиції** (ПІІ), що відіграють надзвичайно важливу та все більшу роль в світовому бізнесі. ПІІ – це транскордонні інвестиції розміром не менше 10 відсотків власного капіталу підприємства, що їх розміщує інвестор для встановлення довгострокової фінансової участі в підприємстві та здійснення певного впливу на його роботу. ПІІ часто називають головним локомотивом економічного зростання; вважається, що вони приносять певну користь національним економікам (InvestorDictionary.com 2006).

Прямі іноземні інвестиції:
Транскордонні інвестиції розміром не менше 10 відсотків власного капіталу підприємства, що розміщуються для встановлення довгострокової фінансової участі в підприємстві та здійснення певного впливу на його роботу

Найбільше зростання ПІІ в період з 70-х років до 1999 року мало місце в країнах, що розвиваються, де щорічні потоки зросли в середньому з 10 млрд доларів США в 70-х та менш ніж 20 млрд у 80-х до 179 млрд доларів США в 1998 році та 208 млрд в 1999 році. На ці потоки прийшлася значна частина світових ПІІ (Graham and Spaulding 2005).

Конференція ООН з торгівлі та розвитку (ЮНКТАД) (2008) повідомляла про глобальне зростання ПІІ в 2007 році. Її звіт фіксує зростання ПІІ в усіх трьох групах: розвинені країни, економіки, що розвиваються, та Південно-Східна Європа й Співдружність незалежних держав (СНД). Результати відображають потужне зростання транснаціональних корпорацій та високі економічні показники в багатьох регіонах світу, а також зростання прибутків корпорацій та надлишок вільних коштів, що різко збільшило вартість транскордонних об'єднань та придбань. На такі операції припадала значна частина потоків ПІІ в 2007 році, хоча вартість об'єднань та придбань в другому півріччі знизилася.

Потоки ПІІ до розвинених країн станом на 2007 рік зростали четвертий рік поспіль, досягнувши 1 трлн доларів США. Надходження від ПІІ в країни, що розвиваються, та в перехідні економіки Південно-Східної Європи й країн СНД зросли на 16 та 41 відсоток відповідно, досягнувши рекордних рівнів.

Приватизація

Приватизація державних підприємств велася особливо активно в 1990-х роках, коли багато країн прагнули переформатувати власну соціалістичну економіку в ринкову. Цей процес і досі триває в багатьох країнах, де держава усе ще володіє та здійснює управління багатьма економічними активами.

«Більше 100 країн з кожного континенту провели приватизацію певної частини або й більшості власних державних компаній практично в кожному секторі інфраструктури, виробництва та сфери послуг. ... За оцінками, в усьому світі було приватизовано 75 000 середніх та великих компаній, а також сотні тисяч менших суб'єктів господарювання...» Вважається, що загалом від приватизації було отримано 735 млрд доларів США (Nellis 1999). Приватизація має неоднозначний характер; продовжується полеміка стосовно того, чи потрібно її проводити, а якщо потрібно, то коли і як. Приватизація – не панацея від економічних проблем, але вона показала себе корисним інструментом для підвищення добробуту та покращення послуг для економіки і суспільства.

Наслідки оцінювання

Яким чином спільнота фахівців з оцінювання розвитку відреагувала на ці ініціативи? Міжнародна фінансова корпорація (МФК) оцінює наслідки інтервенцій на рівні проектів. Вона використовує зрізи стану бізнес-середовища (ВЕ) для «представлення вимірних показників у широкому спектрі проблем бізнес-середовищ з плином часу» (IFC 2007). Цей новий інструмент накопичує різнорідні дані, показники та інформацію щодо проектів, які стосуються бізнес-середовища в країнах. Він забезпечує легкість доступу до таких даних в однорідному форматі. Практикуючі спеціалісти та високопосадовці в галузі оцінки можуть використовувати зрізи стану бізнес-середовища для отримання повної його картини в конкретних країнах. Такі зрізи також можна застосовувати в якості інструменту моніторингу або планування.

Яким чином оцінюються такі види діяльності? На загальному рівні можна розглядати чотири можливих показники:

- ефективність господарчої діяльності;
- економічна стабільність;
- екологічні наслідки економічної діяльності;
- розвиток приватного сектору.

Звіт про світовий розвиток 2005 року значну увагу приділяє дослідженням інвестиційного клімату, стану бізнес-середовища та ефективності діяльності компаній – все це може допомогти виявити шляхи, використовуючи які уряди покращать інвестиційний клімат для підприємств всіх видів. Дослідження охопили 26 000 фірм у 53 країнах, що розвиваються, а також 3000 мікро- та неформальних підприємств в 11 країнах.

Ці дослідження дозволяють порівняти існуючі умови та відстежувати їх зміну з плином часу. Інструментом дослідження стали основний набір питань та кілька модулів, що їх можна використовувати для поглибленого вивчення певних аспектів інвестиційного клімату та їх впливу на продуктивність праці на рівні компаній.

Питання можна поділити на три групи:

- такі, що надають інформацію стосовно підприємств;
- такі, що надають інформацію стосовно інвестиційного клімату, в якому підприємства працюють;
- такі, що надають дані про показники ефективності діяльності підприємств.

До числа згаданих показників увійшли:

- політична невизначеність (значне обмеження: непередбачуваність тлумачення нормативних актів);
- корупція (значне обмеження: хабарі);
- суди (значне обмеження: відсутність впевненості в тому, що суди захистять майнові права);
- злочинність (значне обмеження: втрати через злочинність). Серед інших джерел показників інвестиційного клімату були служба з комерційного ризику, кредитні рейтинги країн (Euromoney Institutional Investor), показники ризику для країн (World Markets Research Centre), служба оцінки ризику для країн (аналітичний відділ журналу Economist), а також *Звіт про глобальну конкурентоспроможність* (Всесвітній економічний форум).

У таких дослідженнях беруть участь міжнародні банки розвитку, міжнародні фінансові установи, організації з розвитку та приватний сектор, надаючи цінну інформацію й рекомендації. В різних країнах світу також

проводяться поточні та періодичні оцінки інвестиційного клімату. Одним з яскравих прикладів є база даних Doing Business, що надає об'єктивні оцінки нормативних документів та їх застосування. Порівняні показники для 145 країн демонструють витрати, що їх несуть підприємства для дотримання нормативно-правових вимог. Ці показники можна використовувати для аналізу конкретних нормативних положень, що сприяють інвестиціям, продуктивності праці та економічному зростанню або ж обмежують їх (World Bank 2007d). Однак звіт Незалежної групи оцінки Світового банку 2008 року окреслив потреби щодо покращення показників Doing Business (IEG 2008).

Екологічна та соціальна сталість

Корпоративна соціальна відповідальність:
Визнання приватними компаніями потреби врахування економічних, екологічних та соціальних впливів і наслідків їхньої діяльності

Корпоративна соціальна відповідальність (КСВ) передбачає активне врахування економічних, екологічних та соціальних впливів і наслідків діяльності підприємств. Компанії приватного сектору, організації та уряди розглядають нові шляхи запобігання тому, щоб комерційна діяльність завдавала шкоди економіці, суспільству або довкіллю у відповідних країнах та секторах. Приміром, британський уряд запровадив різні норми й правила, що стимулюють КСВ в цілому та екологічну й соціальну сталість зокрема.

> Уряд розглядає КСВ як внесок бізнесу в досягнення цілей сталого розвитку. Перш за все, мова йде про те, яким чином бізнес враховує економічні, соціальні та екологічні наслідки своєї діяльності, максимізуючи при цьому вигоди й зводячи до мінімуму негативні аспекти. ... Підхід уряду полягає в стимулюванні запровадження КСВ та формування звітності за ним з використанням передових методів, а у відповідних випадках – в розумному регулюванні та наданні податкових пільг (BEER 2004, p. 1).

Принципи екватора:
Галузевий підхід до визначення і оцінки екологічних та соціальних ризиків у фінансуванні проектів приватного сектору та керування такими ризиками

Прикладом міжнародної діяльності, спрямованої на забезпечення екологічної та соціальної сталості, є **Принципи екватора**, затверджені десятьма західними фінансовими установами в 2003 році. Принципи екватора було розроблено банками приватного сектору. Вони є галузевим підходом до визначення і оцінки екологічних та соціальних ризиків у фінансуванні проектів приватного сектору та керування такими ризиками. В 2006 році було прийнято нову редакцію Принципів екватора. Ця нова редакція відображає зміни, внесені МФК до власних *Стандартів щодо забезпечення соціальної і екологічної сталості*. Дія редакції 2006 року поширюється на всі країни та сектори, а також на всі випадки фінансування проектів, коли капітальні витрати перевищують 10 млн доларів США.

МФК та 61 провідний комерційний банк (з Північної Америки, Європи, Японії, Австралії) добровільно запровадили Принципи екватора в галузі фінансування проектів по всьому світі. Ці установи прагнуть забезпечити розвиток фінансованих ними проектів з дотриманням принципів соціальної відповідальності та екологічно обґрунтованого підходу. Принципи екватора повинні слугувати спільною основою для реалізації окремих внутрішніх екологічних та соціальних процедур і стандартів для фінансування проектів у всіх галузях економіки в світовому масштабі. Приймаючи ці принципи, установи зобов'язуються уважно перевіряти всі пропозиції стосовно проектів, фінансування яких просять надати їхні клієнти. Вони зобов'язуються не надавати кредити на проекти, якщо позичальник не може або відмовляється виконувати встановлені ними екологічні та соціальні практики й правила. Стандарти охоплюють захист довкілля, охорону праці й здоров'я, безпеку, захист корінних народів, природних середовищ перебування видів, а також переселення (Equator Principles 2007).

Публічне зобов'язання дотримуватися принципів – це одне, але добросовісне їх застосування – то зовсім інше. BankTrack, мережа з 18 міжнародних НУО, що спеціалізуються на роботі у фінансовому секторі, відіграла важливу роль в моніторингу реалізації Принципів екватора. Вона критично ставиться до вимог звітності, сформульованих в Принципах екватора, а також до способів, у які відбувається моніторинг фінансових установ (BankTrack 2008).

Глобальні суспільні блага

Економісти визначають **приватні блага** як такі, споживання яких однією особою зменшує кількість благ, доступних іншим, принаймні поки не будуть вироблені додаткові блага (Linux Information Project 2006). Приватні блага повинні бути осяжними. Більшість товарів є приватними благами.

Економісти визначають **суспільні блага** як такі, які люди можуть споживати без обмежень та не зменшуючи при цьому кількість благ, доступних іншим (Linux Information Project 2006). Чисте повітря, наприклад, є суспільним благом, адже дихання чистим повітрям не зменшує кількості чистого повітря, доступного іншим. Суспільні блага зазвичай є нематеріальними; багато з них відносяться до категорії інформації або знань.

Глобальні суспільні блага – це публічні блага, які охоплюють весь світ. Прикладами глобальних суспільних благ є майнові права, безпека, фінансова стабільність та чисте довкілля. При цьому оцінку розвитку також можна вважати різновидом суспільних благ, тому що вона виходить за межі окремої організації. Належне оцінювальне дослідження може мати позитивні супутні наслідки для спільноти розвитку (Picciotto and Rist 1995, p. 23).

Приватні блага:
Блага, споживання яких однією особою зменшує кількість благ, доступних іншим

Суспільні блага:
Блага, споживання яких однією особою не зменшує кількість благ, доступних іншим

Глобальні суспільні блага:
Суспільні блага, що є за своєю природою неконкурентними та невиключними для всього світу, а не лише в межах однієї країни

Глобальні суспільні блага є важливими через те, що після збільшення відкритості національних кордонів поєдналися загальні надбання різних країн. Суспільне благо в одній країні нерідко залежить від внутрішньої політики, заходів і політичних рішень інших країн або міжнародних організацій (gpgNet 2008). Всі залежать від суспільних благ; ні ринки, ні найбагатші особи не можуть обійтися без них.

Оцінка на глобальному рівні майже відсутня:

> Незалежна оцінка програм співпраці, заснованих для створення глобальних суспільних благ, не проводиться, в результаті чого таким програмам часто не вистачає чітких цілей та показників ефективності, що піддавалися б перевірці. Крім цього, вплив політик розвинених країн на країни бідні систематично не оцінюється, незважаючи навіть на те, що допомога, заборгованість, іноземні інвестиції, забруднення довкілля, схеми міграції та режими інтелектуальної власності формуються рішеннями урядів розвинених країн (Picciotto 2002b, p. 520).

Контроль поширення та врешті-решт викорінення ВІЛ/СНІД – це ще один приклад глобального суспільного блага, що посідає чільне місце порядку денного для багатьох міжнародних організацій. Усе ще не оціненим залишається вплив глобалізації на бідних. Оцінювання розвитку має активніше проводитися на місцевому рівні, воно має стати більш глобальним та транснаціональним (Chelimsky and Shadish 1997).

У 2004 році Незалежна група з оцінювання Світового банку опублікувала оцінку участі Банку в глобальних програмах. В рамках звіту «Оцінка підходу Світового банку до глобальних програм: розв’язання викликів глобалізації» було досліджено 26 глобальних програм, що отримують допомогу Банку, і зроблено висновки стосовно проектування, реалізації та оцінки глобальних програм (World Bank 2004a). Звіт особливу увагу приділяє п'яти з 18 своїх висновків:

- Стратегія Банку стосовно глобальних програм є недостатньо визначеною.
- Глобальні програми значно збільшили загальний обсяг допомоги.
- Голоси країн, що розвиваються, представлені недостатньо.
- Глобальні програми виявляють прогалини в інвестуванні та глобальній суспільній політиці.
- Необхідним є нагляд за незалежними глобальними програмами.

У звіті наводяться наступні рекомендації:

- Сформувати стратегічну систему залучення Банку до глобальних програм.

- Пов'язати фінансування з пріоритетами.
- Посилити вибірковість та нагляд стосовно портфеля глобальних програм.
- Вдосконалити управління й керування окремими програмами.
- Провести додаткове оцінювання.

Підсумки

Країни обрали різні підходи до запровадження систем оцінювання. Загальноурядовий підхід передбачає встановлення всеосяжної, комплексної системи на урядовому рівні. Анклавний підхід означає концентрацію на певній частині або секторі уряду. За умов змішаного підходу деякі частини або сектори уряду оцінюються комплексно, в той час як інші розглядаються більш спорадично.

Створення системи оцінювання в країнах, що розвиваються, є складнішим, оскільки їм часто бракує демократичних політичних систем, міцних емпіричних традицій, державних службовців, обізнаних у суспільних науках, а також ефективних адміністративних систем та інститутів. Організації з розвитку концентруються на наданні країнам, що розвиваються, допомоги в нарощуванні потенціалу до створення та підтримки систем оцінювання.

Багато складних проблем розвитку відображаються і на оцінці. До основних рушійних сил розвитку відносяться наступні чинники:

- ЦРТ;
- Монтерейський консенсус;
- Паризька декларація щодо підвищення ефективності зовнішньої допомоги;
- Ініціатива щодо бідних країн з великою заборгованістю;
- роль фондів;
- запобігання конфліктам та відновлення після завершення конфліктів;
- управління;
- боротьба з відмиванням грошей та фінансуванням тероризму;
- перекази коштів працівників;
- гендерні питання;
- розвиток приватного сектору економіки та інвестиційного клімату;
- екологічна та соціальна сталість;
- глобальні суспільні блага.

Цей перелік не є постійним, він змінюється відповідно до нових глобальних подій.

Посилання і рекомендована література

Adams, Richard H. Jr., and John Page. 2003. "International Migration, Remittances and Poverty in Developing Countries." World Bank Policy Research Working Paper 3179, Washington, DC.

BankTrack. 2008. The Equator Principles. http://www.banktrack.org/.

Barslund, Mikkel, and Finn Tarp. 2007. "Formal and Informal Rural Credit in Four Provinces of Vietnam." Discussion Paper 07-07, Department of Economics, University of Copenhagen.

BEER (U.K. Department for Business Enterprise & Regulatory Reform). 2004. "What Is CSR?" London.

http://www.csr.gov.uk/whatiscsr.shtml.

Boyle Richard. 2002. "A Two-Tiered Approach: Evaluation Practice in the Republic of Ireland." In International Atlas of Evaluation, eds. Jan-Eric Furubo, Ray Rist, and Rolf Sandahl, 261–72. New Brunswick, NJ: Transaction Publishers.

———. 2005. Evaluation Capacity Development in the Republic of Ireland. ECD Working Paper Series 14, World Bank, Evaluation Capacity Development, Washington, DC.

Camdessus, Michael. 1998. "Money Laundering: The Importance of International Countermeasures." Address by Managing Director of the International Monetary Fund, Washington, DC.

CGAP (Consultative Group to Assist the Poor). 2003. CGAP Publications on Assessment and Evaluation. http://www.cgap.org/portal/site/CGAP/menuite m.9fab704d4469eb016780801 0591010a0/.

Channah Sorah, Vijaya Vinita. 2003. "Moving from Measuring Processes to Out- comes: Lessons Learned from GPRA in the United States." Paper presented at the joint World Bank and Korea Development Institute conference on "Performance Evaluation System and Guidelines with Application to Large-Scale Construction, R&D, and Job Training Investments," Seoul, July 24–25.

Chelimsky, Eleanor, and William R. Shadish, eds. 1997. Evaluation for the 21st Cen tury: A Handbook. Thousand Oaks, CA: Sage Publications.

Chemin, Matthieu. 2008. "Special Section on Microfinance. The Benefits and Costs of Microfinance: Evidence from Bangladesh." Journal of Development Studies 44 (4): 463–84.

Collier, Paul, V., L. Elliott, Håvard Hegre, Anke Hoeffler, Marta Reynal-Querol, and Nicholas Sambanis. 2003. Breaking the Conflict Trap: Civil War and Development Policy. Washington, DC: Oxford University Press for the World Bank.

Economic Commission for Europe. 1998. Public–Private Partnerships: A New Concept for Infrastructure Development. BOT Expert Group, United Nations, New York. http://rru.worldbank.org/Documents/Toolkits/Highways/pdf/42.pdf.
Equator Principles. 2004. http://www.fatf-gafi.org/dataoecd/14/53/38336949.pdf.
———. 2007. "A Milestone or Just Good PR?" http://www.equator-principles.com/principles.shtml.
FATF (Financial Action Task Force on Money Laundering). 2007. Monitoring the Implementation of the Forty Recommendations.
http://www.fatf-gafi.org/document/60/0,3343,en_32250379_32236920_34039228_1_1_1_1,00.html.
Feuerstein, M. T. 1986. Partners in Evaluation: Evaluating Development and Com munity Programs with Participants. London: MacMillan, in association with Teaching Aids at Low Cost.
Fitzpatrick, Jody L., James R. Sanders, and Blaine R. Worthen. 2004. Program Evaluation: Alternative Approaches and Practical Guidelines. New York: Pearson Education, Inc.
Friis, H. 1965. Development of Social Research in Ireland. Institute of Public Administration, Dublin.
Fukuda-Parr, Sakiko, Carlos Lopes, and Khalid Malik, eds. 2002. Capacity for Development: New Solutions to Old Problems. London: Earthscan Publications.
Furubo, Jan-Eric, and Rolf Sandahl. 2002. "Coordinated Pluralism." In International Atlas of Evaluation, eds. Jan-Eric Furubo, Ray Rist, and Rolf Sandahl, 115–28. New Brunswick, NJ: Transaction Publishers.
Furubo, Jan-Eric, Ray Rist, and Rolf Sandahl, eds. 2002. International Atlas of Evaluation. New Brunswick, NJ: Transaction Publishers.
G-8. 2004a. G-8 Action Plan: Applying the Power of Entrepreneurship to the Eradication of Poverty. Sea Island Summit, June.
———. 2004b. G-8 Plan of Support for Reform. Sea Island Summit, June.
Gerrard, Christopher. 2006. "Global Partnership Programs: Addressing the Challenge of Evaluation." PowerPoint presentation. http://www.oecd.org/secure/pptDocument/0,2835,en_21571361_34047972_36368404_1_1_1_1,00.ppt.
gpgNet. 2008. The Global Network on Global Public Goods. http://www.sdnp.undp.org/gpgn/#.
Graham, Jeffrey P., and R. Barry Spaulding. 2005. Going Global: Understanding Foreign Direct Investment. JPG Consulting.
http://www.going-global.com/articles/understanding_foreign_direct_ investment.htm.
Guerrero, R. Pablo 1999. "Evaluation Capacity Development: Comparative Insights from Colombia, China, and Indonesia." In Building Effective

Evaluation Capacity: Lessons from Practice, eds. Richard Boyle and Donald Lemaire, 177–94. New Brunswick, NJ: Transaction Publishers.

Hatry, Harry P., Elaine Morely, Shelli B. Rossman, and Joseph P. Wholey. 2003. How Federal Programs Use Outcome Information: Opportunities for Federal Managers. IBM Endowment for the Business of Government, Washington, DC.

Hauge, Arild. 2001. "Strengthening Capacity for Monitoring and Evaluation in Uganda: A Results-Based Perspective." ECD Working Paper 8, World Bank, Operations Evaluation Department, Evaluation Capacity Development, Washington, DC.

Hougi, Hong, and Ray C. Rist. 2002. "Evaluation Capacity Building in the People's Republic of China." In International Atlas of Evaluation, eds. Jan-Eric Furubo, Ray Rist, and Rolf Sandahl, 249–60. New Brunswick, NJ: Transaction Publishers.

IDA (International Development Association), and IMF (International Monetary Fund). 2007. Heavily Indebted Poor Countries HIPC Initiative and Multilateral Debt Relief Initiative MDRI: Status of Implementation. http://siteresources.worldbank.org/DEVCOMMINT/Documentation/21510683/DC2007-0021 E HIPC.pdf.

IDS (Institute of Development Studies). 2008. "Impact Evaluation: The Experience of Official Agencies." IDS Bulletin 39 (1). http://www.ntd.co.uk/idsbookshop/details.asp?id=1030.

IEG (Independent Evaluation Group). 2008. "Doing Business: An Independent Evaluation. Taking the Measure of the World Bank–IFC Doing Business Indicators."

IFAC (International Federation of Accountants). 2004. Anti–Money Laundering, 2nd ed. New York.

IFC (International Finance Corporation). 2004. Strengthening the Foundations for Growth and Private Sector Development: Investment Climate and Infrastructure Development. Development Committee, Washington, DC. http://siteresources.worldbank.org/IDA/Resources/PSDWBGEXT.pdf.

———. 2007. Business Environment Snapshots. Washington, DC. http://rru.worldbank.org/documents/BES_Methodology_Note_External.pdf.

IMF (International Monetary Fund). 2009. "A Fact Sheet: Debt Relief under the Heavily Indebted Poor Countries (HIPC) Initiative." http://www.info.org/external/np/exr/facts/hipc.htm.

Investor Dictionary.com. 2006. "Money Laundering." http://www.investordictionary.com/definition/money+laundering.aspx.

Jalan, Bimal. 2000. "Reflections on Gender Policy." In Evaluating Gender Impact of Bank Assistance, 75–76. World Bank, Operations Evaluation Department, Washington, DC.

Joint Progress toward Enhanced Aid Effectiveness, High-Level Forum. 2005. Paris Declaration on Aid Effectiveness: Ownership, Harmonization, Alignment, Results, and Mutual Accountability. http://www1.worldbank.org/harmonization/Paris/FINALPARIS DECLARATION.pdf/.

Katjomulse, Kavazeua, Patrick N. Osakwe, Abebe Shimeles, and Sher Verick. 2007. The Monterrey Consensus and Development in Africa: Progress, Challenges, and Way Forward. United Nations Economic Commission for Africa (UNECA), Addis Ababa.

Kreimer, Alcira, John Eriksson, Robert Muscat, Margaret Arnold, and Colin Scott. 1998. The World Bank's Experience with Post-Conflict Reconstruction. World Bank, Operations Evaluation Department, Washington, DC. http://lnweb90.worldbank.org/oed/oeddoclib.nsf/b57456d58aba40e585256ad400736404/f753e43e728a27b38525681700503796/$FILE/PostCon.pdf.

Kusek, Jody Zall, Ray C. Rist, and Elizabeth M. White. 2004. "How Will We Know Millennium Development Results When We See Them? Building a Results-Based Monitoring and Evaluation System to Give Us the Answer." World Bank Africa Region Working Paper 66, Washington, DC.

Lawrence, J. 1989. "Engaging Recipients in Development Evaluation: The 'Stakeholder' Approach." Evaluation Review 13 (3): 243–56.

Lee, Yoon-Shik. 1999. "Evaluation Coverage." In Building Effective Evaluation Capacity: Lessons from Practice, eds. Richard Boyle and Donald Lemaire, 75–91. New Brunswick, NJ: Transaction Publications.

Linux Information Project. 2006. "Public Goods: A Brief Introduction." http://www.linfo.org/public_good.html.

Mackay, Keith. 2002. "The Australian Government: Success with a Central, Directive Approach." In International Atlas of Evaluation, eds. Jan-Eric Furubo, Ray

C. Rist, and Rolf Sandahl, 157–74. New Brunswick, NJ: Transaction Publishers.

———. 2007. "Three Generations of National M&E Systems in Australia." Power-Point presentation to the Third Latin America and Caribbean Regional Conference on Monitoring and Evaluation, Lima, July 23–24.

———. 2008. M&E Systems to Improve Government Performance: Lessons from Australia, Chile and Columbia. PowerPoint presentation to the High-Level Delegation from the People's Republic of China, Washington, DC, March 6.

Moser, Caroline O. N. 1995. "Evaluating Gender Impacts." New Directions for Evaluation 67 (Fall): 105–17.

NCSTE (China National Center for Evaluation of Science & Technology), and IOB (Netherland's Policy and Operations Evaluation Department). 2004. A Country-Led Joint Evaluation of the ORET/MILIEV Programme in China. http://www.euforic.org/iob/docs/200610201336433964.pdf.

Nellis, John. 1999. “Time to Rethink Privatization in Transition Economies?” IFC Discussion Paper 38, International Finance Corporation, Washington, DC.
OECD (Organisation for Economic Co-operation and Development). 2002a. OECD Glossary of Statistical Terms.
http://stats.oecd.org/glossary/index.htm.
———. 2002b. “Public Management and Governance: Overview of Results-Focused Management and Budgeting in OECD Member Countries.” Paper presented at the 23rd annual meeting of OECD senior budget officials, Washington, DC, June 3–4.
———. 2003. Joint OECD DAC/Development Centre Experts’ Seminar on Aid Effectiveness and Selectivity: Integrating Multiple Objectives into Aid Allocations.
Development Assistance Committee. http://www.oecd.org/document/51/0,2340, en_2649_34435_2501555_119808_1_ 1_1,00.html.
———. 2004a. Final ODA Data for 2003.
http://www.oecd.org/dataoecd/19/52/34352584.pdf.
———. 2004b. The Istanbul Ministerial Declaration on Fostering Growth of Innovative and Internationally Competitive SMEs. http://www.oecd.org/document/16/0,3343,en_2649_201185_32020176_1_1_1_1, 00.html.
———. 2005a. “Aid Rising Sharply, According to Latest OECD Figures.” http://www.oecd.org/dataoecd/0/41/35842562.pdf.
———. 2005b. The Paris Declaration. Development Cooperation Directorate. http://www.oecd.org/document/18/0,2340, n_2649_3236398_35401554_ 1_1_1_ 1,00.html.
———. 2006. Final ODA Data for 2005.
http://www.oecd.org/dataoecd/52/18/37790990.pdf.
———. 2007a. “Development Aid from OECD Countries Fell 5.1 Percent in 2006.” http://www.oecd.org/document/17/0,3343, en_2649_33721_3834126 5_1_1_1_1,00 .html.
———. 2007b. Final ODA Flows in 2006. http://www.oecd.org/ dataoecd/7/20/39768315.pdf.
———. 2007c. “Monitoring the Paris Declaration.” http://www.oecd.org/department/0,3355,en_2649_15577209_1_1_1_1_1,00.html.
OECD, and PUMA (Public Management Committee). 2002. “Overview of Results-Focused Management and Budgeting in OECD Member Countries.” Twenty-third annual meeting of OECD Senior Budget Officials, Washington, DC,
June 3–4.
Oxford Analytica. 2004a. “Foundations Muscle into Aid Arena.” August 10, Oxford.

———. 2004b. “Remittances Fund Investment Growth.” September 7, Oxford.
Patton, M. Q. 1997. Utilization-Focused Evaluation. 3rd ed. Thousand Oaks, CA: Sage Publications.
———. 2006. “Recent Trends in Evaluation.” Paper presented to the International Finance Corporation, Washington, DC, May 8.
Picciotto, Robert. 2002a. “Development Cooperation and Performance Evaluation: The Monterrey Challenge.” World Bank, Operations Development Department, Washington, DC.
———. 2002b. “Development Evaluation as a Discipline.” International Program for Development Evaluation Training (IPDET) presentation, Ottawa, June.
———. 2003. “International Trends and Development Evaluation: The Need for Ideas.” American Journal of Evaluation 24: 227–34.
Picciotto, Robert, and Ray C. Rist. 1995. Evaluating Country Development Policies and Programs: New Approaches and a New Agenda. San-Francisco: Jossey-Bass Publishers.
Pollin, Robert. 2007. Microcredit: False Hopes and Real Possibilities. Foreign Policy in Focus.
http://www.fpif.org/fpiftxt/4323.
Public Services Organisation Review Group. 1969. Report of the Public Services Organization Review Group. Dublin: Stationary Office.
Qureshi, Zia. 2004. Millennium Development Goals and Monterrey Consensus: From Vision to Action. World Bank, Washington, DC. http://wbln0018.worldbank.org/eurvp/web.nsf/Pages/Paper+by+Qureshi/US$File/MOHAMMED+QURESHI.PDF.
Rasappan, Arunaselam. 2007. “Implementation Strategies and Lessons Learnt with Results-Based Budgeting Malaysia.” Training course on program and performance budgeting, ITP Pune, India, October 1–5.
http:// blog-pfm.imf.org/pfmblog/files/rasappan_implementation_strategies_lessons_malaysia.pdf.
Republic of France. 2001. Ministry of the Economy, Finance, and Industry. “Towards New Public Management.” Newsletter on the Public Finance Reform 1 (September), Paris.
Rist, Ray C., and Nicoletta Stame, eds. 2006. From Studies to Streams: Managing Evaluative Systems. New Brunswick, NJ: Transaction Books.
Schacter, Mark. 2000. “Sub-Saharan Africa: Lessons from Experience in Support ing Sound Governance.” ECD Working Paper 7, World Bank, Evaluation Capac ity Department, Washington, DC.
Soros Foundations Network. 2007. “About OSI and the Soros Foundation Network.” http://www.soros.org/about/overview.

Takamasa, Akiyama, and Kondo Masanori, eds. 2003. “Global ODA since the Monterrey Conference.” Foundation for Advanced Studies on International Development (FASID), International Development Research Institute, Tokyo. http://www.fasid.or.jp/english/publication/research/pdf/global.pdf.

Tavistock Institute, in association with GHK and IRS. 2003. The Evaluation of Socio-Economic Development: The GUIDE.
http://coursenligne.sciences-po.fr/2004_2005/g_martin/guide2.pdf.

Thomas, Koshy. 2007. “Integrated Results-Based Management in Malaysia.” In Results Matter: Ideas and Experiences on Managing for Development Results. Asian Development Bank. http://www.adb.org/Documents/Periodicals/MfDR/dec-2007.pdf.

Trosa, Sylvie. 2008. “Towards a Postbureaucratic Management in France.” Politique et management public (2).

Uganda Office of the Prime Minister (OPM). 2007a. National Integrated Monitoring and Evaluation Strategy (NIMES) : 2006–2007 Bi-Annual Implementation Progress Report. Kampala.

———. 2007b. “Working Note: Monitoring and Evaluation of the National Development Plan.” October, Kampala.

UNCTAD (United Nations Conference on Trade and Development). 2008. “Foreign Direct Investment Reached New Record in 2007.” Press Release. http://www.unctad.org/Templates/Webflyer.asp?docID=9439&intItemID= 2068&lang=1.

UNECA (United Nations Economic Commission for Africa). 2007. “Financing Development Section, Trade, Finance, and Economic Commission for Africa, Addis Ababa.” http://www.uneca.org/eca_programmes/trade_and_regional_integration/ documents/MonterreyConsensusMainReport.pdf.

United Nations Office on Drugs and Crime. 2006. The United Nations Convention against Transnational Organized Crime and Its Protocols. http://www.unodc.org/unodc/en/treaties/CTOC/index.html.

U.S. Department of Labor. 1993. Government Performance and Results Act. Employment & Training Administration, Washington, DC. http://www.doleta.gov/performance/goals/gpra.cfm.

U.S. GAO (Government Accountability Office). 2003. Executive Guide: Effectively Implementing the Government Performance and Results Act. http://www.gao.gov/special.pubs/gpra.htm.

World Bank. n.d. “About Private Participation in Infrastructure.” http://www.worldbank.org/infrastructure/ppi/.

———. 1994. Enhancing Women’s Participation in Economic Development. Washing ton, DC: World Bank.
http://books.google.ca/books?id=CJBmEClPci8C&dq=World+Bank.+1994.+ En hancing+Women%E2%80%99s+Participation+in+Economic+Development.+

Washington,+DC:+World+Bank.&printsec=frontcover&source=bn&hl=en&ei= IPy2SdntFYueMoTU_NcK&sa=X&oi=book_result&resnum=4&ct=result#PPA5,M1.
———. 1996b. World Bank Participation Sourcebook. http://www.worldbank.org/wbi/sourcebook/sbhome.htm.
———. 1999. "Monitoring and Evaluation Capacity Development in Africa." Précis Spring (183). http://wbln0018.worldbank.org/oed/oeddoclib.nsf/7f2a291f9f1204c6852568080 06a0025/34b9bade34aca617852567fc00576017/US$FILE/183precis.pdf.
———. 2001. Strategic Directions for FY02–FY04. Washington, DC. http://lnweb18.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c852568080 06a0046/762997a38851fa0685256f82007 77e15/US$FILE/gppp_main_report_ phase_2.pdf#page=21.
———. 2003. Global Development Finance 2003. Washington, DC: World Bank. http://siteresources.worldbank.org/INTRGDF/Resources/GDF0slide0-show103010DC0press0launch.pdf.
———. 2004a. "Evaluating the World Bank's Approach to Global Programs: Address ing the Challenges of Globalization." Independent Evaluation Group, Washington, DC. http://www.worldbank.org/oed/gppp/.
———. 2004b. Global Monitoring Report 2004: Policies and Actions for Achieving the Millennium Development Goals. Washington, DC: World Bank.
———. 2005a. Global Monitoring Report 2005: Education for All. Washington, DC: World Bank.
———. 2005b. World Development Report 2005: A Better Investment Climate for Everyone. Washington, DC: World Bank.
———. 2006a. Global Monitoring Report 2006: Equity and Development. Washington, DC: World Bank.
———. 2007a. Aid Architecture: An Overview of the Main Trends in Official Development Assistance Flows, Executive Summary. http://siteresources.worldbank.org/IDA/Resources/Aidarchitecture-exec summary.pdf.
———. 2007b. Conflict Prevention and Reconstruction. http://lnweb18.worldbank.org/ESSD/sdvext.nsf/67ByDocName/Conflict PreventionandReconstruction.
———. 2007c. The Data Revolution: Measuring Governance and Corruption. http://web.worldbank.org/WBSITE/EXTERNAL/NEWS/0,,contentMDK:20190210~menuPK:34457~pagePK:34370~piPK:34424~theSite PK:4607,00.html.
———. 2007d. Doing Business: Economy Profile Reports. Washington, DC: World Bank. http://rru.worldbank.org/DoingBusiness/.

———. 2007e. The Enhanced HIPC Initiative: Overview. http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTDEBTDEPT/0,,contentMDK:21254881~menuPK:64166739~pagePK:64166689~piPK:64166646~theSitePK:469043,00.html.
———. 2007f. Global Monitoring Report 2007: Development and the Next Generation.
Washington, DC: World Bank.
———. 2008. Online Atlas of the Millennium Development Goals: Building a Better World.
http://devdata.worldbank.org/atlas-mdg/.
Woroniuk, B., and J. Schalwyk. 1998. "Donor Practices: Evaluation and Preparation Tipsheet." OECD, Paris.

Веб-сайти

Фінанси

CGAP (Consultative Group to Assist the Poor). http://www.cgap.org/.
———. Assessment and Evaluation. http://www.cgap.org/publications/assessment_ evaluation.html.
The Equator Principles. http://www.equator-principles.com/index.shtml.
Tedeschi, Gwendolyn Alexander. 2008. Overcoming Selection Bias in Microcredit
Impact Assessments: A Case Study in Peru. http://www.informaworld.com/smpp/content~content=a792696580~db=all~order=page.
World Bank. Doing Business: Economy Profile Reports. http://rru.worldbank.org/ DoingBusiness/.

Гендер

OECD (Organisation for Economic Co-operation and Development). "Gender Tipsheet, Evaluation." Development Assistance Committee. http://www.oecd.org/ dataoecd/2/13/1896352.pdf.

Управління

Transparency International.
http://www.transparency.org/.
World Bank. 2007a. Governance Matters.
http://info.worldbank.org/governance/ wgi2007/ and http://info.worldbank.org/governance/wgi2007/pdf/booklet_ decade_of_measuring_governance.pdf.
———. 2007b. Untitled Video.
http://web.worldbank.org/WBSITE/EXTERNAL/ NEWS/0,,contentMDK:21400275~menuPK:51416191~pagePK:64257043~piPK:437376~theSitePK:4607,00.html.

Цілі розвитку тисячоліття

United Nations. Millennium Development Goals.
http://www.un.org/millennium goals/.

Бідність

Poverty-Environment Web site.
http://www.povertyenvironment.net.
World Bank. AdePT software to make poverty analysis easier and faster.
econ.worldbank.org/programs/poverty/adept.
———. PovertyNet newsletter. http://www.worldbank.org/poverty.

ПІДГОТОВКА ТА ПРОВЕДЕННЯ ЕФЕКТИВНОГО ОЦІНЮВАННЯ РОЗВИТКУ

«Наші задуми провалюються, бо вони не мають мети. Коли людина не знає, до якої гавані прямує, жоден вітер не буде попутним».

—СЕНЕКА

Глава 3: Побудова орієнтованої на результат системи моніторингу й оцінювання

- Важливість орієнтованої на результат системи моніторингу і оцінювання
- Що таке орієнтована на результат система моніторингу і оцінювання?
- Порівняння традиційної та орієнтованої на результат системи моніторингу і оцінювання
- Десять кроків до побудови орієнтованої на результат системи моніторингу й оцінювання

Глава 4: Розуміння контексту оцінювання й програмної теорії змін

- Аналіз вихідних даних
- Визначення головного клієнта та основних зацікавлених сторін
- Розуміння контексту
- Вивчення існуючих знань
- Побудова, використання та оцінка теорії змін

Глава 5: Вибір підходу до оцінювання

- Загальні підходи до оцінювання
- Проблеми і переваги різних підходів

ГЛАВА 3

Побудова орієнтованої на результат системи моніторингу й оцінювання

Уряди країн всього світу намагаються задовольнити вимоги своїх громадян, які прагнуть кращого життя. Внутрішні й зовнішні тиски та вимоги, з якими стикаються уряди та організації з розвитку, спонукають їх шукати нові шляхи покращення державного управління. Це може означати посилення відповідальності та прозорості, а також підвищення дієвості інтервенцій. Орієнтовані на результат моніторинг і оцінювання (МіО) – це управлінський інструмент, що допомагає відстежувати прогрес та демонструвати вплив проектів, програм і політик розвитку.

Ця глава складається з чотирьох основних частин:

- Важливість орієнтованої на результат системи моніторингу й оцінювання
- Що таке орієнтована на результат система моніторингу й оцінювання?
- Порівняння традиційної та орієнтованої на результат системи моніторингу й оцінювання
- Десять кроків до побудови орієнтованої на результат системи моніторингу й оцінювання

Важливість орієнтованої на результат системи моніторингу й оцінювання

У країнах, що розвиваються, все частіше лунають вимоги щодо підвищення ефективності роботи державного сектору. У відповідь на ці вимоги країни розвивають системи управління ефективністю. Ці нові системи передбачають реформування бюджетного процесу, людських ресурсів та організаційної культури. Щоб оцінити, чи працюють заходи державного сектору, потрібно також проводити вимірювання ефективності. Системи МіО відстежують результати (або їх відсутність), досягнуті урядами та іншими сторонами.

Багато ініціатив підштовхують уряди до запровадження систем державного управління, що продемонстрували свою дієвість. Два приклади таких ініціатив – Цілі розвитку тисячоліття (ЦРТ) та Ініціатива для бідних країн з великою заборгованістю.

Окреслена в цій главі стратегія ґрунтується на досвіді розвинених країн – особливо тих, що входять до Організації економічного співробітництва і розвитку (ОЕСР) – але також відображає виклики та складнощі, з якими зіткнулося багато країн, що розвиваються, у своїх спробах запровадити системи вимірювання ефективності. Перед такими країнами постає багато проблем – від браку кваліфікованних працівників до неефективних управлінських структур та непрозорих систем. Хоча основна увага в цій главі приділяється підвищенню ефективності та відповідальності уряду за допомогою дієвої системи моніторингу і оцінювання (МіО), наведені принципи та стратегії є однаково застосовними до організацій, політик, програм і проектів.

Орієнтована на результат система МіО забезпечує надання критично важливої інформації про ефективність роботи державного сектору або організації. Вона допомагає особам, які формують політику та приймають рішення, а також іншим зацікавленим особам відповідати на фундаментальні запитання: чи було виконано обіцянки та досягнуто **довгострокових результатів**[1]. МіО – це засіб для демонстрації покращень або їх відсутності (блок 3.1).

Довгострокові результати:
Вигоди, що їх приносить проект, програма або політика (довгостроковий результат приносить поведінкову або організаційну зміни і не може бути придбаний за гроші)

Повідомляючи про результати різних інтервенцій, уряди та інші організації можуть сприяти підвищенню довіри суспільства до їх роботи. Інформування широкого загалу також є сприятливим для розвитку, в якому зараз спостерігається рух до більшої відповідальності при отриманні допомоги.

Ефективна орієнтована на результат система МіО може бути надзвичайно корисною як інструмент управління та мотивації. Вона допомагає зосереджувати увагу на досягненні довгострокових результатів, важливих для

[1] *В оригіналі це термін Outcome. Тут він перекладається як «довгострокові результати», але широко використовується також переклад «наслідки». Інколи вживають термін «результати», проте такий варіант перекладу може вживатись також і для оригінального терміну Output (примітки).*

організації та зацікавлених сторін, а також надає поштовх до встановлення ключових цілей та завдань, пов'язаних з такими довгостроковими результатами. Крім того, керівники отримують важливу інформацію про те, чи належною, правильною та адекватною є стратегія інтервенції в контексті змін, які ця інтервенція має принести.

Блок 3.1 «Ефект від вимірювання результатів»

Вимірювання результатів є критично важливим через наступні причини:

- Якщо не вимірювати результати, неможливо відрізнити успіх від невдачі.
- Якщо неможливо помітити успіх, неможливо нагородити за нього.
- Якщо неможливо винагородити людей за успішні дії, ймовірно, винагороджується невдача.
- Якщо неможливо помітити успіх, неможливо скористатися позитивним досвідом.
- Якщо неможливо помітити невдачу, неможливо їй запобігти.
- Демонструючи результати, ви отримуєте громадську підтримку.

Джерело: Osborn and Gaebler 1992.

Ефективна система МіО також є надважливим джерелом інформації для оптимізації та вдосконалення інтервенцій, покликаних максимально збільшити ймовірність успіху. Вона допомагає вже на ранніх етапах розпізнати перспективні інтервенції, щоб мати можливість запровадити їх і в інших місцях. Наявність даних про те, як працює певний проект, практика, програма або політика, дозволяє ефективно формулювати та обґрунтовувати бюджетні запити. Також ці дані дозволяють розподіляти обмежені ресурси саме серед тих інтервенцій, що приносять найбільші вигоди.

Після встановлення довгострокових результатів, вибору показників та цілей організація може працювати над їх досягненням. Система МіО здатна надавати оперативну інформацію щодо прогресу та швидко виявляти ознаки проблем, що вимагають певної реакції. Дані моніторингу також дозволяють виокремити об'єкти, що працюють особливо добре або особливо погано. Після цього можна проводити оцінювання для встановлення виявлених відхилень.

Що таке орієнтована на результат система моніторингу й оцінювання?

Інформація про результати може надходити з двох взаємодоповнюючих джерел: системи моніторингу та системи оцінювання (блок 3.2). Обидві системи є критично важливими для дієвого вимірювання ефективності.

Моніторинг з орієнтацією на результат:
Безперервний процес збору та аналізу інформації щодо ключових показників для вимірювання просування до цілей

Ціль/цільове значення:
Вимірний ступінь зміни показника, що має бути досягнутий за визначений термін

Показник/ Індикатор:
Вид даних, що систематично вимірюється і вказує міру просування до цілі

Оцінювання, орієнтоване на результат:
Оцінка запланованої, поточної або завершеної інтервенції з метою визначення її актуальності, ефективності, дієвості, впливу та сталості

Вплив:
Довгострокова вигода (результат), що досягається в разі досягнення певного набору довгострокових результатів (Цілі розвитку тисячоліття заявляють бажані впливи)

Блок 3.2 «Різниця між орієнтованими на результат моніторингом та оцінюванням»

Моніторинг з орієнтацією на результат – безперервний процес збору й аналізу інформації щодо ключових показників та порівняння фактичних результатів з очікуваними для визначення якості реалізації проекту, програми або політики. Це постійний процес вимірювання просування до певних коротко-, середньо- та довгострокових результатів шляхом контролю руху до конкретних, заздалегідь визначених **цілей** за допомогою **показників**. Персонал та особи, які приймають рішення, можуть отримувати завдяки орієнтованому на результат моніторингу зворотний зв'язок стосовно прогресу (або його відсутності) і використовувати цю інформацію задля покращення своєї роботи.

Оцінювання, орієнтоване на результат – оцінка запланованої, поточної або завершеної інтервенції з метою визначення її актуальності, ефективності, дієвості, **впливу** та сталості. Намір полягає в отриманні достовірної й корисної інформації з тим, щоб застосовувати набутий досвід в процесі прийняття реципієнтами рішень. Оцінювання ширше розглядає інтервенцію, запитуючи: наближення до цілі або певний результат спричинені інтервенцією чи зумовлені іншим чинником змін, зафіксованим системою моніторингу. Оцінювання може ставити, зокрема, наступні запитання:

- Чи є актуальними цілі та довгострокові результати?
- Наскільки результативно та ефективно вони досягаються?
- Які неочікувані наслідки були спричинені інтервенцією?
- Чи є інтервенція найбільш економічною та сталою стратегією для задоволення конкретних потреб?

Порівняння традиційної та орієнтованої на результат системи моніторингу й оцінювання

Уряди вже давно проводять традиційні МіО, відстежуючи, зокрема, власні доходи та видатки, укомплектованість штатів та ресурси, програмні та проектні заходи, кількість учасників, обсяги вироблених товарів та наданих послуг. Однак необхідно позначити різницю між традиційними та орієнтованими на результат системами МіО:

- **Традиційні системи МіО** зосереджуються на моніторингу та оцінці залучених ресурсів, заходів та безпосередніх результатів (тобто на реалізації проекту або програми).

- **Орієнтовані на результат системи МіО** поєднують традиційний підхід до моніторингу реалізації з оцінкою довгострокових результатів та впливів або результатів загалом.

Саме цей зв'язок прогресу реалізації з прогресом у досягненні бажаних цілей або результатів урядових політик і програм робить орієнтовані на результат МіО корисним інструментом державного управління. Реалізація системи МіО такого типу дозволяє організації корегувати як теорію змін, так і процеси реалізації для підтримки руху в напрямку бажаних цілей і довгострокових результатів.

Традиційний моніторинг і оцінювання:
Моніторинг і оцінювання, що зосереджуються на реалізації проекту або програми

Орієнтований на результат моніторинг і оцінювання:
Моніторинг і оцінювання, що поєднують традиційний підхід з акцентом на оцінці результатів (наслідків та впливу)

Теорія змін:
Теорія про те, як ініціатива приводить до потрібних результатів

Теорія змін

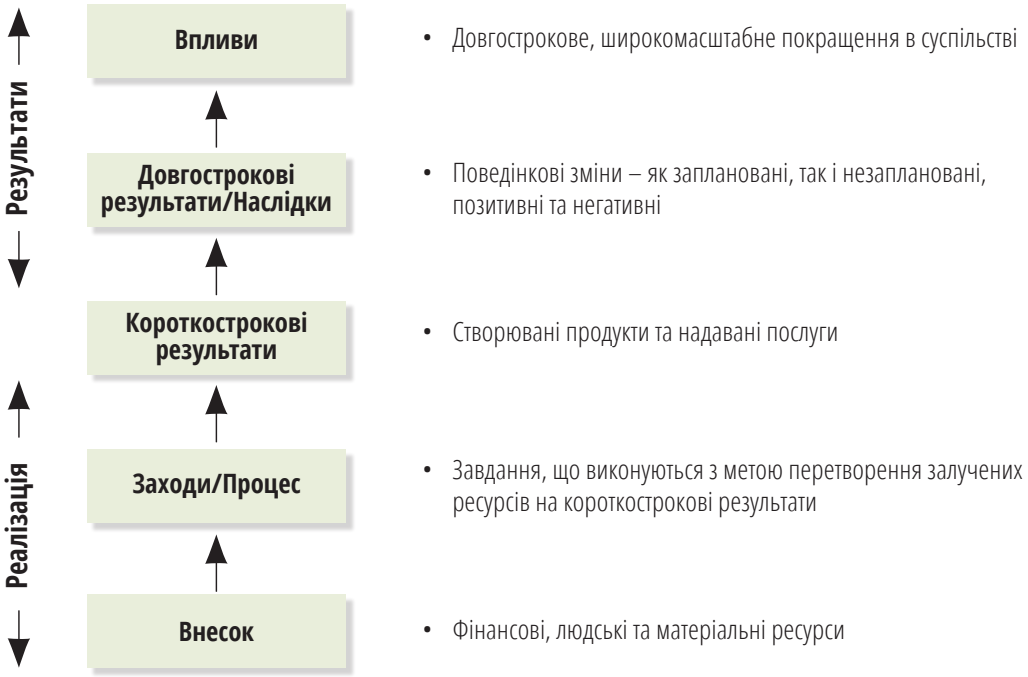
Один зі способів аналізу відмінностей між традиційними та орієнтованими на результат системами МіО полягає у розгляді теорії змін. Згідно з визначенням К'юзека та Ріста (2004), **теорія змін** – представлення того, як, згідно з очікуваннями, інтервенція повинна забезпечити потрібні результати. (додаткова інформація про теорію змін та відповідні визначення наведено в главі 4). Моделі теорії змін зазвичай включають п'ять основних складових: залучені ресурси, заходи, короткострокові результати, довгострокові результати і впливи (табл. 3.1). Деякі моделі теорії змін також включають інші параметри, такі як цільові групи, внутрішні та зовнішні фактори.

Таблиця 3.1. Основні складові теорії змін

Складова	Опис
Внесок /Залучені ресурси	Ресурси, що залучаються до реалізації проекту, програми або політики (кошти, персонал, обладнання, допоміжні матеріали і т. ін.).
Заходи/Процес	Що ми робимо. Заходи – це те, що можна позначити дієсловом («надати», «сприяти», «забезпечити»).
Короткострокові результати/Продукти/Результати	Те, що ми створюємо. Короткострокові результати – осяжні продукти або послуги, що створюються в результаті виконання заходів. Їх зазвичай позначають іменниками. Як правило, ці іменники не супроводжуються жодними модифікаторами. Короткострокові результати є осяжними і їх можна порахувати.
Довгострокові результати/Результати/Наслідки	Чому ми це робимо. Довгострокові результати – поведінкові зміни, спричинені безпосередніми результатами проекту (подолання звички до тютюнопаління, кип'ятіння води, використання надліжкових сіток проти комах). Довгострокові результати можуть бути збільшені, зменшені, скореговані, покращені або підтримувані.
Впливи	Довгострокові зміни, що є наслідком накопичення довгострокових результатів. Можуть бути аналогічними стратегічним завданням.

Джерело: Kusek and Rist 2004.

Рис. 3.1. Програмна теорія змін (логічна модель) для досягнення довгострокових результатів і впливів



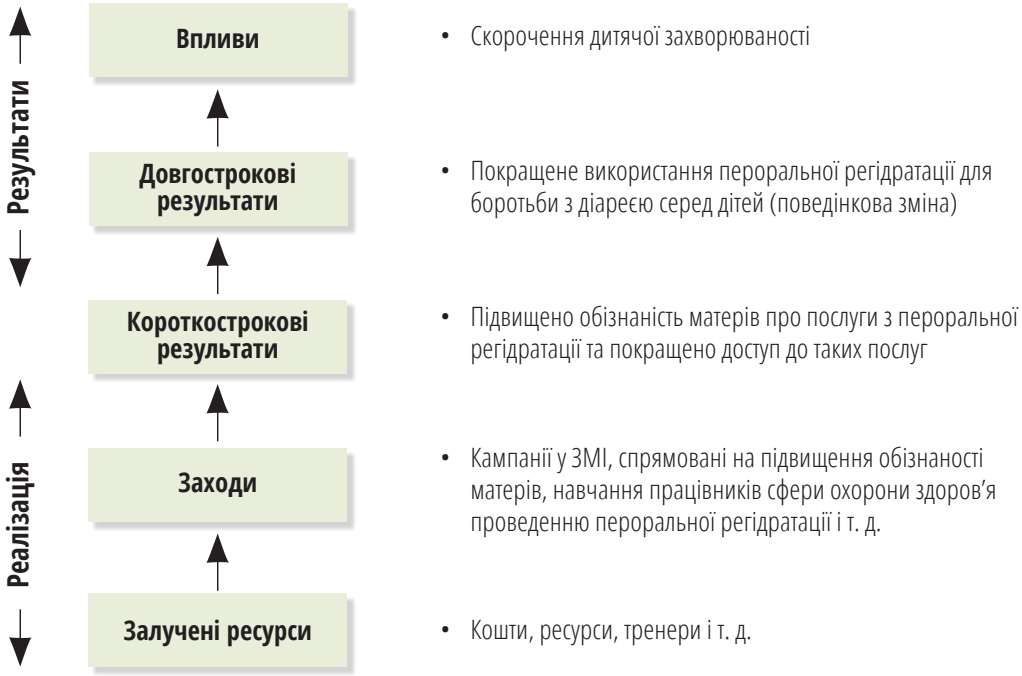
Джерело: Binnendijk 2000.

Теорію змін можна представити графічно (рис. 3.1). Цю модель можна використати для зображення орієнтованого на результат підходу до вирішення проблем – наприклад, скорочення дитячої захворюваності за допомогою пероральної регідратації (рис. 3.2).

Показники ефективності

Частиною моніторингу є вимірювання просування в напрямку досягнення певного довгострокового результату або впливу. Проте виміряти безпосередньо довгостроковий результат неможливо: його спочатку необхідно представити через набір показників, регулярне вимірювання яких надаватиме інформацію про те, чи досягаються довгострокові результати або впливи. Показник ефективності – це «змінна, що дозволяє перевіряти зміни в ході інтервенції розвитку або демонструє результати відносно плану» (OECD 2002, p. 29).

Рис. 3.2. Зразок програмної теорії змін (логічна модель) для скорочення дитячої захворюваності за допомогою пероральної регідратації



Джерело: Взято з Binnendijk 2000.

Наприклад, якщо країна обирає мету – покращення здоров'я дітей шляхом скорочення дитячої захворюваності на 30 відсотків протягом наступних п'яти років, вона повинна спочатку визначити набір показників, що дозволять перекласти зміни у частоті захворювань дітей на інфекційні хвороби у більш конкретний вимір. Серед показників, що можуть допомогти оцінити зміни в дитячій захворюваності, є наступні:

- частота випадків та поширення інфекційних захворювань, таких як гепатит (прямий чинник);
- рівень материнського здоров'я (непрямий чинник);
- ступінь доступності для дітей чистої води.

Саме сукупне представлення набору показників дає керівникам змогу побачити, чи досягає програма результатів. Жодний довгостроковий результат або вплив не можна виміряти за допомогою одного окремого показника.

Вивчення деталізованого набору показників (з розбивкою на складові) дає важливу інформацію щодо того, наскільки дієвими є урядові програми і політики з точки зору досягнення цільового дострокового результату або впливу. Цей підхід також використовується для виявлення місць, де робота ведеться краще або гірше порівняно з середніми показниками (відхилення), а також політик, що працюють добре або погано. Якщо, приміром, набір показників вказує на те, що з плином часу все менше дітей мають доступ до чистої води, уряд може використати цю інформацію для реформування програм, спрямованих на покращення водопостачання, або ж посилити ті програми, що інформують батьків про необхідність обеззараження води перед тим, як давати її дітям.

Інформація, отримана від системи моніторингу, показує ефективність тільки об'єктів вимірювання (хоча її можна також порівнювати з показниками у минулому, а також із запланованими показниками на даний момент та на майбутнє (цілями)). Дані моніторингу не розкривають причин наявних рівнів показників, не надають пояснень щодо змін ефективності у різні періоди або у різних місцях. Ця інформація надходить від системи оцінювання.

Система оцінювання виконує додаткову, але чітко виокремлену функцію в межах системи управління з орієнтацією на результат. Система оцінювання дозволяє:

- глибше дослідити довгострокові результати та впливи;
- використовувати інші джерела даних, крім відстежуваних показників;
- вивчати фактори, які надто складно або дорого контролювати постійно;
- дізнаватися, чому і як відстежувані за допомогою даних моніторингу тенденції мають такий вигляд.

Дані про вплив та причинно-наслідкові зв'язки вимагають серйозного ставлення. Вони можуть відіграти важливу роль в розподіленні стратегічних ресурсів.

Десять кроків до побудови орієнтованої на результат системи моніторингу й оцінювання

Побудова якісної системи МіО з орієнтацією на результат займає 10 кроків (рис. 3.3):

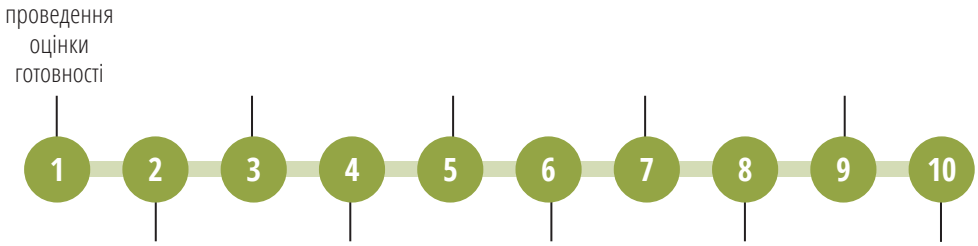
1) проведення оцінки готовності;
2) погодження контрольованих довгострокових результатів/наслідків;
3) вибір ключових показників моніторингу довгострокових результатів/наслідків;

4) збір базових даних за показниками;
5) планування покращень: вибір реалістичних цілей;
6) моніторинг результатів;
7) використання інформації оцінювання;
8) звітування про висновки;
9) використання висновків;
10) підтримування системи МіО в організації.

Рис. 3.3. Десять кроків до побудови та підтримки орієнтованої на результат системи моніторингу й оцінювання

Джерело: Kusek and Rist 2004.

Крок 1: Проведення оцінки готовності



Оцінка готовності – це спосіб визначення здатності та готовності уряду і його партнерів з розвитку до створення орієнтованої на результат системи МіО. В центрі уваги цієї оцінки – такі питання, як наявність або відсутність активних прибічників, а також стимулів, ролей та обов'язків, організаційного потенціалу та перешкод, що заважають розпочати процес.

Стимули

Перша частина оцінки готовності вимагає з'ясування того, які чинники стимулюють рух в напрямку побудови системи МіО (та які чинники можуть гальмувати прогрес). Необхідно розглядати, зокрема, наступні питання:

- Що спричиняє потребу в побудові системи МіО?
- Хто є активними прибічниками побудови та використання системи МіО?
- Що вмотивовує тих, хто активно відстоює побудову системи МіО?
- Хто отримає вигоди від системи?
- Хто не отримає вигід?

Ролі й обов'язки

Далі важливо визначити, хто зараз відповідає за створення даних в організації та інших відповідних організаціях, а також хто є головним споживачем даних. Необхідно розглядати, зокрема, наступні питання:

- Які ролі відіграють центральні та галузеві міністерства в оцінці ефективності?
- Якою є роль законодавчого органу?
- Якою є роль вищого органу фінансового контролю?
- Чи діляться міністерства й відомства інформацією одне з одним?
- Чи спостерігається політизація сформованих даних?
- Хто в країні формує дані?
- На яких рівнях в уряді використовуються дані?

Організаційний потенціал

Ключовими чинниками, що визначають готовність організації до побудови орієнтованої на результат системи МіО, є наявні в організації навички, ресурси та досвід. Необхідно розглядати, зокрема, наступні питання:

- Хто в організації володіє технічними навичками з розробки та впровадження системи МіО?
- Хто володіє навичками з управління такою системою?
- Які системи даних вже існують в організації та якої вони якості?
- Які є технології для підтримки необхідної системи даних? Оцінюватися мають обсяг бази даних, доступність аналізу даних, програмне забезпечення для звітування і т. д.
- Які фіскальні ресурси доступні для розробки й впровадження системи МіО?
- Який досвід має організація стосовно систем звітування про ефективність?

Перешкоди

Як і у випадку з будь-якими організаційними змінами, важливо враховувати перешкоди, що можуть заважати ефективному впровадженню системи. Необхідно розглядати, зокрема, наступні питання:

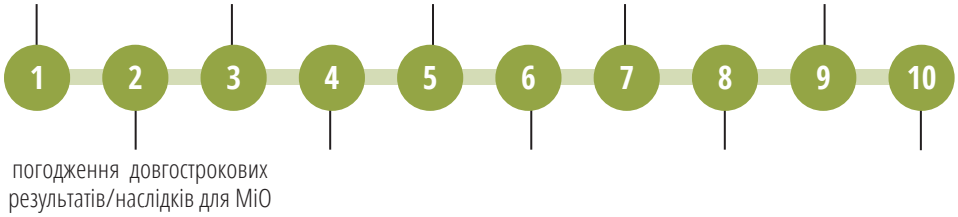
- Чи має місце брак фіскальних ресурсів, політичної волі або активних прибічників системи, орієнтованої на довгостроковий результат стратегії, досвіду?
- Якщо так, то яким чином можна подолати наявні перешкоди?

Передовий досвід свідчить про те, що успіх впровадження дієвої системи МіО визначається різними факторами, серед яких:

- чітке рішення про створення системи МіО, прийняте на національному рівні;
- документи стратегії скорочення бідності, відповідні закони та нормативні акти;
- сильне керівництво та підтримка на найвищих щаблях влади;
- надійна інформація, яку можна використати для прийняття політичних та управлінських рішень;
- громадянське суспільство, що працює з урядом як партнер з побудови системи та відстеження інформації про ефективність;
- інноваційні ділянки, що можуть слугувати початковими практиками або пілотними програмами.

Після оцінки готовності вищі урядовці мають відповісти на запитання: рухатися далі у побудові орієнтованої на результат системи МіО чи ні? Одним словом, треба сказати «так» чи «ні» (тепер, незабаром, згодом).

Крок 2: Погодження довгострокових результатів/наслідків для МіО



Важливо створити зацікавленість в оцінці довгострокових результатів та впливів, яких намагається добитися організація або уряд, замість того, щоб зосереджуватися лише на питаннях реалізації (залучені ресурси, за-

ходи, короткострокові результати). Адже саме довгострокові результати – такі як розширення охоплення дошкільних програм або підвищення якості навчання дітей в початкових школах – дають розуміння того, чи вдалося отримати певні вигоди.

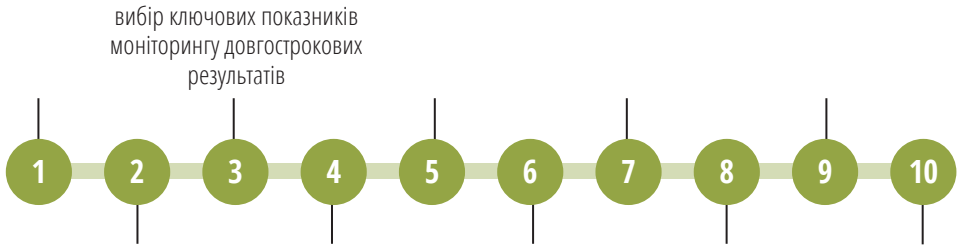
Стратегічні довгострокові результати та впливи повинні визначати, перш за все розподіл ресурсів та заходи уряду і його партнерів з розвитку. Ці впливи мають бути узгоджені зі стратегічними пріоритетами країни. До питань, що мають розглядатися при створенні переліку довгострокових результатів, належать, зокрема, наступні:

- Чи сформульовано національні/секторні цілі (наприклад, Бачення-2016)?
- Чи було зроблено політичні обіцянки з деталізацією покращення показників у певній сфері?
- Чи вказують дані опитувань громадян (наприклад, анкетування) на якісь конкретні актуальні проблеми?
- Чи існує прив'язка донорського кредитування до певних цілей?
- Чи запроваджені закони, що надають відповідні повноваження?
- Чи взяв на себе уряд серйозні зобов'язання стосовно досягнення Цілей розвитку тисячоліття (ЦРТ)?

Погодження довгострокових результатів – це політичний процес, що вимагає отримання згоди основних зацікавлених осіб. Для розуміння їх інтересів застосовуються мозкові штурми, інтерв'ю, фокус-групи та опитування.

Довгострокові результати чітко показують, якими мають бути наслідки заходів («потрібно знати, в який бік рухатись, перш ніж починати рух»). Це ті результати, які сподівається отримати клієнт. Але спочатку вони повинні бути представлені у вигляді набору ключових показників. Чіткий вибір довгострокових цілей з використанням відповідних показників – обов'язкова умова для проектування та побудови орієнтованої на результат системи МіО.

Крок 3: Вибір ключових показників моніторингу довгострокових результатів/наслідків



Зробленим, напевно, можна вважати лише те, що вдалося виміряти. Точне визначення показників, які потрібно контролювати, щоб визначити досягнення довгострокових результатів, не лише допомагає відстежувати прогрес, але й може стати потужним стимулом для зосередження зусиль та узгодження дій всередині організації – якщо визначити показники на достатньо ранньому етапі.

Показник є мірою, яка, за умови систематичного контролю, вказує на просування до цілі (або ж на відсутність такого просування). Він відповідає на запитання: «Як розпізнати успіх, якщо його буде досягнуто?» В нових системах МіО всі показники мають бути кількісними; якісні показники можна вводити пізніше, коли система МіО буде більш зрілою.

Розробка показників – ключовий момент у побудові системи МіО; вона є визначальною для всіх наступних дій зі збору й аналізу даних та звітування про них. Не слід недооцінювати політичні та методологічні проблеми створення достовірних та адекватних показників. Ск'яво-Кампо (1999) зазначає, що показники повинні відповідати критеріям CREAM, тобто бути:

- "clear" – ясними (чіткими та недвозначними);
- "relevant" – актуальними (відповідними предмету оцінки);
- "economic" – економічними (пов'язані з їх відстеженням витрати мають бути поміркованими);
- "adequate" – адекватними (здатними створити базис для оцінки ефективності);
- "monitorable" – контрольованими (повинні піддаватися незалежній перевірці).

Іноді можна знизити витрати, використовуючи вже наявні показники (хоча фахівцям з оцінювання слід мати на увазі ризики, пов'язані з використанням таких показників). Водночас перед цим важливо розглянути актуальність таких показників (як зараз, так і в майбутньому). Деякі показники можуть потребувати адаптації або доповнення додатковими показниками, тісніше пов'язаними з оцінюваним проектом, програмою або політикою.

Кількість показників залежить від того, скільки їх потрібно для відповіді на запитання: «Чи досягнутий потрібний довгостроковий результат?» Зазвичай використовують від двох до семи показників. Вибрані показники не є недоторканними. В ході оптимізації та вдосконалення системи моніторингу одні показники додаються, інші – виключаються.

Обрані показники ефективності, як і стратегії збору даних, що використовуються для збору інформації про ці показники, мають базуватися на реальному стані речей (Kusek and Rist 2004). Потрібно враховувати, зокрема, наступні фактори:

- які системи даних є в наявності;
- які дані можна зараз збирати;
- який є потенціал для розширення та поглиблення збору та аналізу даних.

Заповнивши всі секції матриці, наведеної в табл. 3.2, можна визначити реалістичність фактичного застосування кожного показника. Приклади показників наведені в табл. 3.3.

- Фахівці з оцінювання повинні впроваджувати власні показники в інтересах оцінювання, яке вони проводять.
- З першої спроби обрати належні показники не вдасться. Визначення кінцевого варіанту набору показників потребує часу.
- Всі показники необхідно формулювати нейтрально, а не як «збільшення X» або «зменшення X».
- Фахівці з оцінювання повинні знову й знову проводити пілотні дослідження.

Таблиця 3.2. Матриця для вибору показників

Показник	Джерело даних	Метод збору даних	Хто збиратиме дані?	Частота збору даних	Витрати, пов'язані зі збором даних	Складність зборіу даних	Хто буде аналізувати та звітувати дані?	Хто використовуватиме дані?
1.								
2.								
3.								

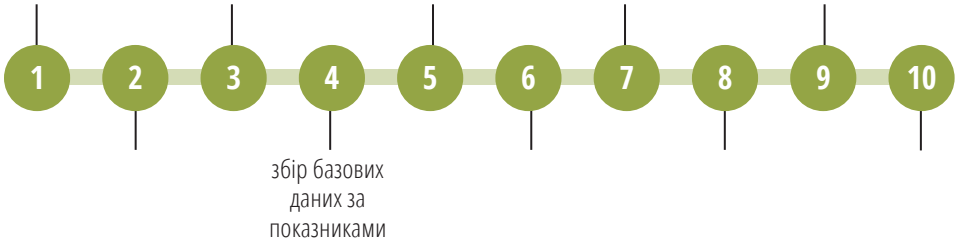
Джерело: Kusek and Rist 2004.

Таблиця 3.3. Зразок системи оцінки результативності роботи: довгострокові результати/наслідки та показники

Довгостроковий результат/Наслідок	Показники	Базові значення	Цілі/Цільові значення
Покращення охоплення дошкільних програм	Відсоток міських дітей, що мають право на дошкільну освіту й отримують її		
	Відсоток сільських дітей, що мають право на дошкільну освіту й отримують її		
Покращена якість початкової шкільної освіти	Відсоток шестикласників, що набирають не менше 70 відсотків балів на стандартних тестах з математики й природничих наук		
	Відсоток шестикласників, що набирають на стандартних тестах з математики й природничих науках більше балів порівняно з базовими результатами		

Джерело: Kusek and Rist 2004.

Крок 4: Збір базових даних за показниками



Вимірювання прогресу (або його відсутності) в досягненні довгострокових результатів/наслідків починається з опису й оцінки початкових умов. Збір базових даних означає, по суті, перше вимірювання показників з тим, щоб дізнатися про поточний стан речей.

Базові значення показників – це інформація (якісна або кількісна) про стан показника на початку інтервенції (або безпосередньо перед такою інтервенцією). Насправді одним із факторів вибору показників є доступність базових даних, що дозволяє проводити оцінку в порівнянні з початковими даними.

Джерела базових даних можуть бути первинними (такими, що збираються конкретно під цю систему оцінки) або вторинними (збираються для інших цілей). Вторинні дані можуть надходити зсередини організації, від уряду або з міжнародних джерел. Використання таких даних може допомогти заощадити кошти й отримати необхідну інформацію. Надзвичайно складно буде отримати первинні базові дані, якщо згодом виявиться, що вторинне джерело не задовольняє потреби оцінювання.

До можливих джерел базових даних відносяться:

- документи (паперові й електронні);
- особи, залучені до проекту, програми або політики;
- громадськість;
- підготовлені спостерігачі;
- механічні заміри й тести;
- географічні інформаційні системи.

Після вибору джерел базових даних щодо показників фахівці з оцінювання вирішують, хто і як має збирати дані. Вони визначають і розробляють інструменти для отримання даних, такі як форми для збору інформації з файлів або записів, протоколів інтерв'ю, опитувань та спостережень. Розробляючи такі інструменти, вони звертають увагу на практичні питання:

- Чи є в наявності (або чи можуть бути легко отримані) якісні дані?
- Чи можна регулярно і своєчасно отримувати дані для відстеження прогресу?
- Чи є заплановане збирання первинних даних раціональним і економічним?

Існує багато шляхів збору даних (розглядаються в главі 9). Ці шляхи можна ранжувати від найменш ретельних, найменш формальних та витратних до найбільш ретельних, формальних та витратних (рис. 3.4).

В табл. 3.4 показано третій крок розробки довгострокових результатів для освітньої політики – встановлення базових значень показників.

Рис. 3.4. Спектр методів збору даних

Джерело: Kusek and Rist 2004.

Таблиця 3.4. Зразок системи оцінки результативності роботи: довгострокові результати/наслідки, показники й базові значення

Довгостроковий результат/Наслідки	Показники	Базові значення	Цілі/Цільові значення
Покращення охоплення дошкільних програм	Відсоток міських дітей, що мають право на дошкільну освіту й отримують її	75 відсотків у міській місцевості станом на 1999 рік	
	Відсоток сільських дітей, що мають право на дошкільну освіту й отримують її	40 відсотків у сільській місцевості станом на 2000 рік	
Покращена якість початкової шкільної освіти	Відсоток шестикласників, що набирають не менше 70 відсотків балів на стандартних тестах з математики й природничих наук	У 2002 році не менше 70 відсотків балів на тестах з математики набрало 47 відсотків школярів, а на тестах з природничих наук – 50 відсотків	
	Відсоток шестикласників, що набирають на стандартних тестах з математики й природничих науках більше балів порівняно з базовими результатами	В 2002 році середній бал з математики серед шестикласників складав 68 відсотків, а середній бал з природничих наук – 53 відсотки	

Джерело: Kusek and Rist, 2004.

Крок 5: Планування покращень: вибір реалістичних цілей/цільових значень

Наступний крок – встановлення цілей/цільових значень – це останній крок в побудові системи оцінки результативності роботи (Performance Framework[2]). «По суті, цілі – це вимірні рівні показників, що їх країна, суспільство або організація прагнуть досягти до певного моменту часу» (Kusek and Rist (2004, p. 91)).

[2] *Можна зустріти інші інтерпретації терміну українською мовою. Наприклад: «рамка досягнень», «таблиця індикаторів», «структура показників результативності» та ін.*

Більшість довгострокових результатів та майже всі види впливів у міжнародному розвитку є складними і досягаються лише через значний проміжок часу. Таким чином, існує потреба встановлення проміжних цілей, що вказують, якого прогресу в напрямку довгострокового результату має бути досягнуто за певний період часу із залученням певних ресурсів. Вимірювання виконання цілей може передбачати використання як прямих, так і непрямих показників, як кількісних, так і якісних даних.

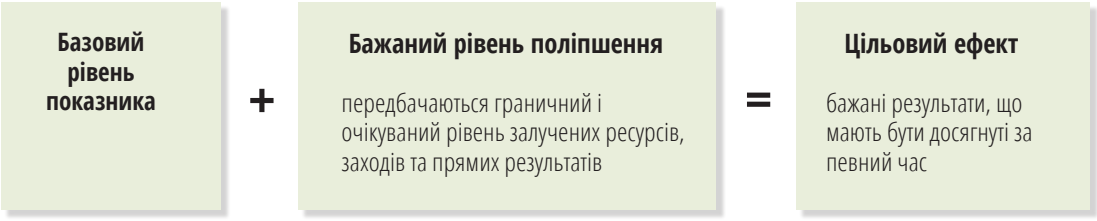
Впливи теорії змін можна розглядати як довгострокові цілі, які є кінцевою метою інтервенції. Довгострокові результати є набором цільових значень показників, послідовних та раціональних (відносно базових значень), що мають бути досягнуті завдяки заходам у відведені для цього реалістичні (з точки зору політики та бюджету) проміжки часу. Інакше кажучи, якщо організація через певний час досягне своїх цілей, вона отримає й потрібний довгостроковий результат (за наявності якісної теорії змін та її успішного втілення в життя).

Встановлюючи цільові значення показників, важливо чітко розуміти наступне:

- базова початкова точка (наприклад, середнє значення за останні три роки, минулий рік, середня тенденція);
- теорія змін та спосіб її розбивки на набір результатів зі вказаними термінами їх досягнення;
- рівні фінансових та кадрових ресурсів, залучених за період часу для досягнення цілі;
- кількість сторонніх ресурсів, що, як очікується, будуть залучені додатково до існуючих ресурсів програми;
- відповідні політичні міркування;
- організаційний (особливо управлінський) досвід реалізації проектів і програм у цій предметній області.

На рис. 3.5 показано, як визначати цілі, що слугують ланками єдиного ланцюга досягнень, які з часом дозволять забезпечити потрібний достроковий результат.

Рис. 3.5. Виявлення очікуваних чи бажаних рівнів поліпшення



Джерело: Kusek and Rist 2004.

Для кожного показника слід задавати лише одне цільове значення. Якщо показник раніше ніколи не використовувався, фахівцю з оцінювання варто уникати визначення конкретної цілі: краще задати ціль у вигляді діапазону. Цілі потрібно встановлювати на середній термін (не більше трьох років). На їх досягнення необхідно відводити реалістичні періоди часу. Більшість цілей встановлюються щороку, але деякі з них можна встановлювати щоквартально або на довші періоди.

В табл. 3.5 показано останній крок розробки довгострокових результатів для освітньої політики: встановлення цільових показників. Тепер ця заповнена матриця демонструє структуру оцінки. Вона визначає довгострокові результати та окреслює план щодо визначення того, чи вдалося програмі цих результатів досягти. Структура визначає схему орієнтованої на результат системи МіО, що почне надавати інформацію про те, як досягаються проміжні цілі на шляху до довгострокового результату.

Таблиця 3.5. Зразок системи оцінки результативності роботи: довгострокові результати/наслідки, показники, базові значення й цілі/цільові значення

Довгостроковий результат	Показники	Базові значення	Цілі/Цільові значення
Покращення охоплення дошкільних програм	Відсоток міських дітей, що мають право на дошкільну освіту й отримують її	75 відсотків в міській місцевості станом на 1999 рік	85 відсотків в міській місцевості станом на 2006 рік
	Відсоток сільських дітей, що мають право на дошкільну освіту й отримують її	40 відсотків в сільській місцевості станом на 2000 рік	60 відсотків в сільській місцевості станом на 2006 рік
Покращена якість початкової шкільної освіти	Відсоток шестикласників, що набирають не менше 70 відсотків балів на стандартних тестах з математики й природничих наук	У 2002 році не менше 70 відсотків балів на тестах з математики набрало 47 відсотків школярів, а на тестах з природничих наук – 50 відсотків	У 2006 році не менше 70 відсотків балів на тестах з математики має набрати 80 відсотків школярів, а на тестах з природничих наук – 67 відсотків
	Відсоток шестикласників, що набирають на стандартних тестах з математики й природничих науках більше балів порівняно з базовими результатами	У 2002 році середній бал з математики серед шестикласників складав 68 відсотків, а середній бал з природничих наук – 53 відсотки	У 2006 році середній бал з математики буде 78 відсотків, а середній бал з природничих наук – 65 відсотків

Джерело: Kusek and Rist 2004.

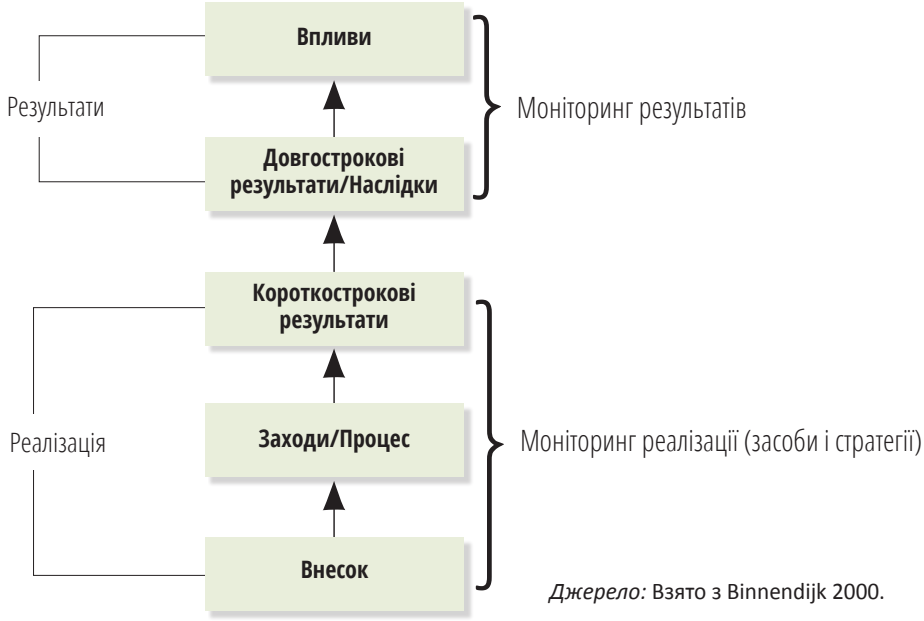
Структура допомагає фахівцям з оцінювання планувати оцінювання. Вона також допомагає управлінцям у питаннях формування бюджету, розподілу ресурсів, кадрового забезпечення та ін. Управлінцям слід регулярно звертатися до сформованої структури з тим, щоб переконатися, що проект, програма або політика рухаються до потрібних довгострокових результатів.

Встановлення цільових показників є критичною умовою досягнення довгострокових результатів політики. Застосування процесу спільної роботи з використанням базових значень показників та цільових рівнів покращення через певний час – основа орієнтованої на результат системи МіО.

Крок 6: Моніторинг результатів

Орієнтована на результат система моніторингу відстежує як саму реалізацію (залучені ресурси, заходи, короткострокові результати), так і результати (довгострокові результати та впливи). На рис. 3.6 показано, яке місце кожен з цих видів моніторингу займає в моделі.

Рис. 3.6. Основні види моніторингу

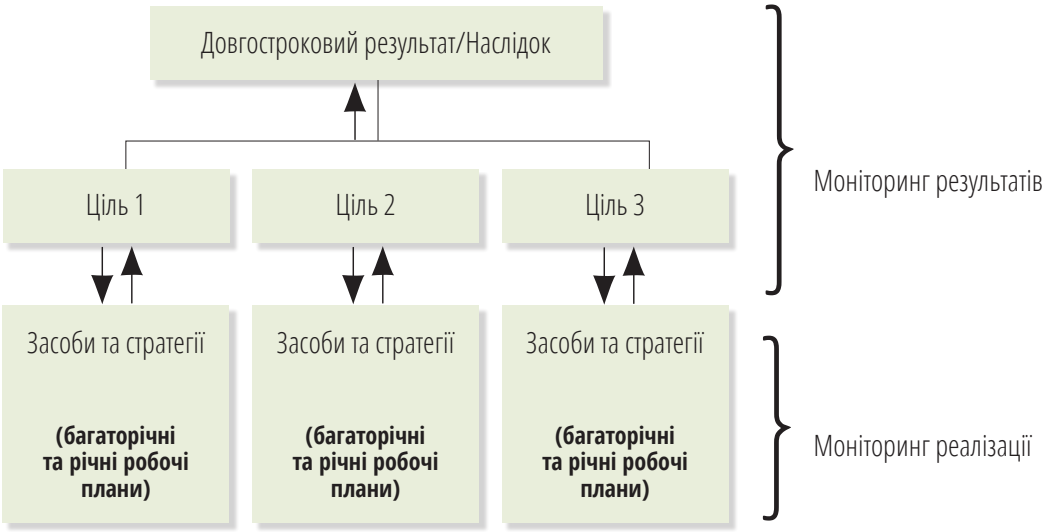


Джерело: Взято з Binnendijk 2000.

Кожен довгостроковий результат визначається низкою показників, для кожного з яких встановлюється цільове значення. Для досягнення цих цільових значень необхідно скоординувати й реалізувати низку заходів і стратегій (Рис. 3.7).

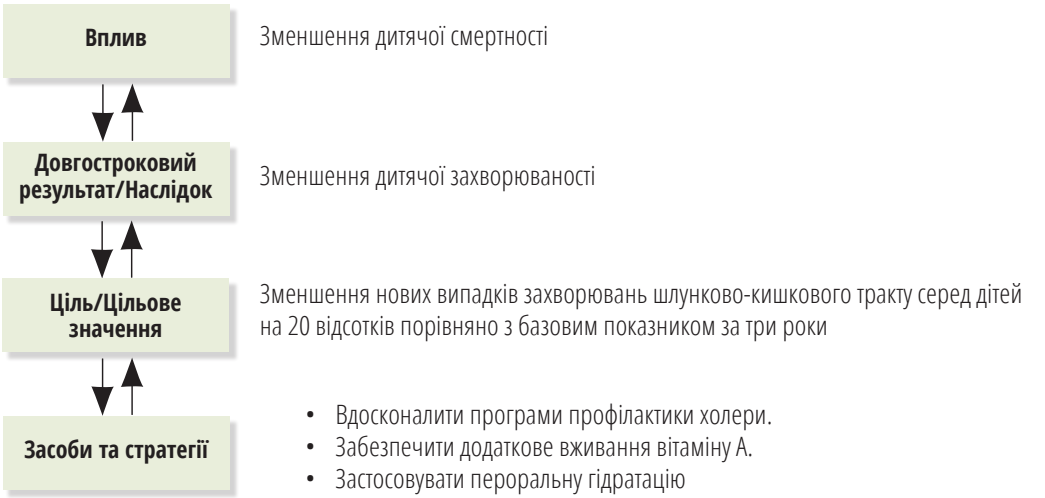
Поєднання моніторингу реалізації з моніторингом результатів є критично важливим. На рис. 3.8 наведено приклад роботи щодо скорочення дитячої смертності.

Рис. 3.7. Зв'язки між моніторингом реалізації та моніторингом результатів



Джерело: Kusek and Rist 2004.

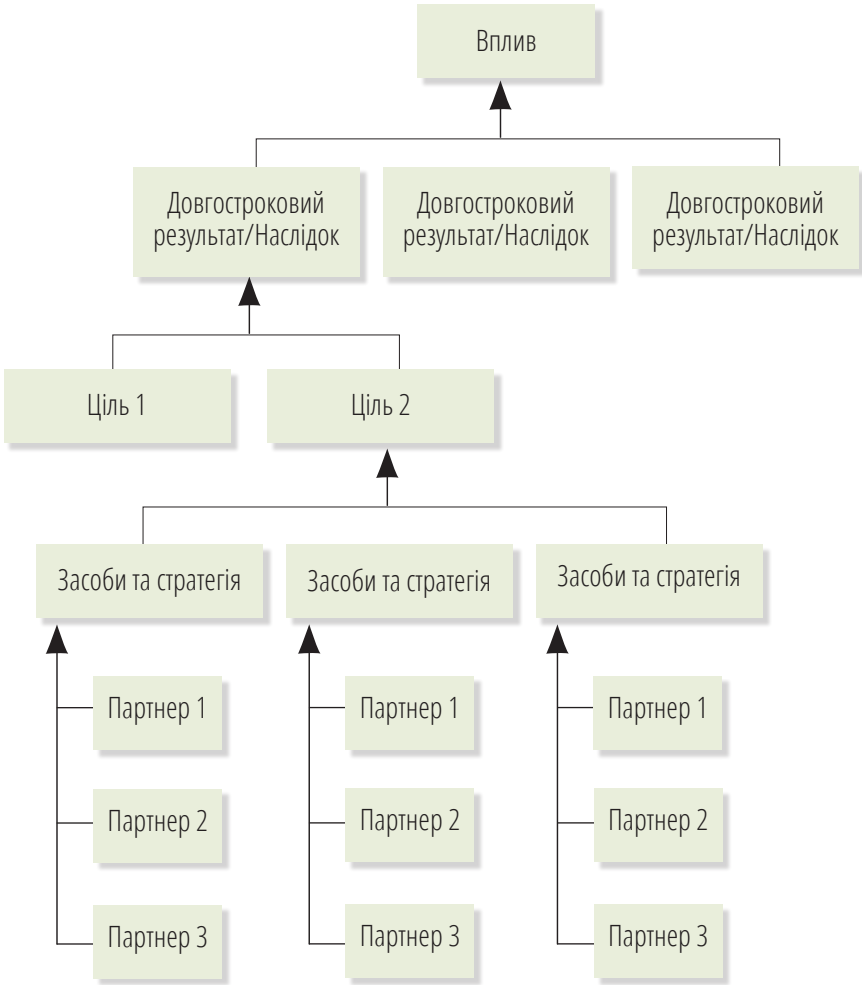
Рис. 3.8. Приклад зв'язку моніторингу реалізації та моніторингу результатів



Джерело: Kusek and Rist 2004.

Робота з партнерами все більше стає стандартом роботи з розвитку. Багато партнерів на найнижчому щаблі цієї ієрархії є потенційними постачальниками ресурсів, заходів та короткострокових результатів, що є частиною стратегії досягнення цілей (рис. 3.9).

Рис. 3.9. Досягнення результатів за рахунок партнерства



Джерело: Kusek and Rist 2004.

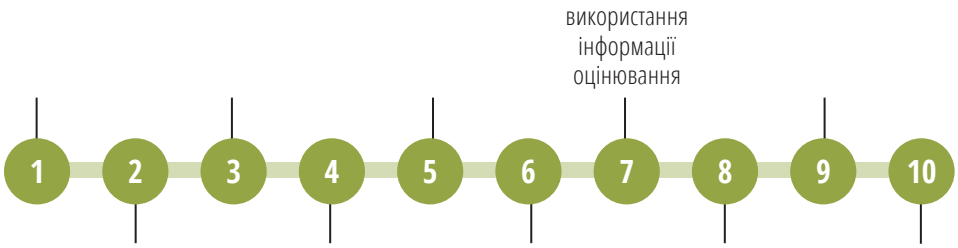
Потужна система МіО повинна підтримуватися бюджетом, планами кадрового забезпечення та плануванням діяльності. Побудова дієвої системи МіО передбачає виконання адміністративних та інституційних завдань, в тому числі:

- впровадження інструкцій зі збору й аналізу даних та звітності;
- призначення відповідальних осіб за різні види діяльності;
- запровадження засобів контролю якості;
- встановлення графіків і планування витрат;
- проробку ролей та обов'язків уряду, інших партнерів з розвитку та громадянського суспільства;
- впровадження інструкцій з прозорості та поширення інформації й аналізу.

Для забезпечення успішності будь-якої системи МіО потрібні:

- відповідальність за неї (володіння нею);
- управління;
- підтримка діяльності;
- достовірність.

Крок 7: Використання інформації оцінювання



Оцінювання відіграє важливу роль у доповненні інформації, що отримується від моніторингу прогресу в напрямку довгострокових результатів і впливів. У той час як моніторинг демонструє наші дії в контексті показників, цілей та довгострокових результатів, оцінка показує:

- чи правильні речі ми робимо (стратегія);
- чи правильно ми робимо те, що робимо (операції);
- чи можливо вдосконалити нашу діяльність (досвід).

Оцінювання здатне розглядати багато важливих питань, що виходять за межі простої системи моніторингу. Наприклад, схеми проведення багатьох інтервенцій базуються на певних припущеннях про причинно-наслідкові зв'язки стосовно певної проблеми або питання. Оцінювання може підтвер-

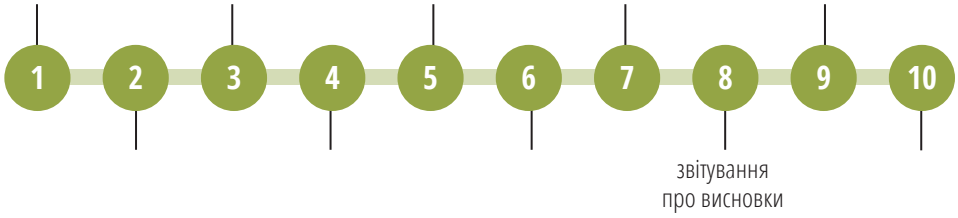
дити або заперечити ці припущення. При цьому застосовуються базована на теорії оцінка та логічні моделі (див. главу 4). Оцінювання може також глибше розглядати результати або тенденції, виявлені системою моніторингу, що викликають цікавість або занепокоєння (наприклад, з'ясовуючи, чому дівчата кидають сільську школу на кілька років раніше, ніж хлопці).

Оцінювання може застосовуватися на додаток до моніторингу:

- у будь-який час, коли виявляється неочікуваний результат або відхилення в показниках, що вимагають додаткового розгляду;
- при розподіленні між проектами, програмами та політиками ресурсів або бюджету;
- коли приймається рішення щодо продовження або припинення пілотного проекту;
- коли протягом тривалого часу не спостерігається жодних покращень і немає чіткого пояснення, чому так відбувається;
- коли аналогічні програми або політики дають різні довгострокові результати, або коли показники того самого довгострокового результату демонструють різні тенденції;
- для отримання уяви про побічну дію інтервенцій;
- коли необхідно оцінити користь, вартість та значення зробленого;
- при вивченні витрат у порівнянні з вигодами.

Покладаючись на інформацію, отриману від системи МіО, уряди й організації стають залежними від достовірності зібраної ними інформації. Неякісна, неточна й упереджена інформація не приносить жодної користі.

Крок 8: Звітування про результати оцінювання



Визначення того, про які результати оцінювання і кому слід повідомляти, в якому форматі та з якою регулярністю – критичний момент для МіО. Аналіз та звітування даних:

- дає інформацію про стан реалізації проектів, програм та політик;
- підказує шляхи вирішення проблем;
- робить можливим розгляд заходів з удосконалення стратегій реалізації;
- дає важливу інформацію про тенденції й напрямки;

- допомагає підтвердити або заперечити теорію змін, на якій ґрунтується проект, програма або політика (аналіз та звітування даних детально розглядаються в наступних главах).

Фахівець з оцінювання повинен обов'язково встановити, коли матимуть місце основні точки прийняття рішень щодо проекту, програми або політики, щоб точно знати, коли результати МіО будуть для осіб, що приймають рішення, найбільш корисними. Якщо дані та їх аналіз будуть отримані запізно, вони не вплинуть на рішення.

Необхідно повідомляти про всі важливі результати – як позитивні, так і негативні (табл. 3.6). Ефективна система МіО повинна включати в себе систему раннього попередження для виявлення проблем і неузгодженостей, а також для демонстрації цінності інтервенції. Робочі звіти мають містити пояснення слабких або незадовільних результатів, а також документувати всі кроки, зроблені для корегування таких результатів.

Аналізуючи та представляючи дані, фахівці з оцінювання повинні:

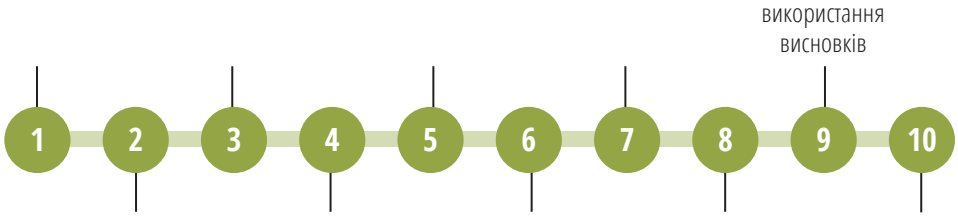
- порівнювати показники з базовими і цільовими значеннями та представляти відповідну інформацію в простому для розуміння графічному вигляді (див. главу 13);
- порівнювати поточну інформацію з даними за минулі періоди, шукати патерни й тенденції;
- уникати радикальних висновків, що ґрунтуються на недостатній інформації (чим більше даних отримано, тим більше певності мають фахівці з оцінювання про реальність тенденцій).
- в жодному випадку не карати людей, які приносять погану звістку (неприємні висновки можуть вказувати на нові тенденції або заздалегідь сповіщати керівників про проблеми, даючи час, необхідний для їх вирішення).

Таблиця 3.6. Зразок таблиці звітування про довгострокові результати/наслідки

Показник довгострокового результату/наслідку	Базове значення	Поточне значення	Ціль	Різниця (цільове значення – поточне значення)
Рівень поширення гепатиту (N = 6 000)	30	35	20	–5
Відсоток дітей, загальний стан здоров'я яких поліпшився (N = 9 000)	20	20	24	–4
Відсоток дітей, що набирають 4 з 5 позитивних оцінок під час фізикального обстеження (N = 3 500)	50	65	65	0
Відсоток дітей, харчування яких поліпшилось (N = 14 000)	80	85	83	2

Джерело: Kusek and Rist 2004.

Крок 9: Використання результатів



Головне для системи МіО – не просто створити орієнтовану на результат інформацію, але й своєчасно надати її відповідним користувачам, щоб ті змогли її врахувати (або усвідомлено не враховувати) в управлінні проектами, програмами або політиками. Партнери з розвитку та громадянське суспільство відіграють важливу роль у використанні інформації для зміцнення відповідальності, прозорості та процедур розподілу ресурсів.

Серед стратегій розповсюдження інформації, що можуть застосовуватися на будь-якому урядовому рівні, є наступні:

- *Надати повноваження ЗМІ*. ЗМІ здатні допомогти у розповсюдженні результатів роботи орієнтованих на результат систем МіО. Вони також можуть викривати корупцію та закликати до вдосконалення державного управління.
- *Прийняти закон про свободу доступу до інформації*. Свобода доступу до інформації – потужний інструмент, що може бути використаний для розповсюдження інформації серед зацікавлених сторін.
- *Впровадити електронне врядування*. Електронне врядування передбачає використання інформаційних технологій для покращення доступності, охоплення, інформування та надання послуг. Електронне врядування дозволяє зацікавленим особам взаємодіяти безпосередньо з урядом для отримання інформації та навіть ведення комерційної діяльності в режимі онлайн.
- *Додати інформацію у внутрішні та зовнішні Інтернет-файли*. Інформацію, так само, як і висновки моніторингу, можна розповсюджувати через внутрішні (відомчі або урядові) та зовнішні веб-сайти. Багато агенцій також створюють бази даних МіО з функцією пошуку.
- *Публікувати річні фінансові звіти*. Найкращий спосіб повідомляти про те, як використовуються гроші платників податків – оприлюднювати бюджетну інформацію. Це дозволяє громадянам контролювати якість та рівень надаваних урядом послуг, а також дізнаватися про те, які послуги або програми уряд вважає пріоритетними.
- *Залучати громадянське суспільство й групи громадян*. Залучення громадянського суспільства й організованих груп громадян сприяє їх більшій активності, відповідальності та узгодженню щодо того, яка інформація є необхідною.

- *Зміцнити законодавчий нагляд*. Законодавчі органи багатьох країн, як розвинених так і тих, що розвиваються, в рамках своєї наглядової функції вимагають надання інформації про результати роботи. Вони прагнуть впевнитись, що бюджети використовуються ефективно.
- *Посилити роль адміністрації генерального ревізора*. Адміністрація генерального ревізора – це головний партнер у визначенні ефективності роботи уряду. Через те, що ревізійні відомства вимагають більше інформації про результати роботи державного сектору, проекти, програми й політики реалізуються ефективніше.
- *Обмінюватися результатами з партнерами з розвитку й порівнювати отримані дані*. В результаті запровадження національних стратегій скорочення бідності й інших аналогічних стратегій і політик, партнери з розвитку (особливо двосторонні та багатосторонні агенції з надання допомоги) обмінюються результатами й висновками, порівнюючи їх.

Розуміння корисності інформації про показники роботи для різних користувачів – головна причина побудови системи МіО як такої. Головні потенційні користувачі в багатьох суспільствах – такі як громадяни, НУО, приватний сектор – нерідко залишаються осторонь інформаційного потоку. Дані МіО можуть використовуватися як для внутрішнього (урядового), так і для зовнішнього (громадського) вжитку. Потреба у цьому має визнаватися та узаконюватися (див. блок 3.3).

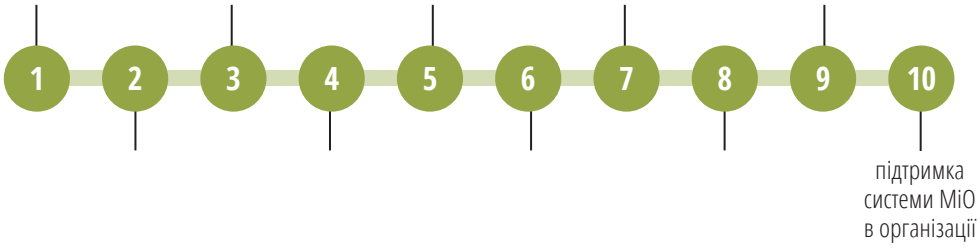
Блок 3.3 «Десять способів використання даних про результати»

Дані про результати можуть використовуватися для:

1. Реакції на вимоги виборних посадовців і суспільства щодо відповідальності.
2. Формулювання й обґрунтування бюджетних запитів.
3. Допомоги у прийнятті рішень про розподіл операційних ресурсів.
4. Початку глибокого вивчення наявних проблем (з теорією змін або реалізацією) та внесення коректив.
5. Додаткової мотивації персоналу до продовження вдосконалення програм.
6. Моніторингу роботи підрядників та реципієнтів (тепер їм недостатньо буде просто складати документи про те, які вони зайняті).
7. Надання даних для спеціальних, глибоких оцінювань програм.
8. Допомоги відстеження надання послуг в порівнянні з точно заданими довгостроковими цілями (чи правильно ми робимо те, що робимо?).
9. Підтримки стратегічного та іншого довгострокового планування (чи правильні речі ми робимо?).
10. Контакту з громадськістю для побудови довіри з боку суспільства.

Джерело: Hatry 1999

Крок 10: Підтримування функціонування системи МіО в організації



Забезпечення довговічності й корисності орієнтованої на результат системи МіО – складне завдання. Для підтримки життєдіяльності системи критично важливими є шість складових:

- попит;
- чітко визначені ролі й обов'язки;
- надійність і достовірність інформації;
- підзвітність;
- потенціал;
- стимули/заохочення.

Для того, щоб з плином часу система залишалася життєздатною, кожна з цих складових має отримувати постійну увагу.

Попит

Попит на МіО можна створити і підтримувати у кілька способів:

- Вбудувати формальну структуру, що вимагає регулярного звітування про результати роботи (приміром, щорічних звітів підрозділів організації).
- Повідомити про доступність цієї інформації через ЗМІ, створивши таким чином попит урядових органів, груп громадян, донорів та широкого загалу.
- Створити практику конвертації стратегії в конкретні цілі, щоб ті, хто цікавиться стратегічним напрямом роботи організації, могли відстежувати її рух до цих цілей.

Чітко визначені ролі й обов'язки

Одна з найважливіших структурних інтервенцій для інституціоналізації системи МіО – це створення чітких, формалізованих правил підпорядкування та обов'язків щодо збору й аналізу інформації про ефективність роботи та звітування про неї. Для цього необхідно:

- видати чіткі інструкції стосовно того, хто за які компоненти системи МіО відповідає, включення відповідальності в огляд результатів роботи окремих працівників;
- побудувати систему, що пов'язує функції централізованого планування й фінансів з лінійними/секторними функціями задля встановлення тіснішого зв'язку між циклами розподілу бюджетних коштів та поданням інформації МіО; тобто мова йде про систему бюджетування за результатами;
- побудувати систему, на кожному щаблі якої існуватиме попит на інформацію (тобто в якій немає жодної частини, де інформація залишалася б невикористаною).

Надійність і достовірність інформації

Система інформації про ефективність роботи має бути здатною надавати як позитивні, так і негативні відомості. Відповідно, ті, хто постачає інформацію, мають бути захищені від політичних утисків. Випущена системою інформація має бути прозорою та піддаватися незалежній перевірці (наприклад, перевірці з боку національного аудиторського управління або незалежної групи університетських викладачів).

Підзвітність

Необхідно знайти шляхи розповсюдження прозорої інформації серед третіх сторін, зацікавлених в ефективності роботи. До основних груп зацікавлених осіб відносяться організації громадянського суспільства, ЗМІ, приватний сектор і уряд.

Потенціал

Проведення оцінки готовності та нарощування потенціалу організації – одні з перших заходів, пов'язаних із побудовою системи МіО. Ключові елементи, на яких має ґрунтуватися процес її побудови – належна технічна компетентність в галузі збору й аналізу даних, управлінські навички зі встановлення стратегічних цілей та організаційного розвитку, існуючі системи пошуку й збору даних, наявність фінансових ресурсів та інституційного досвіду моніторингу й оцінки.

Стимули

Для заохочення використання інформації про результативність роботи слід запровадити належні стимули. Успіх потрібно визнавати та винагороджувати, а проблеми – вирішувати. Не можна карати тих, хто приносить неприємні новини, потрібно цінувати використання досвіду в організації, необхідно ділити заощаджені бюджетні кошти. Не варто очікувати, що корумповані або неефективні системи зможуть видавати якісну інформацію й аналіз.

Заключні коментарі

Побудова системи МіО зовсім необов'язково має проходити у відповідності з цими 10 кроками – можна створювати стратегії з меншою чи більшою кількістю кроків. Завдання полягає в тому, щоб визначити ключові функції й види діяльності, логічно їх згрупувати, після чого послідовно реалізувати.

Орієнтовані на результат системи МіО – потужні управлінські інструменти. Вони допомагають будувати й заохочувати зміни в роботі урядів та організацій. Вони також допомагають створювати базу знань про те, які заходи працюють, а які – ні.

Орієнтована на результат система МіО повинна бути ітеративною. Вона потребує постійної уваги, ресурсів і політичної відповідальності. Для переходу до орієнтації на результат потрібен час, але час, зусилля й винагороди того варті.

Нарощування потенціалу є необхідним завжди; за зупинкою в розвитку неминуче йде спад. Для того, щоб забезпечити встановлення й підтримку системи МіО, можна застосувати кілька кроків:

- Цінуйте своїх активних прибічників й допомагайте їм.
- Переконайте міністерство фінансів і законодавчий орган в тому, що система МіО потребує постійного надання ресурсів так само, як бюджетна система (обсяг ресурсів, що надаються системі МіО, повинен бути приблизно рівним тому, що виділяється на бюджетну систему).
- Шукайте будь-які можливості для того, щоб пов'язувати інформацію про результати з рішеннями про бюджет та розподіл ресурсів.
- Розпочніть з пілотних проектів, щоб продемонструвати дієвість МіО з орієнтацією на результат.
- Краще розпочати з анклавної стратегії (наприклад, побудованої навколо островів інновацій), а не із загальноурядового підходу.
- Відстежуйте як прогрес в реалізації, так і досягнення результатів.
- Доповнюйте моніторинг ефективності роботи оцінками для кращого розуміння результатів у державному секторі.

Після створення структури оцінки (крок 7) її можна використати для формування теорії змін, вибору підходу, написання запитань та вибору схеми оцінювання. Ці питання розкриваються в главах 4–7.

Підсумки

Орієнтована на результат система МіО може бути цінним інструментом допомоги особам, що формують політику та приймають рішення, у відстеженні довгострокових результатів і впливів проектів, програм та політик. На відміну від традиційного оцінювання, МіО з орієнтацією на результат переносить основну увагу із залучених ресурсів та короткострокових результатів на довгострокові результати і впливи. Це ключова архітектура для будь-якої системи вимірювання ефективності роботи.

Орієнтовані на результат системи МіО:

- використовують дані про базові значення показників, щоб охарактеризувати стан проблеми до початку ініціативи;
- відстежують показники, пов'язані із запланованими довгостроковими результатами;
- збирають дані про залучені ресурси, заходи й короткострокові результати, а також про їх внесок у досягнення довгострокових результатів;
- оцінюють простоту, надійність і відповідність застосованої теорії змін;
- систематично звітують перед зацікавленими особами;
- працюють із залученням стратегічних партнерів;
- збирають інформацію про успішність або невдачі стратегії партнерства в контексті досягнення потрібних результатів;
- постійно прагнуть надавати достовірну інформацію, що є корисною в якості управлінського інструмента.

Формування й побудову орієнтованої на результат системи МіО рекомендується проводити у десять кроків:

1. Проведення оцінки готовності.
2. Погодження контрольованих довгострокових результатів/наслідків.
3. Вибір ключових показників моніторингу довгострокових результатів/наслідків.
4. Збір базових даних за показниками.
5. Планування поліпшень: вибір реалістичних цілей.
6. Побудова системи моніторингу.
7. Використання інформації оцінювання.
8. Звітування про результати.
9. Використання результатів.
10. Підтримування життєздатності системи МіО в організації.

Побудова й підтримка орієнтованої на результат системи МіО – непроста справа. Вона вимагає постійної уваги, наявності активних прибічників, витрат часу, зусиль і ресурсів. У процесі можна зіткнутися з організаційними, технічними й політичними проблемами. Для того, щоб підлаштувати систему під потреби організації, може виникнути необхідність кілька разів її переглядати. Але результат того вартий.

Завдання до Глави 3

Практична вправа 3.1: Виправте логіку

Наскільки готовою є ваша організація до проектування й запровадження орієнтованої на результат системи МіО? Оцініть свою організацію за кожним з наступних параметрів та поясніть виставлену оцінку. Обговоріть з колегою існуючі перешкоди на шляху запровадження системи, а також способи їх усунення.

1. Стимули (обведіть потрібну оцінку):

багато стимулів декілька стимулів кілька негативних стимулів

Коментарі:

Стратегії вдосконалення:

2. Ролі й обов'язки (обведіть потрібну оцінку):

повністю зрозумілі скоріше зрозумілі зовсім незрозумілі

Коментарі:

Стратегії вдосконалення:

3. Організаційний потенціал (обведіть потрібну оцінку):

відмінний адекватний слабкий

Коментарі:

Стратегії вдосконалення:

4. Перешкоди (обведіть потрібну оцінку):

серйозних перешкод немає дуже мало перешкод серйозні перешкоди

Коментарі:

Стратегії вдосконалення:

Практична вправа 3.2: Визначення ресурсів/внесків, заходів/процесів, короткострокових результатів, довгострокових результатів/наслідків і впливів

Визначити, до якої категорії відноситься кожне з наступних тверджень (внески, заходи, короткострокові результати, довгострокові результати/наслідки або впливи). Якщо можливо, обговоріть з колегою та поясніть, на чому ґрунтується ваше рішення щодо класифікації кожного з тверджень.

1. Мікропідприємства, що належать жінкам, значною мірою сприяють скороченню бідності в тих спільнотах, де вони працюють.
2. Уряд виділив кошти на кредитування мікропідприємств.
3. Уряд задовольнив 61 заявку від осіб, що пройшли навчання в рамках програми.
4. Міністерство освіти призначило викладачів для викладання курсу.
5. Сімдесят дві жінки пройшли навчання.
6. Доходи осіб, що пройшли навчання, протягом першого року зростають на 25 відсотків.
7. Сто жінок відвідали тренінг з управління комерційною діяльністю мікропідприємства.
8. Спільнотам надається інформація про доступність програм кредитування мікропідприємств.

Практична вправа 3.3: Розробка показників

1. Згадайте програму або політику, про яку Ви знаєте. Якого впливу вона намагається досягти? Які два довгострокових результати Ви б очікували від інтервенції, спрямованої на досягнення такого впливу?

Вплив:_____
Довгостроковий результат/наслідок 1:_____
Довгостроковий результат/наслідок 2:_____

2. Починаючи з довгострокових результатів/наслідків визначте від двох до трьох показників, які б Ви стали використовувати для відстеження їх прогресу.

Довгостроковий результат/наслідок 1:_____
Показник а:_____

Показник b:_____
Показник c:_____

Довгостроковий результат/наслідок 2:_____
Показник a:_____
Показник b:_____
Показник c:_____

Вплив:_____
Показник a:_____
Показник b:_____
Показник c:_____

Примітки

1. Ця глава, в якій активно використовуються матеріали К'юзека й Ріста (2004), присвячена моніторингу. Решта книги – виключно оцінці.

Посилання та рекомендована література

Binnendijk, Annette. 2000. "Results-Based Management in the Development Cooperation Agencies: A Review of Experience. Paper prepared for the OECD/DAC Working Party on Aid Evaluation, Paris, February 10–11 (revised October 2000).

Boyle, R., and D. Lemaire, eds. 1999. *Building Effective Evaluation Capacity*. New Brunswick, NJ: Transaction Books.

IFAD (International Fund for Agriculture Development). 2002. *A Guide for Project M&E: Managing for Impact in Rural Development*. Rome. www.ifad.org/evaluation/guide/.

Furubo, Jan-Eric, Ray C. Rist, and Rolf Sandahl, eds. 2002. *International Atlas of Evaluation.* New Brunswick, NJ: Transaction Books.

Hatry, Harry P. 1999. *Performance Measurement: Getting Results*. Washington, DC: Urban Institute Press.

Khan, M. Adil. 2001. *A Guidebook on Results-Based Monitoring and Evaluation: Key Concepts, Issues and Applications*. Government of Sri Lanka, Ministry of Plan Implementation, Monitoring and Progress Review Division, Colombo.

Kusek, Jody Zall, and Ray C. Rist. 2001. "Building a Performance-Based Monitoring and Evaluation System: The Challenges Facing Developing Countries." *Evaluation Journal of Australasia* 1 (2): 14–23.

———. 2003. "Readiness Assessment: Toward Performance Monitoring and Evaluation in the Kyrgyz Republic." *Japanese Journal of Evaluation Studies* 31 (1): 17–31.

———. 2004. *Ten Steps to Building a Results-Based Monitoring and Evaluation System.* Washington, DC: World Bank.

Malik, Khalid, and Christine Roth, eds. 1999. *Evaluation Capacity Development in Asia.* United Nations Development Programme Evaluation Office, New York.

Osborn, David, and Ted Gaebler. 1992. *Reinventing Government*. Boston: Addison-Wesley Publishing.

OECD (Organisation for Economic Cooperation and Development). 2002. *Glossary of Key Terms in Evaluation and Results-Based Management*. Development Cooperation Directorate and Development Assistance Committee. Paris.

Schiavo-Campo, Salvatore. 1999. "'Performance' in the Public Sector." *Asian Journal of Political Science* 7 (2): 75–87.

UNPF (United Nations Population Fund). 2002. *Monitoring and Evaluation Toolkit for Program Managers.* Office of Oversight and Evaluation. www.unfpa.org/monitoring/toolkit.htm.

Valadez, Joseph, and Michael Bamberger. 1994. *Monitoring and Evaluation Social Programs in Developing Countries: A Handbook for Policymakers, Managers, and Researchers.* Washington, DC: World Bank.

Weiss, Carol. 1972. *Evaluation Research: Methods for Assessing Program Effectiveness.* Englewood Cliffs, NJ: Prentice Hall.

Wholey, Joseph S., Harry Hatry, and Kathryn Newcomer. 2001. "Managing for Results: Roles for Evaluators in a New Management Era." *American Journal of Evaluation* 22 (3): 343–47.

World Bank. 1997. *World Development Report: The State in a Changing World.* Washington, DC.

Веб-сайти

IDRC (International Development Research Centre). 2004. *Evaluation Planning in Program Initiatives.* Ottawa. http://web.idrc.ca/uploads/user-S/108549984812 guideline-web.pdf.

IFAD (International Fund for Agricultural Development). *Practical Guide on Monitoring and Evaluation of Rural Development Projects.* http://www.ifad.org/ evaluation/oe/process/guide/index.htm.

Kellogg Foundation. 1998. *Evaluation Handbook.* http://www.wkkf.org/Pubs/Tools/ Evaluation/Pub770.pdf.

Specialist Monitoring and Evaluation Web Sites. http://www.mande.co.uk/specialist.htm.

Uganda Communications Commission. 2005. "Monitoring and Evaluation." In *Funding and Implementing Universal Access: Innovation and Experience from Uganda.* International Development Research Centre, Ottawa. http://www.idrc.ca/en/ev-88227-201-1-DO_TOPIC.html.

World Bank. *Core Welfare Indicators Questionnaire.* Washington, DC. http://www4.worldbank.org/afr/stats/cwiq.cfm.

———.2001. *Tools: Capturing Experience Monitoring and Evaluation.* Upgrading Urban Communities Resource Framework, Washington, DC. http://web.mit.edu/urbanupgrading/upgrading/issues-tools/tools/monitoring-eval.html# Anchor-Monitoring-56567.

———. 2008. *Online Atlas of the Millennium Development Goals: Building a Better World*. Washington, DC. http://devdata.worldbank.org/atlas-mdg/.

ГЛАВА 4

Розуміння контексту оцінювання й програмної теорії змін

Ця глава – перша з двох глав, що розглядають планування оцінювання. В ній розповідається про те, з чого треба починати. Оцінювання, що починається з якісного планування, має більші шанси бути виконаною вчасно і в межах виділеного бюджету та задовольнити потреби клієнта й інших зацікавлених осіб. Аналіз вихідних даних розглядає попередній досвід і робить з нього висновки, підтверджує або ставить під сумнів теорію, на якій ґрунтується програма та задає контекст, що впливає на програму.

Ця глава складається з п’яти основних частин:

- Аналіз вихідних даних
- Визначення головного клієнта та основних зацікавлених сторін
- Розуміння контексту
- Використання існуючих знань
- Побудова, використання та оцінка теорії змін

Аналіз вихідних даних

З чого почати? Якщо ви прагнете досягти певної цілі, потрібно спочатку визначитися з напрямом руху, а також розглянути досвід тих, хто вже проходив цим шляхом. Вам варто зібрати критичну інформацію для прийняття рішень про часові рамки, витрати, загрози та процеси.

Аналіз вихідних даних (Front-End Analysis): Детальний розгляд питання або проблеми для того, щоб визначити, що про неї відомо та як рухатися далі в розробці оціночного підходу до неї

Аналіз вихідних даних – це детальний розгляд питання або проблеми для того, щоб визначити, що про цю проблему чи питання відомо та як рухатися далі в розробці оціночного підходу до їх вирішення. Це те, що робить фахівець з оцінювання, щоб прийняти рішення про свої наступні дії.

Аналізуючи вихідні дані, фахівець з оцінювання розглядає наступні питання:

- Хто є головним клієнтом оцінювання? Хто є іншими важливими зацікавленими особами? Які питання вони висувають для оцінювання?
- Як вплине на оцінювання вибір часу його проведення в контексті реалізації проекту, програми або політики?
- Скільки часу ви маєте для проведення оцінювання?
- Які є в наявності ресурси та скільки їх?
- Чи є суспільствознавча теорія актуальною для цього оцінювання?
- Що показали оцінювання аналогічних програм? Які питання вони поставили?
- На якій теорії змін ґрунтується проект, програма або політика?
- Які існуючі дані можна використати для цього оцінювання?

Багатьом фахівцям з оцінювання не терпиться закінчити планування оцінювання і якомога скоріше перейти до збору даних. Вони намагаються проводити розвідувальну роботу одночасно зі збором даних. Але якісний аналіз вихідних даних є критичним для отримання відомостей про інтервенцію. Він може зберегти час і гроші, що витрачаються на оцінювання, забезпечити задоволення потреб клієнта, підтримати або створити стосунки не тільки з клієнтом, а й з основними зацікавленими особами. Найважливіше те, що, провівши якісний аналіз вихідних даних, можна гарантувати, що оцінка буде розглядати правильні питання і збирати потрібну інформацію замість того, щоб накопичувати дані, які ніколи не будуть використані.

На початку оцінки деякі фахівці з оцінювання роблять припущення, і частина з цих припущень можуть бути помилковими. Наприклад, фахівець з оцінювання може припустити, що існують великі масиви даних, в той час як насправді в наявності є дуже обмежені обсяги інформації. Він може припустити, що в проведенні оцінки йому допоможуть досвідчені консультанти, які добре знають країну, але згодом виявляється, що люди, на яких він роз-

раховував, зайняті в інших проектах. Період розвідки необхідний для того, щоб більше дізнатися про наявність даних та інших ресурсів.

Визначати, чи можливим та корисним є спільне оцінювання, також варто на початковому етапі. Якщо існує зацікавленість і такий вид оцінки є доцільним, партнерам потрібно визначитися з тим, хто які ролі візьме на себе. Їм також необхідно узгодити такі питання, наприклад, як час проведення оцінювання.

Баланс очікуваних витрат та вигод від оцінювання

Очікувані витрати та вигоди оцінювання, а також способи їх збалансування повинні розглядатися в ході попереднього планування. Серед вигод оцінювання можуть бути такі:

- прийняття рішень на підставі фактів, що дозволяє раціонально вирішувати, що потрібно розширяти або відтворювати, що – покращувати, припиняти або згортати;
- розширення знань про те, що працює (і в яких умовах), а що – ні;
- формування місцевого потенціалу.

Витрати на проведення оцінювання також є важливими. Їх слід розглядати з точки зору:

- вартості програми (немає сенсу витрачати 50 000 доларів на оцінювання програми з бюджетом у 25 000 доларів);
- тягаря, що його покладає оцінювання на бенефіціарів програми та інших осіб, які витрачають час на надання інформаційної або іншої допомоги фахівцям з оцінювання;
- репутаційних збитків для фахівця з оцінювання та спільноти фахівців з оцінювання у випадку, якщо існує висока ймовірність оспорювання результатів через те, що оцінюється дуже політизована, скандальна програма, або через те, що на проведення комплексної оцінки виділяється недостатньо часу.

Труднощі процесу попереднього планування

Впевненість у тому, що все можна передбачити, а також що за наявності попереднього планування оцінювання обов'язково пройде без проблем – це потенційна пастка процесу попереднього планування. До інших таких пасток можна віднести зокрема:

- небажання змінювати початковий план (Leeuw 2003);
- «Макдональдизація» суспільства – «процес, за якого принципи роботи ресторану швидкого харчування охоплюють все більше й більше секторів

як американського суспільства, так і решти світу» (Ritzer 1993, 1) (це особливо актуально тоді, коли рефлективне мислення заміщується контрольними списками, списками запланованих справ і схемами);
- стереотипи («трюїзми»), що раптово виявляються при проведенні попереднього планування («Рандомізовані експерименти? Нізащо: занадто складно, занадто дорого і занадто важко проводити в контексті розвитку»);
- «групове мислення» (погодження з поглядами групи заради того, щоб залишатися її частиною незважаючи на сумніви щодо цих поглядів);
- надання непропорційно великої ваги поглядам людей при владі (автоматичне визначення значущості пропозицій відповідно до статусу осіб, які їх висувають).

Визначення головного клієнта та основних зацікавлених сторін

Важливою складовою попереднього аналізу є визначення головного клієнта та основних зацікавлених осіб для проекту, програми або політики. Визначити основних зацікавлених осіб часто буває непросто.

Головний клієнт

Зазвичай одна ключова зацікавлена особа або група зацікавлених осіб фінансує або замовляє проведення оцінювання і є основним користувачем її результатів. Ця зацікавлена особа або група зацікавлених осіб – головний клієнт оцінювання. Потреби такого клієнта мають великий вплив на оцінювання.

Головний клієнт може:

- санкціонувати й фінансувати програму;
- санкціонувати й фінансувати оцінювання;
- відповідати перед суспільством за інтервенцію;
- бути стороною, перед якою відповідають фахівці з оцінювання.

Головний клієнт – один. Важливо зустрітись з головним клієнтом (або його представниками – наприклад, з радою директорів, якщо мова йде про клієнтську групу) на ранніх етапах роботи, щоб спільно визначити завдання оцінювання з точки зору клієнта. На зустрічі фахівці з оцінювання повинні запитати про часові обмеження та цільове призначення (очікуване використання) оцінювання. Фахівець з оцінювання, який спочатку прислухається до клієнта, розпитує його для визначення причин замовлення оцінювання, може згодом повернутися і надати клієнту більший вибір варіантів проведення оцінювання або ж запропонувати один рекомендований підхід до нього.

Зацікавлені сторони

Зацікавлені сторони/особи – це люди або організації, окрім клієнта, які зацікавлені в інтервенції. Зазвичай це ті, на кого інтервенція здійснює вплив – протягом її реалізації або в наступні роки. До зацікавлених осіб важливо включати тих, кого зазвичай не залучають до участі в оцінюванні.

Зацікавлена сторона/особа: Окрема особа або організація, окрім клієнта, що має зацікавленість в інтервенції

До числа зацікавлених осіб можуть входити:

- учасники;
- прямі бенефіціари;
- непрямі бенефіціари;
- організації з розвитку, що надають фінансування;
- урядовці, виборні посадові особи, державні працівники з відповідними інтересами – наприклад, планувальники та дільничні медсестри;
- директори програм, їх персонал, члени ради директорів, керівники й волонтери;
- особи, що формують політику;
- громада, групи й асоціації за інтересами, в тому числі ті, цілі яких відрізняються від цілей посадовців програми.

Зацікавлені особи часто мають різні точки зору стосовно інтервенції. Донора цікавить належне використання коштів та дієвість інтервенції. Менеджера програми може цікавити належне керування інтервенцією та набуття досвіду. Учасники програми можуть прагнути отримувати послуги в більшому обсязі або кращої якості. Осіб, що формують політику, більш за все має хвилювати те, чи має інтервенція бажаний вплив. Інші учасники спільноти можуть бути зацікавлені у тому, щоб відтворити чи розширити інтервенцію або ж обмежити те, що вони вважають її негативними наслідками. Різноманітність інтересів – це позитивний момент, і початкові обговорення можуть це продемонструвати.

Ролі кожної окремої особи або групи в контексті оцінювання та її потенційного застосування мають бути чітко визначені (див. табл. 4.1).

Визначення та залучення основних зацікавлених сторін

Основних зацікавлених осіб можна визначити, вивчивши документи про інтервенцію та провівши бесіду з головним клієнтом оцінки, спонсорами програми, її персоналом, місцевими посадовцями та учасниками. Із зацікавленими особами можна провести інтерв'ю в індивідуальному порядку або у форматі малих груп. Звертаючись до зацікавлених осіб з приводу оцінювання, планувальник оцінювання повинен чітко визначити мету бесіди з кожним окремо (попередити про майбутнє оцінювання, попросити визначити коло питань, які б мало охопити оцінювання).

Активізація застосування оцінювання – це процес, що починається із зустрічі з головним клієнтом та залучення основних зацікавлених осіб до проек-

тування оцінювання. Якщо звіт з оцінювання завершено і вже має розсилатися, займатися активізацією застосування оцінювання запізно. У певних випадках зустрічі з основними зацікавленими особами проводяться регулярно. Може навіть встановлюватися більш формальна структура взаємодії. Керівник оцінювання може організувати консультативний або керівний комітет.

Таблиця 4.1. Ролі зацікавлених осіб в оцінюванні

Зацікавлена сторона	Формування політики	Прийняття операційних рішень	Надання ресурсів для проведення оцінювання	Реагування	Просто інтерес
Розробники програми					
Спонсори (донори) програми					
Посадовець, рада або агенція, що санкціонує оцінювання					
Постачальники інших ресурсів (приміщень, матеріалів, внесків натурою)					
Голови агенцій, які реалізують оцінювання, та вищі керівники					
Керівники програми, які відповідають за реалізацію					
Персонал програми					
Персонал з моніторингу					
Прямі бенефіціари програми					
Непрямі бенефіціари програми					
Особи, які потенційно також можуть застосовувати програму					
Особи, виключені з програми (наприклад, за критеріями включення)					
Особи, які передбачають негативні наслідки програми або оцінювання					
Особи, які втрачають владу (вплив) в результаті програми					
Особи, які втрачають можливості в результаті програми					
Члени спільноти або широкий загал					
Інші					

Джерело: Автори.

Залучення основних зацікавлених сторін вже на ранніх стадіях забезпечує краще розуміння фахівцем з оцінювання інтервенції, її мети, а також проблем і викликів, пов'язаних з досягненням цієї мети. Команда оцінювання буде краще поінформованою про питання, що мають бути охоплені оцінкою, а також про конкретну інформацію – коли вона має бути надана, а також хто має її надати. Зустріч з основними зацікавленими сторонами допомагає врахувати всі основні критичні питання під час оцінювання. Вона також допомагає сформувати сприйняття оцінки: якщо зацікавлені особи розуміють, що підняті ними питання й теми будуть уважно розглянуті, вони більше схильні підтримувати оцінку та виявляти зацікавленість у ній. Так формується застосування оцінювання.

Ступінь активності зацікавлених осіб в проектуванні та реалізації оцінювання залежить від декількох факторів, у тому числі від ресурсів та стосунків. Приміром, у зацікавлених осіб може не бути вільного часу, і участь вимагатиме від них жертвувати своїми звичайними обов'язками, або ж у них можуть бути політичні причини бажати максимальної незалежності оцінювання.

Проведення аналізу зацікавлених сторін

Існує багато настанов з проведення аналізу зацікавлених осіб. На своєму веб-сайті A Guide to Managing for Quality («Керівництво з управління, орієнтованого на якість») організація «Управлінські науки для охорони здоров'я» та Дитячий фонд ООН (Management Sciences for Health and the United Nations Children's Fund (1998)) описали процес визначення та оцінювання важливості ключових осіб, груп осіб та установ, які можуть значною мірою впливати на успіх оцінювання. Вони також пропонують кілька інших причин для проведення аналізу зацікавлених осіб:

- для визначення осіб, груп та установ, які можуть вплинути на оцінку (позитивно або негативно);
- щоб передбачити, як саме – позитивно або негативно – впливатимуть ці групи на оцінювання;
- для розробки стратегій з отримання максимально дієвої підтримки ініціативи та усунення перешкод на шляху успішної реалізації оцінювання.

У блоці 4.1 дається шаблон проведення аналізу зацікавлених осіб. Хоча він схожий на табл. 4.1, в ньому робиться акцент на формування підтримки оцінювання та зменшення протидії їй.

Важливо, щоб аналіз охоплював широке коло зацікавлених осіб, але при цьому – не занадто широке. Прагнення до залучення осіб з периферії інтересів може викликати лише їх подразнення.

Іноді фахівці з оцінювання прямо залучають одну або декілька зацікавлених осіб до планування та проведення оцінювання (спільне оцінювання розглядається в главі 5). В таких випадках фахівець з оцінювання сприяє залученню зацікавлених осіб до:

- формулювання технічного (проектного) завдання;
- формування команди оцінювання;
- аналізу даних;
- визначення результатів та формулювання висновків і рекомендацій (Mikkelsen 2005).

Блок 4.1 «Як проводити аналіз зацікавлених сторін»

1. Проведіть з колегами мозковий штурм, щоб визначити, які особи, групи та установи впливатимуть на інтервенцію або можуть потрапити в сферу її впливу. Внесіть їх до стовпця «Зацікавлена особа».
2. Визначте конкретні інтереси кожної зацікавленої особи в оцінюванні. Враховуйте такі питання, як потенційні вигоди оцінки для зацікавленої особи, зміни, що їх зацікавленій особі доведеться вносити у зв'язку з оцінкою, а також заходи в рамках проекту, що можуть завдати збитків зацікавленій особі або спричинити конфлікт з нею. Внесіть ці інтереси до стовпця «Інтереси зацікавленої особи у проекті, програмі або політиці».
3. Визначте, наскільки важливими є інтереси зацікавленої особи для успіху оцінювання. Враховуйте як (а) роль, яку основна зацікавлена особа має відіграти для успішності оцінки, та імовірність того, що зацікавлена особа цю роль відіграє, так і (б) ймовірність та вплив негативної реакції зацікавленої особи на оцінку. Свою оцінку кожної зацікавленої особи внесіть у стовпець «Оцінка потенційного впливу», використовуючи наступну шкалу: «А» – надзвичайно важливий, «В» – доволі важливий, «С» – не надто важливий.

Подумайте, які заходи допоможуть вам отримати підтримку зацікавленої особи та зменшити її протидію. Подумайте, які підходи існують до кожної із зацікавлених осіб. Які питання зацікавлена особа може побажати зробити об'єктом оцінки? Наскільки важливим є залучення конкретної зацікавленої особи в процесі планування? Чи є інші групи або особи, що можуть вплинути на підтримку зацікавленою особою вашої оцінки? Внесіть свої стратегії отримання підтримки або усунення перешкод для своєї оцінки в останній стовпець таблиці.

Таблиця блоку. Приклад формату для проведення аналізу зацікавлених сторін			
Зацікавлена сторона	Інтереси зацікавленої сторони в проекті, програмі або політиці	Оцінка потенційного впливу оцінки на зацікавлену сторону та зацікавленої сторони на оцінювання	Потенційні стратегії отримання підтримки та усунення перешкод

Джерело: Management Sciences for Health and the United Nations Children's Fund 1998.

Розуміння контексту

Аналіз вихідних даних також розглядає відносини між етапами програми та загальною метою оцінки. Життя проекту, програми або політики можна розглядати як поступовий процес, на різних етапах якого ставляться різні питання оцінки. Наприклад, немає сенсу з'ясовувати, чи було досягнуто цілей програми всього через кілька місяців після виділення коштів. Більш адекватним питанням на такому ранньому етапі буде те, чи отримала програма ресурси, необхідні для її реалізації. Пенсер та Вестг'юз представляють типологію таких програмних етапів та питань оцінки (табл. 4.2). Наведені питання є лише прикладами; потенційно існує багато можливих питань для кожного етапу.

Таблиця 4.2. Питання, що ставляться на різних етапах оцінювання

Етап розвитку програми	Питання оцінювання
1. Оцінка соціальних проблем і потреб	Якою мірою задовольняються потреби громади?
2. Визначення цілей та мети	Що потрібно зробити, щоб задовольнити ці потреби?
3. Розробка програмних альтернатив	Які послуги можна використати, щоб отримати бажані зміни?
4. Вибір варіанту	Який з можливих програмних підходів є найбільш надійним?
5. Реалізація програми	Як запустити програму?
6. Робота програми	Чи працює програма так, як було заплановано?
7. Довгострокові результати/наслідки/вплив програми	Чи дає програма бажані результати?
8. Ефективність програми	Чи розумною ціною досягаються результати програми?

Джерело: Взято з Pancer and Westhues 1989.

Наступний крок початкового аналізу – визначення політичного контексту. Дослідження може виявити оцінки аналогічних програм. Спочатку фахівець з оцінювання отримує звіти про оцінку та переглядає їх, з'ясовуючи, які проблеми було охоплено, який підхід та інструменти використовувалися, а також вивчає результати. Якщо оцінюється нова інтервенція, може статися так, що оцінку доведеться проводити, не покладаючись на досвід попередніх оцінок. Однак таке трапляється не часто.

Використання існуючих знань

Аналіз вихідних даних розглядає існуючі теоретичні та емпіричні знання про проект, програму або політику. Це також називають використанням **фонду знань**.

Фонд знань:
Сукупність існуючих теоретичних і емпіричних знань про проект, програму або політику

Блок 4.2 «Використання фонду знань щодо профілактики злочинності»

Аналіз даних з профілактики злочинності свідчить про те, що 29 програм спрацювали, 25 – не спрацювали, а 28 були перспективними (інформація ще про 68 програм є неоднозначною). Ці висновки базуються на синтезі більш ніж 600 оцінювань, включаючи оцінювання профілактики злочинності через вплив на школу та сім'ю, програми зі скорочення крадіжок з проникненням у приміщення, арештів, пов'язаних з наркотиками, програми роботи поліції в небезпечних районах, ініціативи замкненого циклу, районних охоронців, програми наставництва, а також різні види тюремних санкцій і програм (управління гнівом, навчальні програми, пізнавальні програми для скорочення рецидивізму, виправні заклади для неповнолітніх і т. д.).

Джерело: Sherman and others 2002.

Обсяг знань, що надходять від оцінювань та інших суспільствознавчих досліджень, включаючи економічні дослідження, зростає з кожним днем. В журналах публікуються статті, що синтезують пояснювальні знання з певної теми, такі як вплив розміру шкільного класу на навчання або програм харчування для вагітних на вагу новонароджених. Проблемно-орієнтовані дослідження функціонування організації поєднують теорії і досягнення таких різних дисциплін, як організаційна соціологія, когнітивна психологія, теорія суспільного вибору, право і економіка (Scott 2003; Swedberg 2003). Такі організації, як Campbell Collaboration (http://www.campbellcollaboration.org/) перевіряють якість оцінювань в певній сфері та синтезують результати

тих з них, що відповідають встановленим критеріям. Репозиторії рандомізованих експериментів у сфері кримінальної юстиції та профілактики злочинності, програм у сферах соціального забезпечення, освіти та охорони здоров'я містять дані про більш ніж 10 000 «експериментів» (Petrosino and others 2003). Під час організації та планування оцінювань надзвичайно важливо виявити та проаналізувати фонд відповідних знань (блок 4.2).

Побудова, використання та оцінка теорії змін

Остання частина аналізу вихідних даних – це створення теорії змін та розуміння того, як її використовувати та оцінювати. Внутрішня логіка, або теорія змін, – важлива тема для оцінювання, незалежно від того, проводиться воно перед дослідженням чи після нього. В цьому розділі ми з'ясовуємо, з якою метою застосовується теорія змін та як її оцінювати.

Хоча існують різні визначення (див. главу 3), одне з них стверджує, що **теорія змін** – це «інноваційний інструмент для проектування та оцінювання ініціатив соціальних змін», свого роду «план структурних елементів», потрібних для досягнення довгострокових цілей ініціативи соціальних змін (ActKnowledge and Aspen Institute 2003). Теорію змін також можна розглядати як відображення того, яким чином організація або ініціатива має досягти результатів, а також визначення зроблених базових припущень.

Теорія змін:
План структурних елементів, потрібних для досягнення довгострокових цілей ініціативи соціальних змін

Теорія змін повинна:

- відображати послідовність залучення ресурсів для реалізації проекту, програми або політики; заходи, для підтримки яких будуть використовуватися залучені ресурси; короткострокові результати, на які спрямовуються кошти проекту, програми або політики (один захід або набір заходів); а також очікувані довгострокові результати й впливи;
- визначити події або умови, що можуть негативно вплинути на досягнення довгострокових результатів;
- визначити, які припущення робляться в рамках програми стосовно причин та наслідків;
- визначити критичні припущення, які зумовлено середовищем та політичним контекстом, а також демонструється аналізом літератури, потребують розгляду в рамках оцінки.

Визначення подій та умов, які можуть негативно вплинути на досягнення інтервенцією бажаних довгострокових результатів – захід, особливо необхідний в умовах взаємозалежності та складності програм розвитку. Міжнародні інститути розвитку кредитують програми, і це дає країнам, що розвиваються, більшу свободу дій, ніж проектне фінансування. Як зазначає Поусон (Pawson 2006):

важливою зміною суспільної політики за останні роки стало зростання складних програм, що мають одразу кілька цілей, охоплюють більш ніж одну територію, проводяться більш ніж однією агенцією, присвячені більш ніж одній темі. ... Причини цього очевидні. Корені соціальних проблем переплетені. Проблеми зі здоров'ям можуть бути спричинені низьким рівнем освіти, нерівністю на ринку праці і станом довкілля, погіршенням житлових умов, відмінними патернами віктимізації жертв злочинів і т. д. Відповідно, особи, які приймають рішення, повинні замислитись, чи не є інтервенції з проведенням лише одного заходу для усунення однієї проблеми спробою лікувати симптоми.

Поусон вважає, що в таких випадках фахівці з оцінювання повинні:

- розуміти внутрішню теорію програми;
- вивчати існуючі дані через синтез досліджень;
- розглядати комплексну програму як ланцюжки інтервенцій, коли одна частина зацікавлених осіб надає ресурси (матеріальні, соціальні, пізнавальні та емоційні) решті зацікавлених осіб, очікуючи, що це викличе зміни в поведінці.

Під час планування оцінювання необхідно приділяти увагу питанню формування й випробування теорії змін, що лежить в основі програми. Для кращого розгляду ключових складових та взаємодій проекту, програми або політики слід використовувати візуальну подачу даних. Такі дані повинні демонструвати причинно-наслідкові зв'язки у проектах, програмах або політиках. Бажаний вплив або ціль зазвичай показують як останню ланку в моделі.

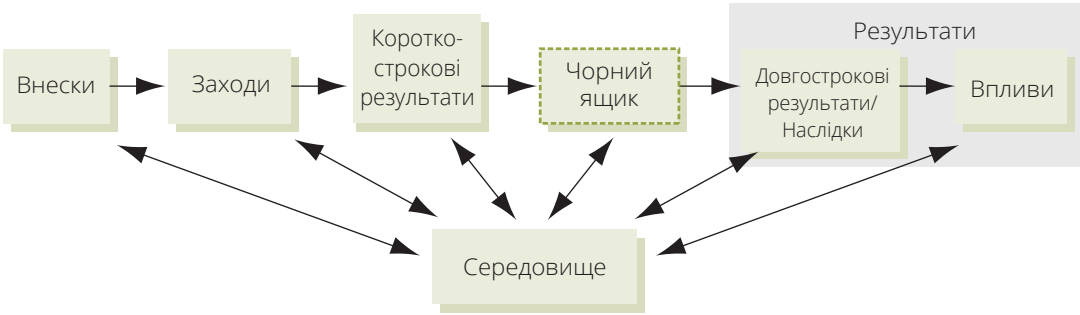
Цінність теорії змін полягає в тому, що вона візуально окреслює коло питань: чому проект, програма або політика має досягти результатів. Теорія змін також визначає складові програми та їх зв'язок. Для виконання організацією заходів, що мають на меті досягнення конкретних цілей, надаються ресурси. Ці ресурси, заходи, короткострокові та цільові довгострокові результати й впливи є взаємопов'язаними.

В певних випадках фахівці з оцінювання можуть виявити, що в інтервенції вже є теорія змін. Якщо це так, цю теорію потрібно уважно проаналізувати. У багатьох випадках доведеться доробити або переробити існуючу теорію змін та узгодити її з безпосередніми учасниками. Якщо теорії змін немає, фахівцю з оцінювання слід створити її та, за можливості, узгодити з керівником і персоналом програми.

У рамках теорії змін потрібно також визначити припущення. Також необхідно виділити найбільш критичні з цих припущень, що мають бути перевірені оцінкою (вони базуються на домінуючому політичному довкіл-

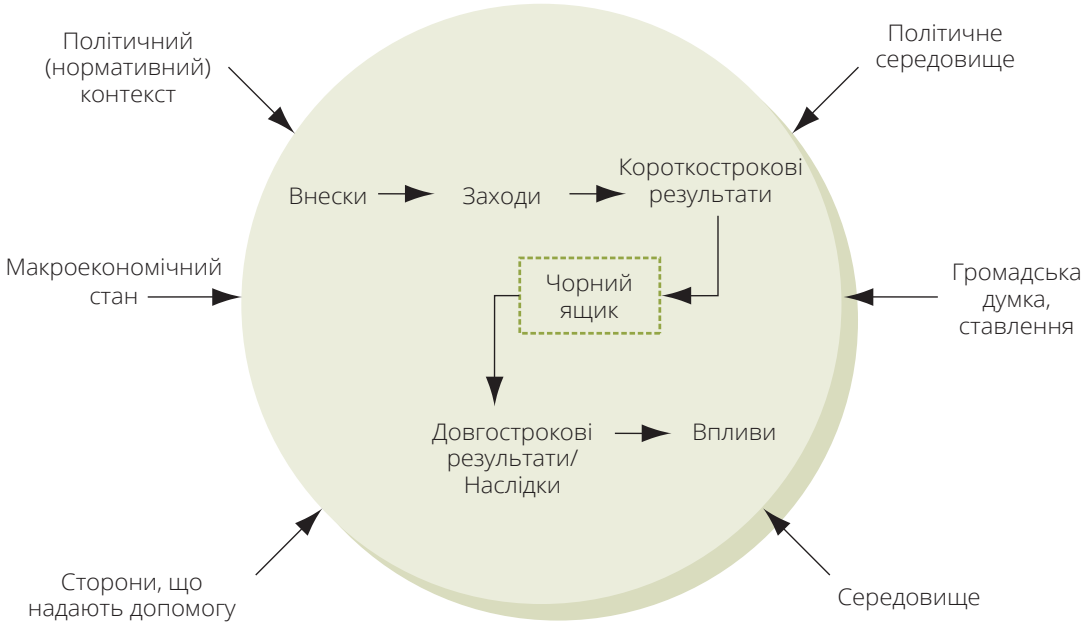
лі, а також вивченні літератури). Теорії змін відкривають «чорний ящик», демонструючи, яким чином інтервенція має перетворити залучені ресурси, заходи і короткострокові результати на довгострокові результати і впливи (рис. 4.1).

Рис. 4.1. Перехід від внесків до результатів



Джерело: Автори.

Рис. 4.2. Потенційні впливи середовища на результати програми



Джерело: Автори.

Важливо визначити, що відбувається в ширшому контексті – тобто в середовищі, в якому працює програма. Це середовище (політичне, макроекономічне і т. д.) впливає на всі частини системи (рис. 4.2).

За кожним проектом, за кожною програмою та політикою стоїть своя теорія змін, яку лише треба сформулювати. Теорію можна представити візуально – різними шляхами, використовуючи різні формати та моделі. Ці представлення іноді називають *моделями теорії*, *логічними моделями*, *структурами змін*, *логічними структурами*, *ланцюговими моделями результатів* та *моделями довгострокових результатів*. Все це – засоби представлення теорії змін. Теорія змін повинна представляти причинно-наслідковий ланцюжок, показувати фактори впливу та визначати ключові припущення.

Навіщо застосовувати теорію змін?

Теорія змін є цінною як для фахівців з оцінювання, так і для зацікавлених сторін, адже вона дозволяє їм разом будувати «однакове уявлення про бачення довгострокових цілей, способів їх досягнення та засобів контролю прогресу» (ActKnowledge and Aspen Institute 2003).

Теорію змін також можна використовувати для звітування про результати оцінювання. Звіт Фонду Келога (Kellogg Foundation 2004) аналізує важливість комунікацій у звітуванні про успіхи й сталість програми. Він визначає три основних шляхи, як представлення теорії змін може підтримувати стратегічну маркетингову роботу:

- опис програм мовою, яка є достатньо зрозумілою і конкретною для розуміння й оцінювання;
- концентрація уваги й ресурсів на пріоритетних програмних операціях та ключових результатах для використання досвіду і вдосконалення програми;
- розроблення цільових комунікаційних та маркетингових стратегій.

Отже, формулювання теорії змін для проекту, програми або політики дає кілька переваг:

- допомагає визначити, які елементи програм є критично важливими для успіху;
- допомагає сформувати спільне розуміння програми й очікувань серед зацікавлених осіб;
- закладає підґрунтя для оцінювання;
- встановлює засоби визначення прогресу в змінних інтервенції, від яких залежать довгострокові результати;
- дозволяє виокремити використовувані припущення, які можна застосувати для формулювання питань оцінювання.

Формування теорії змін

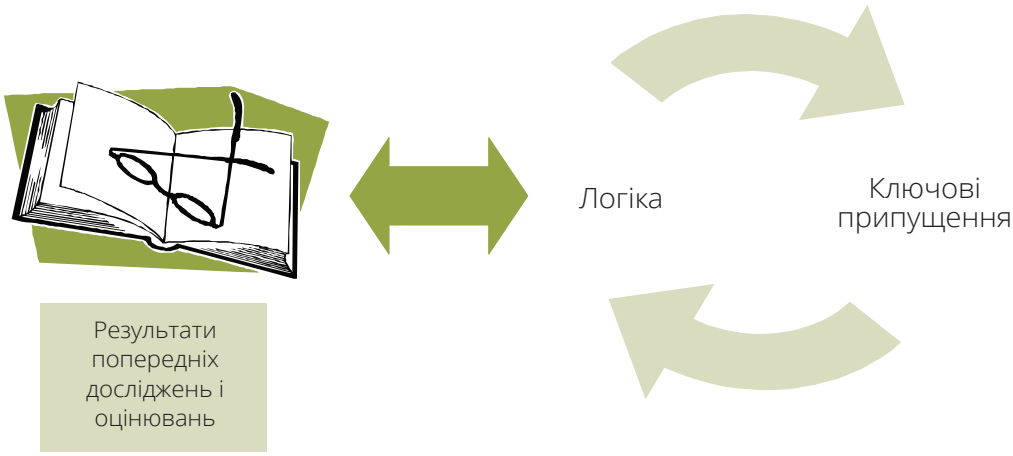
Менеджери звичайно створюють теорію змін, складаючи концепцію проекту, програми або політики. До цього процесу можна залучати зацікавлених осіб. Однак теорія змін не завжди викладається прямо і не завжди або не обов'язково є незмінною з початку до кінця конкретної інтервенції. Певні інтервенції не мають сформульованої теорії змін. В такому випадку її має сформулювати фахівець з оцінювання. Вивчення теорії змін має лягати в основу кожного оцінювання.

Перш ніж аналізувати або формулювати теорію змін, фахівці з оцінювання повинні отримати чітке розуміння призначення й цілей проекту, програми, політики. Варто розглядати три основних питання:

- Чи ґрунтується проект, програма, політика на дослідженнях та оцінюванні?
- Якою є логіка або ланцюжок результатів проекту, програми, політики?
- Які робляться ключові припущення?

Спочатку треба дізнатися якомога більше про пов'язані інтервенції та оцінювання. Після отримання нової інформації розпочинається побудова логіки програми та ключових припущень. Сформована логіка включається в ланцюжок подій і зображується схематично. Після цього визначаються ключові припущення (рис. 4.3).

Рис. 4.3. Процес формування теорії змін



Джерело: Автори.

Чи ґрунтується програма на попередніх дослідженнях і оцінюваннях?

Розробка теорії змін починається з пошуку та аналізу відповідної літератури з досліджень і оцінювань. Наприклад, попереднє дослідження може демонструвати, що за інших рівних умов існує позитивна залежність між академічною успішністю учня та участю батьків у виконанні дитиною домашньої роботи. Або оцінювання попередньої освітньої програми в міських частинах країни, що розвивається, показало помірно успішні результати для компоненту програми, що передбачав відвідування вчителями домівок учнів. Після завершення пошуку літератури та визначення актуальних результатів можна формулювати теорію змін, щоб передбачити, чого варто очікувати від аналогічної інтервенції. Бувають ситуації, коли доводиться формувати теорії змін, не користуючись результатами попередніх досліджень або оцінювань. В такому випадку теорії неминуче будуть слабшими.

Аналіз літератури з оцінювання слід починати з широкого визначення можливих джерел інформації про дослідження та оцінювання. До таких джерел належать:

Блок 4.3 «Огляд літератури щодо програм з поліпшення успішності учнів в південноафриканській країні»

Фахівців з оцінювання організації з розвитку попросили оцінити освітню програму в Африці, метою якої було покращення успішності. Фахівцям з оцінювання повідомили, що відвідування учителями домівок учнів початкових шкіл в сусідній країні дало підвищення успішності. Фахівці з оцінювання почали свою роботу з вивчення літератури про дослідження й оцінювання. Спочатку вони вивчали дослідження освіти й успішності на рівні початкової школи, шукаючи результати аналогічних інтервенцій. Вони зосередилися на матеріалах журналів International Journal of Educational Research, American Education Research Journal, Early Childhood Research and Practice та Journal of Educational Research. Також переглядалися матеріали з оцінювання таких організацій, як Development Gateway, Центр ресурсів оцінки КСР, Інформаційний центр освітніх ресурсів. При цьому фахівці з оцінювання виявили актуальні дослідження та кілька аналогічних проектів і програм. Група оцінювання ознайомилася з дослідницькою та оціночною літературою, щоб дізнатися якомога більше про проблеми й результати аналогічних програм в інших країнах.

Джерело: Автори.

- оціночні дослідження, проведені власне організацією;
- репозиторій публікацій і документів Організації економічного співробітництва і розвитку/Комітету сприяння розвитку (ОЕСР/КСР) та інформації за країнами;
- оцінювальні дослідження, проведені організаціями з розвитку, банками розвитку, неурядовими організаціями (НУО) тощо;
- статті з журналів, присвячених оцінюванню;
- прикладні дослідження певних тем, що публікуються в журналах з психології, соціології, освіти; дослідження теорії розвитку.

Резюме і висновки – ось куди треба дивитись, щоб визначити релевантність звітів оцінювання. Що стосується наукових статей, то можна швидко переглянути анотації до них. Якщо часу обмаль, фахівцям з оцінювання доводиться поспіхом переглядати матеріали в пошуках важливої інформації. Знайшовши результати актуальних досліджень і оцінювань, вони мають уважно їх вивчати (блок 4.3).

Якою є логіка програми?

Логіка програми, політики або проекту орієнтується на завдання та цілі інтервенції. Вона стверджує: «Якщо станеться *X*, то має статися *Y*». Вказуються відомості про природу потрібної зміни – цільову аудиторію, рівень змін, що вважатиметься ознакою успішності, термін, протягом якого планується досягти такого рівня змін. Після цього формується ланцюжок тверджень: «якщо… то…». Для запису тверджень в ланцюжку заходів можна використовувати невеличкі папірці або картки. Їх легко пересувати, додавати або видаляти з ланцюжка.

Фахівці з оцінювання зазвичай вибудовують ланцюжок логіки в зворотному порядку – від довгострокової цілі, вибудовуючи низку подій, що поєднують таку ціль з поточним станом. Якщо теорію змін вже сформовано, фахівець з оцінювання має пройти через аналогічний процес для відтворення цієї теорії змін.

Розглянемо, приміром, інтервенцію з навчання людей методикам оцінювання розвитку (рис. 4.4). Очікувані результати цієї інтервенції – проведення оцінок вищої якості та прийняття краще обґрунтованих рішень відповідальними особами. (Кінцевою метою було б отримання покращень в розвитку, але через те, що на розвиток впливає так багато факторів, фахівці з оцінювання не намагаються вимірювати цей кінцевий довгостроковий результат.) Простий ланцюжок для цієї інтервенції може включати наступні ланки: краще підготовлені фахівці краще проведуть оцінювання, що забезпечить осіб, які формують політику, корисною інформацією, а це, в свою чергу, дозво-

лить приймати обґрунтовані рішення. Корисна інформація має забезпечити прийняття відповідальними особами кращих рішень.

Це була лінійна модель. Якісна теорія змін не передбачає простих лінійних причинно-наслідкових зв'язків; вона показує складні відносини, використовуючи блоки й стрілки, що вказують на зв'язок з попередніми або наступними складовими теорії змін. Вона також детально представляє основні припущення, на яких ґрунтується модель, в тому числі головні контекстуальні чи середовищні чинники або події, що з певною ймовірністю впливатимуть на інтервенцію.

Які ключові припущення?

Початкові логічні схеми часто виявляються лінійними. Коли фахівці з оцінювання розглядають всі ті численні фактори, що взаємодіють з їхніми проектами, програмами та політиками, теорія змін ускладнюється. Визначаючи логіку програми, необхідно також визначити, які припущення вони роблять стосовно процесу змін. Після цього можуть бути вивчені та перевірені в рамках оцінки припущення, що становлять найбільший ризик для успіху інтервенції (ключові припущення).

Рис. 4.4. Проста теорія змін для покращення прийняття рішень підготовленими фахівцями з оцінювання

Джерело: Автори.

Припущення зазвичай можна віднести до однієї з чотирьох груп:

- твердження щодо зв'язків між довгостроковими, проміжними та ранніми результатами на мапі;
- обґрунтування твердження про те, що всі важливі обов'язкові умови успіху виявлено;
- обґрунтування зв'язків між програмними заходами і довгостроковими результатами, які вони мають дати;
- неявне або явне розуміння контекстуальних або середовищних чинників, які, вірогідно, підтримуватимуть рух до поставлених цілей або перешкоджатимуть йому та позитивно чи негативно впливатимуть на обсяг досягнутого;

Фахівці з оцінювання вивчають логіку, що формується, та розглядають свої припущення.

При цьому можна ставити, зокрема, наступні запитання:

- Чи є ця теорія змін правдоподібною? Чи приведе ланцюжок подій до досягнення довгострокової цілі?
- Чи є ця теорія змін здійсненною? Чи є в наявності потенціал та ресурси для втілення стратегій, здатних створити довгострокові результати?
- Чи можна перевірити цю теорію? Чи вказано критерії вимірювання успішності? (Anderson 2004)

Також запитають:

- Що ще відбувається в середовищі такого, що могло б допомогти інтервенції або завадити їй?

Припущення записуються і включаються в ланцюжок подій. При складанні цього ланцюжка можна використовувати невеличкі папірці, які легко пересувати відповідно до змін у теорії, що формується.

Не варто виділяти всі припущення – в такому випадку перелік був би дуже довгим. Однак ключові припущення, помилковість яких означатиме найбільші ризики для успіху програми, обов'язково треба виділити.

У прикладі з програмою навчання до ключових припущень можуть відноситися наступні:

- У фахівців з оцінювання немає наявних джерел навчання для задоволення відповідних потреб.
- Фахівці з оцінювання можуть отримати фінансові ресурси для участі в навчанні.
- Навчання відповідає конкретним потребам фахівців з оцінювання.
- Фахівці з оцінювання цінують навчання і вмотивовані навчатися.

Фахівці з оцінювання отримають підтримку й інші потрібні ресурси, щоб реалізувати на практиці отримані знання.

- Фахівці з оцінювання матимуть навички написання звітів, необхідні для ефективного донесення інформації урядовим агенціям.
- Урядовці, що приймають рішення, використовуватимуть результати оцінок для прийняття краще обґрунтованих рішень.

Щоб цей ланцюжок був дієвим, необхідно працювати над виконанням критичних припущень. Їх можна викласти поряд зі схемою теорії змін або включити в цю схему (рис. 4.5).

Рис. 4.5. Схема простої діаграми теорії змін із зазначеними ключовими припущеннями

Джерело: Автори.

Шаблон теорії змін

Фонд Келога пропонує фахівцям з оцінювання використовувати шаблон для пояснення теорії змін (рис. 4.6).

Починати заповнювати шаблон теорії змін пропонується з середини моделі («Проблема або питання»). Це – центральна частина шаблону і ядро теорії змін. В цьому полі фахівець з оцінювання вписує чітке викладення проблеми або питання, яке має вирішити інтервенція.

На другому кроці («Потреби/активи спільноти») фахівець з оцінювання вказує потреби й активи спільноти або організації. Якщо було проведено оцінку потреб або їх пріоритизацію, відповідна інформація вноситься в цей розділ.

На третьому кроці («Бажані результати») фахівцю з оцінювання пропонується вказати, яких результатів повинна досягти інтервенція в коротко- та довгостроковій перспективі. Це може бути, по суті, баченням майбутнього. Надається перелік короткострокових результатів, довгострокових результатів і впливів.

Рис. 4.6. Шаблон теорії змін

Джерело: Kellogg Foundation 2004.

На четвертому кроці, що називається «Впливові фактори», фахівець з оцінювання має надати перелік потенційних перешкод та позитивних факторів, що можуть мати вплив на бажані зміни. Це можуть бути фактори ризику, існуюче політичне середовище та інші чинники (їх можна отримати в результаті аналізу спеціальної літератури та з інших джерел).

На п'ятому кроці («Стратегії») фахівець з оцінювання має перелічити виявлені в ході дослідження загальні успішні стратегії, що допомогли аналогічним спільнотам або організаціям досягти таких результатів, які є цільовими для проекту, програми або політики («передові практики»).

На останньому кроці («Припущення») фахівець з оцінювання має вказати припущення стосовно того, яким чином і чому визначені стратегії змін будуть дієвими для спільноти або організації. Це можуть бути принципи, переконання або ідеї. Після цього шаблон теорії змін можна використати для графічного представлення теорії змін.

Приклади побудови теорії змін

Два приклади демонструють, яким чином фахівець з оцінювання формує теорію змін. Перший описує програму з покращення академічної успішності учнів шляхом відвідування вчителями домівок учнів. Другий приклад – програма з проведення семінарів для скорочення урядової корупції.

Поліпшення успішності за рахунок відвідування вчителями домівок учнів. Розглянемо наступну ситуацію. Програма працює вже тривалий час і потребує оцінки. Але відповідної теорії змін немає. Фахівцю з оцінювання треба починати формування такої теорії з аналізу спеціальної літератури (як вказано у блоку 4.3). Бажана ціль програми – покращення академічної успішності учнів. Інтервенція – відвідування вчителями домівок учнів. Діє наступна логіка: якщо вчителі (залучений ресурс) відвідуватимуть (захід) домівки своїх учнів (залучений ресурс) і говоритимуть з батьками (короткостроковий результат), вони будуть більш чуйними до дитини (довгостроковий результат). Батьки краще розумітимуть потребу школи в тому, щоб домашні завдання виконувались вчасно, а діти відвідували школу щодня (короткостроковий результат); тому вони докладатимуть зусиль, щоб обидві ці потреби задовольнялись (довгострокові результати). Завдяки тому, що дитина виконує домашні завдання, регулярно відвідує школу та має чулого вчителя, її академічна успішність зросте (вплив).

Фахівець з оцінювання, який створює теорію змін, починає з цільового результату – кращих показників у читанні – і розміщує його внизу схеми (в даному випадку, див. рис. 4.7). Потім вказує інтервенцію: пише зверху «відвідування вчителями домівки учня». Далі він визначає три можливих результати:

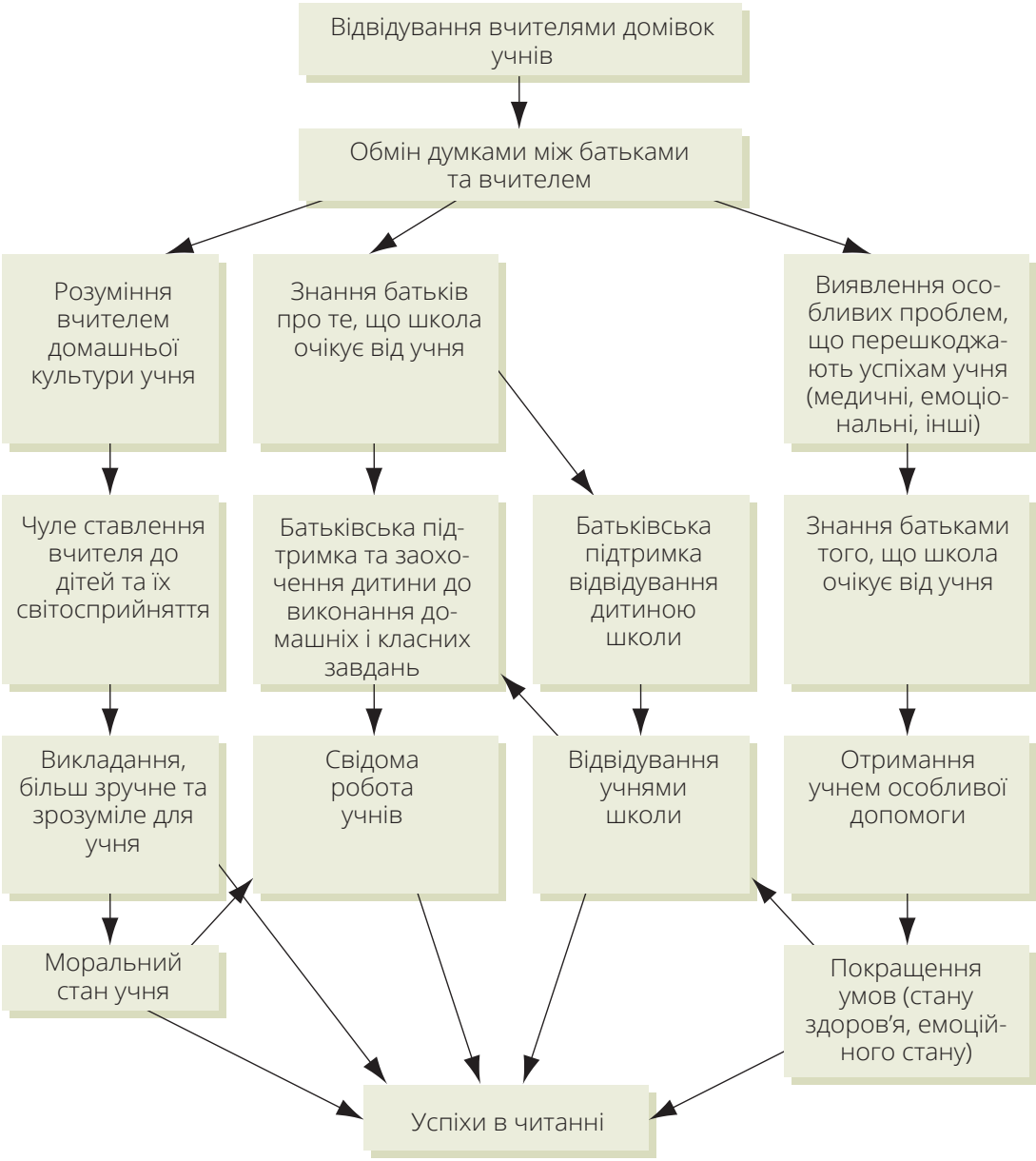
- Вчителі отримують розуміння домашньої культури учня.
- Батьки отримують інформацію про те, чого школа очікує від учнів.
- І вчителі, і батьки можуть працювати над усуненням проблем, що заважають учню добре навчатися.

За кожним з цих трьох можливих результатів фахівець з оцінювання визначає інші результати, формуючи ланцюжки результатів та взаємодію результатів. Наприклад, з результату «розуміння вчителями домашньої культури» фахівець з оцінювання виводить ланцюжок результатів:

- вчителі стають більш чуйними до учнів та їх світосприйняття;
- вчителі починають викладати у більш зручний для учня спосіб;
- мотивація учня зростає.

Тоді покращуються показники учня у читанні.

Рис. 4.7. Теорія змін для програми з поліпшення успішності учнів у читанні шляхом відвідування вчителями домівок учнів



Джерело: Weiss 1972, 50.

Після цього фахівець з оцінювання визначає інші можливі результати кожного з початкових варіантів, кожен з яких завершується досягненнями у читанні. Деякі ланцюжки результатів взаємодіють з іншими можливими результатами.

Ця теорія змін також визначає кілька припущень. В даному випадку вони викладені у вигляді списку, а не намальовані на схемі. Фахівець з оцінювання виокремлює наступні припущення:

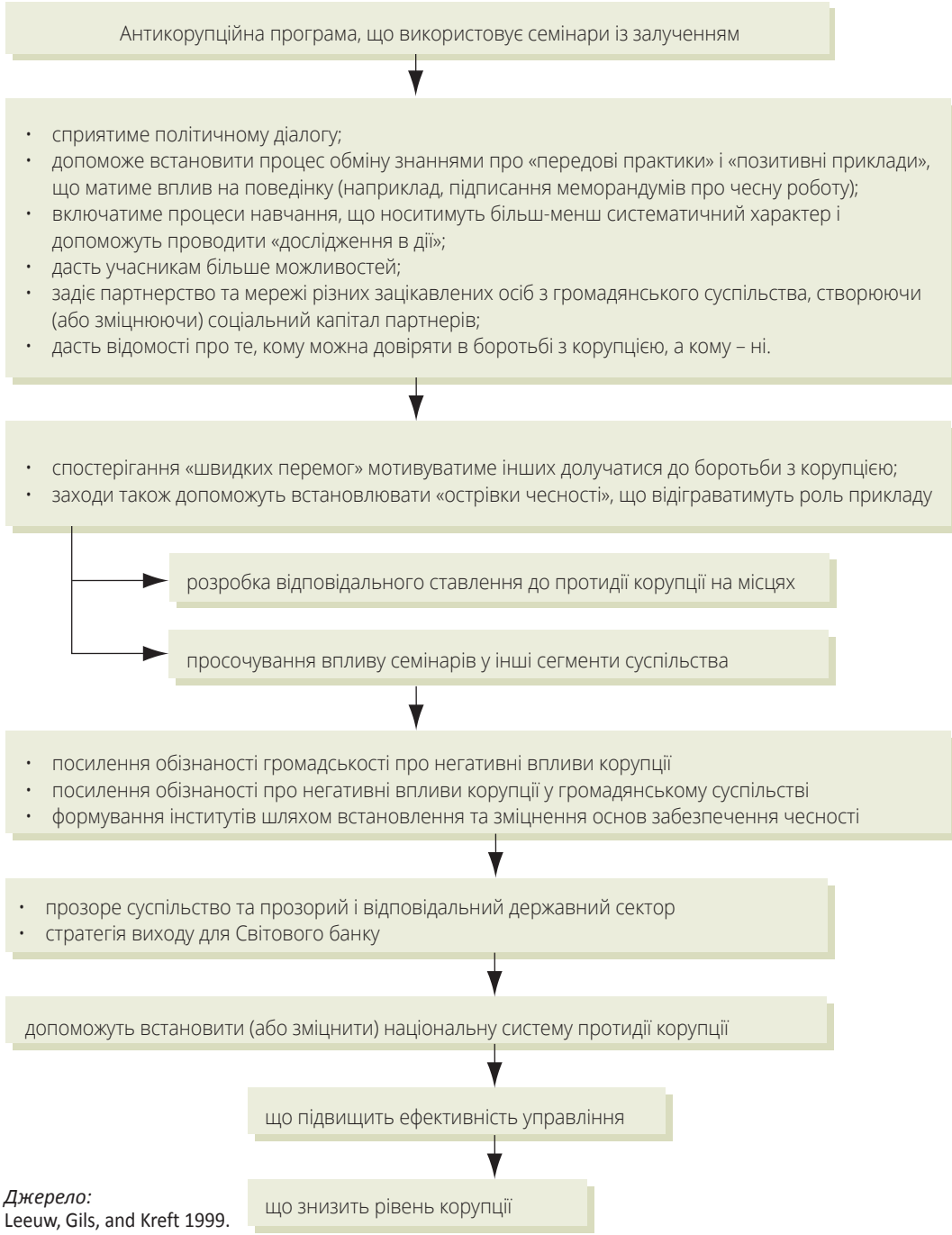
- Діти – з повних сімей з домівками.
- Батьки вільні для розмови та знаходяться вдома тоді, коли вчителі можуть їх відвідувати.
- Вчителі готові відвідувати домівки учнів.
- Батьки прийматимуть вчителів у себе вдома.
- Батьки бажатимуть обговорювати свої погляди на навчання дітей.
- Вчителі краще розумітимуть домашню культуру, а отже, будуть більш чулими до своїх учнів.
- Вчителі корегуватимуть спосіб і манеру викладання, використовуючи нове розуміння.
- Батьки прагнуть брати участь в навчанні дітей та щоб вони відвідували школу, виконували домашні завдання й отримували гарні оцінки.
- У батьків немає потреби у тому, щоб їхні діти працювали.
- У шкільному або домашньому середовищі не відбувається нічого іншого, що б могло викликати покращення академічних успіхів учнів.

Оцінювання формувалося б таким чином, щоб розглянути ті припущення, які, на думку зацікавлених осіб або літературних джерел, є критичними.

Використання семінарів із залученням зацікавлених сторін для боротьби з корупцією.

Розглянемо інший приклад. Програма намагається запровадити семінари для зниження рівня корупції в уряді. Формуючи теорію змін, фахівець з оцінювання починає з того, що в нижній частині схеми вказує довгострокову ціль – скорочення корупції, а у верхній – інтервенцію. Головні події, що, як очікується, матимуть місце, вказуються в ланцюжку подій, розташованому всередині (рис. 4.8) (Haaruis 2005).

Рис. 4.8. Теорія змін для програми з використання семінарів із залученням зацікавлених сторін для скорочення корупції в уряді



Джерело:
Leeuw, Gils, and Kreft 1999.

Хоча було визначено багато припущень, фахівець з оцінювання обмежує припущення, що розглядаються в ході оцінювання, трьома ключовими, вибір яких ґрунтується на аналізі літератури та обговореннях з клієнтом та зацікавленими особами:

- Семінари є ефективними та відповідають потребам учасників і програми.
- Навички, ставлення та мотивація учасників є достатніми, щоб брати участь у семінарах.

В учасників з'явиться відповідальне ставлення до проблеми, яке стане проникатиме й в інші сфери.

Термінологія та графічне представлення теорії змін

Через те, що програмна теорія стала серйозним фактором в оцінюванні, виникли проблеми з термінологією щодо таких термінів, як *логічні моделі*, *моделі довгострокових результатів* та *моделі теорії*. Приміром, Петон (Patton 2002) відокремлює логічну модель від теорії змін, стверджуючи: єдиним критерієм для логічної моделі є те, що вона зображує раціональну, обґрунтовану послідовність переходу від залучених ресурсів і заходів до короткострокових результатів, довгострокових результатів і впливів. Теорія змін, навпаки, повинна також вказувати та пояснювати допущені, гіпотетичні або перевірювані причинно-наслідкові зв'язки. Незалежно від термінології або формату, всі зображення теорії змін повинні представляти ланцюжок причинно-наслідкових зв'язків, показувати наявні впливи, а також визначати ключові припущення.

Моделі теорії змін можна по-різному представляти візуально у вигляді блок-схем. В цьому розділі ми наводимо декілька форматів.

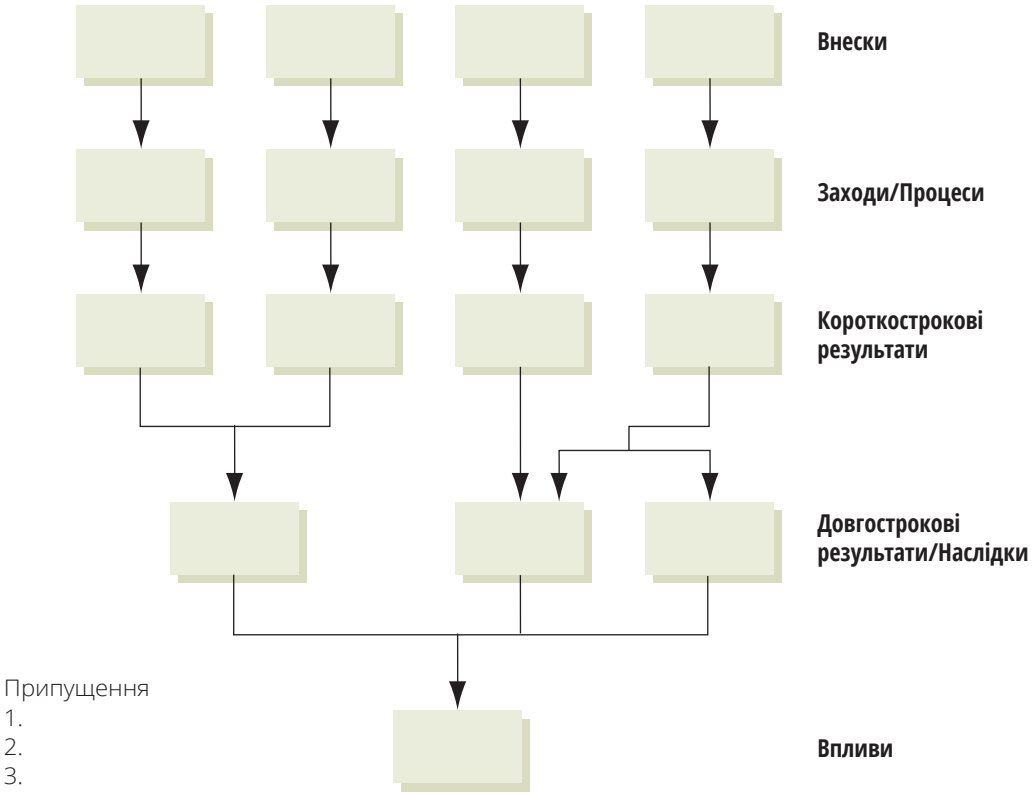
Стандартна блок-схема

Блок-схема:
Схема, в якій показано послідовність наслідків заходів та короткострокових результатів

Блок-схеми – найпопулярніший формат для представлення теорії змін. **Блок-схема** демонструє послідовність наслідків заходів та короткострокових результатів (див. рис. 4.9). Це дуже гнучка логічна модель за умови представлення ключових компонентів – залучених ресурсів, заходів, короткострокових результатів і впливів. Аби блок-схема чітко давала розуміння, яким чином досягаються кінцеві результати, можна представити кілька рівнів результатів (безпосередні, проміжні, кінцеві). Використовуючи цей формат, фахівці з оцінювання мають перелічувати свої припущення, включаючи фактори зовнішнього середовища, що можуть впливати на ступінь досягнення інтервенцією цільового результату.

Причинно-наслідкові зв'язки можна пояснювати за допомогою формулювань «якщо... то...» Наприклад, «якщо реалізується захід, то будуть отримані ці короткострокові результати. Якщо короткострокові результати будуть отримані, то вони приведуть до виходу на перший рівень безпосередніх результатів».

Рис. 4.9. Стандартна блок-схема



Джерело: Автори.

Стандартна блок-схема з акцентом на припущення

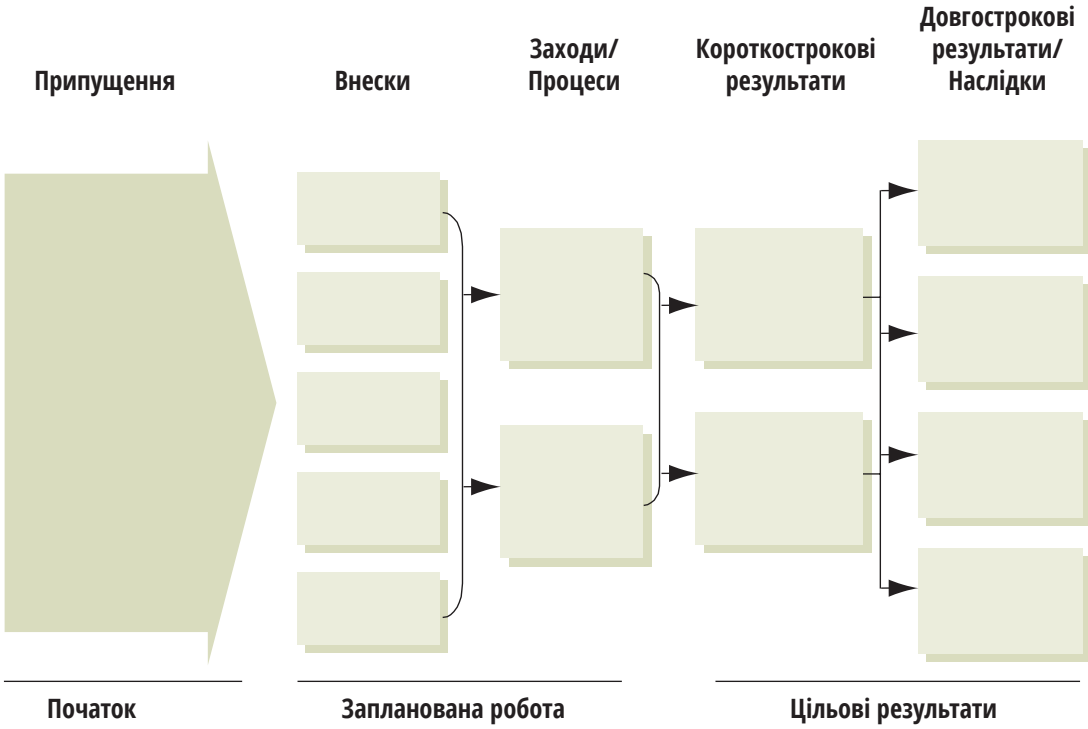
Інший варіант представлення теорії змін показано на рис. 4.10. Ця модель включає припущення, які виступають в якості принципів, що лежать в основі проекту ініціативи.

Стандартний ланцюжок результатів

Ланцюжок результатів схожий з блок-схемою, але в ньому конкретні заходи не обов'язково пов'язуються з конкретними короткостроковими результатами або наслідками. Через відсутність такого рівня деталізації причинно-наслідкового зв'язку заходів, короткострокових результатів і наслідків, який є у блок-схемах, користувачу варто стежити за тим, щоб «стандартизований» підхід не став «спрощенським». Як і у випадку з іншими видами візуального представлення, явно окреслюється вплив зовнішніх факторів. Залучені ресурси, заходи та короткострокові результати часто використовуються як показники ефективності; результати використовуються для визначення дієвості програми (рис. 4.11).

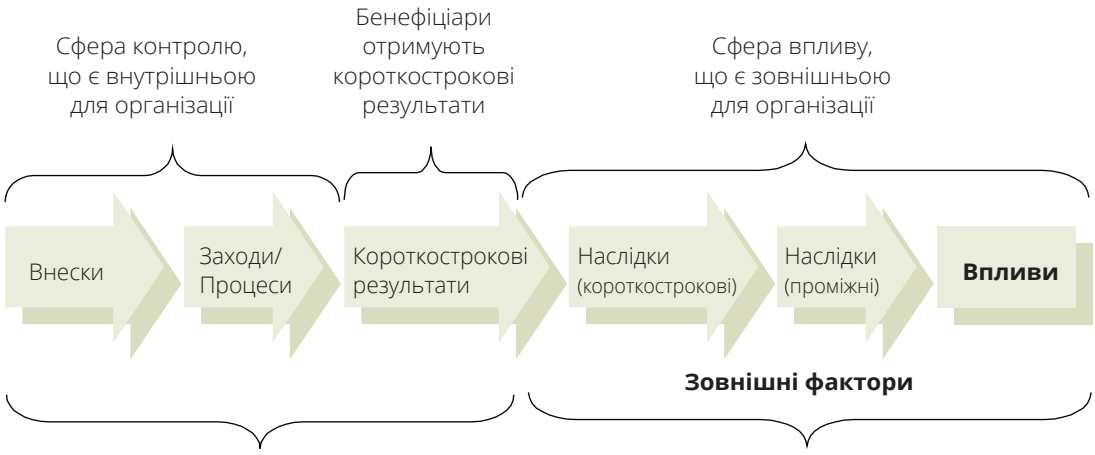
Ланцюжок результатів:
Зумовлена причинно-наслідковими зв'язками послідовність подій в рамках інтервенції розвитку, що забезпечує поступове досягнення цілей, починаючи із залучення ресурсів для проведення заходів для отримання короткострокових результатів і закінчуючи довгостроковими результатами, впливами і зворотним зв'язком

Рис. 4.10. Блок-схема з акцентом на припущення



Джерело: Взято з Kellogg Foundation 2004.

Рис. 4.11. Стандартний ланцюжок результатів



Джерело: Canadian Department of Foreign Affairs and International Trade 2005.

Багато прикладів завершених моделей можна знайти на веб-сайті *University of Wisonsin-Extension*. Посилання на цей сайт наведено в кінці цієї глави.

Логічна структура (логфрейм)

Одним із варіантів моделі теорії змін є **логічна структура**[3], або, як її часто називають, логфрейм. Логфрейм ієрархічно пов'язує заходи, результати, призначення й цілі програми, політики або проекту. Для кожного компонента програми, проекту або політики фахівець з оцінювання визначає потрібні показники ефективності, джерела інформації щодо кожного показника, а також припущення. Логфрейми прояснюють цілі проекту, програми або політики та визначають причинно-наслідкові зв'язки між залученими ресурсами, процесами, короткостроковими результатами, довгостроковими результатами і впливами. Багато організацій з розвитку вимагають використання логфреймів та навчили своїх співробітників користуватися ними.

Логічна структура (логфрейм): Матриця, що ієрархічно пов'язує заходи, результати, призначення й цілі програми, політики або проекту

Логфрейм – це матриця 4 х 4, що стисло представляє критичні елементи проекту, програми або політики. В рамках цього підходу методично розбираються основні питання в контексті причин та наслідків.

Логфрейми можна використовувати для різних цілей:

- покращення якості плану реалізації проекту, програми або політики завдяки встановленню чітких цілей, використанню показників ефективності та оцінці ризиків;
- стисле представлення плану складних заходів;
- допомога персоналу в підготовці деталізованих робочих планів;
- надання об'єктивного підгрунтя для аналізу, моніторингу та оцінки заходів (World Bank 2004).

Логфрейми активно критикуються. Ось деякі з найсерйозніших зауважень:

- При розробці логфрейма фахівець з оцінювання може заплутатися в деталях і випустити з поля зору загальну картину.
- Немає акценту на базових даних.
- Логфрейми нерідко є занадто простими навіть для нескладних проектів. Як зазначає Геспер (Gasper 1997): «Не все важливе можна зафіксувати за допомогою чотирьох-п'ятирівневої схеми, що займає від однієї до трьох сторінок».
- Багато користувачів логфреймів не помічають, що структура включає одні деталі, залишаючи поза увагою інші. Структура слугує допоміжним засобом оцінки, а не її замінником.
- Логфрейм не розглядає неочікувані довгострокові результати; його обсяг обмежується заданими цілями.

[3] *В українському варіанті також зустрічається такий варіант терміну, як «Логічна рамка».*

Після того, як логфрейм сформовано, він зазвичай фіксується і не оновлюється, стаючи «лок-фреймом», тобто непорушним логфреймом (Gasper 1997).

Логфрейм для дитячої консультації міг би включати в якості одного із заходів імунізацію з метою захистити 50 відсотків всіх дітей віком до шести років у певному районі (табл. 4.3). В разі досягнення поставленої мети частота випадків контрольованих дитячих захворювань має знизитися. Це має забезпечити досягнення загальної мети скорочення кількості смертей серед дітей віком до шести років.

Другий стовпець в табл. 4.3 визначає показники, що показують ступінь досягнення кожної цілі. Третій та четвертий стовпці вказують, де було отримано дані для оцінки рівня показників, а також які припущення зроблено щодо природи й доступності цих даних.

Оцінка теорії змін

Після формування теорії змін фахівцям з оцінювання потрібно зупинитися й оцінити якість цієї теорії з різних точок зору. Ці точки зору, або парадигми, передбачають оцінку:

- у контексті соціальних потреб;
- логіки й правдоподібності;
- шляхом порівняння з дослідженнями та практикою;
- шляхом співставлення програмної теорії змін з однією або кількома відповідними науковими теоріями;
- шляхом попереднього спостереження (взято з Rossi, Freeman, and Lipsey 1999).

Теорія змін повинна давати відповіді на наступні запитання:

- Чи є модель точним відтворенням програми?
- Чи всі елементи добре визначено?
- Чи немає пропусків у логічному ланцюжку подій?
- Чи є елементи необхідними й достатніми?
- Чи є відносини правдоподібними й послідовними?
- Чи реалістично очікувати, що програма завершиться переконливим досягненням задекларованих цілей?

Фонд Келога розробив контрольний список для оцінки якості логічної моделі. З цього контрольного списку і взято наступний перелік критеріїв:

- Створено перелік основних заходів, потрібних для реалізації програми.
- Заходи поєднані з програмною теорією.
- Створено перелік основних ресурсів, потрібних для реалізації програми.

Ресурси відповідають типу програми. На всі заходи виділяються належні ресурси в необхідній кількості.

Таблиця 4.3. Логічна структура програми зі зниження дитячої смертності

Компонент	Показники результативності	Перевірка	Припущення
Мета: Покращити соціально-економічне становище жінок та їхніх родин	Підвищення доходу сімей відсоток сімей-учасниць. Вдосконалення засобів оцінки стану здоров'я, якості харчування та освіти	Дослідження домашніх господарств на предмет економічного, соціального та медичного стану всіх членів родини	
Завдання: Надати жінкам можливості заробляти й навчатися, поки за їх дітьми наглядають у центрах денного догляду за дітьми	• Центри денного догляду за дітьми працюють, надаючи доступні технічно й фінансово послуги належної якості у робочий час, забезпечуючи жінкам можливість змін у режимі їх працевлаштування та навчання	Через опитування: зміни у працевлаштуванні та освіті жінок, їх оцінки стосовно наданих послуг з догляду за дітьми. Оцінювання якості послуг догляду за дітьми на підставі спостережень	Інші члени родини працюють і заробляють гроші на незмінному або кращому рівні. Економічні умови залишаються незмінними або покращуються
Короткострокові результати/Продукти: • Підготовлені вихователі, методисти та директори • Центри денного догляду за дітьми модернізовані, вони працюють • Розроблено матеріали • Впроваджено адміністративну систему • Впроваджено систему управління інформацією (СУІ)	• Вихователі пройшли навчання • Центри модернізовано, вони працюють • Матеріали створено й розповсюджено • Працююча СУІ	Дані з СУІ щодо підготовлених спеціалістів, центрів та матеріалів. Оцінки підготовлених спеціалістів. Після початкової підготовки та протягом курсу постійного навчання	Родинні умови дозволяють матерям, які працюють в центрах денного догляду за дітьми, виконувати свої зобов'язання
Заходи/Процеси • Вибрати вихователів і методистів, забезпечити їх початкову підготовку: • Модернізувати центри • Розробити матеріали • Створити адміністративну систему • Надавати послуги денного нагляду за дітьми • Надавати постійне навчання та методичну підтримку • Створити СУІ	*Ресурси* Бюджет Технології Кадри	План дій, бюджети, матеріали бухгалтерського обліку. Дослідження, які показують, що обрана модель і навчальний план працюють. Оцінювання, покликані визначити, що заходи виконувалися ефективно	

Джерело: Inter-American Development Bank (http://www.iadb.org/sds/soc/eccd/6example.html#ex1).

Підсумки

Фахівці з оцінювання повинні протистояти бажанню одразу розпочати оцінювання, не маючи плану. Аналіз вихідних даних – цінна складова процесу планування, що допомагає фахівцям з оцінювання отримати загальне уявлення про проект, програму або політику. Аналіз вихідних даних може дати відповіді на важливі запитання щодо оцінювання, в тому числі про вибір часу для її початку, строки проведення, залучення людських та інших ресурсів, проектування, програмну теорію й логіку, а також про існуючі знання.

Однією із складових аналізу вихідних даних є визначення окремих людей або груп, що беруть участь в інтервенції. Аналіз зацікавлених осіб – один зі способів визначення основних зацікавлених осіб, оцінки та отримання інформації про те, що їм відомо, чим вони можуть допомогти оцінюванню та що очікують отримати від нього.

Інша складова аналізу вихідних даних – розгляд контексту інтервенції. Фахівці з оцінювання мають відшукати пов'язані дослідження та оцінки аналогічних інтервенцій і скористатися ними. Фахівці з оцінювання ставлять різні запитання на різних етапах життєвого циклу інтервенції, отож важливо визначити етап інтервенції.

Формування теорії змін, на якій ґрунтується інтервенція, допомагає фахівцям з оцінювання та зацікавленим особам візуалізувати інтервенцію та визначити припущення, що лежать в основі програми і мають бути перевірені в ході оцінювання. Хоча існують різні шляхи графічного представлення теорії змін, всі вони повинні базуватися на результатах пов'язаних досліджень і оцінок, представляти логічний розвиток, виділяти події, що можуть вплинути на результати, та показувати причинно-наслідковий ланцюжок подій. Зазвичай теорії змін дають уявлення про залучені ресурси, заходи, короткострокові результати, наслідки, а також їх взаємозалежність одне з одним та середовищем, де все відбувається.

Завдання до Глави 4

Завдання 4.1: Застосування теорії змін

Розглянемо програму мікрокредитування, що має на меті створення нових джерел доходу та покращення добробуту домогосподарств за рахунок допомоги жінкам, зокрема бідним сільським жінкам, долучитися до ринку праці та отримати підприємницькі навички, збільшивши таким чином доходи родини. Довгостроковою ціллю є стимулювання розвитку приватного сектора та посилення економічного зростання. Середній розмір кредиту – 225 доларів США, максимальний розмір – 500 доларів. Позики надаються у вигляді одночасних виплат для поповнення оборотного капіталу або як інвестиції в мікропідприємство. Встановлені терміни повернення кредитів – від 1 до 10 років, при цьому середній строк – 2–3 роки. Надається річна відстрочка повернення. Пов'язаний з цією діяльністю проект нарощування навичок охоплює базові теми в сфері бухгалтерського обліку та фінансового менеджменту. На рис. А показано спрощену логічну модель програми мікрокредитування, на рис. В – більш деталізоване графічне відображення теорії змін тієї самої програми. Обведені колами пункти на рис. В – деякі з головних факторів середовища, що можуть вплинути на досягнення цілі програми. Подумайте про те, які припущення лежать в основі цієї програми.

Складіть список з п'яти головних припущень, що лежать в основі програми:

1. _____
2. _____
3. _____
4. _____
5. _____

Рис. А. Проста теорія змін програми мікрокредитування

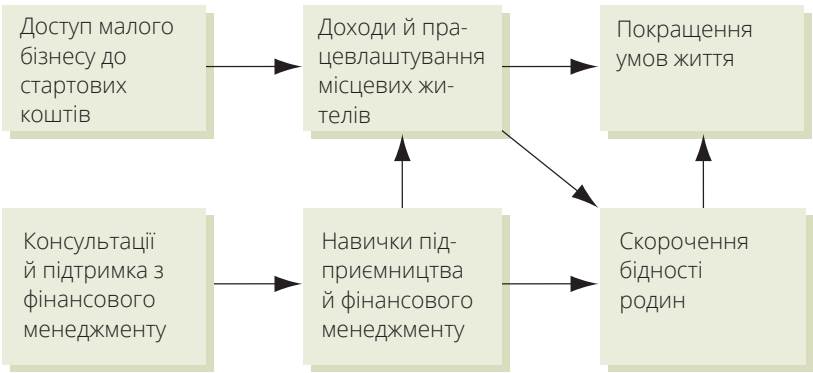
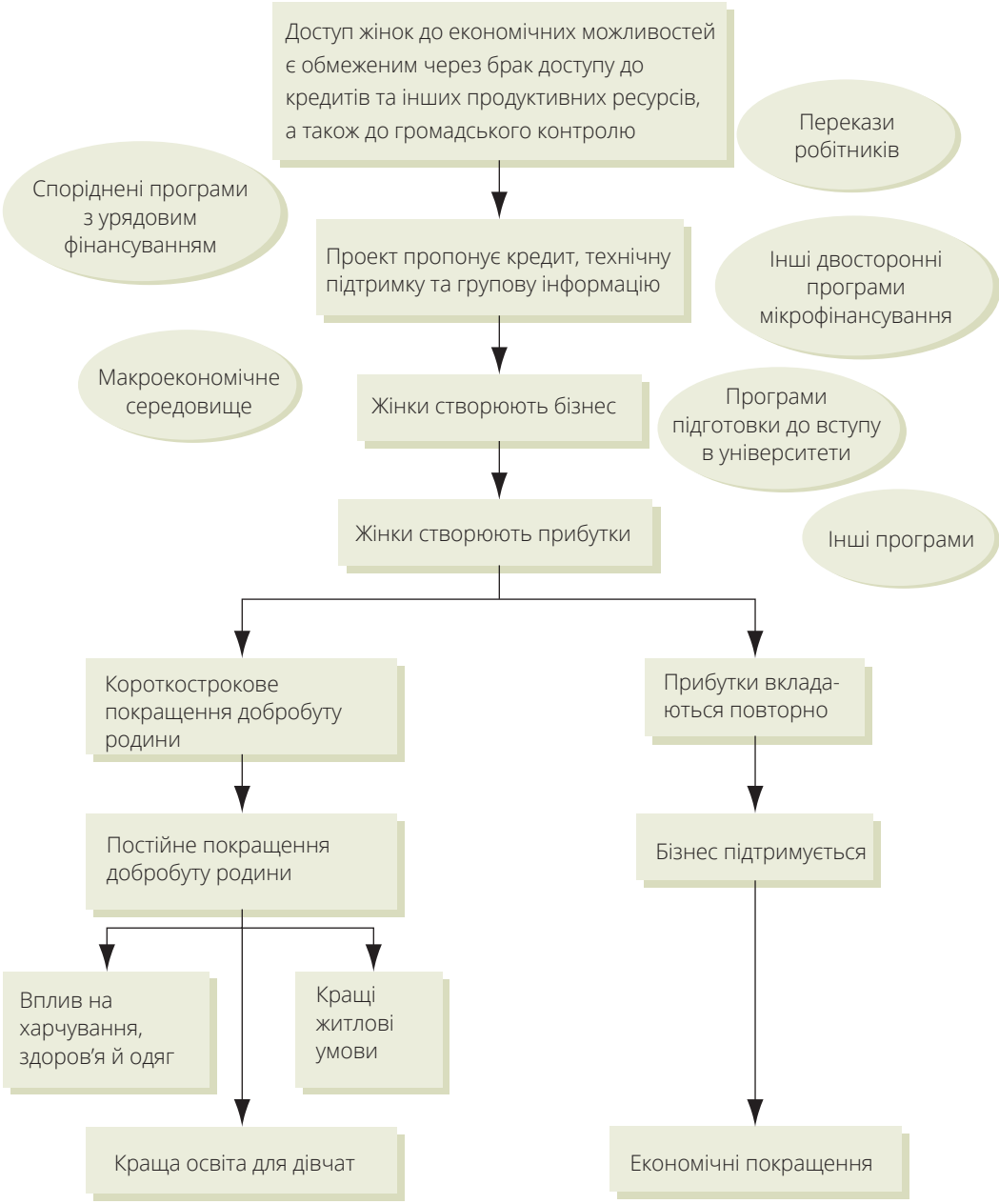


Рис. В. Складніша теорія змін програми мікрокредитування



Практична вправа 4.2: Аналіз вашої програми

Розглянемо програму або проект, над яким ви зараз працюєте або про який вам добре відомо, хоча ви й не берете в ньому участь. Створіть графічну схему теорії змін такої програми або проекту. Постарайтеся визначити припущення, що лежать в основі програми або проекту, особливо ті з них, що стосуються факторів зовнішнього середовища.

Посилання та рекомендована література

ActKnowledge and Aspen Institute. 2003. *Theory of Change.* Roundtable on Community Change.
http://www.theoryofchange.org/.

Anderson, Andrea, A. 2004. "The Community Builder's Approach to Theory of Change: A Practical Guide to Theory Development." Presentation at the Aspen Institute Roundtable on Community Change, New York. http://www.aspeninstitute.org/atf/cf/ percent7BDEB6F227-659B-4EC8-8F84- 8DF23CA704F5 percent7D/rcccommbuildersapproach.pdf.

Bell, P. 1997. "Using Argument Representations to Make Thinking Visible for Individuals and Groups." In *Proceedings of CSCL '97: The Second International Conference on Computer Support for Collaborative Learning,* eds. R. Hall, N. Miyake, and N. Enyedy, 10–19. Toronto: University of Toronto Press.

Bruning, R. H., G. J. Schraw, M. M. Norby, and R. R. Ronning. 2004. *Cognitive Psychology and Instruction.* 4th ed. Upper Saddle River, NJ: Pearson Merrill Prentice-Hall.

Canadian Department of Foreign Affairs and International Trade. 2005. *Six Easy Steps to Managing for Results: A Guide for Managers.* April. Evaluation Division, Ottawa.

CIDA (Canadian International Development Agency). 2005. *Case Study #1: Jamaica Environmental Action Program ENACT.* Caribbean Division, Americas Branch. http://www.acdi-cida.gc.ca/CIDAWEB/acdicida.nsf/En/EMA-218131811-PHY#1.

Eggen, P., and D. Kauchak. 1999. *Educational Psychology: Windows on Classrooms.*
4th ed. Upper Saddle River, NJ: Merrill: Prentice-Hall.

Fitzpatrick, Jody L., James R. Sanders, and Blaine R. Worthen. 2004. *Program Evaluation: Alternative Approaches and Practical Guidelines.* New York: Person Education.

Funnell, S. 1997. "Program Logic: An Adaptable Tool for Designing and Evaluating Programs." *Evaluation News and Comment* 6 (1): 5–7.

Gagne, R. M., and W. D. Rohwer Jr. 1969. "Instructional Psychology." *Annual Review of Psychology* 20: 381–418.

Gasper, D. 1997. "Logical Frameworks: A Critical Assessment Managerial Theory." Pluralistic Practice Working Paper 264, Institute of Social Studies, The Hague.

———. 2000. "Evaluating the 'Logical Framework Approach' towards Learning-Oriented Development Evaluation." *Public Administration Development* 20 (1): 17–28.

Haarhuis, Carolien Klein. 2005. "Promoting Anticorruption Reforms: Evaluating the Implementation of a World Bank Anticorruption Program in Seven African Countries."

http://igitur-archive.library.uu.nl/dissertations/2005-0714-200002/full.pdf.

Healy, A. F, and D. S. McNamara. 1996. "Verbal Learning Memory: Does the Modal Model Still Work?" *Annual Review of Psychology* 47: 143–72.

Heubner, T. A. 2000. "Theory-Based Evaluation: Gaining a Shared Understanding Between School Staff and Evaluators." In *Program Theory in Evaluation: Challenges and Opportunities,* eds. Patricia J. Rogers, T. A. Hacsi, A. Petrosino, and

T. A. Huebner, 79–90. New Directions for Evaluation No. 87. San Francisco, CA: Jossey-Bass.

Kassinove, H., and M. Summers. 1968. "The Developmental Attention Test: A Preliminary Report on an Objective Test of Attention." *Journal of Clinical Psychology* 24 (1): 76–78.

Kellogg Foundation. 2004. *Logic Model Development Guide.* Battle Creek, MI. http://www.wkkf.org/Pubs/Tools/Evaluation/Pub3669.pdf.

Leeuw, Frans, Ger Gils, and Cora Kreft. 1999. "Evaluating Anti-Corruption Initiatives: Underlying Logic and Mid-Term Impact of a World Bank Program." *Evaluation* 5 (2): 194–219.

Leeuw, Frans L. 1991. "Policy Theories, Knowledge Utilization, and Evaluation." *Knowledge and Policy* 4: 73–92.

———. 2003. "Reconstructing Program Theories: Models Available and Problems to Be Solved." *American Journal of Evaluation* 24 (1): 5–20.

Mikkelsen, B. 2005. *Methods for Development Work and Research*: *A New Guide for Practitioners.* Thousand Oaks, CA: Sage Publications.

MSH (Management Sciences for Health), and UNICEF (United Nations Children's Fund). 1998. "Quality Guide: Stakeholder Analysis." In *Guide to Managing for Quality.* http://ERC.MSH.org/quality.

Newman, David Kent. 2007. “Theory: Write-up for Conceptual Framework.” http://deekayen.net/theory-write-conceptual-framework.
Ormrod, J. E. 2006. *Essentials of Educational Psychology.* Upper Saddle River, NJ: Pearson Merrill Prentice-Hall.
Owen, J. M., and P. J. Rogers. 1999. *Program Evaluation: Forms and Approaches.*
Thousand Oaks, CA: Sage Publications.
Pancer, S. Mark, and Anne Westhues. 1989. “A Developmental Stage Approach to Program Planning and Evaluation.” *Evaluation Review* 13 (1): 56–77.
Patton, M. Q. 2002. *Qualitative Research and Evaluation Methods.* 3rd ed. Thousand Oaks, CA.: Sage Publications.
Pawson, Ray. 2006. *Evidence-Based Policy: A Realistic Perspective.* New Brunswick, NJ: Sage Publications.
Petrosino, Anthony, Robert A. Boruch, Cath Rounding, Steve McDonald, and Iain Chalmers. 2003. “The Campbell Collaboration Social, Psychological, Educational, and Criminological Trials Register C20SPECTR.” http://www.campbellcollaboration.org/papers/unpublished/petrosino.pdf.
Porteous, Nancy L., B. J. Sheldrick, and P. J. Stewart. 1997. *Program Evaluation Tool Kit: A Blueprint for Public Health Management.* Ottawa-Carleton Health Department, Ottawa.
http://www.phac-aspc.gc.ca/php-psp/tookit.html.
———. 2002. “Introducing Program Teams to Logic Models: Facilitating the Learn- ing Process.” *Canadian Journal of Program Evaluation* 17 (3): 113–41.
Prensky, Marc. 2001. “Digital Natives, Digital Immigrants.” *On the Horizon* 9 (5). http://www.marcprensky.com/writing/Prensky percent20-percent20Digital percent20Natives,percent20Digital percent20Immigrants percent20-percent20Part1.pdf.
Ritzer, George. 1993. *The McDonaldization of Society.* Rev. ed. Thousand Oaks, CA: Pine Forge Press.
Rogers, Patricia J., T. A. Hacsi, A. Petrosino, and T. A. Huebner, eds. 2000. *Program Theory in Evaluation: Challenges and Opportunities.* New Directions in Evaluation No. 87. San Francisco: Jossey-Bass Publishers.
Rossi, P., H. Freeman, and M. Lipsey. 1999. *Evaluation: A Systematic Approach.*
Thousand Oaks, CA: Sage Publications.
Scott, M. 2003. “The Benefits and Consequences of Police Crackdowns.” *Problem-Oriented Guides for Police, Response Guide 1.* U.S. Department of Justice, Office of Community-Oriented Policing Services, Washington, DC.
Scriven, Michael. 2007. *Key Evaluation Checklist*. http://www.wmich.edu/evalctr/checklists/kec_feb07.pdf.
Shadish, W. R. Jr., T. D. Cook, and L. C. Leviton. 1991. *Foundations of Program Evaluation.* Thousand Oaks, CA: Sage Publications.

Sherman, L. W., D. Farrington, B. C. Welsh, and D. L. MacKenzie, eds. 2002. *Evidence-Based Crime Prevention.* London: Routledge.

Stufflebeam, Daniel L. 2001. *Evaluation Models.* New Directions for Evaluation No.
89. San Francisco, CA: Jossey-Bass.

Stufflebeam, D. L., G. F Madaus, and T. Kellaghan, eds. 2000. *Evaluation Models.*
2nd ed. Boston: Kluwer Academic Publishers.

Suthers, D. D., and A. Weiner. 1995. *Belvédère.*
http://lilt.ics.hawaii.edu/belvedere/index.html.

Suthers, D. D., E. E. Toth, and A. Weiner 1997. "An Integrated Approach to Implementing Collaborative Inquiry in the Classroom." In *Proceedings of CSCL '97: The Second International Conference on Computer Support for Collaborative Learning,* eds. R. Hall, N. Miyake, and N. Enyedy, 272–79. Toronto: University of Toronto Press.

Swedberg, Richard. 2003. *Principles of Economic Sociology.* Princeton, NJ: Princeton University Press.

Taylor-Powell, Ellen. 2005. *Logic Models: A Framework for Program Planning and Evaluation.* University of Wisconsin Extension, Program Development and Evaluation. http://www.uwex.edu/ces/pdande/evaluation/pdf/nutritionconf05.pdf.

U.S. GAO (General Accounting Office). 1991. *Designing Evaluations.* Washington, DC. http://www.gao.gov/special.pubs/10_1_4.pdf.

Weiss, Carol H. 1997. *Evaluation: Methods for Studying Programs and Policies.* Upper Saddle River, NJ: Prentice-Hall.

World Bank. 1996. *World Bank Participation Sourcebook.* Washington, DC: World Bank.
http://www.worldbank.org/wbi/sourcebook/sbhome.htm.

———. 2004. *Monitoring and Evaluation: Some Tools, Methods and Approaches.*
Operations Evaluation Department/Evaluation Capacity Development, Washington, DC.
http://lnweb18.worldbank.org/oed/oeddoclib.nsf/b57456d58aba40e585256ad400736404/a5efbb5d776b67d285256b1e0079c9a3/$FILE/MandE_tools_methods_approaches.pdf.

Worthen, Blaine. R., James R. Sanders, and Jody L. Fitzpatrick. 1997. *Program Evaluation.* New York: Longman.

Веб-сайти

Campbell Collaboration. http://www.campbellcollaboration.org/.
Community Toolbox. *A Framework for Program Evaluation: A Gateway to Tools.*
http://ctb.lsi.ukans.edu/tools/EN/sub_section_main_1338.htm.
Evaluation Center, Western Michigan University. *The Checklist Project.* http://www.wmich.edu/evalctr/checklists/checklistmenu.htm#mgt.
IDRC (International Development Research Centre). 2004. *Evaluation Planning in Program Initiatives* Ottawa, Ontario, Canada. http://web.idrc.ca/uploads/user-S/108549984812guideline-web.pdf.
Suthers, D., and A. Weiner. 1995.*Groupware for Developing Critical Discussion Skills.*
http://www-cscl95.indiana.edu/cscl95/suthers.html.
University of Wisconsin-Extension. *Logic Model*. http://www.uwex.edu/ces/ pdande/evaluation/evallogicmodel.html.
———. *Examples of Logic Models.* http://www.uwex.edu/ces/pdande/evaluation/ evallogicmodelexamples.html.

Theories of Change

Audience Dialogue. 2007a. "Enhancing Program Performance with Logic Models." http://www.uwex.edu/ces/lmcourse/, http://www.wkkf.org/pubs/tools/ evaluation/pub3669.pdf.
———. 2007b. "Program Logic: An Introduction." http://www.audiencedialogue.org/ proglog.htm.
AusAid. 2005. "Section 3.3: The Logical Framework Approach and Section 2.2: Using the Results Framework Approach." http://www.ausaid.gov.au/ausguide/ default.cfm.
BOND. 2001. *Guidance Notes Series I: Beginner's Guide to Logical Framework Analysis.* http://www.ngosupport.net/graphics/NGO/documents/english/273_BOND_ Series_1.doc.
Davies, Rick. 2003. "Network Perspective in the Evaluation of Development Interventions: More than a Metaphor." Paper presented at the EDAIS Conference "New Directions in Impact Assessment for Development: Methods and Practice," November 24–25. http://www.enterprise-impact.org.uk/conference/ Abstracts/Davies.shtml and http://www.mande.co.uk/docs/nape.doc.

Department for International Development. 2002. *Tools for Development: A Handbook for Those Engaged in Development Activity Performance and Effectiveness.* http://www.dfid.gov.uk/pubs/files/toolsfordevelopment.pdf.

den Heyer, Molly. 2001a. *A Bibliography for Program Logic Models/Logframe Analysis.* Evaluation Unit, International Development Research Centre. http://www.idrc.ca/uploads/user-S/10553606170logframebib2.pdf and http://www.mande.co.uk/docs/Phillips.ppt.

———. 2001b. "The Temporal Logic Model™: A Concept Paper." http://www.idrc.ca/ uploads/user-S/10553603900tlmconceptpaper.pdf.

Inter-American Development Bank. http://www.iadb.org/sds/soc/eccd/bexample.html#ex1.

International Fund for Agricultural Development. 2003. *Annotated Example of a Project Logframe Matrix.* http://www.ifad.org/evaluation/guide/3/3.htm and http://www.ifad.org/evaluation/guide/annexb/index.htm.

JISC infoNet. 2001. *Engendering the Logical Framework.* http://www.jiscinfonet.ac.uk/InfoKits/project-management/InfoKits/infokit-related-files/logical- framework-information.

———. 2004. *Logical Framework (LogFRAME) Methodology.* http://www.jiscinfonet.ac.uk/InfoKits/project-management/InfoKits/infokit-related-files/logical- framework-information.

Kellogg Foundation. 2004. *Evaluation Logic Model Development Guide: Using Logic Models to Bring Together Planning, Evaluation, and Action.* http://www.wkkf.org./Pubs/Tools/Evaluation/Pub3669.pdf.

Knowledge and Research Programme on Disability and Healthcare Technology. 2004. *Constructing a Logical Framework.* http://www.kar-dht.org/logframe.html.

McCawley, Paul F. 1997. *The Logic Model for Program Planning and Evaluation.* University of Idaho Extension. http://www.uidaho.edu/extension/LogicModel.pdf.

NORAD (Norwegian Agency for Development Cooperation). 1999. *The Logical Framework Approach: Handbook for Objectives-Oriented Planning.* 4th ed. http://www.norad.no/default.asp?V_ITEM_ID=1069.

PPM&E Resource. 2005. *Logical Framework Approach.* portals.wdi.wur.nl/ppme/ index.php?Logical_Framework_Approach.

Rugh, Jim. "The Rosetta Stone of Logical Frameworks." http://www.mande.co.uk/ docs/Rosettastone.doc.

SIDA (Swedish International Development Agency). 2004. *The Logical Framework Approach: A Summary of the Theory behind the LFA Method.* http://www.sida.se/shared/jsp/download.jsp?f=SIDA1489en_web.pdf&a=2379.

Usable Knowledge, USA. 2006. Online Logic Model Training: An Audiovisual Pre- sentation. http://www.usablellc.net/Logic percent20Modelpercent20percent 28Online percent29/Presentation_Files/index.html.

Weiss, Carol. 1972. *Evaluation Research: Methods for Assessing Program Effectiveness.* Englewood Cliffs, New Jersey: Prentice Hall.

Wikipedia. 2006. "Logical Framework Approach." http://en.wikipedia.org/wiki/ Logical_framework_approach.

Відгуки щодо логічної структури

Bakewell, Oliver, and Anne Garbutt. 2006. *The Use and Abuse of the Logical Framework Approach A Review of International Development NGOs' Experiences.*

Swedish International Development Agency. http://www.sida.se/shared/jsp/download.jsp?f=LFA-review.pdf&a=21025.

MISEREOR. 2005. *Methodological Critique and Indicator Systems.* http://www.misereor.org/index.php?id=4495.

Альтернативні версії логічної структури

Broughton, Bernard. 2005. *A Modified Logframe for Use in Humanitarian Emergen- cies.* http://www.mande.co.uk/docs/EmergencyLogframeBroughton.doc.

Shaikh, Partha Hefaz. 2005. *Intertwining Participation, Rights-Based Approach and Log Frame: A Way Forward in Monitoring and Evaluation for Rights-Based*

Work. http://www.mande.co.uk/docs/PIFA_Article_PDF.pdf. SIDA Civil Society Centre. 2006. *Logical Framework Approach, with an Appreciative Approach.* http://www.sida.se/shared/jsp/download.jsp?f=SIDA28355en_LFA_web.pdf&a=23355.

ГЛАВА 5

Розгляд підходу оцінювання

Якщо раніше розвиток фокусувався на проектах, то зараз він фокусується на програмах та політиках, при цьому особлива увага приділяється їх сталості. В роботі з цими ширшими і складнішими питаннями застосовується широкий спектр підходів до планування й проведення оцінювання. Деякі з цих підходів розбираються в даній главі.

В цій главі розглядаються наступні теми:

- Загальні підходи до оцінювання
- Пов’язані з ними труднощі й переваги

Загальні підходи до оцінювання

Починаючи з 1990-х років міжнародна спільнота стала повільно рухатися від проектів до програм, партнерських підходів і допомоги розвитку (див. табл. 1.3). Партнерські підходи передбачають більшу кількість зацікавлених осіб, що активніше залучаються до складніших операцій – таких як загальносекторне фінансування. Відповідно, оцінювання великих програм проводиться у співпраці з партнерами, хоча об'єднані оцінювання ще не стали нормою. Нові заклики до урядів щодо спільного фінансування розвитку – наприклад ті, що були озвучені в 2008 році в Аккрі – означають, що в майбутньому перед фахівцями з оцінювання постануть серйозніші виклики.

У відповідь на зміни у природі оцінювання розвитку було розроблено широкий спектр підходів. Вибір підходу оцінювання частково залежить від контексту. Підходи не обов'язково є взаємовиключними, і оцінки можуть поєднувати елементи двох або більшої кількості підходів. Виділяють, зокрема, наступні підходи:

- перспективне оцінювання;
- визначення оцінюваності;
- оцінювання з орієнтацією на цілі;
- оцінювання без прив'язки до цілей;
- багатоцентрове оцінювання;
- кластерне оцінювання;
- соціальна оцінка;
- соціально-екологічна оцінка;
- спільне оцінювання;
- картування довгострокових результатів;
- швидка оцінка;
- синтезовані оцінювання й мета-оцінювання;
- інші підходи.

Який би підхід не було обрано, планування проводиться однаково: визначаються питання оцінювання, потім – засоби вимірювання, збираються й аналізуються дані, складаються звіти, застосовуються результати.

Перспективне оцінювання:
Оцінка ймовірних довгострокових результатів запропонованого проекту, програми або політики

Перспективне оцінювання

Перспективне оцінювання проводиться заздалегідь: запропоновану програму аналізують ще до її початку. Це робиться для того, щоб встановити ймовірність успіху, передбачити витрати і проаналізувати альтернативні пропозиції та плани. Перспективне оцінювання проводиться організаціями

оцінки із числа законодавчих органів. Прикладом такої організації є Рахункова палата США (GAO) – отримала таку назву в 2004 році – яка звітує перед Конгресом США. Фахівці з оцінювання GAO іноді допомагають урядовцям, відповідальним за прийняття рішень, надаючи аналітичну інформацію з різних питань та варіантів потенційних програм (U.S. GAO 1990).

Таблиця 5.1. Типи прогнозних питань та шляхи їх вирішення

Мета питання	Аналіз для критики інших	Проведення власного аналізу
Передбачення майбутнього	Наскільки добре адміністрація передбачила майбутні потреби, витрати та наслідки?	Якими є майбутні потреби, витрати та наслідки?
Вдосконалення майбутніх дій	Наскільки ймовірним є успіх адміністративної або законотворчої пропозиції?	Які дії мають найбільший потенціал для забезпечення успіху?

Джерело: U.S. GAO 1990.

GAO часто запитують про ймовірність успіху запропонованих нових програм. Вона розглядає інформацію про альтернативні пропозиції та аналізує результати аналогічних програм – поточних або вже завершених. У табл. 5.1 представлено чотири види прогнозних питань, які ставляться перед GAO.

Більшість перспективних оцінок передбачає ведення чотирьох видів діяльності:

- контекстуальний аналіз запропонованої програми або політики;
- аналіз досліджень оцінки аналогічних програм або політик та синтез результатів і досвіду минулого;
- прогнозування ймовірного успіху або невдачі за умови, що контекст у майбутньому не буде значно відрізнятися від минулого;
- надання рекомендацій щодо зміцнення запропонованої програми й політики в разі, якщо відповідальні особи прагнуть її розвивати (GAO 1990).

(Ресурси й приклади різних видів оцінки наводяться в списку веб-сайтів у кінці цієї глави).

Визначення оцінюваності

Визначення оцінюваності: Коротке попереднє дослідження, яке проводиться для того, щоб вирішити, чи корисною та практичною буде оцінка

Визначення оцінюваності – це коротке попереднє дослідження, яке проводиться для того, щоб вирішити, чи корисним та практичнним буде оцінювання. Цей вид попереднього дослідження допомагає роз'яснити цілі й завдання програми або проекту, визначити, які ресурси є в наявності, виявити прогалини у даних та визначити дані, що потребують розробки,

окреслтити коло ключових зацікавлених осіб та з'ясувати їх потреби в інформації. За результатами визначення оцінюваності можуть бути змінені цілі оцінки та методи її проведення. Розглядаючи інтервенцію в контексті практичної реалізації, враховуючи вибір часу та план оцінки, визначення оцінюваності може заощадити час та запобігти дорогим помилкам.

Хоулі та його колеги розробили підхід до визначення оцінюваності на початку 1970-х років. Вони вважали, що багато оцінювань зазнали невдачі через розходження між «риторикою і реальністю» (Nay and Kay 1982, p. 225), і прагнули це перевірити. Для них визначення оцінюваності було засобом покращення комунікацій між фахівцями з оцінювання та зацікавленими особами. Вони пропонували використовувати визначення оцінюваності, щоб дізнаватися, чи піддається програма оцінюванню, а також для концентрації зусиль фахівців з оцінювання (Fitzpatrick, Sanders, and Worthen 2004). Визначення оцінюваності створювалось як провісник сумативного оцінювання; з тих пір його роль посилилася: тепер вона включає роз'яснення цілей формативного дослідження або виконання функцій інструмента планування (Smith 1989).

Під час визначення оцінюваності фахівці з оцінювання з'ясовують, чи можна проводити оцінювання. Якщо, приміром, пропонується оцінювання з орієнтацією на цілі, на заваді може стати брак чітких програмних цілей або згоди серед зацікавлених осіб. Оцінювання, проведене до роз'яснення цілей та досягнення згоди, буде передчасним. Іноді засобів вимірювання немає або ж вони недоступні. В цьому випадку їх потрібно спочатку розробити, а потім вже проводити оцінювання інтервенції.

Визначення оцінюваності, таким чином, фокусується на практичності проведення оцінювання. Якщо за наявної інформації планувати оцінювання буде нераціонально або інтервенції бракує логічної теорії змін, необхідно провести додаткову підготовчу роботу. Визначення оцінюваності може допомогти роз'яснити інтервенцію й бажані результати. Це часто приводить до встановлення діалогу щодо цілей і завданнь, короткострокових результатів, цільової аудиторії та бажаних довгострокових результатів інтервенції. Такий діалог дозволяє дійти згоди щодо мети, якої потрібно досягти.

Визначення оцінюваності нерідко проводиться групою, до складу якої входять зацікавлені особи – наприклад спеціалісти з реалізації, адміністрування та оцінки. Для визначення оцінюваності група:

- аналізує матеріали, що визначають та описують інтервенцію;
- визначає модифікації інтервенції;
- проводить інтерв'ю з керівниками й співробітниками щодо їх сприйняття цілей і завданнь інтервенції;
- проводить інтерв'ю із зацікавленими особами щодо їх сприйняття цілей і завданнь інтервенції;

- розробляє або вдосконалює модель теорії змін;
- визначає джерела даних;
- визначає людей і організації, які можуть впроваджувати рекомендації, що видаватимуться за результатами оцінки.

Одна з потенційних вигід визначення оцінюваності – можливість забезпечити більш реалістичне й адекватне оцінювання. Сміт (Smith 1989) і Хоулі (Wholey 1987) також зазначають, що воно в змозі покращити:

- здатність відрізнити невдачу інтервенції від невдачі оцінювання;
- прогнозування довгострокових результатів;
- інвестування зацікавлених осіб в інтервенцію;
- ефективність виконання інтервенції;
- навички персоналу з розробки й оцінювання інтервенції;
- наочність інтервенції та відповідальне ставлення до її проведення;
- адміністративне розуміння інтервенції;
- політичний вибір;
- безперервну підтримку інтервенції.

Оцінювання з орієнтацією на цілі

Оцінювання з орієнтацією на цілі[4] (або з орієнтацією на завдання) визначає, наскільки програма або інтервенція виконує чітко окреслені і конкретні завдання (Patton 1990). Оцінювання зосереджується на заявлених довгострокових результатах (цілях або завданнях) проекту, програми або політики. Це типове оцінювання, основа систем оцінки проектів більшості організацій з розвитку.

Оцінювання з орієнтацією на цілі:
Оцінювання, що вимірює, якою мірою програма або втручання/інтервенція досягає чітких і специфічних завдань

Одним з недоліків, який закидають щодо підходу з орієнтацією на цілі, є його зосередженість на економічних та технічних, а не соціальних та гуманітарних аспектах (Hirschheim and Smithson 1988). Другий недолік – те, що подібні оцінки зосереджуються лише на заявлених цілях. Не оцінюється досягнення інших важливих цілей, які можуть бути скоріше опосередкованими, або ж які обговорювалися на засіданнях керівної ради чи підрозділу, але не були включені до переліку заявлених цілей проекту. Нарешті, третій недолік полягає в тому, що оцінювання з орієнтацією на цілі не розглядає непередбачені наслідки (ані позитивні, ані негативні).

Це може бути серйозним упущенням. Приміром, оцінювання нового водоочисного об’єкта може зосереджуватися на заявлених проектних завданнях побудови, експлуатації та обслуговування з метою виробництва певних щоденних об’ємів чистої води для задоволення потреб певної кількості домаш-

[4] *В оригіналі Goal-based evaluation.*

ніх господарств або громад. Наслідки реалізації проекту для людей, що проживають на цій території та мають бути переселені до початку будівництва, нерідко випадають з поля зору. Нездатність сформулювати ціль чи завдання щодо переселення таким чином, щоб забезпечити стабільні засоби існування – приклад серйозного упущення. Якщо оцінювання не ставить питання про переселення через те, що цей аспект не був включений до переліку формальних завдань проекту, проблема лише ускладнюється.

Оцінювання має виявляти неочікувані позитивні та негативні наслідки інтервенції. Навіть за відсутності якихось конкретних завдань фахівцям з оцінювання буде варто задокументувати зменшення захворюваності на хвороби, що передаються через воду, в громадах, що стали отримувати питну воду.

Оцінювання з орієнтацією на результат – метод, який пропагується у цій книзі – звертає увагу на результати незалежно від того, чи були вони означені як цілі або завдання. Оцінку з орієнтацією на цілі можна вдосконалити, забезпечивши відкритість до неочікуваних позитивних або негативних результатів.

Оцінювання без прив'язки до цілей

Оцінювання без прив'язки до цілей було розроблено у відповідь на застосування оцінювання з орієнтацією на цілі (або з орієнтацією на завдання). Майкл Скрайвен (Scriven 1972b), який першим запропонував оцінку без прив'язки до цілей, активно відстоює цей підхід.

Оцінювання без прив'язки до цілей:
Оцінювання, в ході якого фахівці з оцінювання навмисне намагаються виключити будь-які згадки щодо програмних цілей, орієнтуючись в своїй роботі винятково на потреби учасників

В ході **оцінювання без прив'язки до цілей** фахівець з оцінювання навмисне намагається уникати будь-яких згадок про цілі програми. Фахівці з оцінювання не обговорюють цілі з персоналом, не читають брошури чи пропозиції програми. Вони оцінюють лише наочні результати й документальні наслідки, що мають відношення до потреб учасників (Patton 2002). Як зазначає Скрайвен (Scriven 1972b, p. 2):

> Мені здалося, коротко кажучи, що брати до уваги й оцінювати цілі не потрібно, ба, більше – можливо, шкідливо… Чим менше зовнішній фахівець з оцінювання чує про цілі проекту, тим менш обмеженим буде поле його зору і тим більше уваги він приділятиме пошуку фактичних наслідків (не відволікаючись на перевірку наслідків очікуваних).

Оцінювання без прив'язки до цілей збирають дані щодо наслідків та дієвість програм, не обмежуючись вузькою перспективою цілей та завдань. В ході таких оцінювань береться до уваги фактичний досвід учасників програми, викладений їх власними словами. Фахівці з оцінювання повинні ігнорувати те, чого програма намагається досягти, зосередившись на тому, що власне відбувається. Через це даний вид оцінки дуже добре поєднується з якісними дослідженнями, хоча може застосовувати і кількісні, і якісні методи.

Скрайвен (Scriven 1997) пропонує проводити окремі оцінки з орієнтацією на результат задля максимізації сильних сторін та мінімізації недоліків кожного з підходів.

Хоулі, Харті та Ньюкомер (Wholey, Harty, and Newcomer 1994) описують наступні характеристики оцінки без прив'язки до цілей:

- Фахівець з оцінювання уникає інформації про цілі програми.
- Запрограмованим цілям не дозволяється звузити поле оцінювання.
- Фахівець з оцінювання мінімально контактує з керівником та персоналом програми.
- В центрі уваги – фактичні, а не цільові результати програми.

Зазвичай у контексті розвитку важко уникати знання про цілі й завдання інтервенції. Зведення до мінімуму контактів між фахівцем з оцінювання і персоналом програми може бути нераціональним і небажаним. Тим не менш, використовуючи підхід з орієнтацією на результат, фахівець з оцінювання може застосовувати багато елементів оцінки без прив'язки до цілей.

Багатоцентрове оцінювання

Коли мова йде про широкомасштабну інтервенцію, часто буває необхідно розглядати інтервенції, що реалізуються одночасно у різних місцях. Оцінювання такого типу називають **багатоцентровим**. Інтервенція може реалізовуватися всюди однаково або ж із деякими відмінностями. Багатоцентрове оцінювання дає інформацію про загальний досвід інтервенції, а також глибше розуміння відмінностей, що мали місце. Воно може давати відповідь, наприклад, на такі запитання:

Багатоцентрове оцінювання:
Оцінювання, що розглядає інтервенції, які реалізуються у різних місцях

- Які риси інтервенції зустрічаються в усіх місцях її реалізації?
- Які риси є відмінними і чому?
- Чи відрізняються результати залежно від таких відмінностей?

Ключовим моментом є отримання глибинної інформації. Для цього фахівці з оцінювання часто використовують під час багатоцентрового оцінювання вивчення практичних прикладів. Місця для вивчення зазвичай обираються, тому що представляють певні характеристики (наприклад, розмір, етнічну приналежність, соціально-економічний статус), здатні обумовлювати системні відмінності в реалізації й результатах інтервенції. Звичайно, може бути важко визначити, чи зумовлені відмінності в результатах саме відмінностями в інтервенціях. В певних випадках інтервенції демонструють впливи завдяки унікальним відмінностям вихідних умов – наприклад, завдяки ефективному керівництву або активності громадян. В інших випадках зміни пояснюються систематичними відмінностями – наприклад регіо-

нального характеру. Такі відмінності можуть накладати відбиток на рішення щодо відтворення інтервенції деінде.

Багатоцентрове оцінювання повинно враховувати клімат, в якому відбувається інтервенція, а також будь-які культурні, географічні, економічні, масштабні та інші систематичні відмінності, що можуть стати визначальними для процесів та результатів. Важливою є участь зацікавлених осіб, оскільки вони можуть допомогти фахівцю з оцінювання краще зрозуміти місцеві умови.

Багатоцентрове оцінювання зазвичай є потужнішим, ніж оцінка однієї інтервенції в одній місцевості. Воно дає змогу з більшою достовірністю робити висновки для більшої кількості населення, оскільки використовує більшу вибірку й більш різноманітний набір ситуацій інтервенції. Тут застосовуються як внутрішній, так і порівняльний аналіз. Загальні результати оцінювання, а також висновки, зроблені за результатами спостережень за різними інтервенціями, є надійнішим свідченням дієвості інтервенції, ніж дані, отримані лише в одному місці.

Порівняння інтервенцій у різних місцях може дати широкий спектр висновків з досвіду й стратегій вирішення різнопланових проблем. Багатоцентрове оцінювання допомагає також збирати передові практики. Важливо, однак, пам'ятати, що місця для проведення оцінки, обрані за рішенням фахівця з оцінювання (навіть якщо при цьому враховувалася необхідність представлення певних характеристик), не є статистично репрезентативними для населення, і зовсім не обов'язково вони дадуть відомості про всі передові практики.

Фахівці з оцінювання, які проводять багатоцентрове оцінювання, повинні забезпечити максимальну стандартизацію збору даних. Щоб дані були порівнювані, вони повинні всюди збиратися однаково. Збирати їх мають добре підготовлені працівники. Скрізь мають бути доступними однакові види даних. Необхідно заздалегідь отримати достатньо інформації, щоб розробити інструментарій для збору даних.

Кожне місце проведення оцінювання є різним. Деякі показники можуть бути порівнювані, однак у кожному місці в центрі уваги може бути дещо інше. Проектну інтервенцію і, відповідно, оцінку можуть визначати політичні, соціальні, економічні й історичні контексти (Johnston and Stout 1999).

Кластерне оцінювання

Кластерне оцінювання: Оцінювання, що розглядає групи аналогічних або споріднених інтервенцій

Кластерне оцінювання є аналогічнним багатоцентровому, але цілі його відмінні. Зазвичай воно розглядає групи схожих або споріднених інтервенцій, що разом представляють набір інтервенцій. Цей «кластер» інтервенцій розглядається в одному або кількох варіантах обстановки. Як і багатоцентрове, кластерне оцінювання зосереджується на інтервенціях зі спільною місією, стратегією та цільовою аудиторією. На відміну від багатоцентрового оцінювання, кластерне не покликане визначати, чи працює інтервенція, або ж

забезпечувати відповідальність. Воно не оцінює успіх чи невдачу окремих інтервенцій та не визначає, які з них мають бути припинені. Його мета – встановити, що відбулося у кластерах, та перевірити спільні теми й висновки з досвіду. Звітна інформація подається лише в узагальненому вигляді, жоден з проектів не ідентифікується. Як і у випадку з багатоцентровим оцінюванням, участь зацікавлених осіб відіграє ключову роль.

Кластерні оцінювання відрізняються від багатоцентрових тим, що для них не важливими є узагальнюваність та відтворюваність. Відмінності сприймаються позитивно, адже окремі проекти пристосовуються до своїх контекстів. Кластерне оцінювання більше зосереджується на засвоєнні досвіду, ніж на формуванні загальних висновків про якість чи цінність програми.

Хоча спеціальної методології проведення кластерного оцінювання не існує, оцінки такого виду часто використовують якісні підходи для підкріплення кількісних даних. Кластерне оцінювання можна сприймати як вивчення одразу багатьох практичних прикладів з обміном інформацією через мережу конференцій – ось показова характеристика даного підходу.

Недоліком кластерного оцінювання є те, що воно не демонструє результати для окремих місць та не бере до уваги заплановані чи незаплановані відмінності. Дані представляють лише узагальнену інформацію.

Соціальна оцінка

Соціальна оцінка стала важливою складовою багатьох оцінок. Соціальна оцінка розглядає соціальні структури, процеси та зміни в межах групи або спільноти. Вона також може розглядати тенденції, що можуть впливати на групу.

Соціальна оцінка:
Оцінка, що розглядає соціальні структури, процеси та зміни в межах групи або спільноти

Соціальна оцінка – основний інструмент, що використовується для врахування соціальних впливів проектів, програм та політик розвитку. Вона використовується для з'ясування ключових соціальних проблем і ризиків, а також визначення соціальних впливів інтервенції на різних зацікавлених осіб. Зокрема, соціальні оцінки покликані визначати, чи не призведе проект до негативних соціальних наслідків (наприклад, переселення людей для звільнення місця під електростанцію). Виявлені та визнані негативні впливи можна пом'якшити, запроваджуючи відповідні стратегії; надалі такі стратегії можна контролювати й оцінювати в рамках оцінки.

У «Збірнику матеріалів з розвитку Світового банку» (*World Bank Participation Sourcebook* (1996)) подаються наступні цілі соціальної оцінки:

- Визначити коло ключових зацікавлених осіб та запровадити належну схему їх участі у виборі, плануванні та реалізації проекту.
- Забезпечити прийнятність завдань проекту і стимулів змін для тих людей, які мають отримати вигоду від проекту, а також врахування гендерних та інших соціальних відмінностей у плані проекту.

- Оцінити соціальний вплив інвестиційних проектів; у випадку виявлення негативних впливів визначити, яким чином їх можна подолати або принаймні пом'якшити.
- Досягти належного рівня забезпечення участі, розв'язання конфліктів, надання послуг та проведення необхідних заходів з пом'якшення негативних впливів.

«Збірник» визначає наступні спільні запитання, що ставляться в ході соціальної оцінки:

- Хто є зацікавленими особами? Чи відповідають завдання проекту їх потребам, інтересам та можливостям?
- Які соціальні й культурні фактори впливають на здатність зацікавлених осіб брати участь або отримувати вигоди від запропонованих операцій?
- Як впливає проект або програма на різних зацікавлених осіб, зокрема, на жінок та вразливі групи населення? Які соціальні ризики (брак бажання або потенціалу та несумісність з існуючими умовами) можуть негативно вплинути на успішність проекту або програми?
- Які інституційні приготування потрібні для участі й реалізації проекту? Чи існують адекватні плани нарощування необхідного потенціалу?

Інструменти та підходи до оцінювання:

- аналіз зацікавлених осіб;
- гендерний аналіз;
- спільна оцінка вартості сільськогосподарського майна;
- спостереження, інтерв'ю, фокус-групи;
- картування, аналіз завдань та ранжування за рівнем достатку;
- семінари, присвячені плануванню проектів з орієнтацією на завдання.

Приклади основних показників для моніторингу соціального впливу включають, зокрема, наступне:

- рівень участі різних соціальних груп у добровільному тестуванні;
- рівень участі різних соціальних груп в консультаційних заходах;
- звіти про збільшене використання презервативів;
- відсоток членів громади, що беруть участь в догляді за людьми з ВІЛ/СНІД та членами їх сімей.

У блоці 5.1 наводиться приклад застосування соціальної оцінки в рамках проекту.

Блок 5.1 «Соціальна оцінка проекту розвитку й кредитування сільського господарства Азербайджану»

Для відновлення продуктивності селянських господарств Азербайджану було запроваджено проект приватизації селянських господарств, що забезпечував більш гнучку видачу кредитів. Проект зосереджувався на реєстрації нерухомості, розвиткові земельних ринків та наданні кредитів й інформації сільському населенню, особливо людям з низькими доходами.

Метою соціальної оцінки було переконатися, що запропонована інтервенція ґрунтується на відповідальному ставленні зацікавлених осіб, а також що очікувані вигоди є соціально прийнятними. Отримана за результатами оцінки інформація допомогла керівникам програми спланувати процес спільного моніторингу й оцінки.

Перший етап соціальної оцінки складався з наступних компонентів:

- аналіз вторинних даних, включаючи попередні оцінки;
- дослідження домашніх господарств і жінок в трьох з шести регіонів після швидкої оцінки, проведеної із застосуванням якісних методів;
- напівструктуровані персональні інтерв'ю з фермерами, керівниками ферм, безробітними, лідерами громад, групами жінок, місцевими об'єднаннями, техніками, чиновниками;
- спостереження персоналом на місці (член групи жив разом з сім'єю селян для ведення безпосереднього спостереження впливу приватизації ферм);
- п'ять фокус-груп з однорідними групами зацікавлених осіб;
- консультації з особами, що формують політику, та адміністраторами, а також місцевими й міжнародними НУО;
- бесіди з екс-керівниками державних господарств і лідерами громад;
- семінар для зацікавлених осіб.

Оцінка будувалася на чотирьох стовпах:

- Соціальний розвиток: Головна увага приділялася питанням бідності, гендеру й соціального відчуження.
- Інститути: Політична підтримка в сільських районах була нестійкою, що ускладнювало визначення ключових зацікавлених осіб. Було мало інформації стосовно соціальних організацій, бракувало аналізу впливів сільської міграції.
- Участь: Повідомлялось про плутанину й неоднозначності в процесі земельних реформ. Розподіл землі скоротив рівень бідності, обмежив вплив екс-керівників господарств та допоміг селянам отримати нові можливості. Доступ до кредитів покращився, але ставки за ними залишалися високими (15–18 відсотків).
- Моніторинг/оцінка: Для моніторингу реалізації використовувалися показники ефективності. Показники пов'язували залучені ресурси й заходи проекту з кількісно вимірними результатами й впливами.

Оцінка також розглядала впливи у формі збільшення продуктивності, підвищення доходів, скорочення бідності й задоволення учасників.

Джерело: Kudat and Ozbilgin 1999.

Соціально-екологічна оцінка

Організації з розвитку все частіше визнають потребу в тому, щоб програми та проекти вирішували соціально-екологічні проблеми з відповідною оцінкою виконання завдань, що лежать в соціально-екологічній площині. Більшість організацій з розвитку дотримуються набору основних соціально-екологічних стандартів та визначають потенційні соціально-екологічні впливи як чинник процесу затвердження програми або проекту. В разі затвердження проекту або програми ці потенційні впливи контролюються на стадії реалізації та оцінюються по завершенні програми або проекту. Оцінки соціальних та екологічних факторів при цьому розглядаються як такі, що мають проводитися одночасно.

Зараз організації з розвитку визнають роль, яку місцеве населення має відігравати у плануванні, реалізації й моніторингу інтервенцій, що можуть позначитися на стані довкілля та природних ресурсів. Місцеве населення й інші зацікавлені особи є партнерами у збереженні природних ресурсів і управлінні ними.

Соціально-екологічна оцінка може бути єдиною метою оцінювання, або ж лише одним з його компонентів. Соціально-екологічну оцінку слід проводити не лише в проектах, що явно впливають на довкілля (побудова паперових заводів, проклання нафтопроводів), а й на такі інтервенції, як будування нової школи чи відкриття кредитної лінії.

Якщо в організації немає соціально-екологічних стандартів, фахівці з оцінювання можуть використовувати стандарти країни, де працює організація, Принципи екватора, ISO 14031, збірку «Стратегії сталого розвитку» *(Development Strategies: A Resource Book (Dalal-Clayton and Ba 2002)).*

Принципи екватора

Принципи екватора:
Галузевий підхід до визначення і оцінки екологічних та соціальних ризиків фінансування проектів приватного сектору та керування такими ризиками

Принципи екватора – набір принципів, що допомагають фінансовим інститутам проводити визначення, оцінку й роботу з соціально-екологічними ризиками при фінансуванні проектів із сукупними витратами капіталу від 10 млн доларів США (в 2006 році планку було знижено з 50 млн). Ці принципи покликані слугувати загальним знаменником та структурою для впровадження індивідуальних, внутрішніх, екологічних та соціальних процедур і стандартів для проектів розвитку.

ISO 14031

ISO 14031:
Набір міжнародних стандартів екологічного менеджменту, розроблених Міжнародною організацією зі стандартизації

Міжнародна організація зі стандартизації (ISO) розробила й підтримує міжнародні стандарти екологічного менеджменту, відомі як **ISO 14031**. Цей стандарт, вперше запровадженний в 1999 році, є процедурою внутрішнього управління й інструментом, покликаним надавати керівництву надійну та контрольовану поточну інформацію. Він допомагає визначати, чи відповіда-

ють екологічні показники організації критеріям, встановленим керівництвом організації. Оцінка екологічних показників і екологічні аудити допомагають керівникам оцінювати стан екологічної ефективності та визначати напрями вдосконалення (ISO 1999).

Оцінка екологічної ефективності допомагає, встановлюючи порядок:

- вибору показників;
- збору й аналізу даних;
- оцінки інформації за критеріями екологічності (цілі);
- звітування й комунікацій;
- періодичного перегляду й вдосконалення процесу.

Збірка «Стратегії сталого розвитку»

Збірка, видана Організацією економічного співробітництва й розвитку (ОЕСР) і Програмою розвитку Організації об'єднаних націй (ПРООН), містить гнучкі, непримусові настанови з розробки, оцінки та впровадження національних стратегій сталого розвитку відповідно до принципів, викладених в керівництвах зі стратегій сталого розвитку (Dalal-Clayton and Ba 2002). Вона містить ідеї й практичні приклади стосовно основних завдань у стратегічних процесах. Наведені настанови призначені для країн, організацій та окремих осіб, зацікавлених у сталому розвитку на національному або місцевому рівні, а також для міжнародних організацій, що займаються підтримкою такого розвитку.

Оцінювання із залученням зацікавлених сторін

Оцінювання із залученням зацікавлених сторін / Партисипативне оцінювання просуває залучення зацікавлених осіб на якісно новий рівень. Воно передбачає розподіл обов'язків з планування, реалізації й звітування, залучаючи зацікавлених осіб до визначення питань оцінки, збору й аналізу даних, а також складання й перевірки звіту.

Оцінювання із залученням зацікавлених сторін:
Оціннювання, в ході якого повноваження з планування, реалізації і звітування надаються також зацікавленим особам, які можуть допомагати визначати питання оцінки, збирати й аналізувати дані, складати й перевіряти звіт

Полмер (Paulmer (2005, p. 19)) описує оцінювання із залученням зацікавлених сторін як:

> колективну оцінку програми зацікавленими особами й бенефіціарами. Такі оцінки орієнтуються на дії, зміцнюють потенціал зацікавлених осіб та сприяють співпраці й спільному прийняттю рішень з метою активізації застосування результатів оцінки. Існують різні рівні участі бенефіціарів у оцінці.

Перед цим підходом ставлять два основних завдання:

- участь як продукт, де факт участі є метою і одним з показників успіху;
- участь як процес для досягнення заданої мети.

Петон (Patton 1997) наводить наступні базові принципи спільного оцінювання:

- У процесі оцінювання використовуються навички учасників у встановленні цілей, визначенні пріоритетів, виборі питань, аналізі даних та прийнятті рішень на основі даних.
- Учасники відповідально ставляться до оцінки, адже вони приймають рішення й роблять власні висновки.
- Учасники забезпечують зосередженість оцінки на методах і результатах, які вони вважають важливими.
- Люди працюють разом, розвиваючи й підтримуючи групову єдність.
- Всі аспекти оцінки є для учасників зрозумілими й змістовними.
- Високо цінується самостійність в поєднанні з відповідальністю.
- Координатори виступають в якості ресурсів для навчання, учасники – в якості осіб, що приймають рішення, та фахівців з оцінювання.

В ході традиційних досліджень було створено та задокументовано керівництва з конкретних технік з метою підвищення надійності й адекватності даних. Спільне оцінювавання не проводиться за чітко окресленими правилами, що спускаються експертами збирачам даних. Замість цього керівництва зі збору даних розробляються й еволюціонують завдяки консенсусу, осмисленню, діалогу й досвіду (Narayan 1996).

Такий підхід, як оцінювання із залученням зацікавлених сторін, користується в контексті розвитку все більшою увагою, зокрема для проектів розвитку, особливо для ініціатив спільнот. Оцінювання із залученням зацікавлених сторін – це ще один крок від моделі незалежної оцінки або експерта в ролі фахівця з оцінювання.

Спільний підхід визначає, а потім залучає людей, агенції, організації, що мають інтерес у певній сфері. Говорячи про людей, ми маємо на увазі передусім тих, хто відноситься до маргіналізованих груп. Також йдеться про персонал агенцій, осіб, які формують політику, а також всіх, кого стосуються рішення, що приймаються в ході спільного дослідження (Narayan 1996).

Під час спільного оцінювання зацікавлених осіб можуть запропонувати вести щоденники для фіксації вражень про інтервенцію, які можуть допомогти при проведенні інтерв'ю в громаді, а також участі в аналізі даних, інтерпретації результатів та розробці рекомендацій.

Рішення з планування – такі як визначення питань, засобів контролю й стратегій збору даних – приймаються разом з учасниками. Це спільний, горизонтальний процес, на відміну від традиційного, спрямованого зверху вниз (див. табл. 5.2).

Спільний підхід зазвичай підвищує достовірність результатів оцінки в очах персоналу програми, так само як і ймовірність того, що результати бу-

дуть застосовані на практиці. Прибічники спільного оцінювання розглядають його як інструмент для надання учасникам нових можливостей та нарощування місцевого потенціалу для участі в процесі розвитку.

Таблиця 5.2. Риси спільного та традиційного оцінювання

Оціннювання із залученням зацікавлених сторін	Традиційне оцінювання
Центральна фігура – учасник	Центральна фігура – донор
В центрі уваги – навчання	В центрі уваги – відповідальність і винесення суджень
Гнучкий план	Заздалегідь складений план
Більш неформальні методи	Формальні методи
Сторонні координатори	Сторонні фахівці з оцінювання

Джерело: Автори.

Оцінювання із залученням зацікавлених сторін створює серйозні виклики. Воно передбачає вищі трансакційні витрати порівняно з традиційною оцінкою, тому що проведення зборів і донесення очікувань до кожного в зрозумілому вигляді – трудомістка робота, що потребує високої кваліфікації. Перш ніж перейти до виконання нагальних завдань, групи зазвичай проходять через процес врегулювання розбіжностей і формування групових норм. Цей процес групової динаміки іноді називають «формування, буксування, нормування й виконання». Після формування природним чином виникає період конфліктів (буксування). Якщо групі вдається їх подолати, вона встановлює більш конкретні домовленості стосовно того, як члени групи мають співпрацювати (нормування). Після встановлення нових домовленостей учасники переходять до роботи з актуальними завданнями (виконання).

Проблеми також можуть виникати при створенні групи рівноправних учасників у культурі, де різні члени спільнот мають в них різний статус. Фахівець з оцінювання, що бажає провести оцінювання із залученням зацікавлених сторін, повинен вміти залагоджувати конфлікти, координувати та співпрацювати (або передати повноваження особі, що має потрібні здібності). Фахівець з оцінювання також має бути здатний забезпечувати своєчасне навчання базовим навичкам та технікам, пов’язаним з оцінкою та груповими процесами, характерними для цього підходу.

Той, хто навчений традиційному підходу до оцінки і проводить її в традиційний спосіб, може непокоїтися з приводу того, що оцінювання із залученням зацікавлених сторін не буде об’єктивним. Існує ризик того, що най-

тісніше пов'язані з інтервенцією особи не зможуть помічати, що власне відбувається, якщо це не те, що вони очікують побачити. Оцінка і справді може стати упередженою, втратити об'єктивність. Учасники можуть побоюватися критикувати щось через страх відторгнення іншими членами групи або зупинення інтервенції, що призвело б до втрати громадою грошей або ж до того, що вони ніколи б більше не змогли долучитися до роботи організації з розвитку. Підхід до спільного оцінювання з точки зору навчання може допомогти побороти ці страхи. Фахівцям з оцінювання варто враховувати ступінь дискредитації достовірності оцінки (на сторонній погляд) через вибір спільної, а не незалежної оцінки.

Гаріба (Gariba (1998)) описує, наскільки різну реакцію слово «оцінювання» викликає серед донорів та реалізаторів. Донорів може непокоїти, як оцінювання вплине на проект (тобто, воно може призвести до його розширення або зупинення). Реалізатори проекту можуть побоюватися, що оцінювання представить в невигідному світлі їх підходи до управління проектом.

Гаріба описує, як оцінювання із залученням зацікавлених сторін може стати успішним і систематичним способом робити висновки з досвіду. В ході спільного оцінювання партнери в інтервенції розвитку можуть робити висновки зі своєї взаємодії та вживати корективних заходів для підвищення дієвості або ефективності своїх поточних та майбутніх дій.

Гаріба (Gariba 1998) описує три критичних елементи спільного оцінювання:

- *О*цінювання *як інструмент навчання*: Цей принцип формує головну парадигму вибору. Призначення – не розбиратися у наслідках, а створювати для всіх зацікавлених осіб, включаючи донорів, можливість навчатися нового, виконуючи свої ролі в інтервенції розвитку.
- *О*цінювання *як частина процесу розвитку:* Оцінку не можна відокремити від самого процесу розвитку. Результати й відповідні інструменти стають інструментами змін, а не хронологічними звітами.
- *О*цінювання *як партнерство і розділення відповідальності:* Методологія спільного оцінювання впливів передбачає більш-менш рівну вагу для всіх діючих осіб (це різко розходиться з тенденцією, коли фахівці з оцінювання поділяють всіх на «ми»-професіонали й «вони»-діючі особи й бенефіціари проекту). Фахівець із оцінювання з дослідника перетворюється на організатора й учасника.

Згідно з посібником, виданим Канадським агентством міжнародного розвитку (Canadian International Development Agency Guide (CIDA 2004)), якщо зацікавлені особи беруть участь в розробці результатів, вони з більшою імовірністю братимуть участь у реалізації інтервенції. CIDA вважає, що оцінювання із залученням зацікавлених сторін також:

- формує відповідальне ставлення всередині спільнот;
- надає оцінці більш реалістичної орієнтації;
- активізує співробітництво;
- надає місцевим учасникам нові можливості, залучаючи їх до процесу оцінювання.

В блоці 5.2 наведено приклад однієї з технік, що використовуються у спільному оцінюванні.

Блок 5.2 «Побудова довіри за допомогою партисипативного оцінювання»

В Марокко координатор групи формувала довіру серед жінок. Вона попросила їх стати в коло й узятися за руки, потім заплутатися, не відпускаючи рук. Далі «стороння» скомандувала групі розплутатися, знову утворивши коло – на це пішло шість хвилин. Потім вона попросила групу повторити вправу. З другого разу на те, щоб розплутатися, у жінок пішло 10 секунд.

Координатор допомогла групі досягти згоди стосовно ролі «сторонніх» осіб: сторонні особи зазвичай є більш ефективними в ролі координаторів і каталізаторів, ніж в ролі лідерів. Люди на місцях краще знають, як знаходити рішення своїх проблем, тому що живуть з ними.

Джерело: World Bank 1996.

Картування довгострокових результатів/Наслідків

Канадський Міжнародний дослідницький центр з питань розвитку (IDRC) розробив інноваційний підхід до оцінювання. Картування довгострокових результатів не замінює більш консервативні форми оцінювання, а доповнює їх, зосереджуючись на супутній зміні поведінки.

Картування довгострокових результатів/наслідків зосереджено на одному конкретному виді результату – змінах поведінки. В центрі уваги – довгострокові результати, а не досягнення впливів розвитку, яке вважається занадто віддаленим і відображає результати багатьох заходів та інтервенцій. Спроби точно оцінити внесок однієї окремої організації IDRC вважає марними. Замість цього під час картування довгострокових результатів увага зосереджується на поведінці, що визначається зусиллями багатьох сторін, щоб допомогти покращити ефективність проектів, програм і політик.

Картування довгострокових результатів/ Наслідків:
Картування змін поведінки

В ході картування довгострокових результатів визначаються суміжні партнери – окремі особи, групи й організації, що взаємодіють з проектами, програмами й політиками. Картування довгострокових результатів передба-

чає контроль суміжними партнерами змін. Воно також передбачає, що саме їх роль зовнішнього агента дає їм доступ до нових ресурсів, ідей або можливостей у певний період часу. За думкою прибічників картування довгострокових результатів, найуспішніші програми – це ті, що делегують повноваження й відповідальність особам, що діють в рамках проекту або програми.

Саме люди знаходяться в центрі уваги картування довгострокових результатів. Це відхід від оцінки впливу проекту або програми на розвиток до описання змін у поведінці людей – індивідуальній, у складі груп або організацій. Картування довгострокових результатів дає спосіб змоделювати те, що має на меті програма. Така модель відрізняється від більшості традиційних логічних моделей завдяки тому, що визнає: різні суміжні партнери діють в інших системах логіки й обов'язків.

Картування довгострокових результатів пропонує спосіб моніторингу змін у суміжних партнерів та у програмі як організації. Воно заохочує регулярну оцінку програмою шляхів удосконалення власної ефективності. Картування довгострокових результатів також можна використовувати як інструмент оцінки по завершенні програми, коли метою оцінювання є вивчення програми в цілому.

Прибічники картування довгострокових результатів вважають, що багато програм, особливо ті з них, що зосереджуються на зміцненні потенціалу, зможуть краще планувати й оцінювати власні внески у розвиток, якщо зосередять свою увагу на оцінці. Наприклад, метою програми може бути забезпечення громад доступом до чистішої води шляхом встановлення очисних фільтрів. Традиційний метод оцінювання передбачав би оцінку результатів, наприклад, шляхом підрахунку кількості встановлених фільтрів та вимірювання змін концентрації у воді забруднюючих речовин, що сталися після встановлення фільтрів. Картування довгострокових результатів розглядало б поведінку. Для початку було б зроблено припущення, що вода не залишається чистою, якщо люди не підтримують її якість. Потім було б оцінено довгострокові результати програми. Для цього вивчалася б поведінка осіб, що відповідають за якість води – зміни в тому, як вони оволодівають відповідними інструментами, навичками і знаннями та як їх застосовують. Картування довгострокових результатів оцінювало б, як люди контролюють рівень забрудненості, змінюють фільтри або запрошують в разі потреби спеціалістів.

Деякі проблеми, що їх намагається вирішити картування довгострокових результатів, викладені у пісні про цей підхід, написаній Террі Смутило, колишнім директором з оцінювання Міжнародного дослідницького центру з питань розвитку. (Посилання на запис, де Террі Смутило виконує свою пісню *The Output Outcome Downstream Impact Blues* наведено у списку вебсайтів у кінці цієї глави.)

The Output Outcome Downstream Impact Blues
Outputs, Outcomes, Impacts : For Whom, by Whom, Says Who?
Coda
Don't look for *impact* with attribution (4×)
Well there's a nasty little word getting too much use
In development programs it's prone to abuse
It's becoming an obsession now we're all in the act
Because survival depends on that elusive *impact.*

REFRAIN I

Because it's *impact* any place, *impact* any time
Well you may find it 'round the corner or much farther down the line
But if it happens in a way that you did not choose
You get those *Output Outcome Downstream Impact Blues.*
Now when donors look for *impact* what they really wanna see
Is a pretty little picture of their fantasy
Now this is something that a good evaluator would never do
Use a word like *impact* without thinking it through.
But now donors often say this is a fact
Get out there and show us your *impact*
You've got to change peoples' lives and help us take the credit
Or next time you want funding – huh hmm
You may not just get it.

REFRAIN I

Because it's *impact* any place, *impact* any time
Well you can find it 'round the corner or much farther down the line
But if it happens in a way that you did not choose
You get those *Output Outcome Downstream Impact Blues.*
Well recipients are always very eager to please
When we send our evaluators overseas
To search for indicators of measurable *impact*
Surprising the donors what, what they bring back.
Well, *impact* they find when it does occur
Comes from many factors and we're not sure
Just what we can attribute to whom
Cause *impact* is the product of what many people do.

Гіркі думи про багатофакторний вплив і результати
Короткострокові результати, довгострокові результати, впливи: для кого, хто робить, хто замовляє?
Кода
Не намагайтесь визначити *вплив* лише когось одного (4×)
Це слово вживають занадто часто
А в програмах розвитку ним зловживають
Усі ним одержимі, і ми також,
Тому що наше виживання залежить від цього невловимого *впливу*.

ПРИСПІВ 1

Тому що *вплив* усюди, *вплив* завжди
Він може чекати за рогом, а може, вам прийдеться за ним побігати
Але якщо ви на нього натрапите, вибору немає:
Вас очікують гіркі думи про багатофакторний вплив і результати
Коли донори питають про *вплив*, насправді вони хочуть побачити
Милу картинку з їхніх мрій
І ось що розумний фахівець з оцінювання ніколи не робить:
Він не каже «*вплив*», добре не подумавши.
Але донори постійно кажуть нам:
«Давайте, покажіть нам свій *вплив*,
Ви маєте змінити життя людей і прославити нас,
Інакше наступного разу, коли вам знадобляться гроші,
Ви можете їх не отримати».

ПРИСПІВ 1

Тому що *вплив* усюди, *вплив* завжди
Він може чекати за рогом, а може, вам прийдеться за ним побігати
Але якщо ви на нього натрапите, вибору немає:
Вас очікують *гіркі думи про багатофакторний вплив і результати*
Реципієнти завжди добре зустрічають
Наших фахівців з оцінювання, котрих ми посилаємо
Шукати показники вимірного *впливу*
І дивувати донорів своїми знахідками.
Виявляється, що *вплив*, коли він є,
Забезпечують багато факторів, і ми не знаємо напевне,
Який з них що забезпечив.
Тому що *вплив* – це результат роботи багатьох людей.

REFRAIN II

Because it's *impact* any place, *impact* any time
Well you can find it 'round the corner or much farther down the line
But if you look for attribution you're never going to lose
Those *Output Outcome Downstream Impact Blues.*
So donors wake up from your impossible dream
You drop in your funding a long way upstream
Then in the waters they flow, they mingle, they blend
So how can you take credit for what comes
out in the end.

REFRAIN II

Because it's *impact* any place, *impact* any time
Well you can find it 'round the corner or much
farther down the line
But if you look for attribution you're never
going to lose
Those *Output Outcome Downstream Impact Blues.*
Coda (4×then fade)

ПРИСПІВ 2

Тому що *вплив* усюди, *вплив* завжди
Він може чекати за рогом, а може, вам прийдеться
за ним побігати
Але якщо ви бажаєте визначити, хто його створив,
Гіркі думи про багатофакторний вплив і результати
будуть з вами завжди
Отож, донори, облиште свої марні сни.
Ви даєте свої гроші,
Але не ви одні,
Тож ви не можете приписати всі заслуги собі.

ПРИСПІВ 2

Тому що *вплив* усюди, *вплив* завжди
Він може чекати за рогом, а може, вам прийдеться
за ним побігати
Але якщо ви бажаєте визначити, хто його створив,
Гіркі думи про багатофакторний вплив і результати
будуть з вами завжди
Кода (4×)

Швидка оцінка

Цей підхід застосовується, коли оцінку треба провести швидко й недорого. Він особливо корисний в контексті оцінки країн, що розвиваються, де бракує часу, а обмежений доступ до інших ресурсів часто ускладнює проведення більш ретельного дослідження: немає базових даних, низький рівень грамотності вимагає участі працівників у заповненні анкет, бюджети оцінки обмежені (і це лише частина проблем).

Швидка оцінка:
Систематичний, напівструктурований підхід до оцінки, що проводиться в польових умовах, зазвичай силами групи фахівців з оцінювання

Хоча усталеного визначення **швидкої оцінки** не існує, її зазвичай описують як місток між формальним та неформальним збором даних – «досить швидкий і досить чистий» підхід замість «швидкого й брудного». Швидку оцінку можна охарактеризувати як систематичний, напівструктурований підхід, що реалізується в польових умовах, зазвичай силами групи фахівців з оцінювання. В ідеалі ця група складається з різнопланових спеціалістів, що дозволяє враховувати різні точки зору.

Швидка оцінка є найбільш застосовною при розгляді процесів, а не довгострокових результатів та впливів. Зазвичай для неї збирають лише найважливішу інформацію – те, що знати не просто добре, а необхідно. Це інформація, отримана за допомогою як кількісних, так і якісних методів. Основна орієнтація цього підходу в контексті оцінки розвитку – «прагнен-

ня зрозуміти», адже неосудливий підхід має більше шансів забезпечити відкритість і чесність в бесідах.

Проводяться відвідування на місцях, тому що спостереження інтервенції в місці проведення може показувати, наскільки добре ця інтервенція працює. Головним завданням є вибір для інтерв'ю людей з різним досвідом і різними точками зору, особливо тих, які зазвичай не охоплюються оцінюванням. Мала, але дуже різноманітна група ключових інформантів може бути ефективною для отримання цілісної картини. Вкрай важливо вміти слухати.

Швидка оцінка може використовувати ті самі методи збору й аналізу даних, що й будь-яка інша оцінка. Різниця зазвичай в масштабі. Типова швидка оцінка проводиться в обмеженому масштабі – залучаються всього кілька ключових осіб в декількох місцях. Необхідно використовувати більше одного джерела інформації, тому що множинність джерел підвищує правдоподібність, зменшує упередженість та забезпечує цілісну перспективу. Отже, дані, отримані від основних інформантів, слід підкріпляти інформацією попередніх звітів та досліджень, записами і документами, а також даними, зібраними шляхом спостережень, отриманими від фокус-груп, групових інтерв'ю і опитувань. Чим більш узгодженою виявиться інформація з цих джерел, тим більше впевненості в результатах оцінки.

Якщо використовуються якісні методи, необхідними виявляться навички конспектування. Дуже добре, якщо фахівець з оцінювання веде щоденник для занотовування спостережень, тем, думок, інтерпретацій та будь-яких інцидентів, що трапляються під час польової роботи. Такими відомостями слід обмінюватися з іншими членами групи, щоб виявляти подібні випадки.

При проведенні швидкої оцінки фахівці з оцінювання повинні пам'ятати наступні підказки:

- Перед виїздом на місце варто переглянути вторинні дані.
- На місці треба спостерігати, слухати, запитувати й записувати.

Треба вести ретельні записи протягом всього процесу; вони допоможуть не лише скласти звіт, а й інтерпретувати інформацію, зібрану різними членами групи.

Стратегії та досвід, отриманий при проведенні швидких оцінок, вказують на те, що перед швидкою оцінкою слід взяти до уваги наступне:

- Створіть різноманітну, мультидисциплінарну групу, до складу якої входять і чоловіки, і жінки.

- Якщо можливо, залучіть до групи як місцевих (тих, хто знайомий з інтервенцією і знає місцевість), так і сторонніх (вони зможуть оцінити ситуацію свіжим поглядом).
- Краще використовуйте малі, ніж великі групи, щоб максимізувати взаємодії.
- Виділяйте час окремо на збір даних та на їх інтерпретацію.
- Будьте готові йти туди, куди треба (в поля, на базари, в глухі села), а не туди, куди зручно.

Зберігайте гнучкість і адаптивність, тому що нова інформація може змінити план оцінки (FAO 1997).

Синтезовані оцінювання

Синтезоване оцінювання:
Підхід, за якого фахівець з оцінювання шукає спільні моменти у різних інтервенціях, присвячених певній проблемі або темі

Синтезовані оцінювання – корисний підхід у тих випадках, коли вже було проведено багато оцінок результатів схожих інтервенцій. Він дозволяє фахівцю з оцінювання порівнювати різні інтервенції, присвячені певній проблемі або темі, і виявляти спільні моменти. Це корисно, коли оцінка має визначити загальну дієвість інтервенції.

Челімські й Морра (Chelimsky and Morra 1984) перенесли цей метод в ширший політичний контекст, застосувавши його до переважно якісних, а також кількісних оцінок інтервенцій. До цього синтезовані оцінювання використовувалися для роботи з оцінками, що спиралися на переважно кількісний аналіз – наприклад, з оцінками впливу програм харчування молодих жінок на вагу новонароджених і смертність серед немовлят, або ж впливу розміру класу на академічну успішність учнів.

Ідея синтезованого оцінювання полягає в тому, що, хоча окремі оцінки можуть надавати корисну інформацію про конкретну інтервенцію, кожна з них занадто обмежена й прив'язана до контексту, щоб дати змогу зробити загальний висновок стосовно впливу інтервенції. Однак, комбінуючи результати багатьох досліджень, можна робити загальні висновки про вплив інтервенції (і навіть політики).

Однією з переваг синтезованого оцінювання є те, що воно використовує наявні дослідження, завдяки чому є менш витратним, ніж інші види оцінки. Ще одна перевага полягає в тому, що він створює більшу базу для оцінки впливу інтервенції, підвищуючи впевненість стосовно загальних висновків про вплив. Але є й виклики: треба відшукати всі актуальні дослідження, опубліковані та неопубліковані, оцінити якість кожної оцінки; якщо можливо, отримати масиви даних для поєднання для проведення вторинного аналізу.

Синтезоване оцінювання має включати наступне:

- чітко заявлені процедури визначення оцінок та їх масштабу;
- прозорі критерії якості для прийняття рішень щодо включення оцінок до синтезу;
- процедури застосування критеріїв якості (ці процедури часто виконуються комісією, а не окремою особою);
- посилання на всі проаналізовані оцінювання;
- стислі характеристики кожного оцінювання, включеного до синтезу, а також висновки щодо цікавих тем або змінних;
- прогалини або обмеження синтезу.

Іноді використовуються лише окремі елементи синтезованого оцінювання. Приміром, у випадку, описаному в блоці 5.3, якість звітів, очевидно, не перевірялась, що робить результати синтезу дещо сумнівними.

Блок 5.3 «Використання синтезованого оцінювання для вимірювання екологічної користі»

У 2000 році британське Міністерство з міжнародного розвитку (DfID) оприлюднило екологічне дослідження, проведене за допомогою методу синтезованого оцінювання. В 1990-х роках DfID успішно реалізувало цілу низку екологічних проектів, але урядовці мали відчуття, що екологічні ефекти скоріше «малися на увазі, а не вивчалися критично» (p. 1). Було замовлено синтезуюче дослідження з екологічних питань. Воно мало розглянути 49 проектів, що проводилися за підтримки DfID у п’яти країнах. Завданням дослідників було оцінити реалізацію цих проектів і вплив фінансування на покращення екологічної ситуації та захист довкілля. Проекти не були чисто екологічними, але охоплювали широкий спектр екологічних інтервенцій (енергоефективність, заходи у промисловості, лісівництві, сільському господарстві, захист біологічної різноманітності та благоустрій міст).

Розглянувши 49 досліджень, фахівці з оцінювання дійшли висновку, що «існує розрив між високим рівнем пріоритету, що його встановило DfID для екологічних питань… та фактичними позитивними наслідками проектів для довкілля».

Джерело: DfID 2000, 1.

У спеціальній літературі терміни «синтезовані оцінювання» й «мета-оцінювання» іноді вживаються як синоніми. Ми розділяємо ці два терміни. З нашої точки зору, синтезоване оцінювання означає аналітичний огляд результатів набору оцінок, що відповідають стандартам якості. А **мета-оцінювання** – це експертиза одного або кількох оцінювань за професійними стандартами якості.

Мета-оцінювання:
Експертиза однієї або кількох оцінок за професійними стандартами якості

Інші підходи

Для оцінювання розвитку застосовують також інші підходи, теорії та моделі. Більшість з них – це варіації на тему спільного оцінювання. Читачі знайдуть елементи деяких моделей в оцінюванні з орієнтацією на результат, прибічниками якої є автори цієї книги. Йдеться про формулювання й тестування теорії змін, зосередженість на активізації практичного використання оцінювання з самого початку його проведення, – для цього визначаються й залучаються ключові зацікавлені особи.

В цьому розділі характеризуються наступні підходи:

- оцінювання з орієнтацією на практичне застосування;
- наснажуюче оцінювання;
- реалістичне оцінювання;
- інклюзивне оцінювання;
- оцінювання із залученням бенефіціарів;
- горизонтальне оцінювання.

Оцінювання з орієнтацією на практичне застосування

Оцінювання з орієнтацією на практичне застосування:
Оцінка, якість котрої визначається її корисністю й фактичним застосуванням

Оцінювання з орієнтацією на практичне застосування пропонує судити про оцінку з того, наскільки вона корисна і як фактично застосовується. Петон (Patton 2002, p. 173), чия головна книга про оцінку з орієнтацією на практичне застосування пережила вже четверте видання (2008), описує її як таку, що починається з «визначення і організації конкретних, релевантних осіб, які приймають рішення та використовують інформацію (це не абстрактна, пасивна аудиторія), які будуть застосовувати створену оцінкою інформацію». Це оцінювання орієнтоване на те, щоб його результати застосовувалися цільовими користувачами за цільовим призначенням. Цільові користувачі – це ті, хто зобов'язаний застосовувати висновки оцінки та втілювати в життя її рекомендації. Оцінювання з орієнтацією на практичне застосування допомагає цільовим користувачам обирати найбільш адекватні модель, зміст та методи оцінки залежно від конкретної ситуації.

Наснажуюче оцінювання

Наснажуюче оцінювання[5] – це використання концепцій, технік і висновків оцінювання для покращень та самовизначення (Fetterman, Kaftarian, and Wandersman 1996). Воно ґрунтується на глибокій повазі до здатності людей створювати знання про власний досвід і знаходити рішення для проблем, з якими вони стикаються. Створення нових можливостей допомагає досягати цілей – як особистих, так і соціальних. Якість життя покращується, а це дає відчуття добробуту й позитивного зростання (Fetterman and Wandersman 2004).

Наснажуюче оцінювання: Використання концепцій, технік і висновків оцінки для покращень і самовизначення

Фетерман і Вандерсман (Fetterman and Wandersman 2004) характеризують роль наснажуючого оцінювання як «критичного друга». Вони відстоюють думку про те, що члени спільноти повинні займати активну позицію щодо оцінювання; фахівець з оцінювання має відігравати роль координатора, впливаючи на оцінку, а не контролюючи її.

Певні характеристики наснажуючого оцінювання збігаються з характеристиками оцінювання з орієнтацією на практичне застосування. Обидва ці підходи покликані бути корисними й конструктивними на кожному етапі оцінки (Fetterman and Wandersman 2004). Але Петон (Patton 1997) вказує на те, що хоча наснажуюче оцінювання перетинається з підходами спільного оцінювання в сфері відповідального ставлення, актуальності й нарощування потенціалу, головним для нього є сприяння самовизначенню. Наснажуюче оцінювання виходить за межі більшості спільних оцінювань завдяки тому, що фахівець з оцінювання-координатор стоїть на боці людей, які потребують нових можливостей, і відстоює політичні зміни.

Описуючи різницю між наснажуючим оцінюванням та спільним оцінюванням, Елкін і Крісті (Alkin and Christie 2004, p. 56) стверджують: оскільки оцінювання із залученням зацікавлених сторін виникає з системи практичного застосування, його метою є активізація практичного застосування через ці дії (планування, реалізація, аналіз та інтерпретація), а не надання нових можливостей тим, хто піддавався утискам – це вже політичні або визвольні заходи.

Реалістичне оцінювання

Реалістичне оцінювання: Оцінка, що базується на теорії та утворює логічну й узгоджену структуру для роботи з програмами, розглядає зацікавлених осіб як ненадійних експертів і містить елементи інших підходів

Поусон і Тіллі (Pawson and Tilley 2004, p. 1) характеризують **реалістичне оцінювання** як «різновид оцінки, заснованої на теорії». Вони пов'язують її з теорією змін та програмною теорією. Для формування теорії потрібно розуміти загальний контекст, в якому відбувається інтервенція, як ця інтервенція працює в конкретному контексті, а також чи працює вона взагалі.

[5] *В оригіналі Empowerment evaluation.*

Поусон і Тіллі (Pawson and Tilley 2004, p. 22) стверджують, що реалістичне оцінювання створює «логічну й узгоджену структуру» для оцінки програм. Визнається важливість участі зацікавлених осіб у розробці й реалізації програм, але при цьому даний підхід займає позицію між двома протилежними поглядами: зневажати зацікавлених осіб (через їх упередженість) і вважати їх всезнаючими і непогрішимими (завдяки інсайдерським знанням). Зацікавлені особи розглядаються як ненадійні експерти, чиє розуміння потребує формалізації і перевірки (Pawson and Tilley 2004).

Реалістичне оцінювання формується з широкого спектру підходів до дослідження й оцінки. Воно користується елементами (частково або повністю) інших підходів. Реалістичне оцінювання може бути якісним або кількісним; воно часто поєднує обидва методи, але не встановлює експериментально причинно-наслідковий зв'язок між інтервенцією і довгостроковим результатом. Причинно-наслідковий зв'язок визначається через підведення теоретичної основи під роботу інтервенції. Потім в логічній послідовності здійснюються кроки з перевірки гіпотези для визначення того, що може працювати, для кого та за яких обставин. Поусон і Тіллі (Pawson and Tilley 2004) зазначають, що реалістичне оцінювання може бути складним для проведення через те, що жодна проста формула не дає простих рецептів для отримання результату.

Інклюзивне оцінювання

Інклюзивне оцінювання:
Оцінювання, що охоплює найменш захищені верстви населення

Інклюзивне оцінювання зосереджується на охопленні найменш захищених верств населення в рамках систематичного дослідження користі або цінності проекту, програми або політики. Воно базується на даних, але дані надаються найменш захищеними зацікавленими особами – тими, хто зазвичай представлений недостатньо. Інклюзивне оцінювання не охоплює тих, кого традиційно включають до оцінки (Mertens 1999). Як і наснажуюче оцінювання, воно є трансформаційною парадигмою.

Інклюзивне оцінювання ставить запитання типу:

- Які існують важливі відмінності всередині населення, що має отримувати послуги?
- Як надаються послуги всередині різних підгруп?
- Які цінності лежать в основі розподілу послуг?

Оцінка із залученням бенефіціарів:
Інструмент, що використовується для покращення впливу операцій з розвитку завдяки врахуванню думок потенційних бенефіціарів стосовно запланованої або поточної інтервенції

Оцінка із залученням бенефіціарів

Оцінка із залученням бенефіціарів – це «інструмент якісного дослідження, що використовується для покращення впливу операцій з розвитку завдяки врахуванню думок потенційних бенефіціарів стосовно запланованої

або поточної інтервенції» (Salmen 1999, p. 1). Як й інклюзивне оцінювання, цей підхід передбачає охоплення груп, що часто залишаються без уваги. Цей підхід з орієнтацією на проект був розроблений Світовим банком наприкінці 1980-х років, щоб доповнити тодішній більш технічно- та фінансово орієнтований підхід до оцінки. Такий вид оцінювання зазвичай застосовується в проектах з компонентом надання послуг.

Оцінка із залученням бенефіціарів охоплює кінцевого клієнта, бенефіціарів проекту. Ідея полягає в тому, що збільшена участь бенефіціарів у формуванні плану проекту, наданні зворотного зв'язку для моніторингу, а також висловленні думок щодо впливів забезпечує відповідальніше ставлення до інтервенцій і робить бенефіціарів ключовими гравцями в проведенні необхідних і бажаних змін у їх власному розвитку.

Мета оцінки із залученням бенефіціарів – оцінити значимість заходу з точки зору бенефіціарів проекту та інтегрувати її результати в проектні заходи. Оцінка із залученням бенефіціарів відіграє центральну роль в соціальній оцінці, роль містка між культурою та прийняттям рішень (Salmen 1999).

Горизонтальне оцінювання

Горизонтальне оцінювання:
Оцінка, що поєднує процес внутрішнього оцінювання з зовнішньою дружньою перевіркою

Горизонтальне оцінювання – відносно новий підхід до оцінки, що поєднує процес внутрішнього оцінювання з зовнішньою дружньою перевіркою. Це поєднання було покликане нейтралізувати «однобокі владні відносини, що домінують в традиційних зовнішніх оцінках, створюючи більш сприятливу атмосферу для використання досвіду і подальшого вдосконалення програми» (Thiele and others 2006, p. 1).

Горизонтальне оцінювання часто використовувалося для того, щоб більше дізнатися про розроблювані методики дослідження й розвитку та відповідним чином вдосконалити їх. Цей підхід застосовувався в Андській регіональній програмі для розробки нових методик дослідження й розвитку, а також в Уганді для оцінки підходу спільного оцінювання ринкового ланцюжка (Thiele and others 2006, p. 1).

Ключова риса горизонтального оцінювання – дві окремі групи зацікавлених осіб. Перша з них – місцеві учасники, що представляють та критикують процес, що вивчається, та вносять рекомендації щодо його покращення. Друга група – відвідувачі (колеги з інших організацій або проектів, що працюють над схожими темами), які оцінюють процес, виявляють його сильні й слабкі сторони й надають рекомендації щодо покращень (Thiele and others 2006). Одним з компонентів горизонтального оцінювання є семінар, що дозволяє цим двом групам поспілкуватися.

Підсумки

Підхід до оцінювання – це спосіб розглядати оцінювання або представляти його загальну концепцію. Він часто включає філософію і набір цінностей. Деякі підходи використовуються роками, деякі було розроблено або застосовано зовсім недавно для проведення оцінювання розвитку. Головні риси підходів до оцінювання представлені в табл. 5.3.

Охарактеризовані підходи використовуються переважно на рівні окремої інтервенції. В главі 11 аналізуються оцінювання складних інтервенцій розвитку, що все частіше розглядаються як норма.

Таблиця 5.3. Мета, характеристики, сильні й проблемні сторони різних підходів оцінювання

Підхід оцінювання	Ціль/ філософія	Характеристики/ заходи	Сильні сторони	Проблемні сторони
Перспективне оцінювання	Інтервенція розглядається до її початку Даються відповіді на прогнозні питання щодо ймовірності успіху	Виконується контекстуальний аналіз проекту, програми або політики Вивчаються проведені оцінювання аналогічних програм для використання досвіду і виявлення проблем Прогнозується успіх або невдача, пропонуються шляхи зміцнення запропонованої інтервенції, якщо вона буде реалізуватися	Використання вторинного аналізу здешевлює оцінювання З'являється можливість вирішення проблем і зміцнення програми до її початку	Існує залежність від звітів оцінки споріднених програм
Визначення оцінюваності	Визначається, чи корисно та практично проводити оцінку Покращуються зв'язки між фахівцями з оцінювання й зацікавленими особами З'ясовується мета формативного дослідження або функцій як інструмента планування	Розглядаються матеріали й наявні дані Перевіряється спільне розуміння цілей і задач програми Проводяться інтерв'ю з керівниками інтервенції, зацікавленими особами, людьми й організаціями, що впроваджують рекомендації Розробляється модель теорії змін	Оцінка може стати більш реалістичною і адекватною, а програми – краще спланованими Можливе формування підтримки з боку зацікавлених осіб і кращого розуміння цілей і задач Можна запобігти витратам на оцінки, не здатні відповісти на поставлені запитання через брак даних	Якщо завжди застосовувати цей підхід, можливі невиправдані затримки перед проведенням оцінювання

Підхід оцінювання	Ціль/ філософія	Характеристики/ заходи	Сильні сторони	Проблемні сторони
Оцінювання з орієнтацією на цілі	Вимірюється, чи досягнуто цілей, чи виконані завдання Слугує основою для більшості систем оцінки донорських проектів	Визначаються цілі й задачі Оцінюється, чи досягає інтервенція поставлених цілей та чи виконує задачі (нормативна оцінка)	Оцінка спрощується за рахунок простої методики: порівняння фактичних даних і плану	Можуть залишитися непоміченими важливі наслідки, прямо не заявлені як цілі або завдання
Оцінювання без прив'язки до цілей	Відкриває можливість збору даних про наслідки й дієвість програм без обмеження поля зору обраними цілями чи завданнями	Фахівці з оцінювання звільняються від необхідності орієнтуватися на програмні цілі й завдання за рахунок обмеження контактів з керівником і персоналом програми	Більша ймовірність помітити незаплановані побічні ефекти	Дозволяються обмежені контакти з персоналом програми; фахівцям з оцінювання може бути важко уникнути інформації про програмні цілі й завдання
Багатоцентрове оцінювання	Вивчаються інтервенції в стандартній реалізації в усіх місцях та запланованих варіантах; мета – визначити умови, в яких програма найкращим чином досягає цілей і виконує завдання Необхідно оцінити клімат, в якому ведуться інтервенції, а також культурні, географічні, економічні, масштабні та інші системні відмінності, що можуть впливати на інтервенції	Залучаються зацікавлені особи Збирається глибша інформація порівняно з деякими іншими підходами Вибираються місця, що представляють певні характеристики, здатні зумовлювати систематичні відмінності у реалізації й результатах інтервенції Описуються й порівнюються інтервенції в їхніх контекстах, робляться висновки з досвіду, створюються стратегії роботи у різних ситуаціях	Часто виявляються передові практики Надійніші результати, ніж від оцінки однієї інтервенції в одному місці Проводиться локальний і порівняльний аналіз	Необхідний стандартизований збір даних Вимагається підготовка персоналу, доступ в усі місця та завчасне надання інформації для створення інструментів збору даних Можливі нерепрезентативні висновки, актуальні лише в певному контексті
Кластерне оцінювання	Розглядаються групи схожих або споріднених інтервенцій, що разом складають єдиний блок Вивчається «кластер» інтервенцій в одній або кількох ситуаціях	Залучаються зацікавлені особи Менше уваги приділяється узагальнюваності та відтворюваності Відмінності – позитивний факт Використовується багато практичних прикладів з обміном інформацією	В центрі уваги – використання досвіду, а не висновки щодо якості або цінності програми	Лише узагальнені результати, немає даних про результати в окремих місцях

(Продовження)

Підхід оцінювання	Ціль/ філософія	Характеристики/ заходи	Сильні сторони	Проблемні сторони
Соціальна оцінка	Розглядаються соціальні структури, процеси та зміни в межах групи або спільноти Головний інструмент для врахування соціальних впливів інтервенцій розвитку	Вивчається узгодженість між задачами інтервенції та потребами, інтересами й можливостями зацікавлених осіб Розглядаються наслідки соціальних та культурних факторів для здатності зацікавлених осіб брати участь в інтервенції або отримувати від неї вигоди Вивчається відмінний вплив інтервенції на підгрупи зацікавлених осіб	Виявляються потенційні негативні соціальні впливи, завдяки чому можна впроваджувати стратегії їх пом'якшення	З поля зору можуть випускатися екологічні питання, отже, розглядається лише одна частина рівняння
Соціально-екологічна оцінка	Оцінюється виконання соціально-екологічних завдань	В центрі уваги – соціальні та екологічні аспекти всіх інтервенцій, не лише тих, що спрямовані тільки на екологію або соціальну нерівність Є корисною в ситуаціях, коли оцінка цілком присвячена розгляду соціально-екологічних питань	Підкреслюється важливість соціально-екологічних аспектів інтервенцій	Вимагається технічна кваліфікація
Оцінювання із залученням зацікавлених сторін	Розподіл між зацікавленими особами обов'язків з планування оцінки, її реалізації і звітування Участь – продукт, а факт участі є метою і одним з показників успіху Також розглядається участь як процес досягнення заданої мети	Учасники залучаються до постановки цілей, визначення пріоритетів, вибору питань, аналізу даних та прийняття рішень щодо даних Підхід сприяє груповій співпраці, єдності групи і самостійності в поєднанні з відповідальністю Фахівець з оцінювання перетворюється на координатора і порадника, а учасники – на керівників оцінки й осіб, що приймають рішення	Зростає правдоподібність результатів оцінки в очах персоналу програми, а також імовірність практичного застосування результатів Учасники отримують нові можливості, зміцнюється потенціал на місцях	Фахівці з оцінювання повинні вміти координувати, співпрацювати, долати конфлікти, а також навчати учасників технікам оцінки Часто вважається незалежним

Підхід оцінювання	Ціль/ філософія	Характеристики/ заходи	Сильні сторони	Проблемні сторони
Картування довгостроко-вих результатів	Доповнює більш традиційні форми оцінки, зосереджуючись на супутніх змінах поведінки; в центрі уваги – люди й результати, а не досягнення впливів розвитку	Вивчається поведінка, сформована зусиллями багатьох організацій; мета – підвищення ефективності проектів, програм або політик за рахунок надання нових інструментів, технік та ресурсів Визначається коло суміжних партнерів для оцінки їх впливу на зміни Описуються зміни у поведінці окремих особистостей або членів груп чи організацій	Картування допомагає вирватися з пастки Downstream Outcome Blues (див. текст)	Вимагається незалежне й відповідальне оцінювання впливу
Швидка оцінка	Задовольняє потребу в швидкій і недорогій оцінці; загалом збирається лише необхідна інформація для надання відомостей про ефективність	Систематичний, напівструктурований підхід Аналіз документів, інтерв'ю та короткострокові візити Необхідно використовувати більш ніж одне джерело інформації Фахівець з оцінювання повинен вміти добре слухати й конспектувати	Місток між формальним та неформальним збором даних; найкраще підходить для розгляду процесів і проблем	Зазвичай проводиться в малому масштабі Надає обмежену, переважно описову інформацію
Синтезоване оцінювання	Відшукуються всі релевантні оцінки, оцінюється їх якість, результати комбінуються або підсумовуються; корисний підхід в ситуації, коли вже проведено багато оцінок інтервенції	Відшукуються всі релевантні дослідження Встановлюються критерії визначення якості досліджень Використовуються лише якісні дослідження Комбінуються результати або іншим чином синтезуються висновки різних досліджень щодо ключових питань Синтез може використовуватися з якісними або кількісними даними	Використовуються наявні оцінки й дослідження, що дозволяє заощадити Створюється більша й міцніша база для оцінки впливу	Можуть виникнути складнощі з пошуком всіх релевантних досліджень і отриманням даних Існує певний ризик упередженості при виборі «якісних» оцінок (ризики можна зменшити, довіряючи відбір комісії)

(Продовження)

Підхід оцінювання	Ціль/ філософія	Характеристики/ заходи	Сильні сторони	Проблемні сторони
Оцінювання з орієнтацією на практичне застосування	Інтервенція оцінюється за корисністю та фактичним застосуванням	Спочатку – визначення й організація конкретних релевантних осіб, що приймають рішення, а також користувачів інформації (а не абстрактної пасивної аудиторії), які використовуватимуть інформацію з цієї оцінки	Підхід забезпечує застосування оцінки, з самого початку орієнтуючись на цільове використання цільовими користувачами	Акцент робиться на головному клієнті; може не охоплювати інших зацікавлених осіб
Наснажуюче оцінювання	Використовуються концепції, техніки і висновки оцінки для сприяння покращенням і самовизначенню	Йде далі, ніж спільна оцінка: виказується глибока повага до здатності людей створювати знання про власний досвід і знаходити рішення для проблем, з якими вони стикаються	Відстоюються інтереси незахищених груп	Фахівець з оцінювання стає захисником (оцінка може розглядатися як політична і упереджена)
Реалістичне оцінювання	Формується логічна й узгоджена структура для оцінки програм через опис контексту; розглядається, як повинна працювати програма та чи існує причинно-наслідковий зв'язок між програмою та цільовим результатом	Використовуються різні підходи дослідження й оцінки; частково або повністю використовуються елементи інших підходів Використовуються якісні дослідження, кількісні дослідження або їх поєднання	Теоретично пояснюється, яким чином могла б працювати інтервенція, за яких умов і для кого	Зрозуміло, що модель кожної програми повинна визначатися індивідуально
Інклюзивне оцінювання	Охоплюються найменш захищені зацікавлені особи в рамках систематичного вивчення користі або цінності проекту, програми або політики	Отримуються дані від найменш захищених зацікавлених осіб; не охоплюються ті, кого традиційно включають в оцінку	В центрі уваги – найменш захищені зацікавлені особи та відмінності у наданні послуг	Фахівець з оцінювання стає захисником (оцінка може розглядатися як політична і упереджена)
Оцінювання із залученням бенефіціарів	Покращується вплив операцій розвитку завдяки збору думок цільових бенефіціарів стосовно запланованої або поточної інтервенції	Охоплюються бенефіціари проекту, посилюючи в них почуття відповідальності	Співпраця формує відповідальне ставлення і зміцнює потенціал. Цінність оцінки зростає завдяки врахуванню думок бенефіціарів	Екологічні проблеми відокремлюються від їхніх соціальних наслідків
Горизонтальне оцінювання	Поєднує процес внутрішнього оцінювання з зовнішньою дружньою перевіркою для нейтралізації однобоких владних відносин, що домінують в традиційних зовнішніх оцінках	Використовується, щоб більше дізнаватися про методики оцінки й розвитку та покращувати їх	Поєднує сильні сторони самооцінки та зовнішньої дружньої оцінки	Дружні оцінки можуть бути взаємно позитивними

Джерело: Автори.

Завдання до Глави 5

Практична вправа 5.1: Вибір найбільш адекватного підходу оцінювання

Оберіть підхід оцінювання для аналізу кожного з описаних нижче завдань. Охарактеризуйте підхід, назвіть вигоди, які він дає, та виклики, які з ним пов'язані. Поясніть, чому обрали саме цей підхід.

1. Оцінити стратегічний фокус технічної допомоги країні на основі результатів п'яти досліджень, виконаних різними організаціями розвитку.
2. Виявити успішні освітні інтервенції, що втілювалися в рамках проектів і програм вашої організації з метою покращення освітніх систем в регіоні.
3. Оцінити найзначніші питання стосовно природних ресурсів країни й екологічного сектору.
4. Оцінити розвиток рисівництва в країні, у тому числі важливість рису в поточному культурному, соціальному та економічному контекстах; системи виробництва рису; обмеження, з якими стикаються рисівники; проведені дослідження й розроблені технології, а також майбутні пріоритети для подальшого розвитку галузі.
5. Оцінити оцінки, проведені організацією з розвитку, яка за останні 30 років отримала мільйони доларів на міжнародні сільськогосподарські дослідження.

Примітки

1. Значну частину інформації в цьому розділі взято з Earl, Carden, and Smutylo (2001).

Посилання та рекомендована література

Alkin, Marvin, and Christina Christie. 2004. "An *Evaluation Theory Tree." In Evaluation Roots: Tracing Theorist Views and Influences*, ed. M. Alkin. 12–65, Thousand Oaks, CA: Sage Publications.

Chambers, R. 1991. "Shortcut and Participatory Methods for Gaining Social Information for Projects." In *Putting People First: Sociological Variables in Rural Development,* 2nd ed., ed. M. M. Cernea, 515–37. Washington, DC: World Bank.

Chelimsky, E., and L. G. Morra. 1984. "Evaluation Synthesis for the Legislative User." In *Issues in Data Synthesis,* ed. W. H. Yeaton and P. M. Wortman, 75–89. New Directions for Program Evaluation No. 24. San Francisco: Jossey-Bass.

Christie, Christina, and Marvin Alkin. 2004. "Objectives-Based Evaluation." In *Encyclopedia of Evaluation*, ed. Sandra Mathison. 281–85. Thousand Oaks, CA: Sage Publications.

CIDA (Canadian International Development Agency). 2004. *CIDA Evaluation Guide 2004.* Ottawa.

Cousins, J. B., and L. M. Earl, eds. 1995. *Participatory Evaluation in Education.* Bristol, PA: Falmer Press.

Dalal-Clayton, Barry, and Stephen Ba. 2002. *Sustainable Development Strategies: A Resource Book.* Sterling, VA: Earthscan Publications. http://www.nssd.net/res_book.html#contents.

DFID (Department for International Development). 2000. *Environment: Mainstreamed or Sidelined?* Environmental Evaluation Synthesis Study EV626, January, London.

http://www.dfid.gov.uk/aboutdfid/performance/files/ev626s.pdf.

Duignan, Paul. 2007. *Introduction to Strategic Evaluation: Section on Evaluation Approaches, Purposes, Methods, and Designs.* http://www.strategicevaluation.info/se/documents/104f.html.

Earl, Sarah, Fred Carden, and Terry Smutylo. 2001. *Outcome Mapping: Building Learning and Reflection into Development Programs.* International Development Research Centre, Ottawa. http://www.dgroups.org/groups/pelican/docs/Mapping_M&E_capacity_ 080606.pdf.

ECDPM (European Centre for Development Policy Management). 2006. *Study on Capacity, Change and Performance: Mapping of Approaches Towards M&E of Capacity and Capacity Development.* Maastricht.

Eerikainen, Jouni, and Roland Michelevitsh. 2005. "Environmental and Social Sustainability. Methodology and Toolkit: Various Approaches." International Program for Development Evaluation Training (IPDET) presentation, Ottawa, July.

FAO (Food and Agriculture Organization of the United Nations). 1997. "Rapid Rural Appraisal." In *Marketing Research and Information Systems Marketing and Agribusiness.* Rome.

http://www.fao.org/docrep/W3241E/w3241e09.htm.

Fetterman, David M. 2001. *Foundations of Empowerment Evaluation.* Thousand Oaks, CA: Sage Publications.

Fetterman, David M., S. Kaftarian, and A. Wandersman, eds. 1996. *Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability.* Thousand Oaks, CA: Sage Publications.

Fetterman, David M., and Abraham Wandersman. 2004. *Empowerment Evaluation Principles in Practice.* New York: Guilford Publications.

———. 2007. "Empowerment Evaluation: Yesterday, Today, and Tomorrow." *American Journal of Evaluation* 28 (2):179–98. http://homepage.mac.com/profdavidf/documents/EEyesterday.pdf.

Fitzpatrick, Jody L., James R. Sanders, and Blaine R. Worthen. 2004. *Program Evaluation: Alternative Approaches and Practical Guidelines.* New York: Pearson.

Gariba, Sulley. 1998. "Participatory Impact Assessment: Lessons From Poverty Alleviation Projects in Africa." in *Knowledge Shared: Participatory Evaluation in Development Cooperation*, ed. Edward T. Jackson and Yussuf Kassam, 64–81. Bloomfield, CT: Kumarian Press.

Gill, M., and A. Spriggs. 2002. *The Development of Realistic Evaluation Theory through the Evaluation of National Crime Prevention Programmes.* http://www.evaluationcanada/distribution/20021010_gill_martin_spriggs_angela. pdf.

Glass, Gene V., and Mary Lee Smith. 1979. "Meta-Analysis of Research on Class Size and Achievement." *Educational Evaluation and Policy Analysis* 1 (1): 2–16.

Hirschheim, R., and S. Smithson. 1988. "A Critical Analysis of Information Systems Evaluation." In *IS Assessment: Issues and Changes*, eds. N. Bjorn-Andersen and

G. B. Davis. Amsterdam: North-Holland.

ISO (International Organization for Standardization). 1999. *Environmental Management: Performance Evaluation Guidelines.* ISO 14301, Geneva.

Johnston, Timothy, and Susan Stout. 1999. *Investing in Health: Development in Health, Nutrition, and Population Sector.* World Bank, Operations Evaluation Department, Washington, DC. http://wbln0018.worldbank.org/oed/oed-doclib.nsf/6e14e487e87320f7852568080 06a001a/daf8d4188308862f852568 420062f332/US$FILE/HNP.pdf.

Khon Kaen University. 1987. *Proceedings of the 1985 International Conference on Rapid Rural Appraisal.* Rural Systems Research and Farming Systems Research Projects, Khon Kaen, Thailand.

Kretzmann, John P., John L. McKnight, and Deborah Puntenney. 1996. A Guide to Massing Local Business Assets and Modulizing Local Business Capabilities. Skokie, IL: ACTA Publications.

Kudat, Ayse, and Bykebt Ozbilgin. 1999. *Azerbaijan Agricultural Development and Credit Program.*

http://lnweb18.worldbank.org/ESSD/sdvext.nsf/61ByDocName/Azerbaijan AgriculturalDevelopmentandCreditProject/US$FILE/AzerbaijanAgricultural DevelopmentandCreditProject424KbPDF.pdf.

Kumar, Krishna, ed. 1993. *Rapid Appraisal Methods.* Washington, DC: World Bank.

Light, R. J., and D. B. Pillemer. 1984. *Summing Up: The Science of Reviewing Research.* Cambridge, MA: Harvard University Press.

Mertens, D. 1999. "Inclusive Evaluation: Implications of Transformative Theory for Evaluation." *American Journal of Evaluation* 20 (1): 1–14.

Narayan, Deepa. 1996. *Toward Participatory Research.* World Bank Technical Paper 307, World Bank, Washington, DC.

http:// www-wds.worldbank.org/external/default/WDSContentServer/WDSP/ IB /1996/04/01/000009265_3980625172923/Rendered/PDF/multi0page.pdf.

Nay, J., and P. Kay. 1982. *Government Oversight and Evaluability Assessment.* Lexington, MA: Heath.

OECD (Organisation for Economic Cooperation and Development). 1997. *Searching for Impact and Methods: NGO Evaluation Synthesis Study.* Development Assistance Committee, Paris.

http://www.eldis.org/static/DOC5421.htm.

Patton, Michael Q. 1990. *Qualitative Evaluation and Research Methods*. 2nd ed. Thousand Oaks, CA: Sage Publications.

———. 1997. *Utilization-Focused Evaluation: The New Century Text.* Thousand Oaks: CA: Sage Publications.

———. 1997. "Toward Distinguishing Empowerment Evaluation and Placing It in a Larger Context." *Evaluation Practice* 18 (2):147–63.

———. 2002. *Qualitative Evaluation and Research Methods*, 3rd ed. Thousand Oaks, CA: Sage Publications.

———. 2008. *Utilization-Focused Evaluation*. 4th ed. Thousand Oaks, CA: Sage Publications.

Paulmer, Hubert E. 2005. “Evaluation Guidelines of International Development Aid Agencies: A Comparative Study.” International Rural Planning and Development, School of Environmental Design and Rural Development, University of Guelph, Ontario.

Pawson, Ray, and Nick Tilley. 2004. “Realistic Evaluation.” In *Encyclopaedia of Evaluation*, ed. Sandra Matthieson, 359–67. Thousand Oaks, CA: Sage Publications.

Picciotto, Robert. 2007. “The New Environment for Development Evaluation.” *American Journal of Evaluation* 28: 509–21.

Preskill, Hallie, and Darlene Russ-Eft. 2005. *Building Evaluation Capacity: 72 Activities for Teaching and Training.* Thousand Oaks, CA: Sage Publications.

Salmen, Lawrence F. 1999. *Beneficiary Assessment Manual for Social Funds.* World Bank, Social Protection Team Human Development Network, Washington, DC. http://lnweb18.worldbank.org/ESSD/sdvext.nsf/07ByDocName/Beneficiary AssessmentManualforSocialFunds/US$FILE/ percent5BEnglish percent5D+ Beneficiary+Assessment +Manual.pdf.

Sanders, J. R. 1997. “Cluster Evaluation.” In *Evaluation for the 21st Century: A Handbook*, eds. E. Chelimsky and W. R. Shadish Jr., 396–401. Thousand Oaks, CA: Sage Publications.

Scrimshaw, N., and G. R. Gleason. 1992. *Rapid Assessment Procedures: Qualitative Methodologies for Planning and Evaluation of Health Related Programs.* Boston: International Nutrition Foundation for Developing Countries.

Scriven, Michael. 1972a. “Objectivity and Subjectivity in Educational Research.” In *Philosophical Redirection Educational Research: The Seventy-First Yearbook of the National Society for the Study of Education*, ed. L. G. Thomas. Chicago: University of Chicago Press.

———. 1972b. “Pros and Cons about Goal-Free Evaluation.” *Evaluation Comment* 3: 1–7.

———. 1991. *Evaluation Thesaurus.* 4th. ed. Thousand Oaks, CA: Sage Publications. Smith, M. F. 1989. *Evaluability Assessment: A Practical Approach.* Boston: Kluwer Academic Press.

Smith, Mary Lee, and Gene V. Glass. 1980. “Meta-Analysis of Research on Class Size and Its Relationship to Attitudes and Instruction.” *American Educational Research Journal* 17: 419–33.

Smith, Nick L., and Paul R. Brandon, eds. 2007. *Fundamental Issues in Evaluation.* New York: Guilford Press.

Smutylo, Terry. “The Output Outcome Downstream Impact Blues.” http://www.idrc.ca/en/ev-65284-201-1-DO_TOPIC.html.

Thiele, G., A. Devaux, C. Velasco, and D. Horton. 2007. “Horizontal Evaluation: Fostering Knowledge Sharing and Program Improvement within a Network.” *American Journal of Evaluation* 28 (4): 493–508.

Thiele, G., A. Devaux, C. Velasco, and K. Manrique. 2006. “Horizontal Evaluation: Stimulating Social Learning among Peers.” International Potato Center, Papa Andina Program, Lima, Peru. Draft of May 18. http://www.dgroups.org/groups/pelican/docs/Hor_Evaln_18_05.doc?ois=no.

Tilley, Nick. 2000. *Realistic Evaluation: An Overview.* Paper presented at the Founding Conference of the Danish Evaluation Society, Nottingham Trent University, September. http://www.danskevalueringsselskab.dk/pdf/Nick percent20Tilley.pdf.

Turpin, R. S., and J. M. Sinacore, eds. 1991. *Multisite Evaluations.* New Directions for Program Evaluation No. 50. San Francisco, CA: Jossey-Bass.

U.S. GAO (General Accounting Office) 1990. *Prospective Evaluation Methods: The Prospective Evaluation Synthesis.* Washington, DC. http://www.gao.gov/special.pubs/10_1_10.PDF.

———. 1992. *The Evaluation Synthesis.* Washington, DC. http://www.gao.gov/special.pubs/pemd1012.pdf.

Wandersman, A., J. Snell-Johns, B. E. Lentz, D. M. Fetterman, D. C. Keener, M. Livet, P. S. Imm, and P. Flaspoler. 2005. “The Principles of Empowerment Evaluation.” In *Empowerment Evaluation Principles in Practice,* eds. D. M. Fetterman and A. Wandersman, 27–41. New York: Guilford.

Wholey, J. S. 1987. “Evaluability Assessment: Developing Program Theory.” In *Using Program Theory in Evaluation*, ed. L. Bickman, 77–92. New Directions for Program Evaluation No. 33. San Francisco: Jossey-Bass.

Wholey, J. S., H. P. Hatry, and K. E. Newcomer, eds. 1994. *Handbook of Practical Program Evaluation.* San Francisco: Jossey-Bass.

World Bank. 1996. *World Bank Participation Sourcebook.* Washington, DC. http://www.worldbank.org/wbi/sourcebook/sba108.htm#D.

———. 2004. “Social Assessment.” I*n Turning Bureaucrats into Warriors.* 135–38.Washington, DC.

http://www.worldbank.org/afr/aids/gom/manual/GOM-Chapter percent2024.pdf.

Веб-сайти

Equator Principles. http://www.equator-principles.com/, http://www.ifc.org/ ifcext/equatorprinciples.nsf/Content/ThePrinciples.

Frechtling, Joy, and Laure Sharp Westat, eds. 1997. *User-Friendly Handbook for Mixed-Method Evaluations.* Washington, DC: National Science Foundation. http://www.nsf.gov/pubs/1997/nsf97153/start.htm.

IFC (International Finance Corporation). *Environmental and Social Policies and Guidelines.* http://www.ifc.org/ifcext/enviro.nsf/Content/PoliciesandGuidelines.

ISO (International Organization for Standardization). *ISO 14031.* http://www.iso.org/iso/catalogue_detail?csnumber=23149, http://www.iso-14001.org.uk/ iso-14031.htm, and http://www.altech-group.com/ftp/EPEarticle.pdf.

IUCN (World Conservation Union). *Sustainability Assessment.* http://www.iucn.org/themes/eval/search/iucn/sustassess.htm.

World Bank. 2009. *Social Analysis.* http://web.worldbank.org/WBSITE/ EXTERNAL/TOPICS/EXTSOCIALDEVELOPMENT/EXTSOCIALANALYSIS/ 0,,menuPK:281319~pagePK:149018~piPK:149093~theSitePK:281314,00.html.

Кластерне оцінювання

AusAid. 2004. *Governance in PNG: A Cluster Evaluation of Three Public Sector Reform Activities.* Evaluation and Review Series 35, Sydney. http://www.ausaid.gov.au/publications/pdf/governance_in_png_qc35.pdf.

Наснажуюче оцінювання

Empowerment Evaluation Blog. http://www.eevaluation.blogspot.com/.

Соціально-екологічна оцінка

Lanco Amarkantak Thermal Power. *Environmental & Social Review.* Report prepared for the IFC Board of Directors. http://www.ifc.org/ifcext/ spiwebsite1.nsf/1ca07340e47a35cd85256efb00700ceeC126975A64D3306E8 52572A00048 07BD.

Визначення оцінюваності

Lewin Group. 1997. *An Evaluability Assessment of Responsible Fatherhood Programs.*
http://fatherhood.hhs.gov/evaluaby/intro.htm.

Синтезоване оцінювання

Kruse, S. E., T. Kyllönen, S. Ojanperä, R. C. Riddell, and J. Vielaj. *Searching for Impact and Methods: NGO Evaluation Synthesis Study.* Institute of Development Studies, University of Helsinki. http://www.eldis.org/static/DOC5421.htm.

Оцінювання з орієнтацією на цілі

IFAD (International Fund for Agricultural Development). *Country Program Evaluation of the People's Republic of Bangladesh.* http://www.ifad.org/evaluation/ public_html/eksyst/doc/country/pi/bangladesh/bangladesh.htm.

Оцінювання без прив'язки до цілей

Evaluation Center, Western Michigan University.
http://www.wmich.edu/evalctr/ project-pub.html.

Мета-оцінювання

Campbell Collaboration Online Library.
http://www.campbellcollaboration.org/ frontend.aspx.

Багатоцентрове оцінювання

Australian HIV Test Evaluation Group. 1995. *Multisite Evaluation of Four Anti-HIV-1/HIV-2 Enzyme Immunoassays.*
http://www.ncbi.nlm.nih.gov/entrez/ query.fcgi?cmd=Retrieve&db=PubMed&list_uids=7882108&dopt=Abstract/.

SRI International. 2001. *A Multisite Evaluation of the Parents as Teachers (PAT) Project.* http://policyweb.sri.com/cehs/projects/displayProject.jsp?Nick=pat.

Картування наслідків

African Health Research Fellowship Program. *Strengthening Leadership Capacity to Improve the Production and Use of Health Knowledge in Africa.* http://www.idrc
.ca/fr/ev-34425-201-1-DO_TOPIC.html.

Оцінювання із залученням зацікавлених сторін

Contra Costa Health Services. 2009. *Preventing Chronic Disease: A Participatory Evaluation Approach.* http://www.cchealth.org/groups/chronic_disease/guide/ evaluation.php.

Community Integrated Pest Management. 2001. *Picturing Impact: Participatory Evaluation of Community IPM in Three West Java Villages.* http://www.c ommunityipm.org/docs/Picturing percent20Impact/Picturing percent20 Impact percent20top percent20page.html.

Перспективне оцінювання

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 1998. *Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya.* http://www.econ.yale.edu/~egcenter/infoconf/kremer_paper.pdf.

Швидка оцінка

UNICRI (United Nations Interregional Crime and Justice Research Institute), and ACI (Australian Institute of Criminology). *Global Programme against Trafficking in Human Beings, Rapid Assessment: Human Smuggling and Trafficking from the Philippines.* http://www.unodc.org/pdf/crime/trafficking/ RA_UNICRI.pdf.

Соціальна оцінка

World Bank. 1996. “Morocco: Fez Medina Rehabilitation Project.” In *World Bank Participation Sourcebook.* Washington, DC: World Bank. http://www.worldbank.org/wbi/sourcebook/sba108.htm#D.

Оцінювання з орієнтацією на практичне застосування

Evaluation Center, Western Michigan University. *Utilization-Focused Evaluation Checklist.* http://www.wmich.edu/evalctr/checklists/ufechecklist.htm.

ПЛАНУВАННЯ І ПРОВЕДЕННЯ

«Справи можуть ускладнюватися нескінченно, адже щось одне завжди тягне за собою інше».

—Е.Б. Уайт

Глава 6: Розробка питань оцінювання і початок заповнення матриці плану

- Джерела питань
- Види питань
- Визначення й вибір питань
- Як ставити хороші питання
- Планування оцінювання

Глава 7: Вибір дизайну причинно-наслідкових, нормативних та описових питань оцінювання

- Встановлення зв’язку між питаннями й планом
- Планування причинно-наслідкових питань
- Планування описових питань
- Планування нормативних питань
- Потреба у ретельнішому плануванні

Глава 8: Вибір і створення інструментів для збору даних

- Стратегії збору даних
- Характеристики хороших засобів контролю
- Кількісні і якісні дані
- Інструменти для збору даних

Глава 9: Вибір стратегії формування вибірки

- Введення до формування вибірки
- Види вибірок: випадкові та невипадкові
- Визначення розміру вибірки

Глава 10: Планування і проведення аналізу даних

- Стратегія аналізу даних
- Аналіз якісних даних
- Аналіз кількісних даних
- Поєднання кількісних і якісних даних

ГЛАВА 6

Розробка питань оцінювання і початок заповнення матриці плану

Це перша з п’яти глав, в яких обговорюються конкретні кроки з планування оцінювання. В цій главі розглядаються види питань оцінювання та пояснюється, коли варто застосовувати кожне з них. Також описується, як складати і структурувати хороші питання.

Ця глава складається з п’яти основних частин:

- Джерела питань
- Види питань
- Визначення й вибір питань
- Як ставити хороші питання
- Планування оцінювання

Джерела запитань

Чому вибір питань оцінювання є настільки важливим? Одна з причин полягає в тому, що питання задають напрям оцінювання та його форму (це питання повніше розглядається в главі 6). Питання оцінювання є критичним елементом, що допомагає ключовим особам і групам покращувати ефективність роботи, приймати рішення й інформувати суспільство. Визначення й формулювання питань, на які має відповісти оцінка – критично важливий процес. Він потребує проведення ретельних досліджень (Fitzpatrick, Sanders, and Worthen 2004).

Фахівці з оцінювання ставлять питання для того, щоб отримати інформацію про проект, програму або політику, яку вони оцінюють. Розповсюджена помилка при розробці питань – вважати, що всі залучені особи однаково розуміють цілі інтервенції. Наприклад, у питанні «Чи допомогла програма учасникам?» різні зацікавлені особи можуть по-різному тлумачити слова «допомогла» і «учасники». Цю проблему можна вирішити, узгодивши теорію змін (розглядалася в попередній главі).

Фітцпатрик, Сандерс і Вортен (Fitzpatrick, Sanders, and Worthen 2004) наводять перелік джерел, використання яких дозволяє фахівцям з оцінювання враховувати різноманітні точки зору:

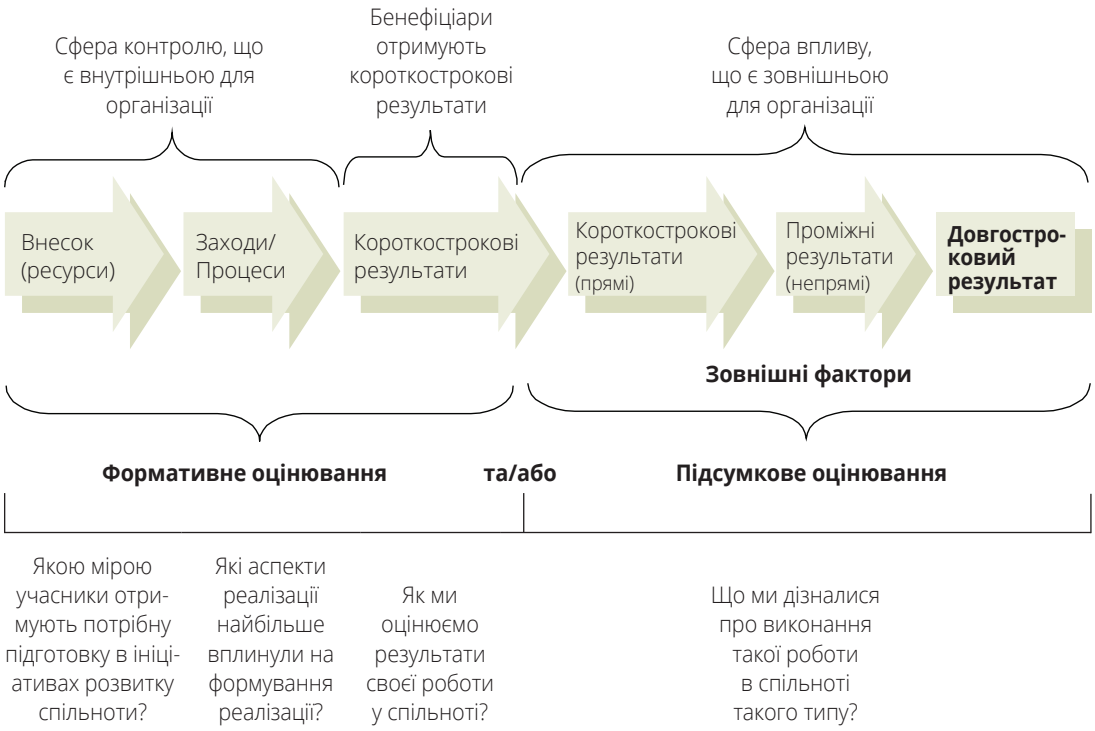
- питання, інтереси й цінності зацікавлених осіб;
- моделі оцінювання;
- схеми й підходи, в тому числі евристичні (метод проб і помилок);
- результати досліджень і оцінювань, а також важливі проблеми, що порушуються в літературі;
- професійні стандарти, контрольні списки, настанови, інструменти та критерії, розроблені або використані деінде;
- думки та знання експертів-консультантів;
- власні професійні судження фахівця з оцінювання.

В главі 4 розповідалося про те, як виявляти коло зацікавлених осіб та разом з ними визначати, які питання потребують оцінки. Також згадувалося про важливість вивчення попередніх досліджень і оцінок для формулювання питань. Було приділено багато уваги розробці та використанню моделей теорії змін для визначення пріоритетів оцінки. Питання зумовлюються головними припущеннями, на яких ґрунтується модель.

На рис. 6.1 показано види питань оцінювання, які слід ставити на різних етапах. Загальні питання в нижній частині схеми показують, що формативні питання можна виводити із заходів і безпосередніх результатів, а сумативні – з проміжних та довгострокових результатів. Питання, виведе-

ні з короткострокових результатів, можна записувати як формативні або сумативні. Питання повинні випливати з основних припущень, зроблених в рамках логічної моделі про те, як працюватиме програма, та які вигоди або довгострокові результати будуть отримані. Ще одним джерелом питань є вивчення проведених оцінювань аналогічних програм, а також різних точок зору зацікавлених осіб на проект, програму або політику (див. главу 4).

Рис. 6.1. Використання логічної моделі для формулювання питань оцінювання



Види питань

В ході планування оцінювання можна розглянути багато питань. Всі вони повинні бути чітко визначені і вимірні.

Питання можна поділити на три групи: описові, нормативні та причинно-наслідкові. Види поставлених питань (а також наявні дані, час і кошти) визначатимуть план оцінювання.

Описові питання

Описове питання:
Питання про те, що є

Описові питання – це питання про те, що є. Описуються різні аспекти процесу, умови, парадигми або комплексу організаційних відносин чи мереж. Петон (Patton 2002) називає описові питання фундаментом оцінювання.

Описові питання:

- спрямовані на розуміння або опис програми чи процесу;
- фіксують поточний стан справ;
- прямі (хто, що, де, коли, як, скільки);
- можуть використовуватися для описання залучених ресурсів, заходів та короткострокових результатів;
- часто використовуються для збору інформації про думки клієнтів програми. Приклади описових питань:
- Якими є цілі програми з точки зору різних зацікавлених осіб?
- Якими є основні програмні заходи?
- Як люди стають учасниками програми?
- Де було впроваджено програму?
- Які послуги програма надає чоловікам? Які послуги вона надає жінкам?
- Якими є наслідки програми для учасників?
- Якою мірою план програми враховує досвід попередніх аналогічних програм?
- Наскільки помітними є відмінності у реалізації програми в різних місцях?
- Яким є фаховий рівень постачальників послуг?
- Коли було впроваджено програму?
- Скільки жінок брали участь у програмі?
- Наскільки високою є вартість програми порівняно з вартістю аналогічних програм?
- Які існують неформальні канали комунікацій всередині організації?
- Наскільки корисною, на думку учасників, була для них програма?

Оціночні питання про формування політики часто є описовими. Ріст (Rist 1994) описує процес «політичного циклу» з трьома етапами, кожному з яких властиві свої запитання (table 6.1).

Нормативні питання

Нормативне питання:
Питання, яке порівнює те, що є, з тим, що має бути

Нормативні питання порівнюють те, що є, з тим, що має бути. Вони порівнюють поточний стан речей зі встановленою ціллю, метою або орієнтиром. Такі питання своєю орієнтованістю на виконання завдань схожі з питаннями, які ставляться під час аудиту ефективності. Зокрема:

- Чи робимо ми те, що маємо робити?
- Чи досягаємо ми мети?
- Чи вдалося нам досягти того, чого ми прагнули?

Якщо програма використовує систему моніторингу з орієнтацією на результат, якщо визначені показники, цільові значення і строки, нормативні питання можна ставити, щоб отримати інформацію про залучені ресурси, заходи і короткострокові результати.

Іноді програма має завдання, але відсутні чіткі критерії щодо того, як вони мають вимірюватися чи досягатися. Показники або цілі не визначені. Тоді у фахівця з оцінювання є кілька варіантів, і деякі з них кращі за інші (блок 6.1).

Пошук стандартів або критеріїв зазвичай починається з критеріїв, наведених в керівних документах програми – таких як закони або дозвільні документи, видані керівним органом. Критерії також можуть вказуватися як показники з певними цільовими значеннями в орієнтованих на результат системах управління. Інші органи, що можуть встановлювати стандарти – це системи акредитації, повноважні комісії, професійні організації та інші комісії.

Таблиця 6.1. Якісні питання, що ставляться на різних етапах політичного циклу

Етап	Приклади
Формулювання політики	• Що робилося раніше у відповідь на таку ситуацію або проблему? • Яка є інформація про попередні дії, що допомогла б особі, яка приймає рішення, зробити вибір з наявних варіантів?
Впровадження політики	• Яка інформація про проблему або обставину зумовила вибір політики чи програми? • Що зробила організація або установа у відповідь на ініціативу? • Якою є кваліфікація відповідальних за впровадження? • Наскільки зацікавленими є керівництво й персонал? • Які діють засоби управління стосовно розподілу ресурсів? • Чи адекватно відображає організаційна структура вимоги проведення ініціативи? • Які засоби для вибору між конкуруючими вимогами є в організації? • Які впроваджено системи зворотного зв'язку для допомоги керівникам?
Політична підзвітність	• Наскільки відповідає програма або політика обставинам? • Наскільки гострою була увага до відповідального ставлення?

Джерело: Взято з Rist 1994.

Блок 6.1 «Оцінювання інтервенції, в якої відсутні стандарти»

Розглянемо багатосекторну програму з наступними завдання:

- Покращити оцінки за читання серед дітей у вибраних шкільних округах.
- Підвищити обізнаність про ВІЛ/СНІД та способи його профілактики в регіоні.
- Збільшити кількість мікропідприємств у селищі та підвищити їх прибутковість.

Фахівець з оцінювання прагне знати, як конкретно трактувати ці завдання (яка частина дітей має покращити свої оцінки, чия обізнаність про ВІЛ/СНІД збільшиться та як ми дізнаємось про кількість мікропідприємств, чий бізнес розширився, а також про обсяги такого розширення?), але в програмі такої інформації немає. Що може зробити фахівець з оцінювання? Один підхід – попрацювати з відповідальними за програму, тобто посадовцями, що займаються адмініструванням програми або її впровадженням. Ці посадовці мають бути здатні вказати обґрунтовані рівні показників, яких повинна досягти програма.

Проблема полягає в тому, що одна група може не сприймати встановлені іншою групою стандарти. Наприклад, працівники з наглядовими функціями можуть не погоджуватися зі стандартами, запропонованими реалізаторами програми. Вони можуть вважати, що ці стандарти є занадто низькими.

Інший підхід – запросити одного або кількох експертів для того, щоб вони узгодили спільний стандарт. Потенційна проблема цього підходу в тому, що такий стандарт може відображати особисту упередженість експерта. Але можна запросити кількох експертів. В такому випадку важливо, щоб група експертів сприймалася як політично нейтральна або врівноважена та щоб експерти попередньо не були залучені до програми.

Найслабша та найризикованіша альтернатива – це встановлення стандарту самим фахівцем з оцінювання. Цього підходу слід уникати, адже він створює для фахівця з оцінювання проблеми. Працівники програми можуть вважати, що стандарти зависокі або занизькі, або ж що не можна встановлювати стандарти для оцінки ефективності пост-фактум.

Джерело: Автори.

Приклади нормативних питань:

- Чи освоїли ми бюджет?
- Чи досягли ми цілі прийняти 5000 учнів за рік?
- Чи вдалося нам вакцинувати 80 відсотків дітей, як планувалося?
- Чи виконали ми завдання з осушення 100 000 гектарів землі?
- Чи був процес вибору учасників справедливим і неупередженим?

Причинно-наслідкові питання

Причинно-наслідкові питання:
Питання, що визначає, які зміни забезпечила інтервенція

Причинно-наслідкові питання визначають, які зміни забезпечила інтервенція. Такі питання ще нерідко називають питаннями про довгострокові результати, про впливи або атрибутивними. Вони спрямовані на виявлення змін, забезпечених інтервенцією. Причинно-наслідкові питання мають з'ясовувати наслідки проекту, програми або політики. Це питання про те, що трапилось. Причинно-наслідкові питання – це питання про те, чи досягла програма бажаних результатів.

Моделі програмної теорії змін зображають бажані довгострокові результати і впливи інтервенції, але довгострокові результати не обов'язково визначаються причинно-наслідковими питаннями. Приміром, у програмі з представлення фермерам нових, покращених сортів зернових питання про довгостроковий результат – це питання «Чи збільшився урожай?» Це описове питання: йдеться про те, наскільки збільшився врожай зернових. Якщо ж у ході оцінки питається, чи збільшився врожай в результаті реалізації програми, а не, наприклад, завдяки надзвичайно сприятливій погоді, тоді це буде причинно-наслідкове питання. Причинно-наслідкові питання передбачають порівняння за одним або кількома показниками не тільки до і після інтервенції, а й за умови її наявності й відсутності.

Приклади причинно-наслідкових питань:

- Чи ефективною була стратегія партнерства трьох країн для збереження біорізноманіття регіону і підтримання засобів до існування?
- Чи отримали учасники програми фахової підготовки більш високооплачувані роботи, ніж мали б без навчання?
- Чи скоротила програма підтримки мікропідприємств рівень бідності в населених пунктах, де вона працювала?
- Чи допомогло підвищення податку на бензин покращити якість повітря?
- Чи допомогло збільшення урядом розміру штрафів для компаній-порушниць скоротити використання дитячої праці у швейній промисловості?
- Які інші впливи або побічні ефекти (позитивні або негативні) мала ця інтервенція для широкого загалу?

Фахівці з оцінювання мають ставити такі питання в розрізі причин і наслідків. Через те, що одночасно проводиться велика кількість заходів, важко продемонструвати, що довгострокові результати повністю (або принаймні переважно) зумовлені інтервенцією. Шукаючи відповіді на причинно-наслідкові питання, фахівці з оцінювання мають з великою обережністю виключати інші можливі пояснення змін.

В главі 7 розглядаються способи пошуку відповідей на причинно-наслідкові питання та види аналізу, необхідні для встановлення причинно-наслідкового зв'язку. Через те, що відповідати на цей вид питань важче, ніж на описові або нормативні питання, використовувати його варто лише тоді, коли це справді необхідно.

Багато оцінок застосовують лише описові та нормативні питання, особливо якщо це формативні оцінки, в центрі уваги яких знаходиться реалізація інтервенції. Оцінки, зосереджені на впливі, ставлять причинно-наслідкові питання, але вони зазвичай також включають деяку кількість описових та нормативних питань. Блок 6.2 демонструє, як оцінка може включати різні види питань.

Блок 6.2 «Оцінювання політик та інтервенцій за допомогою причинно-наслідкових питань»

Поліпшення профілактичних медичних послуг

Політика: Забезпечити надання всім дітям профілактичних медичних послуг.

Ціль: Зменшити смертність серед немовлят і дітей дошкільного віку.

Питання оцінювання:

1. Який відсоток дітей отримав профілактичні медичні послуги з початку програми? (описове питання).
2. Чи отримали профілактичні медичні послуги діти із цільових груп з низькими доходами? (нормативне питання).
3. Чи знизився завдяки програмі рівень дитячої смертності? (причинно-наслідкове питання).

Підготовка школярів середньої школи до ринку праці

Політика: Забезпечити навчання учнів середніх шкіл знанням і навичкам, потрібним для працевлаштування на місцевих ринках.

Ціль: Забезпечити, щоб випускники могли виконувати добре оплачувану кваліфіковану роботу.

Питання оцінювання:

1. Як середні школи готують учнів для виконання тієї роботи, що її пропонує місцевий ринок праці? (описове питання).

2. За рік по тому, чи отримують випускники вищі зарплати, ніж ті школярі, що залишили програму? (описове питання).
3. Наскільки враховують актуальний стан ринку середні школи, приймаючи рішення щодо того, в яких сферах треба готувати учнів? (нормативне питання).
4. Наскільки більше заробляють випускники порівняно з тим, скільки вони б заробляли без програми?

Надання безкоштовної імунізації від кору

Інтервенція: Сімейні клініки надають безкоштовну імунізацію від кору всім дітям до п'яти років у трьох регіонах країни протягом одного року.

Питання оцінювання:

1. Як клініки інформували батьків про можливість безкоштовної імунізації дітей? (описове питання).
2. Чи досягла програма цілі з надання імунізації від кору 100 відсоткам дітей віком до п'яти років у трьох регіонах протягом минулого року? (нормативне питання).
3. Чи використовувала програма інноваційні способи для охоплення дітей з категорій найвищого ризику? (описове питання).
4. Чи зменшилася за результатами програми пропорція дітей, які хворіють на кір? (причинно-наслідкове питання).
5. Чи вдалося програмі забезпечити зниження дитячої смертності, пов'язаної з ускладненнями кору? (причинно-наслідкове питання).

Впровадження шкільної програми з урахуванням потреб ринку

Інтервенція: Три середні школи в трьох містах впроваджують шкільну програму з урахуванням потреб ринку.

Питання оцінювання:

1. Наскільки відмінною є навчальна програма порівняно з програмами шкіл, що не брали участі в програмі? (описове питання).
2. Чи достатньою мірою шкільна програма враховувала потреби ринку? (нормативне питання).
3. Наскільки вдалося випускникам цих шкіл отримати високооплачувані робочі місця? (описове питання).
4. Чи отримали в результаті інтервенції випускники шкіл з новою програмою більш високооплачувану роботу порівняно з тим, якби програма залишалася незмінною? (причинно-наслідкове питання).

Джерело: Автори.

Відношення між видами питань і теоріями змін

Як співвідносяться види питань з теоріями змін, що розглядалися в попередній главі? Питання про те, чи були залучені ресурси та досягнуті цільові короткострокові результати, є зазвичай нормативними: чи було на витрачені кошти придбано необхідну кількість товарів та послуг за вказаний період? Питання щодо досягнення цільових довгострокових результатів також є нормативними. Питання про те, наскільки покращилися оцінки за читання за певний період часу – описове; питання про те, чи піднялися оцінки за читання до заданого керівниками програми рівня – нормативне. Питання, що перевіряють, чи принесла інтервенція зміни, яких за її відсутності не сталося б – причинно-наслідкові питання. Питання про те, чи привели ці зміни до інших проміжних результатів або впливів – також причинно-наслідкові.

Питання про зміни в довгострокових результатах – або описові, або погано сформульовані причинно-наслідкові. Якщо питання задумувалося як причинно-наслідкове, його варто переформулювати, щоб було зрозуміло: це питання не лише про те, які зміни відбулися, а й про те, чи зумовлені вони інтервенцією (тобто зміна має бути більшою, ніж вона була б за відсутності інтервенції). Інші довгострокові результати та впливи також можуть бути віднесені на рахунок інтервенції відповідно до теорії змін.

Визначення й вибір питань

Як фахівцю з оцінювання вирішити, які з безлічі можливих питань слід обрати? Кронбах (Cronbach 1982) пропонує визначати й обирати питання в два етапи – етап накопичення й етап звуження.

Етап накопичення

Етап накопичення[6]:
Етап оцінки, коли фахівець з оцінювання складає вичерпний список потенційно важливих питань

На **етапі накопичення** фахівець з оцінювання складає вичерпний список потенційно важливих питань. Відкидаються лиш деякі питання; використовується багато джерел. Кронбах (Cronbach (1982)) підсумовує сутність етапу накопичення в плануванні оцінювання наступним чином:

> Перший крок – відкрити свій розум питанням, що можуть хоч якось пролити світло на предмет вивчення. Цей етап вже сам по собі є оцінкою, що вимагає збору даних, обґрунтованого аналізу та прийняття рішень.

[6] *В оригіналі термін звучить як Divergent phase. В україномовній літературі можна зустріти варіант «дивергентна фаза», «фаза розходження».*

Дуже невелика частина цієї інформації та аналізу є кількісною. Дані отримуються у неформальних бесідах, від випадкових спостережень та перегляду існуючих записів. Природничі та якісні методи набагато краще підходять для такої роботи через те, що, враховуючи погляди учасників і зацікавлених сторін, вони дозволяють фахівцю з оцінювання виявляти сподівання й страхи, що могли ще не проявитися як політичні питання.

Фахівцю з оцінювання слід подивитись на програму очима різних представників спільноти, що приймає рішення, в тому числі фахівців, які будуть проводити програму в разі її прийняття, а також громадян, які мають отримувати її послуги.

У певний момент нові питання перестають поставати. Тоді фахівцю з оцінювання варто зупинитися, переглянути список питань і зайнятися їх упорядкуванням.

Етап звуження

Етап звуження[7]: Етап оцінки, коли фахівець з оцінювання звужує перелік питань, визначених на етапі накопичення, щоб залишити найкритичніші з них

На **етапі звуження** фахівець з оцінювання звужує перелік питань, визначених на етапі накопичення, щоб залишити найкритичніші з них. Як фахівець з оцінювання визначає, які саме питання є найкритичнішими? Фітцпатрик, Сандерс і Вортен (Fitzpatrick, Sanders, and Worthen 2004) пропонують наступні критерії для визначення того, які із запропонованих питань оцінки слід розглядати:

- Хто застосовуватиме інформацію? Кому це цікаво? Хто буде найбільше розчарований, якщо це питання буде виключене з оцінювання?
- Чи зменшить відповідь на це питання непевність, чи надасть інформацію, якої зараз немає?
- Чи дасть відповідь на питання важливу інформацію? Чи матиме це вплив на хід подій?
- Це питання становить для когось лише скороминущий інтерес, чи воно стосується критичних вимірів постійної зацікавленості?
- Чи стане оцінка помітно менш комплексною, якщо видалити це питання?
- Чи раціональним буде пошук відповіді на це питання з урахуванням фінансових та людських ресурсів, часу, методів та технологій?

Цей список критеріїв можна звести у матрицю, щоб допомогти фахівцю з оцінювання та клієнту звузити початковий перелік питань до зручного в роботі набору (див. табл. 6.2).

[7] *В оригіналі термін звучить як Convergent phase. В україномовній літературі можна зустріти варіант «конвергентна фаза», «фаза звуження».*

Фахівець з оцінювання особливу увагу має приділяти питанням, що їх ставлять клієнт і головні зацікавлені особи. Якщо існують розходження у поглядах на питання, важливо буде подолати їх вже на цьому ранньому етапі. Даний процес допомагає фахівцю з оцінювання й клієнту, а також іншим ключовим зацікавленим особам отримати відчуття спільної відповідальності або партнерства, що може виявитися дуже цінним на пізніших етапах оцінки.

Таблиця 6.2. Матриця вибору й ранжування питань оцінювання

Питання оцінювання										
Чи питання оцінювання:	1	2	3	4	5	6	7	8	9	10
буде цікавим для цільової аудиторії?										
зменшить існуючу невизначеність?										
надасть важливу інформацію?										
викликає постійний (не скороминущий) інтерес?										
буде критичним для охвату та комплексності дослідження?										
матиме вплив на хід подій?										
є таким, на яке можна відповісти за наявних фінансових і людських ресурсів, часу, методів і технологій?										

Джерело: Взято з Fitzpatrick, Sanders, and Worthen 2004.

Як ставити хороші питання

Щоб поставити хороші питання, фахівець з оцінювання спершу визначає головні проблеми, які має вирішувати проект, програма або політика. Як вже зазначалося вище, головні проблеми зазвичай визначаються шляхом аналізу відповідної літератури, включаючи оцінювання аналогічних програм, теорії змін та програмні документи, а також в ході бесід із зацікавленими особами програми і замовником оцінювання.

Наведемо приклади головних проблем, що можуть вивчатися в ході оцінки програми, спрямованої на зниження дитячої смертності:

- численні причини дитячої смертності;
- конкуруючі поточні програми;
- дієвість методів, що застосовуються для охоплення матерів з низькими доходами;

- обсяг і природа використання харчових додатків не за цільовим призначенням.

Після визначення цих проблем фахівець з оцінювання може ставити питання, що допоможуть вирішити, чи вплинула на них політика або інтервенція. Ось приклади питань, що використовуються для отримання інформації щодо проблем:

- Які методи охоплення аудиторії використовувалися програмою?

Які методи охоплення були найефективнішими?

- Яким був рівень захворюваності на небезпечні для життя дитячі хвороби протягом дії програми?
- Наскільки знизився рівень смертності?
- Які інші заходи з покращення здоров'я матерів проводилися?

Слід уникати питань, що охоплюють більш ніж одну проблему (наприклад: «Скільки жінок пройшли медичне обстеження і отримували харчові добавки?»). Замість цього питання слід розбити на два: «Скільки жінок пройшли медичне обстеження?» та «Скільки жінок отримували харчові добавки?»

Проблему можна розглядати за допомогою всіх трьох видів питань, змінюючи формулювання. Наприклад, оцінювання програми, покликаної зменшити кількість поранень і смертей від протипіхотних мін, може передбачати наступні питання:

- Де частішше за все трапляються інциденти з протипіхотними мінами? (описове).
- Чи досяг проект цілі – знищення 1000 протипіхотних мін в регіоні за визначений проміжок часу? (нормативне).
- Чи зменшилася в результаті проведення інтервенції кількість людей, які отримали поранення або були вбиті внаслідок підривання протипіхотних мін? (причинно-наслідкове).

Наступні рекомендації можуть допомогти фахівцям з оцінювання скласти кращі питання.

- Встановіть чіткий зв'язок між кожним питанням оцінки та метою дослідження.
- Впевніться, що питання оцінювання торкаються найважливіших проблем.
- Переконайтеся, що на всі питання можна дати відповідь.
- Оберіть реалістичну кількість питань.
- Зосередьтеся на важливих питаннях, на які необхідно отримати відповідь, а не на тих, які було б цікаво дослідити.
- Переконайтеся, що на питання можна дати відповіді з урахуванням строків і наявних ресурсів оцінки.

- Подумайте про час проведення оцінювання в контексті циклу програми. Наприклад, питання про вплив краще ставити після того, як інтервенція в повному обсязі ведеться протягом кількох років.

Питання оцінювання можуть стосуватися проекту, програми, всеосяжної політичної проблеми, конкретної проблеми або конкретної інтервенції, пов'язаної з політикою. Наприклад, якщо загальною метою (політичною проблемою) є боротьба з бідністю, може бути започатковано цілу низку програмних інтервенцій. Кожна політика перетворюється на дію за допомогою інтервенції, покликаної досягти певних цілей. Кінець кінцем, якщо політика й інтервенції виявляються дієвими, а теорія змін – правильною, загальні цілі також будуть досягнуті. Якщо це не так, потрібно переглянути як інтервенції, так і політику. Можеливо, необхідно змінити один з цих елементів або й обидва.

Планування оцінювання

Так само, як архітектор проектує будівлю, фахівець з оцінювання планує оцінку. План оцінювання повинен вказувати на те, з яких елементів складатиметься оцінка. Це ще не повноцінний робочий план дослідження.

План оцінки складається з наступних елементів:

- головна проблема або питання оцінювання;
- загальний підхід оцінювання;
- конкретні питання й підпитання оцінювання;
- формування концепції (виміри або показники), джерела даних та методологічні стратегії для того виду збору даних, що буде використовуватися;
- запланований аналіз;
- стратегія розповсюдження.

Петон (Patton 1997) виділяє два напрями планування: концептуальний і технічний. Концептуальне планування зосереджується на тому, що учасники думають про оцінювання. Розглядаються такі проблеми, як визначення мети оцінювання та його основних зацікавлених осіб, а також політичних моментів, які треба мати на увазі.

Технічне планування охоплює збір та аналіз даних. Ці технічні питання лежать в основі матриці плану, яка має складатися для кожного оцінювання. Для кожного питання або, що частіше, підпитання матриця плану вимагає:

- визначити вид питання або підпитання, що ставиться (описове, нормативне, причинно-наслідкове);
- вказати міру (показник або змінну), за якої буде відшукуватися відповідь на питання або підпитання (наприклад, відсоткове зростання забезпеченості житлом в регіоні або кількість вакцинованих дітей);

- визначення методологічної схеми, що забезпечить достатню інформацію для відповіді на описові, нормативні або причинно-наслідкові питання;
- визначення джерел даних для кожного питання або підпитання;
- рішення щодо того, чи потрібна схема вибірки, і якщо так, то якою вона буде;
- визначення виду інструмента(-ів) для збору даних, що використовуватиметься для кожного питання або підпитання;
- визначення шляхів аналізу й представлення даних.

Іноді міра, за якої має даватися відповідь на питання – це узгоджений показник з чітко визначеним цільовим значенням і граничним терміном. Це ідеальна ситуація; так частіше за все трапляється, коли для інтервенції було створено систему моніторингу й оцінки, або коли інтервенція є частиною більшої системи моніторингу сектора чи, приміром, міністерства. Незалежно від того, чи встановлено ціль, необхідно вказати наявність або відсутність базових рівнів.

Заповнена матриця оцінки є планом оцінювання. Це ще не повноцінний робочий план: він не включає всі завдання, не вказує, хто й коли буде виконувати кожне із завдань. Повний план робіт описується в главі 12.

Етапи процесу планування оцінювання

В ідеалі процес оцінювання треба починати заздалегідь, з початкового плану програми. Після цього воно проходить кілька окремих і важливих етапів.

Етап 1: Початкове планування або визначення масштабів оцінювання

Етап початкового планування або визначення оцінювання слугує для роз'яснення природи й масштабу оцінки. Протягом цього етапу визначаються головна мета оцінювання, коло зацікавлених осіб, з якими треба провести консультації, особа, що проводитиме оцінку, а також терміни надання результатів. Це період розвідувальної роботи. Виявляються ключові проблеми з точки зору головного клієнта та інших зацікавлених осіб, аналізується література, визначається, які інтервенції можуть вплинути на програму. Розробляються або вдосконалюються теорія змін та припущення, на яких вона ґрунтується.

Етап 2: Планування оцінювання

В кінці етапу початкового планування або визначення масштабів повинно бути накопичено достатньо знань про контекст оцінювання для того, щоб прийняти рішення про загальний підхід. Етап планування оцінювання є центральним у його підготовці, а кульмінацією – матриця плану оцінювання.

Недосконалий загальний план створить обмеження для можливості робити висновки щодо ефективності інтервенції.

Хорошою практикою є представлення загального плану спонсору оцінювання (клієнту) та іншим ключовим зацікавленим особам з відповідним його обговоренням до формулювання кінцевої версії. Це дає змогу уникнути несподіванок, сприяє формуванню позитивного сприйняття і підтримки оцінки. План оцінювання також варто перевірити із залученням консультативної групи та застосуванням дружньої перевірки. В особливо відповідальних випадках проекти плану можна розміщувати на веб-сайті для отримання коментарів.

Матриця плану може використовуватися як основа для розробки технічного завдання (ТЗ). ТЗ може слугувати основою для запиту пропозицій або як керівництво для групи оцінювання, якщо проводиться внутрішня оцінка. Якщо визначення масштабу і підготовча робота проводяться зовнішнім консультантом, до переліку документів, які він має надати за результатами своєї роботи, варто включити матрицю. Згодом можна розробити інше ТЗ для впровадження плану оцінювання.

Етап 3: Проведення оцінювання

«Фаза дій» оцінювання передбачає збір та аналіз даних. Як правило, якщо потрібно зібрати різні види даних (або однакові види даних з різних джерел), необхідно розробити і випробувати різні інструменти. Аналіз часто проводиться паралельно зі збором даних.

Пройшовши приблизно дві третини процесу збору даних, група оцінювання має провести узгоджувальну нараду для розгляду наявних результатів та виявлення нових тем і головних висновків. **Узгоджувальна нарада** – ефективний спосіб вже на ранніх етапах досягти згоди стосовно трьох–п'яти головних висновків. Мета проведення узгоджувальної наради – одразу домовитися про головні теми і переконатися, що головна проблема або головне питання оцінки не залишилося поза увагою. (Хоча звіт можна вибудувати навколо питань оцінки, формування заключного звіту і спілкування з особами, що приймають рішення, з опорою на тему або висновки може бути більш дієвим підходом, адже, скоріше за все, не всі питання оцінки викликають однаковий інтерес.)

Узгоджувальна нарада:

Збори, на яких група оцінювання обговорює і узгоджує головну тему, а також перевіряє, що головна проблема, яку розглядає оцінка, не залишилась поза увагою

Етап 4: Звітування про результати оцінювання

На етапі складання звітів можна поділитися першими висновками або констатаціями з особами, що відповідають за програму, і обговорити їх. Це дозволить виправити фактичні помилки і перевірити нові відомості до складання звіту і розробки рекомендацій. Після завершення аналізу підсумовуються результати, перевіряються проекти звітів, у звітах враховуються

відповідні зауваження, і клієнту та головним зацікавленим особам надається заключний звіт.

Звіт зазвичай містить відомості про вихідні дані й контекст оцінювання, вказує його мету, описує масштаб і методологію, а також викладає результати (вказуються як заплановані, так і незаплановані результати). Він також включає інформацію про набутий досвід і рекомендації. Розуміти, що і чому працює погано – не менш важливо, ніж розуміти, що і чому працює добре, тож необхідно з'ясувати обидва ці аспекти. Звіт слід складати з урахуванням цільової аудиторії; він має бути позбавлений вузькофахових термінів і легко читатися (написання звітів розглядається в главі 13).

Етап 5: Поширення результатів оцінювання й подальша робота

Планування оцінювання означає і планування спілкування, причому не лише з клієнтом та ключовими зацікавленими особами, а й усередині групи оцінювання. Оцінка не може вважатися завершеною, поки не буде завершене її поширення: таким чином, розробка плану поширення також є частиною процесу планування. Результати не обов'язково представляти в друкованому вигляді. Дієвим інструментом для поширення поточних результатів ще не завершеного оцінювання є брифінги – зустрічі для стислого інформування. Вони можуть бути особливо актуальні, якщо результати є неочікуваними або критично важливими.

Результатом багатьох оцінювань є дії із:

- модифікації інтервенції;
- усунення виявлених в ході оцінювання перешкод;
- інформування майбутніх політик або інтервенцій (доопрацювання теорії змін);
- надання рекомендацій іншим спеціалістам у зв'язку з набутим досвідом;
- зміни уявлень про сутність проблеми.

Багато організацій використовують системи подальшої роботи для відстеження формальних рекомендацій та аналізу набутого досвіду. Пошук в базі даних за темою, сектором, регіоном і датою підвищує його ефективність. Незалежно від того, чи є такі системи в наявності, фахівцям з оцінювання варто подумати про те, щоб надіслати електронну копію із захистом від редагування до відповідних баз знань оцінки. Також варто представити результати своєї роботи на конференціях з оцінки або опублікувати відповідні статті в професійних журналах.

Зв'язок між етапами

Характеристики різних етапів зведено в блоці 6.3. Зв'язки між різними компонентами представлено на рис. 6.2.

На рис. 6.2 показано центральну роль, що її відіграє активізація використання оцінки для всього процесу оцінювання. Всі інші процеси й етапи оцінювання повинні орієнтуватися на забезпечення активного застосування результатів на практиці. Це важливо.

Блок 6.3 «П'ять етапів процесу оцінювання»

Етап 1: Початкове планування або визначення масштабів оцінювання

Отримайте повне розуміння програми, проекту або політики:

- Проведіть зустріч з головним клієнтом оцінювання.
- Визначте коло інших ключових зацікавлених осіб і проведіть зустрічі з ними.
- Дослідіть контекст програми і зберіть вихідні матеріали.
- Знайдіть релевантні пов'язані оцінки.
- Вивчіть попередні оцінювання й використані питання, плани і стратегії збору даних.
- Зустріньтеся із персоналом програми (якщо ви – зовнішній фахівець з оцінювання).
- Перегляньте і вдоскональте або створіть теорію змін для програми.

Етап 2: Планування оцінювання

Визначте питання й проблеми:

- Разом з клієнтом визначте головну мету оцінки, проблемні аспекти й критичні часові обмеження.
- Визначте коло інших ключових зацікавлених осіб і зустріньтеся з ними, щоб визначити питання й проблеми, що можуть бути охоплені оцінкою.
- Визначте, які є ресурси (бюджет на залучення консультантів і відрядження, члени групи, набір навичок і т. д.).
- Оцініть потреби зацікавлених осіб, включаючи часові обмеження.

Підготуйте технічне завдання й матрицю оцінки:

- Визначте вид оцінювання.
- Визначте конкретні питання й підпитання оцінювання.
- Оберіть засоби вимірювання для кожного питання й підпитання.
- Визначте джерела даних для роботи з кожним питанням або підпитанням.
- Виробіть належне формулювання для кожного питання або підпитання.
- Створіть стратегію збору даних, включаючи інструменти й методи визначення вибірки для кожного питання або підпитання.

- Розробіть стратегію аналізу даних.
- Визначте потреби в ресурсах і часі.

Етап 3: Проведення оцінювання

- Поінформуйте клієнта й ключових зацікавлених осіб про план оцінювання.
- Підготуйте робочий план, включаючи аналіз і випробування методології, інструменти попереднього випробування, навчання працівників, що збирають дані, а також розробку протоколу.
- Зберіть дані.
- Підготуйте дані для аналізу, розробивши шаблонні таблиці (якщо це не виконано в рамках планування оцінювання) і очистивши дані.
- Проаналізуйте дані.
- Створіть графічне представлення.
- Сформулюйте результати.

Етап 4: Звітування про результати оцінювання

- Проведіть узгоджувальну нараду.
- Визначте головні висновки й теми: що працює, що ні, а що потребує вдосконалення. Напишіть звіт.
- Поінформуйте клієнта про результати і підсумки.
- Поінформуйте посадовців і ключових зацікавлених осіб про результати і підсумки, в разі необхідності внесіть корективи.
- Дайте посадовим особам програми перевірити й прокоментувати проект звіту.
- Розробіть ясні й конкретні рекомендації, в яких має бути зазначено, хто, що й коли має робити.
- Переконайтеся, що рекомендації ґрунтуються на фактах.

Етап 5: Поширення результатів оцінювання й подальша робота

- Визначте, хто і який вид інформаційних продуктів буде отримувати (наприклад, брифінг, стислий виклад на 2-4 сторінки, повний звіт, всебічний семінар) і реалізуйте план.
- Окресліть висновки з набутого досвіду й механізми їх використання та поширення.
- Працюйте над впровадженням наданих рекомендацій.
- Розмістіть електронний файл у форматі, захищеному від редагування, в репозиторіях знань з оцінювання.
- Розгляньте можливість подальшого поширення даних через професійні організації і журнали.

Джерело: Автори.

Рис. 6.2. Підхід до оцінювання розвитку

Джерело: Автори.

Матриця плану оцінювання

Матриця плану оцінювання:
Матриця, що організує питання оцінки й плани збору інформації для відповіді на них

Матриця плану оцінювання – організаційний інструмент, який ми наполегливо рекомендуємо застосовувати при плануванні оцінювання. Вона дозволяє впорядковувати питання оцінювання й плани збору інформації для відповіді на них. Матриця зв'язує описові, нормативні й причинно-наслідкові питання з планом і методиками. Окрім безпосередньої користі, яку вона приносить в якості інструменту планування, матриця може допомогти активізувати практичне застосування оцінювання й покращити співпрацю між фахівцями з оцінювання й персоналом програми.

Фахівцям з оцінювання потрібен інструмент для визначення необхідних елементів оцінки та їх чіткого пов'язування на кожному кроці. Який інструмент вони використовують для аналізу програми; якими є її контекст, вимірні цілі, збір даних і аналіз; які стратегії будуть застосовуватися – ці фактори будуть відрізнятися. Деякі фахівці з оцінювання можуть створити власні інструменти.

Мета матриці плану оцінювання – упорядкувати мету оцінки й питання та підібрати до предмета оцінки відповідні техніки збору даних. Матриця плану оцінювання зазвичай містить наступні пов'язані елементи:

- головна проблема оцінювання;
- загальний підхід;
- питання й підпитання;
- види питань і підпитань;
- засоби контролю або показники;
- цілі або стандарти (для нормативних підпитань);
- наявність або відсутність базових даних;
- стратегія планування;
- джерела даних;
- вибірка або перепис;
- інструмент збору даних;
- аналіз і графічне представлення даних;
- коментарі.

Протоколи збору даних і оцінка робочих завдань і графіків їх виконання, технічне завдання, плани комунікативної співпраці можуть додаватися до матриці або залишитися окремими, проте пов'язаними інструментами.

Метод збору даних може охоплювати більш ніж одне питання, і навпаки – кілька методів збору даних можуть відноситися до одного питання. Матриця плану оцінювання містить дані про відомі й заплановані джерела

інформації. З переходом процесу від планування до реалізації джерела можуть розширюватися й уточнюватися.

Матриця оцінки не є чимось незмінним. Як і будь-який інший інструмент планування, вона, безсумнівно, потребуватиме модифікації в міру проведення оцінки. В ході оцінки фахівці з оцінювання можуть переглядати матрицю, оновлювати її і керуватися нею в ході реалізації оцінки. Хоча планування заздалегідь має звести до мінімуму виникнення нових проблем, навіть найкращий план не може гарантувати відсутність несподіванок. Шаблон матриці плану показано на рис. 6.3.

У матриці плану оцінювання питання розбиваються на стільки підпитань, скільки буде необхідно. У кожному випадку необхідно вказати: вид підпитання (описове, нормативне або причинно-наслідкове), міру (це може бути змінна або показник), що використовується для відповіді; ціль або стандарт для порівняння, якщо це – нормативне підпитання; наявність чи відсутність базових даних; джерело або джерела, що використовуватимуться для відповіді на підпитання; чи буде застосовуватися вибірка, і якщо так, то якого типу; який інструмент використовуватиметься для збору даних; аналіз, який планується провести; будь-які коментарі. Як приклад коментарів можна навести побажання перевірити якість набору даних, вказати обмеження плану та представити дані в графічному вигляді.

Матриця часто подається на аркуші розміром 21х35 см або ж на двох аркушах паперу, розташованих поруч. Деякі фахівці з оцінювання віддають перевагу вертикальному заповненню матриці для кожного питання. Та який би формат не використовувався, фахівці з оцінювання повинні заповнити всі клітинки для кожного підпитання.

Заповнена матриця плану оцінювання займе багато сторінок. Саме цей документ дозволяє особі, яка приймає рішення, зрозуміти, що потрібно зробити та як мають даватися відповіді на питання оцінювання. В наступних главах детально розглядається заповнення стовпців.

Зараз ті читачі, які намагаються застосувати матрицю для планування оцінювання програми, політики або проекту, можуть визначити питання й підпитання та вказати види підпитань. Також можна заповнити стовпці, в яких йдеться про наявність чи відсутність базових даних, засоби контролю або показники, цілі (якщо підпитання нормативне).

Підпитання описових питань також будуть описовими. Підпитання нормативних питань можуть бути описовими, але щонайменше одне з них має бути нормативним. Причинно-наслідкові питання містять щонайменше одне причинно-наслідкове підпитання, але можуть містити описові або нормативні підпитання.

Рис. 6.3. Шаблон матриці плану

Матриця плану для: _____
Головна проблема оцінювання: _____ Загальний підхід оцінювання: _____

Питання	*Підпитання*	*Вид підпитання*	*Міра або показник*	*Цільове значення або стандарт (для нормативного підпитання);*	*Базові/ вихідні дані?*	*Джерело даних*	*Дизайн*	*Вибірка або перепис*	*Інструмент збору даних*	*Аналіз даних*	*Коментарі*

Підсумки

Фахівцям з оцінювання необхідно разом з головним клієнтом і основними зацікавленими особами визначати можливі питання. Після завершення попереднього дослідження й проведення зустрічей з клієнтом і основними зацікавленими особами, розробки програмної теорії змін та визначення основних припущень, що лежать в основі програми, фахівець з оцінювання може починати формулювати, а потім обирати питання оцінювання з великого списку сформульованих. Питання оцінювання слід перевіряти на відповідність головній проблемі оцінки.

Фахівці з оцінювання використовують описові, нормативні та причинно-наслідкові питання. Важливим є формулювання кожного питання, адже воно допомагає визначити засоби пошуку відповіді на питання.

Рекомендований підхід до упорядкування оцінки – застосування матриці плану оцінювання. Ця матриця допомагає упорядкувати, зокрема, питання, плани, стратегії збору і аналізу даних. В наступних главах наводиться покрокове керівництво із заповнення матриці плану.

Завдання до Глави 6

Практична вправа 6.1: Види питань

Класифікуйте кожне з наступних питань про ініціативу з медичної профілактики для жінок з сільської місцевості за видами (описові, нормативні, причинно-наслідкові). Якщо для того, щоб вид якихось питань став яснішим, потрібно їх переформулювати, переформулюйте (таке нерідко трапляється в реальному житті).

1. Чи було надано в рамках ініціатив необхідні консультації, підтримку та інші послуги 30 жінкам з сільської місцевості протягом першого місяця роботи, як планувалося?
2. Чи дозволили місце й час надання послуг охопити максимальну кількість жінок, які змогли взяти участь?
3. Які були найкращі методи охоплення жінок у віддалених районах і забезпечення їм доступу до програми?
4. Чи помічалися раніше проблеми медичного характеру серед жінок з сільської місцевості, особливо серед тих, які брали участь в даній ініціативі?
5. Скільки жінок, рахуючи з початку програми, отримали які види послуг?
6. Наскільки дієвою є ініціатива з медичної профілактики для жінок порівняно з іншими інтервенціями, спрямованими на покращення стану здоров'я жінок із сільської місцевості?
7. Який вплив справила медична ініціатива на жінок, їхні родини та загальну спільноту селян, в якій вони живуть?
8. Наскільки задоволені учасниці отриманими консультаціями, інформацією, підтримкою та іншими послугами?
9. Чи відповідає медична ініціатива для жінок з сільської місцевості урядовим стандартам ефективності?
10. Що кажуть учасниці з приводу того, як впливає на них програма?
11. Наскільки жінки, що отримують послуги, задовольняють критеріям участі?
12. Чи досягла програма цілі підвищення знань жінок щодо методів профілактики?

Практична вправа 6.2: Модифікація видів питань

Напишіть одне описове, одне нормативне й одне причинно-наслідкове питання для кожної з наступних програм:

1. Програма професійного навчання для молодих чоловіків.
2. Програма побудови шляху, що зв'язує три громади з центральним ринком.

Посилання та рекомендована література

Cronbach, L. J. 1982. *Designing Evaluations of Educational and Social Programs*. San Francisco: Jossey-Bass.

Feuerstein, M. T. 1986. *Partners in Evaluation: Evaluating Development and Community Programs with Participants*. London: MacMillan, in association with Teaching Aids at Low Cost.

Fitzpatrick, Jody L., James R. Sanders, and Blaine R. Worthen. 2004. *Program Evaluation: Alternative Approaches and Practical Guidelines*. New York: Pearson Education Inc.

Human Rights Resource Center. 2000. "Questions about Evaluation." In *The Human Rights Education Handbook: Effective Practices for Learning, Action, and Change.* University of Minnesota, Minneapolis. http://www1.umn.edu/humanrts/edumat/hreduseries/hrhandbook/part6B.html.

Lawrence, J. 1989. "Engaging Recipients in Development Evaluation: The 'Stake- holder' Approach." *Evaluation Review* 13 (3): 243–56.

OECD (Organisation for Economic Co-operation and Development). 2007. *Criteria for Evaluating Development Assistance*. Development Assistance Committee. http://www.oecd.org/document/22/0,2340, en_2649_201185_2086550_1_1_1_1, 00.html.

Patton, Michael Quinn. 1997. *Utilization-Focused Evaluation*. 3rd ed. Thousand Oaks, CA: Sage Publications.

———. 2002. *Qualitative Research and Evaluation Methods*. 3rd ed. Thousand Oaks, CA: Sage Publications.

Rist, R. C. 1994. "Influencing the Policy Process with Qualitative Research." In *Handbook of Qualitative Research*, ed. N. K. Denzin and Y. S. Lincoln, 545–57. Thousand Oaks, CA: Sage Publications.

Shadish, William 1998. "Some Evaluation Questions." *Practical Assessment, Research & Evaluation* 63.

http://PAREonline.net/getvn.asp?v=6&n=3. Also available from ERIC/AE Digest at http://www.ericdigests.org/1999-2/some.htm.

Веб-сайти

Kellogg Foundation. 2004. *Evaluation Logic Model Development Guide*. http://www.wkkf.org/pubs/tools/evaluation/pub3669.pdf.

World Bank. 1996. *Participation Sourcebook*. http://www.worldbank.org/wbi/ sourcebook/sbhome.htm.

ГЛАВА 7

Вибір дизайну для причинно-наслідкових, нормативних та описових питань оцінювання

Після вибору питань оцінювання фахівець з оцінювання обирає підхід планування оцінювання, що є найбільш відповідним з урахуванням конкретного питання. В цій главі наводяться певні рекомендації щодо планування та вказуються сильні й слабкі сторони різних варіантів. При цьому важливо пам'ятати, що кожна ситуація – унікальна. Не буває ідеального плану, а отже, не існує «єдино вірного» підходу до питання оцінювання.

Ця глава складається з п'яти основних частин:

- Встановлення зв'язку між питаннями й дизайном
- Дизайни для причинно-наслідкових питань
- Дизайни описових питань
- Дизайни для нормативних питань
- Потреба у ретельнішому плануванні

Встановлення зв'язку між питаннями та дизайном

Проводячи оцінювання, ми намагаємося відповісти на питання. В главі 6 зазначено, що існує три основних види питань: описові, нормативні та причинно-наслідкові. Обраний план має відповідати виду питання. Наприклад, якщо ставляться причинно-наслідкові питання, план має передбачати можливість відповіді саме на них.

Організації розвитку шукають вирішення проблем розвитку так само, як селянам трапляється шукати розв'язання проблеми, що переслідує їхнє село (див. блок 7.1). Однак і фахівці з оцінювання, і селяни час від часу припускаються помилок, намагаючись відповісти на питання.

Перша можлива помилка в пошуку відповіді на питання – починати з вибору стратегії збору даних. Відповідаючи дівчинці, яка запитала про слонів, старійшина села може сказати: «Давай проведемо опитування й дізнаємось, що, на думку селян, змушує слонів іти геть». Вибір стратегії збору даних вже на першому кроці майже гарантовано не дасть змоги отримати потрібну інформацію.

Друга можлива помилка – вважати, що кожна оцінка проводиться за одним планом. Зазвичай оцінка охоплює кілька питань, кожне з яких вимагає окремого планування. Як правило, при оцінюванні доводиться шукати відповіді на описові й нормативні питання, та іноді – на причинно-наслідкові. Фахівцю з оцінювання варто уникати вибору питань на основі методик, які він прагне застосувати (на які питання ми можемо отримати відповідь, якщо проведемо опитування або фокус-групу?). Не слід також думати, що розгляд причинно-наслідкового питання відміняє необхідність розгляду питань описових та нормативних.

У прикладі зі слонами можливо, що поглиблене дослідження практичного прикладу переміщень слонів покаже, що тварини залишили село через причини, не пов'язані ані з калатанням у казани, ані з пилом у повітрі. Може виявитися, що село просто опинилося на шляху міграції слонів.

Блок 7.1 «Чому втекли слони?»

Двічі на рік з протилежних боків через африканське село проходили слони, топчучи все на своєму шляху. Всі селяни відганяли слонів. Хтось калатав у казани, інші свистіли, вигукували, кричали. Дехто здіймав пил і бігав довкола, намагаючись відстояти свою територію.

Після того, як слони покинули село, одна дівчинка спитала: «Чому слони пішли?» Селяни відповіли: «Тому що ми прогнали їх геть». Тоді дівчинка спитала: «Але що змусило їх піти: калатання у казани, свист та інші голосні звуки чи пил, який ми здіймали?»

Широкі категорії дизайнів

Фахівці з оцінювання можуть вибирати з трьох широких категорій дизайнів: експериментальні, квазі-експериментальні та неекспериментальні. Характеристики кожної з цих категорій наведено нижче.

Експериментальний дизайн

Чимало фахівців з оцінювання вважають **експериментальний дизайн** – також відомий як рандомізований або істинний експеримент – найбільш досконалим та надійнішим дизайном оцінювання. Однак багато інших фахівців з оцінювання з цим не погоджуються. Як уже зазначалося в главі 6, фахівцям з оцінювання слід прагнути максимальної ретельності в роботі, але правильний дизайн – це той, що найбільше підходить для відповіді на питання оцінювання. Використовуючи істинний експериментальний дизайн, фахівці з оцінювання повинні доводити, що без проведення інтервенції бажаного результату досягти не вдалося б. Якщо повернутися до нашого прикладу зі слонами, то фахівець з оцінювання повинен довести: якби селяни не калатали у казани, не свистіли і т. д., слони б не пішли з села.

Експериментальний дизайн: Різновид дизайну попереднього оцінювання, що вимагає випадкового розділення аудиторії щонайменше на дві групи, щоб кожен член досліджуваної сукупності осіб мав рівні шанси бути охопленим інтервенцією (експериментальна група) або не охопленими нею (контрольна група)

Щоб продемонструвати, що саме інтервенція є причиною отриманих результатів, експериментальний дизайн передбачає порівняння результатів двох груп – охопленої та не охопленої інтервенцією. Основний критерій, що вирізняє експериментальні дизайни серед інших – випадкове розподілення піддослідних по групах. За випадкового розподілення особи у групах мають бути схожими з точки зору вихідних умов, статі, контексту, часу і т. д.

Тут наша аналогія закінчується, адже селяни не можуть розділити слонів на дві або три випадково утворені групи, які можна охопити дією різних інтервенцій або залишити як є. Але, задля наочності, уявімо, що уряду відомо про 30 сіл у провінції, які повідомляють про проблеми зі слонами, що руйнують домівки й нищать урожай. Уряд прагне допомогти вирішити цю проблему таким чином, щоб слони й селяни жили в гармонії. Було найнято експерта з поведінки тварин, але він може одночасно працювати лише з 20 селами. Прийнято рішення випадковим чином вибрати 20 сіл, яким допомагатиме експерт. Решта 10 сіл будуть контрольною групою. Для збору базових даних про думки щодо слонів та оцінку втрат урожаю й інших збитків для експериментальної і контрольної груп організуються опитування дорослих селян в усіх 30 селах. Села, що будуть охоплені інтервенцією, вибираються наступним чином: назви всіх 30 сіл пишуться на папірцях і кладуться у чашу. 10 найстарших із старійшин сіл по черзі (всліпу) дістають з чаші по одному зі згорнутих папірців. Написана на папірці назва села зачитується вголос і записується на дошці. Обрані таким чином 10 сіл утворюють контрольну групу. Мета поведінкової інтервенції – розв'язати проблему, покращити від-

носини і зменшити збитки та супутні витрати. Отже, якщо робота експерта буде ефективною, а всі інші фактори залишаться незмінними, то після заданого проміжку часу мають бути помітні відмінності між селами експериментальної та контрольної груп в контексті цієї мети. В даному випадку саме експерт виступає в ролі інтервенції.

Квазі-експериментальний дизайн

Квазі-експериментальний дизайн: Різновид дизайну оцінювання, за якого охоплені та не охоплені інтервенцією групи (порівняльні групи) формуються попередньо або постфактум, але без випадкового розподілення на групи, або коли протягом певного часу проводяться повторні вимірювання

Квазі-експериментальний дизайн аналогічний експериментальному, але він не забезпечує випадкового розподілення осіб по групах. Дизайни такого типу часто є практичнішими порівняно з експериментальними, тому що експериментальні дизайни складніші для виконання, а квазі-експериментальні не потребують рандомізації: необхідно лише створити порівняльні групи.

Квазі-експериментальний дизайн можна застосовувати для порівняння схожих, але не еквівалентних груп. Групи можуть бути з різних, але схожих сіл, але одну й ту саму групу можна використовувати багато разів (до і після). Приміром, у нашій історії зі слонами квазі-експериментальний дизайн міг би передбачати вибір двох сіл з одного району, з однаковим кліматом, з однаковою кількістю слонів у даній місцевості, кількістю селян та будинків і т. д. Одне село буде охоплено інтервенцією (калатання в казани), друге – ні.

Окрім того, ці дві групи можна сформувати, змінюючи поведінку селян під час появи слонів. Коли слони приходять уперше, селяни калатають у казани. Під час другої появи слонів у селі селяни у казани не калатають. Таким чином, ми отримуємо два варіанти розвитку подій для порівняння поведінки слонів.

В обох прикладах дві групи є схожими, але не однаковими. В першому прикладі на результати можуть вплинути відмінності між двома групами чи середовищем, в якому вони існують, у другому відмінною для двох ситуацій є лише поведінка селян. Застосовуючи квазі-експериментальний дизайн, фахівці з оцінювання не можуть чітко пов'язати інтервенцію з вирішенням проблеми або показати причинно-наслідковий зв'язок з тією ж точністю, що її забезпечує справжня рандомізація, але вони можуть отримати багато нової інформації та зробити припущення щодо ймовірних причин і наслідків.

Неекспериментальний дизайн

Неекспериментальний (описовий) дизайн: Різновид дизайну оцінювання, за якого не створюються окремі групи, охоплені та не охоплені інтервенцією, а основна увага приділяється опису

Неекспериментальні, або **описові, дизайни** не передбачають порівняння груп. Замість цього виконується детальний опис відносин між інтервенцією та її наслідками. Застосовуючи неекспериментальне дослідження, фахівець з оцінювання вибирає, наприклад, коли проводити вибірку, кого та де до неї включати. Створення двох або декількох еквівалентних чи схожих груп навіть не розглядається.

Неекспериментальне оцінювання може, наприклад, використовувати аналіз існуючих даних чи інформації, опитування чи фокус-групу для збору відповідних даних, актуальних для питань оцінювання. В центрі уваги неекспериментальних дизайнів – характеристики, частота і зв'язки (Project STAR 2006).

В табл. 7.1 стисло представлено ключові характеристики кожної з трьох категорій дизайнів.

Таблиця 7.1. Дві ключові характеристики експериментальних, квазі-експериментальних та неекспериментальних дизайнів

Різновид дизайну оцінювання	Випадково сформована контрольна група	Невипадково сформована порівняльна група	Повторення вимірювань
Експериментальний	Так	Ні	Так
Квазі-експериментальний	Ні	Імовірно	Імовірно
Неекспериментальний	Ні	Ні	Ні

Джерело: Автори.

Нотація дизайну

Дизайни оцінювання іноді записуються за допомогою символів «X» та «O». В цих записах «X» означає інтервенцію, а «O» – спостереження. Кожній інтервенції та кожному спостереженню присвоюється індекс. Наприклад, дизайн оцінювання, в якому за однією інтервенцією слідує одне спостереження, записується наступним чином:

$$X\ O_1$$

Запис дизайну оцінювання, в якому за одним спостереженням слідує інтервенція, а далі – ще два спостереження, має наступний вигляд:

$$O_1\ X\ O_2\ O_3$$

Кожна група записується окремим рядком. Наступний запис описує дизайн оцінювання, що передбачає використання двох груп, одна з яких охоплюється дією інтервенції, а інша – ні. Обидві групи один раз спостерігаються до початку інтервенції в експериментальній групі і двічі – після:

$$O_1\ X\ O_2\ O_3$$

$$O_1\ O_2\ O_3$$

Планування причинно-наслідкових питань

Причинно-наслідкові питання становлять найбільшу проблему; відповідь на них потребує особливо ретельно продуманого планування. В оцінюваннях, що ставлять причинно-наслідкові питання (тобто питання про впливи і причини таких впливів), дизайн спрямований на виключення інших раціональних пояснень спостережних результатів, окрім інтервенції. Це робиться для того, щоб можна було зробити висновок, що саме інтервенція спричинила вплив. Отже, неохідно вирішити – відносити спостережені зміни (впливи) на рахунок інтервенції чи іншого фактора.

Під час розгляду причинно-наслідкових питань дизайни оцінювання можна формувати таким чином, щоб можна було відповісти на питання: «Як би склалася ситуація, якби інтервенція не проводилася?» Може не бути можливості точно виміряти результати відсутності інтервенції, але можна приблизно оцінити, що сталося б, якби інтервенції не було.

Експериментальний дизайн

Коріння експериментальної моделі в медичних дослідженнях, де вона часто використовувалася для тестування препаратів та протоколів лікування. Застосовуючи цей дизайн до оцінки в сфері охорони здоров'я, організація, що прагне скоротити захворюваність на малярію в регіоні, може поставити питання: «Що є найкращим для скорочення захворюваності на малярію в регіоні?» Можливе підпитання: «Чи дозволяє використання спеціально оброблених надліжкових сіток проти комах скоротити захворюваність на малярію в цьому регіоні?»

Експериментальний дизайн перетворює питання на пропозицію. У прикладі з малярією пропозиція виглядатиме наступним чином: якщо людям видати спеціально оброблені надліжкові сітки проти комах, в регіоні поменшає випадків захворювання на малярію.

Як уже зазначалося в главі 4, для пропозиції необхідно розробити теорію змін. Рандомізація – найважливіший фактор в експериментальному дизайні. З числа добровольців можна вибрати групу пацієнтів для випадкових випробувань. Такі пацієнти повинні мати однакову стадію захворювання, бути однакової статі тощо. Після цього їм довільно призначається один з кількох режимів медикаментозного лікування. Одна підгрупа приймає ліки, вибір яких заснований на існуючих знаннях, друга – інший, перспективний препарат. В ході досліду з подвійним сліпим методом ні учасники дослідження, ні медичний персонал не знають, який препарат яким пацієнтам дається.

Існує слабка, але зростаюча тенденція переходу оцінювання розвитку до використання експериментальних дизайнів. Рух у цьому напрямку викликаний розчаруванням від того, що, незважаючи на роки роботи з оцінювань, досі достеменно невідомо, що працює в контексті розвитку та за яких умов.

Класичний експеримент складається з шести кроків:

1. Формулювання гіпотези.
2. Отримання базових даних (вимірюється залежна змінна).
3. Випадковий розподіл випадків по групах, охоплених і не охоплених інтервенцією (експериментальна та контрольна групи).
4. Активація інтервенції або незалежної змінної.
5. Повторне вимірювання залежної змінної (перевірка після експерименту).
6. Розрахунок відмінностей між групами та перевірка статистичної значущості.

У прикладі з малярією стоїть питання про те, чи зменшує використання спеціально оброблених надліжкових сіток проти комах захворюваність на малярію в регіоні. В цьому випадку означені шість кроків виглядатимуть наступним чином:

1. Формулювання гіпотези: Використання в домівках надліжкових сіток, оброблених антимоскітним засобом, знижує захворюваність на малярію.
2. Отримання базових даних: Протягом двомісячного періоду в регіоні на малярію захворіло 172 особи.
3. Випадкове розподілення випадків по групах, охоплених і не охоплених інтервенцією (експериментальна й контрольна групи).
4. Активація заходів: Треба видати надліжкові сітки одній групі (експериментальна група), не видаючи їх іншій (контрольна група).
5. Повторне вимірювання залежної змінної (перевірка після експерименту): За два місяці інтервенції в експериментальній групі на малярію захворіло 65 осіб, а в контрольній групі – 118 осіб.
6. Розрахунок відмінностей між групами: В експериментальній групі захворіло на 53 особи менше, ніж у контрольній.

Запис дизайну класичного експерименту з однією експериментальною і однією контрольною групами має наступний вигляд:

$$O_1 \; X \; O_2$$

$$O_1 \; O_2$$

В деяких випадках фахівці з оцінювання також пишуть на початку рядка кожної групи літеру R, показуючи, що ця група сформована довільно. У випадку експериментального дизайну – такого, як описано вище (з довільним визначенням учасників, яким видається надліжкова сітка проти комах) – запис має наступний вигляд:

$$R\ O_1\ X\ O_2$$

$$R\ O_1\ O_2$$

Поведінка людини складна, а соціальні ситуації важко аналізувати. Експериментальний дизайн добре працює при тестуванні медичних препаратів на людських організмах, які є більш-менш однаковими. Але, розглядаючи складну людську поведінку, будьте готові до хибних позитивних і хибних негативних висновків. Хибні позитивні висновки трапляються, якщо дослідження показує, що певний ефект був спричинений інтервенцією, хоча насправді це не так. Хибні негативні висновки трапляються, якщо дослідження показує відсутність зв'язку між інтервенцією та успіхом, в той час як насправді такий зв'язок існує.

Хибні позитивні висновки можуть виникати через те, що дані надаються людьми самостійно. Жителі регіону можуть знати про експеримент і про те, які домівки охоплені дією інтервенції, та не повідомляти про якусь кількість випадків захворювання на малярію, щоб догодити представникам уряду.

Хибні негативні висновки зустрічаються частіше. Відсутність помітних відмінностей між двома групами може бути спричинена тим, що двомісячний період вимірювань припав на сухий сезон, коли захворюваність на малярію загалом є низькою. Хибний негативний висновок може бути зроблений через те, що учасники експериментальної групи не кожної ночі використовують надліжкові сітки, або через те, що учасники контрольної групи самі купляють собі такі ж сітки.

Якби в ході дослідження не ставилися додаткові питання, було б важко інтерпретувати результати навіть цієї простої інтервенції з випадковим формуванням груп. Для пояснення результатів фахівцям з оцінювання варто було би поставити деякі з наступних питань:

- Яку інформацію отримали жителі щодо використання надліжкових сіток?
- Як визначався склад сімей-учасниць програми та хто фактично спав під сіткою?
- На які реалізаційні проблеми звернули увагу учасники експериментальної та контрольної груп?
- Як співвідносяться показники захворюваності на малярію протягом двох двомісячних періодів з історичними даними?
- Чи проводилися в регіоні інші заходи з профілактики малярії під час проведення даної інтервенції?

Всі ці питання – описові, а отже, простіші для відповіді.

Контрольні групи

Експериментальний дизайн спрямований на виключення або контроль інших факторів, які можуть виступати альтернативними поясненнями результатів експерименту. Використовуючи експериментальний дизайн, фахівці з оцінювання порівнюють еквівалентні групи. **Контрольна група** – це група, для якої зберігаються звичайні умови; її члени не охоплюються дією інтервенції. Група, на яку поширюється дія інтервенції, може називатися **експериментальною групою**. Використання контрольної групи дозволяє порівнювати групи, охоплені дією інтервенції, з тими, на які дія інтервенції не поширювалася.

Контрольна група:
Група експерименту, члени якої не охоплюються дією інтервенції

Експериментальна група:
Група експерименту, члени якої охоплюються дією інтервенції

Використання контрольних груп нерідко означає неучасть в заходах інтервенції тих, хто цього потребує. Іноді такий підхід є виправданим через брак ресурсів. Тоді можливе лише обмежене охоплення аудиторії дією інтервенції, або ж програма розтягнеться у часі. В інших випадках ефект від інтервенції є не перевіреним, тому незрозуміло, наскільки цінними є ті заходи, масштаб яких обмежується. Якщо дієвість інтервенції підтверджується, буде важко пояснити, чому на частину людей вона не поширюється (Patton 2008).

Випадковий розподіл

Експериментальний дизайн передбачає випадковий розподіл потенційних учасників програми по групах, охоплених та не охоплених дією інтервенції. Це робиться для того, щоб групи були максимально ідентичними (неупереджений розподіл) з точки зору факторів, здатних впливати на результати програми або інтервенції. До таких факторів можуть відноситися вік, стать, освіта, погляди, історія учасників і т. ін.

В ідеалі людина здатна довільно приймати рішення про те, хто буде брати участь в інтервенції, а хто – не буде. В реальному ж житті визначення надійної контрольної групи є проблемою. Сформувати таку групу можна, наприклад, довільно розподіливши ресурси проекту або програми. Бенефіціари такого проекту або програми в цьому випадку стають випадковою вибіркою зі всієї аудиторії. Потім цю вибірку можна порівнювати з іншою випадково сформованою вибіркою осіб, які не є бенефіціарами (тобто з контрольною групою) (White 2007).

Випадковий розподіл дає можливість ефективно виміряти вплив. Він також може бути більш справедливим, ніж проведений іншим способом, адже коли експериментальна та контрольна групи формуються повністю довільно, такі фактори, як упередженість або фаворитизм, виключаються.

Незважаючи на те, що випадковий розподіл є більш застосовним до інтервенцій розвитку, ніж здається, використання його не завжди можливе. Як і у випадку з медичними випробуваннями, обмеження масштабу інтервенції для створення контрольної групи буває неетичним, але іноді може не вистачати ресурсів для залучення всіх, хто хоче взяти участь в інтервенції. Це означає, що деякі учасники програми залишаються поза заходами. У керівника проекту може з'явитися бажання залучити до участі в інтервенції тих, в кого є найвищі шанси отримати від неї користь. Це допоможе максимально корисно використати обмежені кошти програми. Однак, з точки зору оцінювання, такий підхід призведе до спотворення результатів, адже вибірки створюватимуться без дотримання принципу неупередженості й випадковості.

Якщо випадковий відбір виявляється неможливим, можна, наприклад, збирати дані про фактори, які є різними в двох групах та можуть мати вплив на результати. Ці змінні можна використати в ході аналізу даних в якості контрольних змінних. Використання **контрольних змінних** дозволяє фахівцю з оцінювання виключити деякі альтернативні пояснення навіть у тих випадках, коли випадкове розподілення неможливе.

Контрольна змінна:

Зовнішній фактор, який утримується на постійному рівні для того, щоб звести до мінімуму його можливий вплив на результат інтервенції

Під час формування груп фахівцям з оцінювання слід враховувати проблему помилки вибірки. **Проблема вибірки** має місце, якщо різниця між учасниками заходів і тими, хто в них участі не бере, викликана непомітними відмінностями між групами, а не впливом інтервенції.

Похибка вибірки[8]:

Спотворення фактів або даних про результати програмної інтервенції, викликане систематичними відмінностями у характеристиках підмножини аудиторії, охопленої дією інтервенції, а також підмножини тієї ж аудиторії, на яку дія інтервенції не поширюється

> Процес рандомізації забезпечує статистичну усереднену рівнозначність всіх характеристик експериментальної та контрольної груп перед початком інтервенції. Рандомізовані експерименти вирішують проблему вибірки, формуючи контрольну групу з осіб, які б взяли участь у програмі, але випадково були з неї виключені. Випадковий розподіл осіб у групах не виключає проблему вибірки, а лише врівноважує упередженість у формуванні експериментальної та контрольної груп, скасовуючи її дію при розрахунку розміру впливу. Тоді будь-які відмінності в усереднених довгострокових результатах для двох груп по закінченню інтервенції можна віднести саме на рахунок цієї інтервенції (World Bank 2008).

Проблема вибірки може виявлятися у два способи. По-перше, учасники можуть самостійно приймати рішення про участь у програмі. По-друге, керівники програми можуть обирати тих учасників, які мають найвищі шанси на успіх. Згадаємо для прикладу програму використання спеціально оброблених надліжкових сіток проти комах. Якби програмна інтервенція передбачала виведення на ринок таких сіток з дуже низькою вартістю, виникла б проблема вибірки: в експериментальну групу потрапили б лише ті, хто

[8] *В оригіналі «Selection bias».*

може собі дозволити придбати надліжкові сітки і дійсно це зробив. Ті ж, хто не має змоги придбати сітки або не знає про них, в експериментальну групу не потрапили б. Більше того, було б дуже складно порівняти експериментальну й контрольну групи через відсутність інформації про те, хто використовував надліжкові сітки, а хто – ні.

Для зменшення ймовірності помилкових позитивних висновків фахівці з оцінювання використовують методи, запозичені із суспільствознавчих наук. Використовуючи експериментальний дизайн оцінювання, вони докладають всіх зусиль для контролю реалізації програми, політики або проекту та середовища, в якому ця реалізація відбувається. Якщо вдається належно контролювати всі паралельні фактори, фахівці з оцінювання можуть впевнено відносити спостережені відмінності на рахунок інтервенції.

Розглянемо оцінку інтервенції із застосування добрива для підвищення врожайності кукурудзи. Проект використовує теплицю, в якій селяни контролюють температуру, полив і стан ґрунту. В рамках дизайну оцінювання в теплиці створено дві окремі ділянки. Довільно одну з них обрано для внесення добрива; інша ділянка виконує роль контрольної. На обох ділянках підтримується однакова температура, освітленість і поливний режим, а кукурудза росте в абсолютно однаковому ґрунті. Потім вимірюються обсяги врожаю. Якщо врожай на експериментальній ділянці більший, ніж на контрольній, фахівець з оцінювання може стверджувати, що добриво було ефективним.

А тепер уявімо, що інтервенція проводилася на полях, а не в контрольованому середовищі теплиці. Що буде, якщо дві ділянки розташовані поблизу одна від одної і частина добрива потрапить на контрольну ділянку? Контрольну ділянку можна перенести на іншу частину поля, але ґрунт, освітлення, режим опадів там можуть дещо відрізнятися. Двом полям також може приділятися різна увага. Хоча оцінка все ще може виміряти вплив, важко буде стверджувати, що збільшення врожаю зумовлене лише дією добрива.

У складному, мінливому світі, де проводяться інтервенції розвитку, стає все важче виокремити вплив якогось одного з великої кількості факторів. Розвиваючи тему сільського господарства, уявімо, що в ідеальних погодних умовах, що також збіглися у часі з високим попитом на сільськогосподарську продукцію, було проведено меліоративну інтервенцію. Протягом останніх років доходи в регіоні проведення інтервенції зростали. Але чи стане подальше зростання доходів результатом інтервенції? Можливо, воно було викликане іншими факторами – такими як збільшення рівня опадів, сприятливий стан економіки або незвичний період політичної стабільності? В ідеалі треба взяти ділянки, що знаходяться в регіоні, де проводиться меліоративна інтервенція, і довільно розбити їх на дві групи, після чого в одній

групі провести інтервенцію, а в іншій – не проводити. Але які є варіанти, якщо випадкове розподілення неможливе, а отже, не можна застосувати експериментальний дизайн?

У багатьох подібних випадках використовують квазі-експериментальні дизайни. В рамках квазі-експериментальних дизайнів проводять порівняння, але без випадкового розподілу учасників по групах. Четерджі (Chatterji 2007), Петон (Patton 2007), а також Бамбергер і Вайт (Bamberger and White 2007) написали чудові праці, в яких описали обмеження, з якими стикаються експериментальні дизайни на практиці.

Внутрішня валідність

Внутрішня валідність:
Здатність дизайну виключати всі інші можливі альтернативні фактори або пояснення наявних результатів, окрім інтервенції

Говорячи про виключення інших можливих пояснень, ми маємо на увазі внутрішню валідність. **Внутрішня валідність** означає здатність дизайну виключати інші пояснення спостережених результатів. Дизайн оцінювання, якому властива сильна внутрішня валідність, зміцнює впевненість фахівців з оцінювання у своїх висновках щодо того, чи стала інтервенція причиною результатів, які вони спостерігають. Якщо внутрішня валідність дизайну слабка, переконати інших в тому, що саме інтервенція стала причиною наявних результатів, буде важче. Питання внутрішньої валідності піднімається через те, що за відсутності рандомізації обґрунтованість результатів оцінки може бути легко скомпрометована.

Ці загрози внутрішній валідності результатів оцінювання – просто можливі альтернативні пояснення; насправді їх може не існувати. Таким чином, внутрішня валідність залежить від контексту. Квазі-експериментальні дизайни мають долати загрози внутрішній валідності.

Британське товариство оцінювання (United Kingdom Evaluation Society 2003) визначає внутрішню валідність наступним чином:

> Це впевненість, яку можна мати щодо висновків про дієвість інтервенції. Загроза внутрішній валідності – це твердження, що дизайн оцінювання залишає місце для сумнівів щодо причинно-наслідкового зв'язку між інтервенцією і наявними результатами. Можна подати це у вигляді питання: «Чи не може щось інше, окрім інтервенції, бути причиною різниці в ситуації до проведення інтервенції та після неї?»

У своїй класичній праці Кук та Кемпбел (Cook and Campbell 1979) визначають кілька розповсюджених загроз внутрішній валідності:

- історія
- дозрівання
- повторне тестування
- помилки вибірки
- вибування
- регресія до середнього значення
- інструментарій

Ефект історії. **Ефект історії** означає можливість впливу на результати подій, які мали місце в ході інтервенції або між регулярними вимірюваннями, і при цьому не є частиною інтервенції. Історія – завжди актуальна загроза при довгостроковому дослідженні. Цю загрозу, певно, найскладніше виявити, адже фахівцю з оцінювання доводиться вивчати події, що відбулася протягом періоду інтервенції і могли вплинути на результати. Наявність історичних подій, що впливають на результати, є досить імовірною, якщо говорити про результати для окремої особи. Персональна історія включає низку подій, деякі з яких можуть бути надзвичайно значущими. Якщо ж говорити про групу осіб, варто виявляти подію, яка одночасно вплинула на більшість або принаймні частину осіб достатньо, щоб змінити вимірювану характеристику. Як зазначають Бросарт, Клей і Вілсон (Brossart, Clay, and Willson 2002):

Ефект історії:
Ефект, що його мають не пов'язані з інтервенцією події на її результат

> Якщо всі особи є членами однієї групи, шукати таку подію можна за допомогою інтерв'ю або спостережень; якщо учасники є незалежними, імовірність того, що така подія одночасно змінить якусь характеристику, є низькою. Виключенням може бути ситуація, коли подія відбувається в місці, де знаходяться одночасно всі учасники (наприклад, у лікарні).

Уявімо, що під час проведення програми, спрямованої на молодь, яка знаходиться в зоні ризику, неповнолітнім було скоєно резонансний злочин. З'являються заклики до жорсткіших заходів по відношенню до проблемних підлітків. Ця ситуація може змінити спектр осіб, охоплених програмою, та, імовірно, вплинути на результати. Опитування суспільної думки особливо чутливі до таких впливів, адже на погляди можуть сильно впливати недавні події та висвітлення певних питань у ЗМІ (Office of Juvenile Justice and Delinquency Prevention 1989).

Розглянемо для прикладу таку інтервенцію, як надання провінційним фермерам нового насіння або навчання їх покращеним методам вирощування рослин. Залежною змінною, або виміром довгострокового результату, може бути збільшення доходів від врожаю порівняно з попереднім роком. Але може статися так, що фермери продовжать використовувати старі методи вирощування рослин та старі, випробувані сорти насіння, навіть якщо попри це їхній середній дохід збільшиться. Поглиблений аналіз питання може показати, що причиною змін стали сприятливі для врожаю погодні умови. Саме клімат, а не інтервенція, став причиною отриманих результатів, тобто події поза межами інтервенції. Оцінка, що базується на порівнянні стану речей до та після інтервенції, часто зазнає впливу ефекту історії.

Ефект дозрівання:
Вплив на результати, що його має дорослішання або розвиток

Ефект дозрівання. **Ефект дозрівання** має місце, якщо результати спричинено дорослішанням або розвитком. Дорослішаючи, люди стають більш зрілими. При цьому вони можуть по-новому сприймати ситуації та реагувати на них. До змін, природним чином спричинених плином часу, відносяться дорослішання, розумнішання та набуття досвіду. Цей ефект поширюється як на окремих осіб, так і на групи. Наприклад, діти за дворічний термін зазвичай починають краще читати навіть без додаткового навчання. Організації також розвиваються і змінюються. Ці зміни можуть бути частиною природного циклу зростання й розвитку і ніяк не пов'язані з інтервенцією. Оцінювання, що базується на порівнянні стану до та після інтервенції, часто є слабким саме через ефект дозрівання.

Можна сказати, що дозрівання має дві форми: короткострокову та довгострокову. Приклади короткострокового дозрівання – втома й навчання. Довгострокове дозрівання включає психофізичний розвиток, культурні зміни, а також зміни середовища, здатні впливати на психологічні концепції. Якщо вимірювання розділяють декілька місяців, потенційно важливим є довгострокове дозрівання.

Приміром, оцінювання може розглядати вплив дворічної програми навчання читанню на оцінки з читання серед учнів початкової школи. За два роки когнітивні навички дітей покращаться незалежно від участі у програмі. Як фахівець з оцінювання може впевнитися, що кращі оцінки з читання відображають результати програми, а не ефект дозрівання?

Ефект повторного тестування:
Вплив на результати, спричинений тим, що особи проходять тест кілька разів

Ефект повторного тестування. **Ефект повторного тестування** (короткостроковий) трапляється тоді, коли особи проходять той самий тест до та після інтервенції, або ж просто проходять його кілька разів. Вони можуть навчитися відповідати на питання, а це знизить валідність результатів.

Розглянемо, наприклад, інтервенцію, спрямовану на підвищення кваліфікації вчителів сільських шкіл. Наприкінці кожного тижня вчителі проходять тестування. Члени комітету з оцінювання визначають результати тестів за допомогою контрольного списку і надають вчителям зворотний зв'язок щодо їхніх показників. А покращення показників за стандартним контрольним списком може бути просто результатом того, що вчителі вчаться краще проходити повторювані тести.

Похибка вибірки:
Спотворення фактів або даних про результати програмної інтервенції, викликане систематичними відмінностями у характеристиках підмножини аудиторії, охопленої дією інтервенції, а також підмножини тієї ж аудиторії, на яку дія інтервенції не поширюється

Похибка вибірки. **Похибка вибірки** утворюється, якщо група людей, які за власною ініціативою долучилися до участі в програмі, порівнюється з групою людей, які цього не зробили. Дві групи не є еквівалентними, тому що ініціативні особи з більшою ймовірністю навіть без програмної інтервенції покращать свої навички або змінять погляди порівняно з тими, хто не взяв участь у програмі. Помилка вибірки може мати місце навіть тоді, коли

частина людей вирішила відповісти на питання опитування, а інша частина відповідей не надала. Помилка, пов'язана з ініціативністю, може виникнути в будь-якій програмі, де передбачена реєстрація з власної ініціативи. Це один із ризиків, пов'язаних з квазі-експериментальними дизайнами.

Ефект вибування. Говорячи про **ефект вибування,** ми маємо на увазі вибування учасників. Втрата учасників може створити хибне враження про дієвість інтервенції. Диференційоване вибування учасників може стати джерелом упередженої оцінки так само, як і формування вибірки. Хоча й існує спокуса представити результати лише щодо тих, хто успішно пройшов програму, таке представлення буде необ'єктивним, адже ті, хто вибув із програми, ймовірно, мали гірші показники, ніж ті, хто її пройшов до кінця. І хоча проходження програми й отримання повного ефекту від неї – важливі моменти з точки зору оцінки, вони не повинні порушувати порівняння з показниками порівняльної групи.

Ефект вибування:
Вплив, що його має на визначення результатів інтервенції вибування учасників програми

Розглянемо, наприклад, програму навчання викладачів з 400 учасниками, одним з показників ефективності якої є підсумкові бали випускників. В ході трирічної програми 25 учасників померли від СНІДу. Втрата цих учасників неприродно знизила середній підсумковий бал, створивши враження, що програма була менш успішною, ніж насправді.

Розглянемо інший приклад, пов'язаний з тією ж програмою освіти викладачів. Уявімо, що правила коледжу, який спонсорує програму, забороняють вагітним жінкам відвідувати заняття або здавати екзамени. Тоді з підсумкових балів будуть виключені показники вагітних жінок.

Ефект регресії до середнього значення. Явище регресії відображає природну схильність осіб, які набирають дуже високі або дуже низькі бали, при повторному тестуванні показувати результати, більш наближені до середніх. Цей ефект відомий як **регресія до середнього значення**. Якщо вимірювання ненадійне, результати кількох вимірювань будуть різнитися. Імовірно, що показники при повторному вимірюванні змістяться до середніх значень, а не до крайніх. Таким чином, в програмах, де особи або групи відбираються на основі крайніх значень їх показників, можна очікувати змін, викликаних «міграцією» крайніх груп до середини незалежно від того, чи отримали їхні представники користь від програми.

Регресія до середнього значення:
Природна схильність осіб, що набирають дуже високі або дуже низькі бали, при повторному тестуванні показували результати, більш наближені до середніх

Уявімо, наприклад, що програма з покращення навичок ведення бухгалтерського обліку в ході інтервенції мікрокредитування обирає учасників на основі результатів тестування арифметичних здібностей. При цьому відбираються особи з найвищими балами. Якщо дати цим учасникам пройти той самий тест після інтервенції, їхні результати можуть знизитися, наблизившись до середнього рівня.

Ефект інструментарію:
Вплив на результати оцінювання, спричинений браком надійних інструментів для їх вимірювання

Ефект інструментарію. **Ефект інструментарію** має місце в разі зміни надійності інструменту. Зміни можуть бути пов'язані зі змінами мети вимірювання (наприклад, в результаті похибок калібрування). Приміром, оцінювання програми, спрямованої на збільшення маси тіла дорослих людей за рахунок надання інформації про правильне харчування, може не показати помітного ефекту, якщо ваги, що використовуються для зважування людей, не було відкалібровано, або якщо різні ваги калібрувалися в різних місцях та за різними методиками.

Квазі-експериментальні дизайни

Квазі-експериментальний дизайн:
Дизайн, за якого порівнюються групи зі схожими характеристиками, або ж протягом певного часу проводяться численні вимірювання показників однієї групи

Методи **квазі-експериментального дизайну** можуть використовуватися для проведення оцінок тоді, коли неможливо довільно створити експериментальну й контрольну групи. До квазі-експериментальних відносяться такі дизайни, що використовують порівняння показників груп зі схожими характеристиками або численні вимірювання показників однієї групи, без утворення порівняльних груп. Фахівці з оцінювання формують групи якомога рівнішими з точки зору важливих характеристик (стать, доходи, соціально-економічне становище). Потім порівнюються показники цих еквівалентних груп, які іноді називають *порівняльними групами* (див. блок 7.2). Іноді фахівець з оцінювання може формувати порівняльну групу, підбираючи її за ключовими характеристиками. В інших випадках він формує порівняльну групу, що відрізняється від групи, охопленої дією інтервенції, але достатньо схожу для цілей порівняння.

Вибір квазі-експериментального дизайну заздалегідь є бажаним, але не обов'язковим. Наявність порівняльної групи зазвичай забезпечує більшу надійність оцінювання, ніж використання повторних вимірювань показників однієї групи. В ідеалі базові дані щодо порівняльної групи отримуються одночасно з базовими даними щодо групи, охопленої дією інтервенції, та до початку програмної інтервенції. У випадку використання квазі-експериментальних дизайнів для оцінки постфактум порівняльна група формується після початку програми або навіть після її завершення. Це, вочевидь, набагато складніша для фахівця з оцінювання ситуація.

Запис квазі-експериментального дизайну аналогічний запису дизайну експериментального. Різниця в нерівнозначному розподіленні учасників по групах, яке має місце у випадку з квазі-експериментальними дизайнами. Іноді на початку рядка для запису нерівнозначних груп ставиться літера N. Таким чином, базовий квазі-експериментальний дизайн з експериментальними та порівняльними групами записується наступним чином:

$$N\ O_1\ X\ O_2$$

$$N\ O_1\ O_2$$

Блок 7.2 «Чи ефективно працюють комунальні школи? Оцінка сальвадорської програми EDUCO»

Реалізована в Сальвадорі програма комунальних шкіл (EDUCO) була спрямована на стрімке розгортання системи сільської освіти після громадянської війни. Оцінка програми була покликана виміряти ефект децентралізації освіти з передачею відповідальності за неї на місця, а також вплив шкіл на результати учнів.

Питання було в тому, чи не довелося заради швидкості розгортання шкільної системи пожертвувати якістю викладання. Оцінювання порівнювало показники довгострокових результатів (це результати учнів на стандартних тестах з математики й мови) третьокласників з EDUCO та з традиційних шкіл, звертаючи увагу на характеристики студентів та виключаючи помилку вибірки засобами статистичного контролю. Оскільки результати тестів можуть бути непоказовими в короткостроковій перспективі, фахівці з оцінювання також враховували навчальні дні, пропущені через відсутність викладача.

Різниця в освітніх результатах може бути зумовлена багатьма факторами. Фахівцям з оцінювання потрібно було визначити, чи відображає різниця в результатах тестування відмінності між видами шкіл, а також інші фактори, такі як сімейні умови (освіта, розмір сім'ї, доходи), характеристики учня (стать, вік, кількість братів та сестер), школи (набір учнів, кваліфікація викладачів, технічна база, фінансування) й викладачів (освіта, стаж).

У створенні моделі для визначення чистого впливу виду школи (контролюючи інші фактори) фахівці з оцінювання спиралися на дані, зібрані в ході досліджень, проведених Міністерством освіти. За рахунок складного статистичного моделювання, що дозволило врахувати всі зазначені вище фактори, фахівці з оцінювання дійшли висновку, що показники успішності учнів шкіл програми EDUCO та звичайних шкіл були приблизно однаковими. Швидке розгортання шкільної системи не мало негативного впливу на навчальний процес. Інакше кажучи, комунальні школи не поступалися в ефективності звичайним.

Щоб визначити, чи справила інтервенція вплив, оцінювання має продемонструвати зміну ключових показників в результаті дії інтервенції. Існує багато варіацій квазі-експериментальних дизайнів; деякі з них надійніші за інші. Нижче розглядається вісім варіантів квазі-експериментальної оцінки:

- порівняння стану до й після інтервенції без порівняльної групи;
- дизайн нерівнозначного порівняння до й після інтервенції;

- дизайн нерівнозначного порівняння після інтервенції;
- дизайн багаторазового порівняння з перервою;
- дизайн довгострокового оцінювання;
- вибірковий дизайн;
- кореляційний дизайн із застосуванням засобів статистичного контролю;
- зіставлення із групою з іншими характеристиками.

Порівняння стану до й після інтервенції без порівняльної групи

Порівняння стану до й після інтервенції:
Дизайн оцінювання, за якого значення показників, отримані після початку інтервенції, порівнюються зі значеннями, отриманими до її початку

Порівняння стану до й після інтервенції – спосіб вимірювання змін. Значення ключових показників після початку інтервенції зіставляються зі значеннями, отриманими до її початку. Стандартними способами вимірювання показників до та після інтервенції є попередні тести й тести постфактум (дані, отримані до початку інтервенції, нерідко називають базовими). Збір базових даних іноді називають дослідженням вихідних умов.

Просте порівняння стану до й після інтервенції – це не дуже надійний квазі-експериментальний дизайн, слабкість якого полягає в тому, що одне лише вимірювання показників до початку інтервенції та після її завершення не дає підстав стверджувати, що саме інтервенція зумовила зміни. Люди можуть змінювати поведінку через те, що опиняються під спостереженням. Зміни також можуть спричиняти інші події, що відбуваються паралельно з інтервенцією.

Якщо спостерігаються незначні зміни, фахівцю з оцінювання не слід квапитися з висновком про те, що інтервенція не спрацювала. Розглянемо для прикладу інтервенцію, спрямовану на скорочення бідності. Район проведення інтервенції був такий бідний, що всі жителі мали право на участь в програмі, отже, контрольної групи не було. Через 10 років пропорція бідних не змінилася. Фахівець з оцінювання не може зробити висновок, що інтервенція зі скорочення бідності не спрацювала, адже може бути, що без неї (тут знову постає питання про відсутність втручання) в бідності жило б іще більше людей.

Порівняння стану до й після інтервенції зазвичай вважається ненадійним, адже воно не враховує впливу на довгострокові результати інших факторів; порівнюються лише загальні умови. Порівняння груп, охоплених та не охоплених дією інтервенції, не проводиться. Тому цей елемент має використовуватися лише в поєднанні з іншими.

Оцінювання з використанням порівняння стану до й після інтервенції записується наступним чином:

$$O_1 X O_2$$

Дизайн нерівнозначного порівняння до й після інтервенції

Хоча учасники квазі-експериментальних оцінок розподіляються по групах не випадково, порівняльну групу все ж можна використовувати. Така порівняльна група називається нерівнозначною. Її можна використовувати для порівняння, але фахівцям з оцінювання слід уважно враховувати описані вище загрози внутрішній валідності.

Щоб забезпечити більшу рівнозначність груп, фахівці з оцінювання намагаються формувати їх за однаковими характеристиками. До таких характеристик відносять демографічні показники, результати перевірки кваліфікації, показники роботи, оцінку суджень і т. ін. Фахівці з оцінювання можуть проводити попереднє тестування всіх учасників і формувати групи за його підсумками. Приміром, в одній інтервенції, спрямованій на підвищення обізнаності про гендерні проблеми, було проведено попереднє тестування щодо понять і принципів гендерної обізнаності. Оцінки було відранжовано в порядку спадання. Одного з двох учасників, оцінки яких були найкращими, включили до однієї групи, а другого – до іншої. Ця процедура повторювалася до тих пір, доки всі учасники не були розподілені по двох групах, одна з яких пройшла додаткове навчання. На певному етапі обидві групи пройшли ще одне тестування. Це було необхідно, щоб дізнатися, чи будуть оцінки групи з додатковим навчанням відрізнятися від оцінок другої групи.

Запис нерівнозначного порівняння записується наступним чином:

$$N\ O_1 \quad X \quad O_1$$

$$N\ O_2 \quad O_2$$

Дизайн нерівнозначного порівняння після інтервенції

Дизайн нерівнозначного порівняння після інтервенції менш надійний, ніж дизайн нерівнозначного порівняння до й після інтервенції. Як і в попередньому випадку, створюється порівняльна група. Таким чином збираються дані за підсумками інтервенції. Хоча це краще, ніж оцінка без порівняльної групи, серйозною проблемою є те, що стартові умови в експериментальній та контрольній групах можуть відрізнятися. Тобто, знаючи про підсумкові показники обох груп, ми нічого не знаємо про вихідні. Відмінності між експериментальною та порівняльною групами можуть бути зумовлені відмінностями в початкових умовах, а не впливом інтервенції. Тим не менш, це може бути оптимальним варіантом у випадках, коли оцінку доводиться проводити постфактум:

$$\begin{array}{lcl} N\ O_1 & X & O_2 \\ N & & O_2 \end{array}$$

Дизайн багаторазового порівняння з перервою

Дизайн багаторазового порівняння з перервою передбачає неодноразове вимірювання показників однієї групи до початку інтервенції, проведення інтервенції та неодноразове вимірювання показників тієї ж групи після закінчення інтервенції. Дизайн багаторазового повторного порівняння в межах однієї групи записується наступним чином:

$$O_1\ 0_2\ 0_3\ X\ O_4\ O_5\ O_6$$

Під перервою мається на увазі перерва у послідовних вимірюваннях показників, що вводиться на час інтервенції. Від класичного порівняння до й після інтервенції цей підхід відрізняє багаторазове повторення вимірювань. Також цей дизайн може передбачати залучення однієї або кількох порівняльних груп.

Дизайн довгострокового оцінювання (лонгітюдний дизайн, проспективний дизайн)

Дизайн довгострокового оцінювання:
Дизайн, що передбачає відстеження ситуації для окремих осіб у кількох точках тривалого проміжку часу

Під час реалізації **дизайну довгострокового оцінювання** (різновид багаторазового порівняння) показники піддослідних оцінюються у кількох точках тривалого проміжку часу. Метою дизайну є виявити, як змінюється стан речей з плином часу. Наприклад, у сфері охорони здоров'я може існувати зацікавленість щодо медичних проблем дітей, матері яких інфіковані ВІЛ/СНІД і приймали ліки для попередження передачі вірусу. Дизайн довгострокового оцінювання дозволяє відстежувати стан дітей протягом тривалого часу. Вивчення результатів дозволяє виявити, чи спостерігаються у таких дітей схожі проблеми зі здоров'ям.

Довгострокове оцінювання дає змогу зібрати багато інформації, яку не можна отримати з інших джерел. Проте це дорогий і складний підхід, для якого також характерні проблеми з вибуванням учасників (які помирають або втрачають контакт з фахівцями з оцінювання).

Дизайн довгострокової оцінки записується наступним чином:

$$X\ O_1\ O_2\ O_3 \ldots$$

Вибірковий дизайн (панельний дизайн)

Вибірковий дизайн:
Дизайн, що передбачає відстеження ситуації для вибірки учасників у кількох точках тривалого проміжку часу

Одним з різновидів дизайну довгострокової оцінки є **вибірковий дизайн**. Замість відстеження показників окремих осіб (як під час звичайної довгострокової оцінки) вибірковий дизайн відстежує стан певної вибірки учасників. Наприклад, програма може розглядати процес зміни поглядів і патернів поведінки стосовно гендеру серед учнів окремої шко-

ли. В цьому випадку вибірковий дизайн передбачатиме збір інформації про гендерні погляди кожного учня одного й того ж класу з першого по шостий роки навчання.

Вибірковий дизайн записується наступним чином:

$$X\ O_1\ O_2\ O_3\ O_4\ O_5\ O_6 \ldots$$

Кореляційний дизайн із застосуванням засобів статистичного контролю

Іноді етичні або практичні проблеми унеможливлюють використання експериментального дизайну оцінки. В найпростішому випадку **кореляційний дизайн** розглядає залежність між двома змінними, які не можна регулювати (або більшою їх кількістю), і вважається неекспериментальним дизайном. Саме з такими простими кореляційними дизайнами часто асоціюється фраза «кореляція не означає причинності». Проте важливо пам'ятати, що для існування причинно-наслідкового зв'язку дві змінні повинні бути пов'язаними. На сьогоднішній день широкого застосування набули складні аналітичні методики з використанням форм множинної регресії. Кореляційні методи часто використовують для формування статистично коректних порівняльних груп. Джонсон (Johnson 2002) вказує на те, що такі порівняльні групи використовуються для відповіді на питання про стосунки, зв'язки, причини й наслідки. Узгодження дизайну довгострокової оцінки з частково кореляційними методами – ефективний шлях до виділення причинних наслідків.

Кореляційний дизайн:
Розглядається залежність між двома або більшою кількістю змінних, які не можна регулювати

Наприклад, може вивчатися залежність між обсягами виконаних домашніх завдань та результатами іспитів за умови однакового рівня мотивації усієї цільової аудиторії. Завдання – встановити кореляцію між часткою виконаних домашніх завдань, не пов'язаною з мотивацією, та часткою набраних на іспитах балів, також не пов'язаною з мотивацією. Якщо можна пов'язати або передбачити залежну змінну за допомогою трьох незалежних змінних – має місце причинний зв'язок (Garbin 2009; psych.unl.edu/psycrs/942/q2/control.ppt).

Розглянемо для прикладу оцінку, що знаходить кореляцію між кількістю жінок на державних посадах та чесністю врядування. Щоб визначити, чи існує залежність між двома змінними, можна зібрати дані про частку жінок на державних посадах країни в різних сферах, а також про зареєстровані масштаби корупції. Звичайно, свідчення про кореляцію не означають існування причинності; навіть якщо уряди з більшою кількістю жінок в своєму складі корелюють з меншою корумпованістю, необхідно виключити будь-які ймовірні альтернативні пояснення цього факту.

Оскільки кореляційний дизайн може організовуватися по-різному, записувати його також можна у кілька різних способів. Перша *(a)* з наведених нижче форм запису відповідає дизайну з трьома групами й одним спостереженням, друга *(b)* показує дві групи, одна з яких є експериментальною. Третя форма запису *(c)* демонструє три різних види заходів (X, Y та Z), після кожного з яких проводиться спостереження:

(a)	O_1	***(b)***	O_1	***(c)***	X	O_1
	O_2		X O_2		Y	O_2
	O_3				Z	O_3

Зіставлення із групою з іншими характеристиками

Зіставлення із групою з іншими характеристиками: Дизайн, що використовується для вимірювання впливу інтервенції на учасників проекту порівняно з тими, хто не бере участі в проекті, шляхом прогнозування імовірності участі в певній групі (наприклад, експериментальній або контрольній) на основі наявних чинників, зазвичай за допомогою логічної регресії для створення гіпотетичної групи. Також використовується для підбору або в якості незалежної змінної – самостійно або разом з іншими змінними

Зіставлення із групою з іншими характеристиками використовується для вимірювання впливу програми на учасників проекту порівняно з тими, хто не бере участі в проекті, але має аналогічні характеристики (White and Masset 2005). Для використання цього методу фахівці з оцінювання мають спочатку зібрати базові дані. Після цього вони повинні визначити наявні характеристики, що можуть бути пов'язані з питанням оцінки (наприклад: «Чи вищі підсумкові оцінки дівчат, що живуть поблизу школи, порівняно з дівчатами, яким доводиться йти до школи більше п'яти кілометрів?»). До наявних характеристик можуть відноситися стать, вік, шлюбний стан, відстань від дому до школи, умови проживання й харчування, кількість братів та сестер, що випускаються з середньої школи, а також черговість народження. Після вибору змінних можна сформувати експериментальну й порівняльну групи, підібравши кожному учаснику експериментальної групи особу з найбільш схожими характеристиками в контрольну групу. Результатом будуть пари окремих осіб або сімей, максимально схожих одна на одну за всіма характеристиками окрім цільової змінної інтервенції (White 2007).

Існують спеціальні комп'ютерні програми для реалізації даного підходу. Найбільш популярна з них «Stata» (Aliendo and Kopeinig 2005).

Неекспериментальні дизайни

Неекспериментальний дизайн: Різновид дизайну оцінювання, за якого не створюються окремі групи, охоплені та не охоплені інтервенцією, а основна увага приділяється опису

Простий крос-секційний дизайн

Простий крос-секційний дизайн оцінювання дає можливість визначити поточний стан справ у конкретний момент часу. Такий вид дизайну часто використовується спільно з опитуванням. Фахівці з оцінювання цікавляться відповідями підмножини загальної вибірки. Така підмножина

може бути сформована на основі таких характеристик, як вік, стать, рівень доходів, освіта, етнічна приналежність або ж ступінь залучення до інтервенції. Мета дизайну – систематично розібрати на складові підмножини вибірки для того, щоб фахівці з оцінювання могли їх детально вивчити.

Простий крос-секційний дизайн:
Дизайн, що фіксує поточний стан характеристик певної підмножини аудиторії в конкретний момент часу; іноді ще називається «разовим»

Дослідження з охопленням різних груп формує вибірку громадян, бенефіціарів вибірки або колишніх учасників інтервенції в певний момент часу. Після цього від них отримуються дані, які включаються у звіт (блок 7.3). Іноді питання може бути спрямоване на визначення поточного становища осіб, які брали участь в інтервенції кілька років тому.

Простий крос-секційний дизайн може давати відповіді, зокрема, на наступні питання:

- Чи розходяться погляди учасників з різним освітнім рівнем на цінність навчання?
- Чи отримали жінки та чоловіки різні навчальні послуги?

Наприклад, питання оцінювання може зосереджуватися на тому, чи задоволені підгрупи громадян чи бенефіціарів інтервенції отриманими послугами або чому вони не користуються послугами. Фахівці з оцінювання використовують такий дизайн, щоб дізнатися про різницю між підгрупами в контексті таких змінних, як отримані послуги, використання послуг або думки про послуги в певний момент часу.

Блок 7.3 «Використання крос-секційного дизайну для відповіді на описові питання»

Фахівці з оцінювання хотіли дізнатися про думки жінок-учасниць щодо програми, покликаної надати їм більше економічних можливостей для започаткування власного малого бізнесу. Їхні відгуки могли пролити світло на те, чи допомогли їм отримані знання ефективно розпочати власну справу, в яких сферах вони працюють, а також чи виявилися надані їм знання корисними при веденні бізнесу. Враховуючи обмеженість ресурсів, фахівці з оцінювання вирішили провести коротке опитування серед недавніх випускниць програми (разовий дизайн). Опитувальник містив питання про демографічні характеристики учасниць, що дозволило порівнювати відповіді в розрізі освітнього рівня, вікових груп та етнічної приналежності.

Дизайн оцінки з охопленням різних груп записується наступним чином:

$$X \qquad O_1$$
$$O_2$$
$$O_3$$
$$''$$

Разовий дизайн

Разовий дизайн: Дизайн, що передбачає вивчення групи, охопленої дією інтервенції, в один певний момент часу після закінчення цієї інтервенції

Разовий дизайн передбачає вивчення групи, охопленої дією інтервенції, в один певний момент часу після закінчення цієї інтервенції. Якщо наводити аналогії, то разовий дизайн – це як фотознімок, на якому надрукована дата. Такий дизайн можна використовувати для пошуку відповіді на питання: «Скільки жінок пройшли навчання?» або «Скільки учасників отримали консультації з працевлаштування протягом вказаного періоду?»

Фахівці з оцінювання можуть використовувати разові дизайни для того, щоб ставити учасникам питання про те, чи сподобалася їм програма або для того, щоб з'ясувати, звідки вони дізналися про запропоновані послуги. Разовий дизайн записується наступним чином:

$$X\ O_1$$

Стратегії відстеження причин

Для встановлення факту зв'язку між інтервенцією та наявними змінами нерідко потрібно застосовувати структуровану та кількісно обчислювану стратегію збору даних. Якщо фахівець з оцінювання проводить швидку оцінку або ж оцінює невелику або нову інтервенцію, такі підходи можуть бути недоступними та небажаними. Хоча завжди, якщо це можливо, слід обирати надійний дизайн оцінювання, у згаданих випадках єдиним доступним варіантом можуть бути слабші підходи з використанням стратегій відстеження причин.

Що можна зробити, якщо вибірка невелика, стратегії збору даних сформульовані дуже вільно, а проводити складний статистичний аналіз недоцільно? Для виключення конкуруючих гіпотез можна використати один або декілька з восьми логічних аргументів. Ці аргументи стосуються представлення логіки причинності й називаються **стратегіями відстеження причин.**

Стратегії відстеження причин: Різновид неекспериментальних дизайнів, що складається з аргументів щодо причинних зв'язків, заснованих на моделях теорії змін і логічному виключенні альтернативних або конкуруючих пояснень

1. *Висновок зі списку причин.* Ми знаємо, що певний результат завжди спричинено фактором A, B, C або D. Якщо жоден з факторів B, C і D не мав місця, ми можемо бути майже повністю впевнені, що причиною став фактор A. Хоча ми не можемо застосувати рандомізацію, але можемо використовувати досвід досліджень, в яких її застосовували.

   У наведеному в блоці 7.1 прикладі селяни знають, що слони втікають, якщо свистіти у свистки (A), калатати в казани (B), кричати (C) або бігають довкола, здіймаючи пил (D). Якщо вони використовують лише підхід A і в результаті слони йдуть геть, селяни можуть бути майже повністю впевнені, що свистіння змушує слонів втікати.

2. *Висновок зі способу дій.* Цей метод корисний, якщо мають місце кілька можливих чинників. Припустимо, нам відомо, що результат майже завжди спричиняється фактором A, B, C або D і в даному випадку C і D були виключені, що звужує вибір до A й B. Окрім того, спостерігався причинний ланцюжок/спосіб дій/патерн подій, характерний для A. Цей висновок підкріплюється, якщо спосіб дій для чинника A значно відрізняється від способу дій для чинника B.

   Якщо селяни дізнаються від жителів іншого села, що слони не втекли, коли ті бігали довкола і здіймали пил, вони можуть бути практично впевнені, що слони втікають від свисту. Цей висновок підкріплюється тим, що свистіння по суті своїй дуже відрізняється від здіймання пилу.

3. *Передування в часі.* Певний ефект спостерігався тільки після початку інтервенції, а не раніше.

   Якщо слони прийшли, а потім селяни почали свистіти й після цього слони залишили село, вони вважатимуть, що свистіння може бути якось пов'язане зі втечею слонів. Якщо селяни спочатку свистіли, а потім слони тим не менш прийшли у село, то, напевне, не свист змусив тварин піти.

4. *Постійний зв'язок.* Ефект спостерігався всюди, де проводилася інтервенція.

   Скажімо, селяни зустрілися з іншими селянами з усього регіону й обмінялися своїми гіпотезами щодо того, що свистіння змушує слонів втікати. В інших селах пробують застосовувати цей підхід, після чого слони йдуть геть. Тоді можна бути практично впевненими, що свист змушує слонів залишати села.

5. *Сила зв'язку.* Наявні зміни були набагато сильнішими там, де реалізовувалася програма (порівняно з місцями, де діяли інші можливі чинники).

   Якщо в регіоні застосовують багато різних методів відлякування слонів, і ті села, що застосовували свист, найуспішніше впоралися з цим завданням, втечу слонів можна пов'язати зі свистінням.

6. *Біологічний градієнт.* Що інтенсивніша дія інтервенції, то значніші спостерігаються зміни.

   Уявімо, що селяни, відганяючи слонів від села, використовують кілька методів. Коли вони дуже голосно свистять у багато свистків, слони йдуть геть. Коли вони свистять лише в один свисток, слони не йдуть. Можна зробити висновок про зв'язок між голосним свистінням і втечею слонів.

7. *Зв'язність.* Зв'язок між інтервенцією і наявними змінами логічно узгоджується з іншими фактами, які ми знаємо про інтервенцію й результати.

   Небезпечні тварини – гіпопотами, крокодили та гієни – втікають, коли селяни свистять у свистки. Можна зробити логічний висновок, що свистіння допомагає вигнати небезпечних тварин з села. Отже, цей підхід можна застосувати і до слонів.

8. *Аналогія.* Форма зв'язку між інтервенцією та наявними змінами нагадує добре відомий зв'язок між аналогічною інтервенцією та її наслідками.

   Селяни дізнаються про те, що у південноамериканському селі за допомогою пронизливого свистіння відганяють пум. Вони вважають, що пума боїться таких звуків. Тоді африканські селяни побачать аналогію і дійдуть висновку, що голосні, різкі звуки можуть відганяти слонів.

У кожному з цих випадків діє той самий принцип: дослідник систематично виключає альтернативні пояснення, доки не переконається, що наявні зміни найімовірніше спричинені (або ні) в основному (або певним чином) саме інтервенцією.

Розробляючи стратегію збору даних, фахівці з оцінювання мають визначити, які з наведених вище фактів можна й необхідно збирати, а потім планують, як відповідні факти будуть отримані. Не всі ці дані є необхідними для встановлення причинно-наслідкових зв'язків; фахівці з оцінювання збирають якусь частину, що дає найґрунтовніші підстави для висновків з урахуванням рішень, які мають бути прийняті за результатами оцінки. Але одночасне використання одразу кількох стратегій дає міцніше підґрунтя для висновків щодо причинно-наслідкових зв'язків.

Дизайн з вивченням практичного прикладу (дизайн вивчення випадку)

Вивчення випадків:
Неекспериментальний дизайн, що передбачає глибокий, комплексний опис і розуміння інтервенції в цілому і в контексті

Вивчення практичного прикладу – це різновид неекспериментальної оцінки. Він не передбачає випадкового розподілу або створення порівняльної групи. Вивчення практичного прикладу часто використовується в тих випадках, коли фахівець з оцінювання прагне добитися глибокого розуміння процесу, події або ситуації й пояснити отримані результати. Цей підхід стає у нагоді, якщо питання ставляться щодо того, як щось працює або чому щось відбувається. Він особливо корисний, якщо інтервенція має інноваційний або експериментальний характер або її сутність не до кінця зрозуміла. Вивчення практичних випадків виходить за межі простого описання, воно також охоплює тлумачення ситуацій найбільш обізнаними про них особами.

Вивчення практичних прикладів часто використовується для оцінки інтервенцій розвитку. Такий підхід особливо добре підходить для описання сприйняття реалізації інтервенції учасниками, а також причин подій, що мали при цьому місце. Описове вивчення практичного прикладу може застосовуватися для дослідження крайніх варіантів реалізації програми або типової інтервенції.

В ході вивчення практичних прикладів можуть застосовуватися кількісні та якісні методи збору даних – як окремо, так і разом. Вивчатися може один або кілька прикладів. Мета і завдання – зосередитися на поглибленому розумінні наслідків інтервенції для організацій, спільнот, програм, міст або країн.

Наприклад, для оцінки роботи громадського транспорту в країні достатньо порівняти поточні значення ключових показників з базовими та цільовими. Якщо за ключові показники взято загальний кілометраж пробігу транспортних засобів громадського транспорту, кількість користувачів системи та отримані доходи, можна виконати оцінку за допомогою загальнонаціонального дослідження. Однак якщо актуальними є інші питання, що вимагають збору більш прихованих даних, необхідне вивчення практичних прикладів.

Наприклад, якщо потрібно оцінити програму з покращення роботи транспорту в сільській місцевості, фахівець з оцінювання може вивчати, що обирають люди, які використовують громадський транспорт. Дизайн може передбачати отримання даних безпосередньо від сільських жителів. Для збору таких даних в загальнонаціональному масштабі потрібно багато ресурсів. Простіше збирати їх в межах вузько визначеної географічної області (практичний приклад).

Крім того, фахівці з оцінювання можуть вивчати кілька практичних прикладів. У нашому випадку це означає збір даних з кількох місцевостей. Практичні приклади можуть обиратися випадково або цілеспрямовано, за певними критеріями (приклади найкращого випадку, типового випадку, найгіршого випадку, включаючи тільки ізольовані сільські регіони, а також сільську місцевість поблизу великих міст). Ті самі стратегії збору даних, що використовуються для вивчення одного практичного прикладу, можна застосовувати і для вивчення кількох прикладів.

Вивчення практичних прикладів в розвитку є актуальним у випадках, коли потрібно отримати розуміння конкретної ситуації для корегування політики або практики. Вивчення практичних прикладів не просто ефективніше за великі загальнодержавні дослідження, воно також надає поглиблену інформацію, яка часто є корисною для осіб, що приймають рішення (блок 7.4). Приміром, порівняльне дослідження використання клінік безкоштовної імунізації може допомогти з'ясувати, чому один підхід ефективніший за інший.

Дизайн оцінювання з вивченням практичних прикладів записується наступним чином:

O_1

O_2

O_3

Блок 7.4 «Приклад дизайну опису випадків для описових питань»

Дослідження з вивчення програми мікрокредитування в Індії мало за мету дослідити способи, що пояснювали процес формування маркетингових ідей жінками та їх втілення. Для опису випадків обрали п'ять жінок та їхні проекти і далі відслідковували прогрес упродовж трьох років.

Дизайни для описових питань

До описових відносяться питання типу «скільки?» Вони можуть виявляти сприйняття або думки. Описові питання зазвичай спираються на описові або неекспериментальні дизайни. Забезпечуючи відповіді на описові питання, ці дизайни не передбачають створення порівняльної групи, яка не охоплюється дією інтервенції. Вони зосереджуються лише на тих, на кого поширюється дія інтервенції. Деякі дизайни, що використовуються для роботи з описовими питаннями – ті ж самі, що й для причинно-наслідкових питань.

До найчастіше застосовуваних дизайнів для відповіді на описові питання належать наступні:

- прості дизайни оцінюванняи з охопленням різних груп;
- разові дизайни;
- порівняння стану до й після інтервенції;
- дизайни багаторазового порівняння з перервою;
- лонгітюдні;
- вивчення практичних прикладів.

Далі ми розглянемо, як деякі з цих дизайнів працюють з описовими питаннями.

Порівняння стану до й після інтервенції

Дизайни з порівняння стану до й після інтервенції розглядалися в попередньому розділі під час обговорення дизайнів для причинно-наслідкових питань. Ці дизайни також можна використовувати для відповіді на описові питання. Дизайн порівняння стану до й після інтервенції передбачає, що фахівці з оцінювання ставитимуть питання про характеристики групи до початку інтервенції та після її завершення; порівняльної групи немає (блок 7.5). Наприклад, можна поставити питання, чи покращили учасники програми свої знання з догляду за дітьми, а потім провести відповідне тестування перед початком програми та після її завершення. Оцінка з використанням порівняння стану до й після інтервенції записується наступним чином:

$$O_1 \; X \; O_2$$

Блок 7.5 «Використання дизайну порівняння стану до й після інтервенції для відповіді на описові питання»

Дизайн порівняння стану до й після інтервенції може розглядати розміри заробітної плати вибірки учасників програми професійної освіти до початку інтервенції та через два роки після її завершення, щоб встановити середній розмір збільшення доходів. Цей дизайн також легко можна трансформувати в дизайн порівняння до й після інтервенції з охопленням різних груп. Для цього треба ставити питання підгрупам осіб, які працюють в різних сферах. Це дасть інформацію про зв'язок між підвищенням заробітної плати та різними видами діяльності.

Дизайн багаторазового порівняння з перервою

Дизайни багаторазового порівняння з перервою були представлені раніше як дизайни для причинно-наслідкових питань. Дизайн багаторазового порівняння з перервою також можна використовувати для відповіді на описові питання (блок 7.6). Дизайни цього виду розглядають зміни, що відбуваються з плином часу, зазвичай з метою виявлення тенденцій. Якщо такі дизайни використовуються для відповіді на описові питання, метою є дослідження й опис змін з плином часу до й після інтервенції. Таким чином, дизайни багаторазового порівняння з перервою можуть використовуватися для виявлення тенденцій. Дизайн багаторазового порівняння з перервою записується наступним чином:

$$O_1 \; O_2 \; O_3 \; X \; O_4 \; O_5 \; O_6 \ldots$$

Блок 7.6 «Приклад дизайну багаторазового порівняння з перервою для описових питань»

Можуть розглядатися показники дитячої смертності до й після інтервенції з надання харчових добавок для матерів. Інший варіант: можуть розглядатися зміни відносно учасників програми кредитування мікропідприємств до жінок-підприємців до й після цієї програми.

Лонгітюдний дизайн

Лонгітюдний дизайн: Різновид дизайну багаторазового порівняння з перервою, що передбачає повторювані вимірювання однієї змінної для одних і тих же піддослідних

Лонгітюдний дизайн – це різновид дизайну багаторазового порівняння з перервою, що передбачає повторювані вимірювання однієї змінної для одних і тих же піддослідних. Якщо лонгітюдний дизайн використовується для відповіді на описові питання, він може застосовуватися, наприклад, для з'ясування того, чи зберігаються з часом отримані знання в дітей-учасників освітньої програми.

Для відповіді на описові питання також можна використовувати панельний дизайн, який є різновидом лонгітюдного, що передбачає відстеження меншої групи осіб у кількох точках часу з детальною фіксацією їхнього досвіду. Панельні дизайни майже завжди застосовують якісні методи (відкриті запитання в анкетах, поглиблені інтерв'ю, спостереження), а також кількісні дані. Вони дають можливість глибше дослідити зміни, які люди переживають внаслідок інтервенції. Лонгітюдний дизайн записується наступним чином:

$$X\ O_1,\ O_s,\ O_3 \ldots$$

Дизайни для нормативних питань

Логіка для нормативних питань є аналогічною логіці описових питань. Відмінність полягає в тому, що нормативні питання завжди оцінюються в порівнянні з певним критерієм або стандартом. Висновки порівнюються з таким стандартом, який може спиратися на певні показники й цілі. Як правило, однакові дизайни підходять і для нормативних, і для описових питань.

Аудит ефективності розглядає деякі аспекти ефективності роботи організації (Mayne 2005, 2006). Аудити ефективності можуть бути дуже схожими на нормативні оцінки. Барзлі (Barzelay 1997) визначає сім видів аудиту ефективності, ґрунтуючись на дослідженні країн-членів Організації економічного співробітництва й розвитку (ОЕСР). В табл. 7.2 наведено чотири найбільш актуальних види.

Блок 7.7 ілюструє згадане вище твердження: якщо неможливо створити експериментальний дизайн оцінювання з випадково сформованими контрольними групами, можна використати квазі-експериментальний дизайн з порівняльними групами.

Дизайни для причинно-наслідкових питань

Для роботи з причинно-наслідковими питаннями зазвичай використовуються експериментальні дизайни.

Потреба у ретельнішому плануванні оцінювання

Які дизайни є дієвими для відповіді на причинно-наслідкові питання?

Лідери в сфері міжнародного оцінювання розвитку останнім часом обговорюють потребу ретельнішого планування оцінювання програм (Davidson 2006; Scriven 2006; Bamberger and White 2007). Вони зазначають, що більшість оцінювань, проведених офіційними установами розвитку – це оцінки процесу, зосереджені на тому, наскільки добре програма функціонує. Збільшення кількості спільних оцінювань дозволило враховувати думки бенефіціарів, але «не надало даних, достатніх для кількісного аналізу впливу» (Bamberger and White 2007, p. 58).

Таблиця 7.2. Чотири види аудиту результативності роботи

Вид	Одиниця аналізу	Мета
Аудит ефективності	Організація або юрисдикція; елемент процесу або програми	Визначити можливості скорочення бюджетних витрат, пов'язаних з отриманням результатів програми
Аудит дієвості/ результативності	Політика, програма або значний елемент програми	Оцінити вплив державних політик; оцінити дієвість політики або програми
Аудит потенціалу управління ефективністю	Проблема управління організацією або юрисдикцією	Оцінити потенціал систем і процедур юрисдикції, організації або програми в контексті досягнення поставлених цілей.
Аудит інформації про ефективність	Організація	Перевірити якість інформації про ефективність, наданої організацією

Джерело: Взято з Barzelay 1997.

Блок 7.7 «Вплив програм з професійного навчання на звільнених робітників»

Багато країн, що розвиваються, стикаються з проблемою перепідготовки робітників після скорочення кадрів на державних підприємствах. Оцінювати такі програми непросто, тому що вони часто складаються з кількох різних компонентів. Вони можуть слугувати різним структурам і по-різному вимірювати результати (працевлаштованість, самозайнятість, доходи за місяць, доходи за годину). Для оцінювання програм професійної підготовки фахівці з оцінювання поставили наступні питання:

- Чи успішнішим було повернення на ринок праці учасників програми порівняно з тими, хто не брав у ній участі?
- Наскільки економічно ефективною була кожна з програм?

Для групи учасників підібрали аналогічну групу працівників, які не брали участі в програмі. Для визначення впливу програми підготовки були використані адміністративні дані, дані опитувань і статистичні методи. Існує ймовірність, що учасники програми чимось відрізнялися від неучасників, що зумовило різницю в результатах. Можливо, наприклад, вони були більш вмотивованими, ніж люди з порівняльної групи, або мали більший досвід роботи, що полегшило для них процес працевлаштування. Для підвищення надійності цього дизайну буде необхідно довільно розподілити тих, хто має право на участь в програмі, щоб виключити можливість домінування різних характеристик в групах.

Джерело: Автори.

Підходи з орієнтацією на результат та підвищена увага до Цілей розвитку тисячоліття (ЦРТ) спричинили нові заклики до демонстрації впливу. Заклики до оцінювання впливу, що його забезпечив проект, програма або політика, надходять з різних джерел, серед яких:

- Монтерейська конференція з фінансування розвитку 2002 р., яка закликала до активнішого застосування агенціями розвитку управління з орієнтацією на результат;
- Паризькі домовленості 2005 р., що стимулювали співпрацю донорів у просуванні оцінювання впливу;
- організація Poverty Action Lab, яка відстоює використання рандомізованих дизайнів та пропонує навчальні програми, присвячені таким дизайнам, для країн, що розвиваються;

- Центр глобального розвитку (ЦГР), який активно відстоює необхідність застосування ретельнішого планування оцінювань, зокрема у своїй публікації «Та коли ми врешті навчимось?» (When Will We Ever Learn?) (ЦГР, 2006). ЦГР також закликав незалежні агенції розвитку забезпечити більшу незалежність і ретельність при оцінюванні розвитку (Bamberger and White 2007).

Нові уявлення про міжнародний розвиток і оцінювання розвитку сприяють активнішому застосуванню оцінювання впливу. Зусилля уряду Іспанії демонструють це нове мислення і стимулюють рух до ретельнішого планування оцінювання (блок 7.8). Говорячи про «ретельність», часто мають на увазі експеримент з використанням рандомізованої контрольної групи. Равальйон, директор дослідницького відділу Світового банку, відзначає, що рандомізація, тим не менш, не завжди є адекватним рішенням, особливо в контексті розвитку (Ravallion 2009). Наприклад, можна забезпечити рандомізацію для певних інтервенцій, пов'язаних з розвитком. Але як забезпечити випадковість розміщення інфраструктурних проектів і пов'язаних з ними програм? Равальйон обговорює те, як рандомізовані експерименти можуть стикатися з серйозними проблемами узагальнення через те, що їх часто проводять лише для вузьких і відокремлених проектних інтервенцій в певних умовах. Він також описує супутні наслідки. Таким чином, ті, хто обирає рандомізовані дизайни, часто припускають, що програма не впливає на тих, хто не бере в ній участі. Проте відомо, що супутні наслідки виходять за межі інтервенцій розвитку. Це пов'язано з кількома факторами: рухом суб'єктів, які потрапляють в сферу дії інтервенції та виходять з неї, довгостроковим наслідуванням (одна група наслідує дії іншої, якщо вони видаються ефективними), а також зі втручанням місцевих урядів, які спрямовують вивільнені завдяки роботі організацій розвитку на проведення аналогічної інтервенції для обраної контрольної групи.

Блок 7.8 «Підтримка Іспанією ретельних дизайнів оцінювання»

Іспанія прагне покращити дієвість допомоги, стимулюючи застосування управління з орієнтацією на результат у власній агенції розвитку, а також у країнах-партнерах. Було запроваджено нову програму для підтримки Світового банку в оцінюванні впливу інноваційних програм з метою покращення результатів гуманітарного розвитку. Програма «Іспанський довірчий фонд Світового банку для оцінювання впливу» (SIEF) – найбільший довірчий фонд в історії, коли-небудь створений Світовим банком – зосереджувалася на оцінюванні впливів і результатах (World Bank 2007).

Дизайни оцінювання мали бути відповідними у кожній ситуації та до кожного виду питань, які ставляться в ході оцінки. Петон (Patton 2007) розглядає відповідність дизайнів різним видам інтервенцій розвитку, відзначаючи, що «різні ситуації впливу та різні цілі оцінки потребують різних методів» (ст. 1). Фахівці з оцінювання мають починати зі з'ясування ситуації; дизайн виникне з неї. Як зазначає Петон: «Існує багато ситуацій впливу розвитку, а отже, багато підходів до оцінки впливу» (ст. 2).

Бамбергер і Петон (Bamberger and Patton 2007) пропонують наступні рекомендації, спрямовані на збільшення надійності дизайну оцінки й подолання часових і бюджетних обмежень:

1. Будуйте дизайн оцінювання на основі якісної теоретичної моделі програми. Це допоможе пояснити зв'язки в причинному ланцюжку та визначити припущення. Крім того, такий підхід дає змогу визначити місцеві економічні, політичні, інституційні, екологічні та соціально-культурні фактори, що пояснюють відмінності в ефективності та результатах.
2. Затвердіть ефективний дизайн, в якому використовується кілька методів, та поєднуйте якісний і кількісний підходи:
   - використовуйте якісні дані для тріангуляції, щоб отримати додаткову базу на підтримку кількісних результатів, сформувати дослідження та витлумачити кількісні результати;
   - максимально використовуйте вторинні дані, в т. ч. дані моніторингу проекту;
   - завжди, якщо дозволяють час і бюджет, збирайте дані в додаткових точках циклу проекту;
3. Спрощуйте інструменти збору даних.
4. Творчо працюйте зі вторинними даними, зокрема, використовуючи дані з попередніх досліджень в якості базових або для формування контрольних чи порівняльних груп.
5. Розгляньте можливість скорочення розміру вибірки, якщо це детермінована вибірка.
6. Скорочуйте витрати на збір даних, залучаючи дешевших інтерв'юерів, віддаючи прямому спостереженню перевагу перед квартирним опитуванням, об'єднуючи або синхронізуючи зусилля з іншими оцінками, спільно проводячи опитування.

Фахівці з оцінювання повинні розглядати всі можливі дизайни, щоб отримати якомога надійніші результати (табл. 7.3). Вибір правильного дизайну є критично важливим, тому що «різні методи можуть давати зовсім різні результати. Проблема в тому, щоб визначити, який дизайн та які методи є найбільш адекватними, продуктивними та корисними в даній ситуації» (Patton 2002, p. 255).

Оцінювання – це одночасно мистецтво й наука. Приймаючи рішення щодо дизайну, фахівці з оцінювання мають пам'ятати, що ідеального дизай-

ну не існує; всі оцінки передбачають компроміси щодо часу, витрат і практичності. Ці компроміси варто визнавати, оцінюючи їх можливий вплив на результати й висновки.

Таблиця 7.3. Переваги й недоліки популярних експериментальних дизайнів

Вид дизайну	Переваги	Недоліки
Експериментальний		
Порівняння	Контролюються внутрішні загрози валідності	Важко проводити в державному секторі
	Цей дизайн корисний для розгляду відмінностей між групами; контролюються фактори історії й дозрівання, якщо порівняльна група підібрана вдало	Вибір та вибування учасників – фактори ризику
Квазі-експериментальний		
Порівняння стану до й після інтервенції	Цей дизайн корисний для отримання контексту для вимірювання змін	Загрозою можуть бути ефекти тестування, інструментів, регресії до середнього значення, вибування учасників, історії та дозрівання
Неекспериментальний		
Разовий дизайн	Цей дизайн корисний для роботи з описовими й нормативними питаннями; багато разових дизайнів закладають основи для вивчення ситуації	Низька ефективність з точки зору відповіді на причинно-наслідкові питання

Джерело: Автори.

Равальйон (Ravallion 2009) рекомендує орієнтувати підхід за питаннями оцінювання. Найважливішими можуть бути такі питання: «Хто виграє і хто програє від інтервенції? Чи виглядає інтервенція на практиці так само, як на папері? Якою була мінливість реалізації? Які варіанти реалізації слід розвивати? Для кого та за яких умов інтервенція є дієвою?»

В цілому кожне питання оцінювання потребує дизайну – експериментального, квазі-експериментального або неекспериментального. Експериментальний дизайн спрямований на контроль всіх факторів «експерименту» з метою визначення або передбачення розвитку подій. Експериментальний дизайн використовує рандомізований розподіл піддослідних по щонайменше двох групах – експериментальній та контрольній.

Квазі-експериментальний дизайн схожий з експериментальним у тому, що використовує дві або більше груп, але не передбачає випадкового розподілу учасників по групах. Зазвичай створюється порівняльна група для того, щоб мати можливість вивчати одну групу, охоплену дією інтервенції, та ще одну, аналогічну, на яку така дія не поширюється. Неекспериментальні дизайни носять більш описовий характер. Вони не передбачають ні рандомізованого розподілу, ні використання порівняльної групи.

Для більшості інтервенцій розвитку важко створити дизайн, який дає відповіді на причинно-наслідкові питання. Це пов'язано зі складністю ситуації. Важко «довести», що саме інтервенція створює спостережений ефект.

Дизайни оцінювань здатні допомогти визначити вплив програми, надаючи фахівцю з оцінювання можливість контролю реалізації та вимірювання програми. Метою є виключення інших можливих пояснень спостережуваних подій.

Розглядаючи причинно-наслідкові питання, фахівець з оцінювання має звернути увагу на один або кілька з наступних видів дизайнів оцінюванння:

- дизайн нерівнозначного порівняння;
- дизайн багаторазового порівняння з перервою;
- кореляційний дизайн із застосуванням засобів статистичного контролю;
- лонгітюдний дизайн;
- панельний дизайн;
- порівняння стану до й після інтервенції;
- кроссекційний дизайн;
- зіставлення із групою з іншими характеристиками;
- відстеження причин.

Описові питання зазвичай оцінюються з використанням описових або неекспериментальних дизайнів. Дизайни для описових питань зосереджуються лише на відповіді на питання, що не пов'язані зі встановленням впливових факторів або зіставленням показників з певними критеріями. Деякі дизайни, що використовуються для роботи з описовими питаннями – ті ж самі, що й для причинно-наслідкових питань.

Розглядаючи описові питання, фахівець з оцінювання має звернути увагу на один або кілька видів дизайнів оцінювання:

- разові дизайни;
- кроссекційний дизайн;
- порівняння стану до й після інтервенції;
- дизайн багаторазового порівняння з перервою;
- лонгітюдний дизайн;
- вивчення практичних прикладів.

Логіка для нормативних питань є аналогічною логіці питань описових. Відмінність полягає в тому, що нормативні питання завжди оцінюються в порівнянні з певним критерієм або стандартом.

Багато лідерів міжнародного оцінювання розвитку закликають до більшої ретельності у плануванні оцінювань. Спостерігається рух проектів, програм і політик в напрямку управління з орієнтацією на результат. Відповідно, можна очікувати збільшення кількості дизайнів оцінювань, спрямованих на встановлення впливових факторів.

Завдання до Глави 7

Практична вправа 7.1: Вибір дизайну оцінювання

Вас попросили виміряти вплив побудови комунальної клініки для навчання батьків лікуванню поширених сімейних захворювань та виявленню проблем, що можуть виявитися серйозними. Цілі програми – підвищити кількість батьків з базовою обізнаністю про медичну профілактику, надання першої допомоги та стратегії раннього лікування, а також зменшити кількість дітей і літніх людей, захворювання яких переходять у фазу загострення.

1. Якими є цільові довгострокові результати?
2. Сформулюйте причинно-наслідкове, нормативне й описове питання для цієї оцінки.
3. Який би дизайн ви використали для відповіді на ці питання? У чому полягають сильні сторони й обмеження цього дизайну? Чим обраний вами дизайн кращий за інші можливі дизайни?

Практична вправа 7.2: Вибір дизайну оцінювання й стратегії збору даних

Вас попросили створити дизайн оцінювання для шестимісячного дослідження з метою оцінки дієвості кампанії інформування про медичну профілактику в вашій країні. Передбачається, що кампанія складатиметься з дводенних семінарів, які будуть проводитися професійними медиками в різних громадах по всій країні. Мета вашого оцінювання – визначити, чи забезпечила інформаційна кампанія покращення практики в сфері охорони здоров'я. У вашому розпорядженні – помірний бюджет і шість асистентів, які мають допомагати вам планувати й проводити оцінювання.

1. Чи є ваше основне питання оцінювання описовим, нормативним або причинно-наслідковим? Дайте пояснення.
2. Чи повинна ваша стратегія збору даних бути структурованою, вільною або ж поєднувати обидва підходи? Чому?
3. Як би ви визначили найбільш важливі для вимірювання довгострокові результати, та як би ви їх вимірювали?
4. Який дизайн/дизайн оцінювання ви б використали?
5. У чому полягають сильні й слабкі сторони вашого дизайну?

Додаток 7.1. Стислий виклад ключових термінів і ознак різних видів дизайнів

Терміни IPDET	Види дизайнів	Візуальний запис	Основні переваги	Основні недоліки
Експериментальний	**Для експериментальних дизайнів характерне випадкове розподілення при формуванні контрольної та експериментальної груп**			
	Рандомізований Порівняльна група	O_1 X O_2 O_1 O_2	Висока внутрішня валідність, виявляє зміни з плином часу за наявності та відсутності охоплення дією інтервенції	Значні витрати, етичні міркування, складність узагальнення
	Тільки після інтервенції Рандомізований Порівняльна група Без порівняння до тесту	O_3 X O_4	Висока внутрішня валідність, трохи практичніший, корисний для порівняння довгострокових результатів	Не визначає зміни з плином часу
Квазі-експериментальний	**Всі квазі-експериментальні дизайни дещо слабші за експериментальні з точки зору валідності, або ж мають низьку внутрішню валідність. Квазі-експериментальні дизайни передбачають порівняння, але без рандомізованого розподілу по групах**			
Порівняння стану однієї групи до й після інтервенції	Всередині групи Дизайн порівняння стану до й після інтервенції	O_1 X O_2	Практичний, необхідно враховувати контекст	Загрози тестування, інструментарію та регресії
Дизайн нерівнозначного порівняння до й після інтервенції	Порівняння стану до й після інтервенції Нерівнозначне порівняння груп	N O_1 X O_1 N O_2 O_4	Необхідно враховувати контекст, більша впевненість порівняно з порівнянням груп	Виключається ефект історії, важко контролювати всі змінні, які зумовлюють нерівнозначність груп
Дизайн нерівнозначного порівняння після інтервенції	Порівнюються дані тільки після завершення інтервенції	N O_1 X O_2 N O_2	Може бути найкращим дизайном для ситуацій, коли оцінка проводиться постфактум	Немає відомостей про вихідні дані щодо групи, охопленої дією інтервенції
	Порівняння з нерівнозначною порівняльною групою тільки після завершення інтервенції	O_1 X O_2	Практичний, необхідно враховувати контекст, контролюються ефекти тестування, інструментарію, регресії, історії	Етичні міркування, вибір може негативно вплинути на валідність
	Порівняння з різними видами інтервенцій після їх закінчення	X O_1 Y O_2 Z O_3	Є можливість порівняння інтервенцій, необхідно враховувати контекст	Залишається багато загроз

Дизайн багаторазового порівняння з перервою (добре підходить для описових питань)	Багаторазове порівняння (всередині групи)	$O_1 O_2 O_3 X O_4 O_5 O_6$	Частково контролюється загроза ефекту історії, контролюється дозрівання	Загроза упередженості тестування
	Багаторазове порівняння груп (нерівнозначних)		Виключаються загрози ефекту історії, знижується регресія до середнього значення	Дорогий і трудомісткий метод, складно відстежувати людей протягом тривалого часу
Лонгітюдний дизайн	Немає вихідних даних	X $O_1 O_2 O_3$...	Окремі особи відстежуються протягом тривалого часу	Дорогий метод, складно відстежувати окремих людей протягом тривалого часу
Панельний дизайн	Протягом певного часу відстежується одна група	X $O_1 O_2 O_3 O_4 O_5$...	Детальна інформація	Метод може бути витратним
Кореляційний дизайн із застосуванням засобів статистичного контролю		O_1 O_2 O_3	Використовує статистику для визначення кореляції між окремими випадками з метою виокремлення потенційних загроз, визначає важливі відносини та змінні, що можуть вплинути на вимірювання	Вимагає використання великих вибірок не дає змоги зробити висновок щодо причини, носить гіпотетичний характер
Зіставлення із групою з іншими характеристиками	Учасники інтервенції порівнюються зі схожими на них неучасниками		Використовується для оцінювання програм з добровільною участю – надійніша оцінка впливу проекту або програми на учасників	Вимагає значних наборів даних та обчислювальних потужностей
Неекспериментальний	**Ідеальний для опису. Всі неекспериментальні дизайни поступаються експериментальним**			
Порівняння з охопленням різних груп	Порівняння всередині групи, порівняння різних груп	X O_1 O_2 O_3 "	Чіика картина у певній точці	Немає чітких даних про те, що відбувається з плином часу
Разовий дизайн		X O_1	Легкість, практичність	Багато загроз для валідності, слабкий дизайн
Стратегії відстеження причин	За допомогою моделей теорії змін і логіки встановлюються причинно-наслідкові зв'язки			
Вивчення практичного прикладу		O_1 O_2 O_3	Детальна контекстуальна інформація	Трудомісткий підхід з низькою внутрішньою валідністю

Посилання та рекомендована література

Aliendo, Marco, and Sabine Kopeinig. 2005. *Some Practical Guidance on the Implementation of Propensity Score Matching.* Discussion Paper 1588, IZA, Institute for the Study of Labor, Bonn.http://ftp.iza.org/dp1588.pdf.

Bamberger, Michael, and Howard White. 2007. "Using Strong Evaluation Designs in Developing Countries: Experience and Challenges." *Journal of Multidisciplinary Evaluation* 4 (8): 58–73.

Barzelay, M. 1997. "Central Audit Institutions and Performance Auditing: A Comparative Analysis of Organizational Strategies in the OECD." *Governance:*
An International Journal of Policy and Administration 103: 235–60.

Boruch, Robert. 2004. "Ethics and Randomized Trials." International Program for Development Evaluation Training (IPDET) presentation, Ottawa.

Brossart, Daniel F., Daniel L. Clay, and Victor L. Willson. 2002. "Methodological and Statistical Considerations for Threats to Internal Validity in Pediatric Outcome Data: Response Shift in Self-Report Outcomes." *Journal of Pediatric Psychology* 27 (1): 97–107.

Brown, Randall S., and Ellen Eliason Kisker. 1997. "Nonexperimental Designs and Program Evaluation." *Children and Youth Services Review* 19 (7): 541–66. http://www.aei.org/publications/pubID.17770/pub_detail.asp.

Campbell, D. T., and J. C. Stanley. 1963. "Experimental and Quasi-Experimental Designs for Research." In *Handbook of Research on Teaching,* ed. N. L. Cage. Chicago: Rand-McNally.

CGD (Center for Global Development). 2006. *When Will We Ever Learn? Improving Lives through Impact Evaluation.* Washington, DC.

Chatterji, M. 2007. "Grades of Evidence: Variability in Quality of Findings in Effectiveness Studies of Complex Field Interventions." *American Journal of Evaluation* 283: 239–55.

Cohen, M. 2001. "Evaluating Microfinance's Impact: Going Down Market." In *Evaluation and Poverty Reduction,* eds. O. N. Feinstein and R. Picciotto, 193–203. New Brunswick, NJ: Transaction Publishers.

Cook, T. D., and D. T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis for Field Settings.* Boston: Houghton Mifflin.

Davidson, E. J. 2000. "Ascertaining Causality in Theory-Based Evaluation." *New Directions for Evaluation* 87: 17–26.

———. 2006. "The RCT's Only Doctrine: Brakes on the Acquisition of Knowledge?"
Journal of Multidisciplinary Evaluation 6: ii–v.

Garbin, Cal. 2009. *Statistical Control.* Lincoln, NE: University of Nebraska. psych.unl.edu/psycrs/942/q2/control.ppt.

Grembowski, D. 2001. *The Practice of Health Program Evaluation.* Thousand Oaks, CA: Sage Publications.

Homer-Dixon, Thomas. 1995. *Strategies for Studying Causation in Complex Ecological Political Systems.* Occasional Paper; Project on Environment, Population, and Security. American Association for the Advancement of Science, Washington, DC, and the University of Toronto. http://www.library.utoronto.ca/pcs/eps/method/methods1.htm.

Mayne, John. 2005. "Ensuring Quality for Evaluation: Lessons from Auditors." *Canadian Journal of Program Evaluation* 20 (1): 37–64.

———. 2006. "Audit and Evaluation in Public Management: Challenges, Reforms, and Different Roles." *Canadian Journal of Program Evaluation* 21 (1): 11–45.

Miles, M. B., and A. M. Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook.* 2nd ed. Thousand Oaks, CA: Sage Publications.

NIOSH (National Institute for Occupational Safety and Health). 1999. *A Model for Research on Training Effectiveness TIER.* Centers for Disease Control, Atlanta. http://www.cdc.gov/niosh/99-142.html.

Office of Juvenile Justice and Delinquency Prevention. 1989. *Evaluating Juvenile Justice Programs: A Design Monograph for State Planners.* Report prepared for the U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention by Community Research Associates, Inc.

Patton, Michael Q. 2002. *Qualitative Research and Evaluation Methods.* 3rd ed. Thousand Oaks, CA: Sage Publications.

———. 2005. "The Debate about Randomized Controls in Evaluation: the Gold Standard Question." International Program for Development Evaluation Training (IPDET) presentation, Ottawa, July.

———. 2007. *Design Options and Matching the Type of Impact Evaluation and Attribution Issue to the Nature of the Intervention: Background Discussion on Impact Evaluation for International Development Efforts.* November.

———. 2008. "The Logic of Experimental Designs and 10 Common Criticisms: The Gold Standard Debate." In *Utilization-Focused Evaluation.* 4th ed. Thousand Oaks, CA: Sage Publications.

Powell, Keith D. 2004. "Research Methods on Psychological Science: Psychology 242." Powerpoint presentation, November 8. Department of Psychology, University of Chicago, Chicago, IL.

Prennushi, Giovanna, Gloria Rubio, and Kalanidhi Subbarao. 2002. *PRSP Sourcebook Core Techniques.* Washington, DC: World Bank. http://go.worldbank.org/3I8LYLXO80.

Project STAR. 2006. *Study Designs for Program Evaluation.* Aguirre Division, JBS International, Inc. Bethesda, MD. http://www.nationalserviceresources.org/filemanager/download/performance Measurement/Study_Designs_for_Evaluation.pdf.

Ravallion, Martin. 2009. "Should the Randomistas Rule?" *The Economists Voice* 6(2): 1–5.

Schweigert, F. J. 2006. "The Meaning of Effectiveness in Assessing Community Initiatives." *American Journal of Evaluation* 27: 416. http://aje.sagepub.com/cgi/content/abstract/27/4/416.

Scriven, Michael. 2006. "Converting Perspective to Practice." *Journal of Multidisciplinary Evaluation* 6: 8–9.

———. 2007. *Key Evaluation Checklist.* February. http://www.wmich.edu/evalctr/checklists/kec_feb07.pdf.

Stake, R. E. 1995. *The Art of Case Study Research.* Thousand Oaks, CA: Sage Publications.

Stufflebeam, Daniel L. 2004. *Evaluation Design Checklist.* Western Michigan University, Evaluation Center, Kalamazoo, MI. http://www.wmich.edu/evalctr/checklists/evaldesign.pdf.

Stufflebeam, D. L., G. F. Mdaus, and T. Kellaghan, eds. 2000. *Evaluation Models: Viewpoints on Educational and Human Services Evaluation.* Boston: Kluwer.

Trochim, W. M. "The Research Methods Knowledge Base. http://www.socials researchmethods.net/kb.

Trochim, W., and D. Land. 1982. "Designing Designs for Research." *Researcher* 1 (1): 1–16.http://www.socialresearchmethods.net/kb/desdes.htm.

United Kingdom Evaluation Society. 2003. *Glossary of Evaluation Terms.* http://www.evaluation.org.uk/Pub_library/Glossary.htm.

Wadsworth, Y. 1997. *Everyday Evaluation on the Run.* St. Leonards, New South Wales, Australia: Allen and Unwin.

White, Howard. 2007. "Challenges in Evaluating Development Effectiveness." Working Paper, World Bank, Washington DC.

White, Howard, and Edoardo Masset. 2005. "Quasi-Experimental Evaluation." PowerPoint presentation, February 16.

World Bank. 1998. *Do Community-Managed Schools Work? An Evaluation of El Salvador's EDUCO Program.* Impact Evaluation of Education Reforms Paper 8, Development Research Group, Washington, DC.

———. 2004. *PovertyNet: Evaluation Designs.* Washington, DC: World Bank. http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,contentMDK:20188242~menuPK:412148~pagePK:148956~piPK: 216618~theSitePK:384329,00.html.

———. 2006a. *Conducting Quality Impact Evaluations under Budget, Time, and Data Constraints.* Independent Evaluation Group, Washington, DC.
———. 2006b. *Impact Evaluation: The Experience of the Independent Evaluation Group of the World Bank.* Independent Evaluation Group, Washington, DC.
———. 2007. PowerNet. Spanish Impact Evaluation Fund. http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,contentMDK:21419502~menuPK:384336~pagePK:148956~piPK: 216618~theSitePK:384329,00.html.
———. 2008. *PovertyNet: Impact Evaluation, Methods and Techniques, Evaluation Designs.* Washington, DC: World Bank. http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,contentMDK:20188242~menuPK:415130~pagePK:148956~piPK: 216618~theSitePK:384329,00.html.
Yin, R. K. 1984. *Case Study Research.* Thousand Oaks, CA: Sage Publications.

Веб-сайти

Campbell Collaboration. http://www.campbellcollaboration.org/.
Schweigert, F. J. 2006. "The Meaning of Effectiveness in Assessing Community Initiatives." *American Journal of Evaluation* 27: 416. http://aje.sagepub.com/cgi/ content/abstract/27/4/416.
Scriven, Michael 2007. *Key Evaluation Checklist.* http://www.wmich.edu/evalctr/ checklists/kec_feb07.pdf.
Stufflebeam, Daniel L. 2004. *Evaluation Design checklist.* Western Michigan University, Evaluation Center, Kalamazoo, MI. http://www.wmich.edu/evalctr/ checklists/evaldesign.pdf.
University of Northern Iowa, Department of Psychology. http://www.psych.uni.edu/psycrs/457/e2/control.ppt.
World Bank. *Doing Impact Evaluation Series.* http://web.worldbank.org/WBSITE/ EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,menuPK:384336~pagePK:149018~piPK:149093~theSite PK:384329,00.html#doingIE.

ГЛАВА 8

Вибір і створення інструментів для збору даних

У попередніх главах розглядалися питання оцінювання й відповідні їм дизайни. Цю главу присвячено способам збору даних з метою відповіді на питання оцінювання.

Ця глава складається з чотирьох основних частин:

- Стратегії збору даних
- Характеристики хороших вимірів
- Кількісні й якісні дані
- Інструменти для збору даних

Стратегії збору даних

Дані можна отримувати з багатьох джерел, в тому числі з існуючих документів, від електромеханічних вимірювань, спостережень, опитувань, фокус-груп та експертних суджень. Неможливо виокремити якийсь найкращий спосіб. Як показано в табл. 8.1, рішення про вибір методу залежить від наступних факторів:

- про що потрібно дізнатися;
- де знаходяться дані;
- доступні ресурси й час;
- складність даних, які мають бути зібрані;
- частота збору даних;
- цільові форми аналізу даних.

Вибір методів частково залежить від того, на яке питання оцінювання необхідно дати відповідь, наскільки добре фахівці з оцінювання розуміють інтервенцію, а також від наявності ресурсів і часу. Встановлюється компроміс між поглибленим розумінням, що його забезпечує, приміром, вивчення практичних прикладів (інтенсивний збір даних), та валідністю результатів, забезпечуваною опитуваннями (екстенсивний збір даних). Для інтенсивного збору даних зазвичай використовуються напівструктуровані підходи, що дозволяють давати гнучкі відповіді. Для екстенсивного збору даних зазвичай потрібні структуровані підходи, що дозволяють ефективно опрацьовувати інформацію, отримувану від великої кількості респондентів.

Таблиця 8.1. Приклад таблиці прийняття рішень щодо методу збору даних

Якщо треба дізнатися:	Треба розглядати наступні методи:
Чи стали малоосвічені селяни з низьким рівнем грамотності після участі у програмі писати краще, ніж ті, що в ній участі не брали?	• Збір зразків письма до й після інтервенції • Використання результатів тестування до й після інтервенції
Чи стали учасники активніше брати участь в навчанні своїх дітей?	• Спостереження взаємодій між батьками та дітьми до й після інтервенції • Опитування дітей, батьків і вчителів щодо участі батьків у навчальному процесі до й після інтервенції
Чи були учасники програми задоволені якістю семінарів з підвищення грамотності та наступних заходів?	• Використання структурованих інтерв'ю з учасниками • Опитування з метою встановлення, чи є рівень грамотності достатньо високим

Джерело: Автори.

Блок 8.1 «Правила збору даних»

Збираючи дані, фахівці з оцінювання мають дотримуватися наступних правил:

- Якомога частіше використовувати одночасно кілька методів збору даних.
- Якщо можливо, використовувати наявні дані (це швидше, дешевше і легше, ніж отримувати нові).
- У разі використання наявних даних слід дізнаватися, яким чином попередні фахівці з оцінювання збирали ці дані, визначали змінні та забезпечували достовірність. Перевірити обсяг відсутніх даних.
- Якщо необхідно зібрати вихідні дані – встановити і виконувати відповідні процедури; забезпечити точне фіксування визначень та кодування; постійно проводити попереднє тестування; перевіряти точність кодування і використовуваних даних.

Джерело: Автори.

Для визначення типу даних, що мають збиратися, фахівцям з оцінювання потрібно визначити, що в оцінюванні найважливіше для клієнта. Клієнт більше зацікавлений у числових даних щодо стану шкіл у країні чи у поглибленому розумінні ситуації в найбідніших міських районах? Важливими можуть бути й обидва питання, але наявні ресурси вимагають надання пріоритету лише одному з них. Незалежно від того, які використовуються види даних, фахівці з оцінювання мають застосовувати певні правила (див. блок 8.1).

Структурований підхід

Структурований підхід до збору даних вимагає, щоб дані збиралися в абсолютно однаковий спосіб. Це особливо важливо для багатоцентрових та кластерних оцінок. В ході таких оцінок фахівцям з оцінювання потрібно мати змогу порівнювати результати з різних місць, щоб робити висновки про те, що і де працює (блок 8.2). Структурованість також важлива при порівнянні альтернативних інтервенцій для визначення найбільш економічно ефективної з них.

Структурований збір даних:
Підхід до збору даних, за якого всі дані збираються абсолютно однаково

Структуровані підходи використовуються для отримання кількісних даних, якщо фахівець з оцінювання:

- має дати відповіді на питання про масштаби;
- має справу з великою вибіркою або популяцією;
- знає, що саме потрібно вимірювати;
- повинен представити результати у числовому вигляді;
- має проводити порівняння між різними місцями або інтервенціями.

Блок 8.2 «Застосування структурованого підходу до оцінювання сільськогосподарської інтервенції»

Розглянемо приклад оцінювання сільськогосподарської інтервенції. Для відповіді на одне питання спеціалісти, які оцінювали успішність дренування земель, обрали за показник вологість ґрунту. Потім вони вирішили зібрати дані про вологість ґрунту з багатьох місцевостей регіону до й після проведення дренування протягом однакового періоду часу (а також за однакових погодних умов).

Для відповіді на друге питання фахівці з оцінювання використовують структуроване керівництво з проведення інтерв'ю: фермерів запитують про їхню думку щодо наслідків проекту. За підсумками інтерв'ю створюється звіт про відсоток респондентів, які висловлюють ті чи інші погляди (такі результати представляються у вигляді таблиці й зазвичай називаються підрахунком частотності). Питання для структурованого інтерв'ю мають бути вузькоспрямованими й точно сформульованими; необхідно також надати варіанти відповідей. Всім респондентам необхідно в абсолютно однаковий спосіб ставити однакові запитання, пропонуючи на вибір один і той самий набір можливих відповідей.

Для відповіді на третє питання фахівці з оцінювання планують використовувати відомості про розміри врожаю й ціни протягом певного часу до інтервенції й після неї, в дренованій місцевості та в іншій місцевості, де дренування не проводилося. З метою вивчення четвертого питання вирішили опитати 100 із 2600 учасників про їхню думку щодо проекту та його наслідків. Для цих інтерв'ю фахівці з оцінювання планують використати напівструктуровані питання. Планується також ставити уточнюючі питання, щоб краще зрозуміти думки учасників.

Джерело: Автори.

Напівструктурований підхід

Напівструктурований підхід до збору даних: Підхід до збору даних, за якого дані не завжди збираються в один спосіб

Напівструктурований підхід до збору даних може бути систематичним та виконуватися відповідно до загальних процедур, але дані при цьому не завжди збираються в один спосіб. Напівструктуровані інтерв'ю часто ґрунтуються на попередньо визначеному наборі питань, але порядок, в якому вони ставляться, може залежати від обставин. Більше того, після деяких з наданих відповідей ставляться уточнюючі питання, які формулюються в ході інтерв'ю. Такі підходи є більш відкритими та гнучкими, ніж структуровані. Вони дозволяють респондентам розповідати те, що вони бажають, і так, як бажають.

Напівструктуровані методи збору даних зазвичай є якісними. Вони використовуються тоді, коли фахівець з оцінювання:

- проводить розвідувальну роботу в новій області розвитку;
- прагне досягти розуміння певних тем або питань;
- прагне почути розповіді учасників або отримати поглиблену інформацію;
- прагне отримати поглиблену, детальну, «інсайдерську» інформацію;

- прагне зрозуміти результати структурованого збору даних, якщо вони є неочікуваними та не дуже зрозумілими, або ж просто хоче мати яскраві приклади для підкріплення висновків, отриманих після структурованого збору даних.

Наприклад, в ході оцінки муніципального проекту розвитку фахівці з оцінювання можуть обрати напівструктурований підхід до збору даних. Оскільки такі програми передають контроль над рішеннями з планування групам на місцях, фахівцю з оцінювання варто використовувати напівструктурований підхід, якщо він хоче більше дізнатися про те, як приймаються рішення, а також отримати інформацію про погляди членів громади щодо процесу і довгострокових результатів проекту.

Характеристики хороших вимірів

Фахівці з оцінювання вимірюють переконання, ставлення, думки, знання, навички, ефективність та звички. Визначаючи, як виміряти актуальну змінну та зібрати дані щодо неї, фахівцям з оцінювання варто пам'ятати про чотири основних моменти:

- Чи є вимірювання релевантним? Чи контролюється таким чином справді важливий фактор, а не той, що його найлегше контролювати?
- Чи є вимірювання достовірним? Чи дасть воно відомості про фактичну ситуацію?
- Чи є вимірювання валідним? Чи відображає воно те, що прагне виміряти фахівець з оцінювання?
- Чи є вимірювання надійним? Якщо завжди в однаковий спосіб збирати дані з того самого джерела, щоразу використовуючи однакові правила для прийняття рішень, чи будуть результати однаковими?

Говорячи про *релевантність*, ми маємо на увазі, наскільки важливими є виміряні показники. Фахівцям з оцінювання слід уникати випадків, коли вони вимірюють те, що легше, замість того, що потрібно. Їм також варто уникати спроб вимірювати занадто багато речей. Матриця дизайну – це інструмент, що допоможе забезпечити релевантність зібраних даних.

Достовірними називають дані, що заслуговують на довіру та видаються цільовій аудиторії правдоподібними. Приміром, думку викладачів не можна вважати найдостовірнішим показником, якщо ви намагаєтесь встановити, чому такий значний відсоток студентів залишає навчання. Більш достовірними показниками будуть думки самих студентів, яким залишили навчальний заклад, а також думки їхніх друзів.

Валідність – це термін, який означає, що вимірювання дійсно дає ті дані, які повинно давати. Чи дають питання змогу отримувати точну інформацію? Візьмемо для прикладу листи очікування. Вони мають низьку валідність як показники потреби впровадження навчальних програм для дітей, тому що часто є застарілими, а також через те, що батьки реєструють своїх дітей одночасно у кількох листах очікування. Після того, як ім'я дитини внесено до одного листа очікування, воно часто не виключається з інших листів.

Існує два види валідності – зовнішня й змістовна.

Зовнішня валідність:
Те, наскільки схоже, що зміст тесту або процедури вимірює те, що має вимірювати

Змістовна валідність:
Те, наскільки адекватно зміст тесту або процедури вимірює цільову змінну

- ***Зовнішня валідність*** – це те, наскільки схоже, що зміст тесту або процедури вимірює те, що має вимірювати. Приміром, якщо оцінка спрямована на визначення фізичної підготовки особи, час, за який ця особа пробігає 100 м, може видаватися валідним показником.
- ***Змістовна валідність*** – це те, наскільки адекватно зміст тесту або процедури вимірює цільову змінну. Наприклад, якщо фахівці з оцінювання намагаються виміряти стан здоров'я людини, вони мають консультуватися з медиками, щоб підтвердити, що обраний показник має високу змістовну валідність. Приміром, пропорція маси жиру в загальній масі тіла є зазвичай більш валідним показником, ніж звіт особи про власні звички в сфері здорового харчування. Тестування знань про здорове харчування може бути більш валідним, ніж дані, надані самостійно. Однак таке тестування не може бути показником з високою валідністю, адже респондент може й не застосовувати знання про здорове харчування на практиці. Результати можуть виглядати оптимістичніше за справжню картину.

Надійність:
Те, наскільки однаково з точки зору об'єкта й методів проводяться вимірювання

Надійність – це термін, що використовується для опису стабільності вимірювань, тобто того, наскільки однаково з точки зору об'єкта й методів проводяться вимірювання. Наприклад, інструменти вимірювань для спортивних подій мають бути надійними. Рулетка, за допомогою якої визначають довжину стрибка, має кожного разу вимірювати цей показник однаково. Якщо це так – вимірювання надійне, якщо ж ні, то результати можуть бути хибними, що ставить під сумнів підсумки змагань в цілому.

Зважування новонароджених – приклад надійного вимірювання (за умови належної повірки ваг). Відвідуваність занять у школі – приклад показника з низькою надійністю, тому що його значення може коливатися залежно від того, коли саме під час навчального року проводяться вимірювання.

Кількісні й якісні дані

Дані можна поділити на кількісні або якісні. **Кількісні дані** – дані, представлені в числовій формі, **якісні дані** – представлені не в числовій формі.

Кількісні дані: Дані, представлені в числовій формі

Якісні дані: Дані, представлені не в числовій формі

Кількісні дані – це дані, які можна точно виміряти. Прикладами їх є вік, вартість, довжина, висота, площа, обсяг, вага, швидкість, час і температура.

Якісні дані – це опис. Такі дані можуть отримуватися за результатами спостережень або ж надаватися інформантами, але не обов'язково точно вимірюватися. Прикладами якісних даних є дані про стосунки та поведінку.

Розглянемо оцінювання програми мікрокредитування. Кількісні дані для цієї програми можуть включати: кількість учасників з розбивкою за статтю, віком та кількістю дітей; дохід; обсяги складських запасів; вартість продукції; обсяг продажів. Якісні дані для цієї програми можуть включати опис продукції, сімейних стосунків, поведінки учасників, стосунків з громадою, розширення можливостей.

Петон (Patton 2002) виокремлює три методи збору даних, що можуть дати якісні результати:

- поглиблені неструктуровані інтерв'ю;
- безпосередні спостереження (з використанням детальних описів);
- аналіз письмових документів.

Дослідник описує види інформації, що її отримують фахівці з оцінювання, використовуючи кожен з цих трьох методів:

- Неструктуровані інтерв'ю дають прямі відомості про враження, думки, почуття й знання.
- Безпосередні спостереження можуть дати детальні описи діяльності, поведінки, дій, а також повного спектру міжособистісних взаємодій і організаційних процесів.
- Аналіз документів дає витяги, цитати або цілі уривки із записів, меморандумів та кореспонденції, офіційних публікацій та звітів, щоденників, а також письмові відповіді на відкриті питання анкет і опитувань.

Найбільше якісних даних дає перебування в досліджуваному середовищі. Фахівець з оцінювання безпосередньо спостерігає за діяльністю і взаємодіями, іноді беручи в них участь. В ході збору даних створюється масив записів – це вихідні дані. Потім вони структуруються: формуються зв'язні детальні описи, в яких можна виділити головні теми, категорії, показові практичні приклади (Patton 2002).
Якість зібраних якісних даних залежить від фахівця з оцінювання.

Петон зазначає:

> Систематичне й сумлінне спостереження вимагає значно більшого, ніж просто бути присутнім і роздивлятися довкола. Вміло проводити інтерв'ю – це набагато більше, ніж просто ставити питання. Аналіз змісту вимагає значно більшого, ніж просто читання документів. Формування корисних і достовірних якісних висновків за результатами спостережень, інтерв'ю й аналізу змісту потребує дисципліни, знань, підготовки, практики, творчого підходу й важкої праці (2002 р., ст. 5).

Блок 8.3 «Контрольний список із 20 якісних питань Петона»

1. Чи робить програма акцент на окремих довгострокових результатах, тобто чи очікується, що різні учасники зазнаватимуть впливів у якісно різний спосіб? Чи існує потреба або бажання описувати та оцінювати ці індивідуалізовані результати?
2. Чи зацікавлені особи, які приймають рішення, в з'ясуванні та розумінні внутрішньої динаміки програм, тобто їх сильних і слабких сторін, а також загальнопрограмних процесів?
3. Чи існує потреба в детальній, поглибленій інформації про певні практичні випадки у клієнтів або про місця реалізації програми (наприклад, особливо успішні випадки, незвичайні невдачі або критично важливі випадки), зумовлена програмними, фінансовими або політичними причинами?
4. Чи існує зацікавленість в особливій увазі, що приділяється розбіжностям, особливостям та унікальним характеристикам окремих клієнтів та програм (це підхід, протилежний порівнянню всіх клієнтів або програм на основі стандартизованих, однорідних показників)?
5. Чи потрібна інформація про деталі реалізації програми: Що відчувають учасники програми? Які послуги надаються клієнтам? Як організовується програма? Що роблять її співробітники? Чи потрібно особам, які приймають рішення, знати, що відбувається в рамках програми та як вона розгортається?
6. Чи існує в співробітників програми та інших зацікавлених осіб інтерес до збору детальної описової інформації про програму для її покращення (тобто чи існує зацікавленість у формативному оцінюванні)?
7. Чи існує потреба в інформації про нюанси якості програми – описової інформації про якість програмних заходів і результатів, а не лише про їх рівень, обсяг чи кількість?
8. Чи потребує програма індивідуалізованої системи забезпечення якості?
9. Чи існує серед законодавців або інших осіб, які приймають рішення або надають кошти, зацікавленість у відвідуванні фахівцями з оцінювання місць реалізації програми (через те, що особи, які приймають рішення, не мають можливості відвідувати ці місця особисто і не володіють навичками збору інформації, притаманними кваліфікованим фахівцям з оцінювання)? Чи потрібно проводити законодавчий моніторинг у кожному окремому випадку?

Чи є проблемою нав'язливість оцінки? Чи буде застосування стандартизованих інструментів оцінки (анкет і тестів) занадто нав'язливим порівняно зі збором даних за допомогою природних спостережень і неструктурованих інтерв'ю? Чи буде збір якісних даних менш обтяжливим для учасників, ніж збір кількісних даних? Чи існує потреба в ненав'язливих спостереженнях?

10. Чи існують потреба й бажання персоналізувати процес оцінки шляхом використання методів, що ґрунтуються на персональному, очному контакті з програмою – методів, які можна назвати «гуманістичними» та особистими, тому що при цьому учасники не нумеруються й не класифікуються, а почуваються природно, неформально й просто.
11. Чи варто збирати перш за все описові дані та складати звіти з урахуванням точок зору різних зацікавлених осіб на основі інформації, отриманої від безпосереднього, особистого контакту з цими особами?
12. Чи є цілі програми розпливчатими, загальними та неконкретними? Тоді варто проводити оцінку без прив'язки до цілей, яка дозволить отримати інформацію про фактичні наслідки програми, не зосереджуючись на вимірюванні виконання завдань.
13. Чи існує ймовірність того, що програма здійснює непередбачений негативний вплив на клієнтів або учасників та/або має побічні ефекти, що свідчить про потребу застосування підходу, що виявлятиме інші наслідки окрім формально заявлених як цільові (це, знов-таки, свідчить про потребу оцінювати без прив'язки до цілей)?
14. Чи має місце брак перевірених кількісних інструментів для оцінки важливих результатів програми? Чи стоїть стан теоретичної бази на заваді запровадженню або розробці валідного, надійного й достовірного стандартизованого інструментарію для кількісного вимірювання конкретних довгострокових результатів програми?
15. Чи має оцінка розвідувальний характер? Чи знаходиться програма на етапі попереднього оцінювання, якщо цілі й зміст програми все ще розробляються?
16. Чи потрібно проводити визначення оцінюваності для формування сумативного дизайну?
17. Чи існує потреба в поглибленні, більшій деталізації та розширенні значення статистичних висновків або результатів опитувань?
18. Чи став збір даних кількісної оцінки настільки рутинним, що вже ніхто не звертає на результати особливої уваги? Це свідчить про те, що, можливо, потрібно відійти від стандартної процедури та застосувати нові методи для отримання нової цінної інформації про програму.
19. Чи існує потреба в розробці програмної теорії, що ґрунтується на спостереженнях за заходами та впливами програми, а також за зв'язком між заходами й довгостроковими результатами?

Джерело: Patton 1987.

Петон (Patton 1987) розробив контрольний список із 20 питань, що допомагає фахівцю з оцінювання вирішити, чи є якісні методи належною стратегією оцінки (див. блок 8.3). Якщо на будь-яке з питань ви відповідаєте «так», то, ймовірно, збір принаймні деяких якісних даних є доречним.

Збір даних зазвичай охоплює як кількісні, так і якісні дані, але один з підходів може бути домінуючим. Ці два підходи можна охарактеризувати наступним чином:

Кількісний підхід:

- більш структурований;
- робить акцент на надійності;
- важчий у розробці;
- легший для аналізу.

Якісний підхід:

- менш структурований;
- легший у розробці;
- може давати «багаті дані» (характерні щодо кожного об'єкта вивчення);
- характеризується більшою працемісткістю збору й аналізу даних;
- робить акцент на валідності.

Який підхід буде обрано – залежить від цілей оцінки (табл. 8.2). На практиці кількісні та якісні дані пов'язані. За словами Трокіма (Trochim 2006), «всі кількісні дані ґрунтуються на якісних судженнях, а всі якісні дані можуть описуватися й оброблятися із застосуванням числових методів». Комп'ютеризований аналіз (аналіз змісту) письмових документів зосереджується на частоті вживання різних слів, видах слів та структурах, перетворюючи якісні дані на кількісні.

Обструктивний метод:
Метод, який явно використовується для спостереження за учасниками

Дані можуть збиратися в обструктивний або необструктивний спосіб. **Обструктивні методи** – це спостереження за учасниками, про які цим учасникам відомо. Такі методи використовуються для вимірювання вражень, думок і поглядів шляхом проведення інтерв'ю, опитувань і фокус-груп. Спостереження, що проводяться помітно для об'єктів спостереження, також є обструктивними.

Таблиця 8.2 Коли використовувати кількісний та якісний підходи

Якщо ви...	Використовуйте наступний підхід:
Бажаєте провести статистичний аналіз Бажаєте точності Знаєте, що саме хочете виміряти Бажаєте охопити велику групу	Кількісний
Бажаєте почути розповіді учасників або отримати поглиблену інформацію Не впевнені, що саме бажаєте виміряти Не потребуєте кількісного визначення результатів	Якісний

Джерело: Автори.

Блок 8.4 «Вимірювання популярності музейних експозицій за допомогою нав'язливих та ненав'язливих методів»

Директор музею хотіла дізнатися, яка з експозицій її закладу найпопулярніша. Для цього вона могла опитати людей на виході з музею. Однак такі добровільно надавані відомості є ненадійними, адже респонденти можуть називати найвідоміші експозиції тільки тому, що пам'ятають лише їх. Також вони можуть надавати неправдиві відомості, прагнучи позитивного сприйняття з боку особи, яка проводить інтерв'ю. Щоб уникнути цих проблем, одне відоме дослідження оцінювало популярність експозицій музею, вимірюючи потертість підлоги перед експонатами. Цей ненав'язливий метод дозволив запобігти спотворенню даних, характерному для добровільного надання відомостей.

Якщо в ході оцінки для збору даних використовуються анкети, учасники дослідження знають, що їх вивчають. Це може призвести до отримання ненадійних результатів. Петон (Patton 1987, 33) стверджує: «Сам інструмент здатний викликати реакцію, і його нав'язливий характер та втручання в нормальну взаємодію програми й клієнта унеможливлює точне відображення фактичних результатів цієї програми».

Ті, кого вивчають, можуть змінити власну поведінку або відповіді. Наприклад, викладач, за лекцією якого веде спостереження керівник освітньої системи, може викладати предмет інакше, ніж якщо таке спостереження відсутнє.

Необструктивні методи – це спостереження за учасниками, про які їм не відомо. До прикладів необструктивних методів відноситься використання даних з документів або архівів, а також непомітне спостереження за учасниками (блок 8.4).

Необструктивний метод: Метод, що непомітно використовується для спостереження за учасниками

Інструменти для збору даних

Вибір техніки збору даних залежить від конкретної ситуації. Незважаючи на те, який метод обирається для збору даних, вся зібрана інформація потенційно може зазнати впливу **викривлення/систематичної похибки**. Викривлення інформації означає, що коли респондентів просять надати інформацію про себе або інших людей, вони можуть – навмисно або ні – приховати якісь відомості. Правдиві дані можуть бути спотворені через те, що респонденти щось забули або побоюються наслідків надання правди-

Систематична похибка/ Викривлення інформації: Навмисне або ненавмисне спотворення даних в ході їх збору, аналізу або формування звіту

вої відповіді. Вони також можуть бути збентежені або почуватися ніяково, визнаючи щось, що вважають соціально неприйнятним. Всі добровільно надавані дані не захищені від викривлення.

Респондентам може бути ніяково правдиво відповідати на питання, приміром, про використання захисних засобів під час статевого акту або про дату останнього відвідування лікаря. Вони можуть говорити те, що, як їм здається, хоче почути фахівець з оцінювання.

Також може мати місце помилка вибірки – той факт, що люди, які вирішили взяти участь у програмі, можуть відрізнятися від тих, які участі не брали. Помилка вибірки розглядалася у главі 5. Це часто є проблемою для опитувань, інтерв'ю та фокус-груп. Добровільні учасники можуть системно відрізнятися від тих, хто відмовляється від участі.

Зазвичай використовують більш ніж один підхід для збору даних: для відповіді на різні питання оцінки або ж формування кількох джерел даних при розгляді одного питання оцінки. В ході оцінки, приміром, можуть збиратися наявні дані про врожай, зібраний фермерами, проводитися інтерв'ю зі споживачами сільськогосподарської продукції, а також опитування фермерів. Іноді фахівці з оцінювання використовують фокус-групи або вивчають практичні приклади для розробки тем для опитувальника або тлумачення результатів опитувань.

Тріангуляція методів:
Збір однієї й тієї самої інформації з використанням різних методів для підвищення точності даних

Збір однієї й тієї самої інформації з використанням різних методів для підвищення точності даних називається ***тріангуляцією методів***. Фахівці з оцінювання використовують тріангуляцію методів для підкріплення результатів. Чим більше різних методів використано для отримання інформації, що підтверджує висновки, тим достовірнішими ці висновки вважаються.

Тріангуляція джерел:
Збір однієї й тієї самої інформації з використанням різних джерел для підвищення точності даних

Тріангуляція методів – не єдиний вид тріангуляції. Дензін (Denzin 1978) визначає кілька видів тріангуляції, зокрема **тріангуляцію джерел**. Ще одним видом тріангуляції є **тріангуляція фахівців з оцінювання**, що передбачає залучення більш ніж одного фахівця з оцінювання.

Тріангуляція фахівців з оцінювання:
Збір однієї й тієї самої інформації більш ніж одним фахівцем з оцінювання задля підвищення точності оцінки

Учасники дослідження – не єдині люди, які можуть вплинути на результати оцінки. Фахівці з оцінювання й умови проведення оцінювання також можуть здійснювати певний вплив. Наприклад, жінки можуть давати різні відповіді інтерв'юерам чоловічої та жіночої статі; їхні відповіді також можуть відрізнятися, якщо інтерв'ю проходять наодинці та в присутності їхніх чоловіків.

У наступних розділах наводиться опис дев'яти інструментів для збору даних:

- Інструмент 1: Спільний збір даних
- Інструмент 2: Аналіз документів та вторинний аналіз
- Інструмент 3: Спостереження

- Інструмент 4: Опитування й інтерв'ю
- Інструмент 5: Фокус-групи
- Інструмент 6: Щоденники, журнали та опитувальні листи для самостійного заповнення
- Інструмент 7: Експертні судження
- Інструмент 8: Метод Дельфі
- Інструмент 9: Інші інструменти вимірювання

Інструмент 1: Партисипативний (із залученням зацікавлених сторін) збір даних

Партисипативний збір даних передбачає активне залучення груп або спільнот. Прикладами технік партисипативного збору даних є громадські збори, картування, а також спостережні прогулянки.

Громадські збори

Один з найпоширеніших методів партисипативного збору даних – проведення громадських зборів. Такі збори дозволяють членам громади ставити питання, робити зауваження, а також обговорювати важливі для громади теми.

Для того, щоб збори давали дані, придатні для практичного використання, їх треба добре організовувати. Фахівець з оцінювання і зацікавлені особи мають погодити мету зборів та зобов'язатися бути присутніми. До початку зборів фахівець з оцінювання має визначити та оголосити базові правила. Варто охопити наступні питання: як ідентифікувати виступаючих, скільки часу надається на виступ, в якому форматі мають ставитися питання й даватися відповіді. Якщо аудиторія є достатньо писемною, базові правила слід викласти письмово і запропонувати для ознайомлення тим, хто не був присутній з початку події.

Необхідно активно анонсувати проведення громадських зборів за допомогою листівок, оголошень в газетах та на радіо. Члени громади також можуть взяти на себе обов'язки щодо поширення відомостей про зібрання. Фахівцям з оцінювання не варто покладатися на місцевих посадовців, які мають запросити людей, тому що вони можуть запрошувати вибірково, і на збори прийдуть лише ті учасники, які поділяють їхні погляди. Місце проведення зборів слід обирати так, щоб на них прийшло якомога більше членів громади, з урахуванням зручності, доступності та безпеки учасників зборів (Minnesota Department of Health 2007).

Проведення громадських зборів має кілька переваг:

- Ці збори здатні підвищити довіру до оцінювання завдяки більшій відкритості процесу та активнішому залученню людей.
- Проводити громадські збори недорого й відносно просто.
- Збори дозволяють залучати широкий спектр учасників.
- Менш формальна обстановка може сприяти активнішій участі громади.
- Громадські збори здатні підвищити рівень обізнаності населення про оцінку, допомогти краще її зрозуміти. Це, в свою чергу, сприяє більшій її підтримці.
- Збори можуть дати фахівцю з оцінювання відомості про важливі питання програми.
- Вони можуть виявити питання, що потребують детальнішого розгляду.

Громадські збори мають і свої недоліки (див. табл. 8.3). Наприклад, члени громади, які беруть участь у зборах, можуть не представляти всю громаду, оскільки деякі люди, які мають гарні ідеї або чітке розуміння проблем, можуть не захотіти відвідувати такі заходи чи виступати на них. При цьому відвідувати збори можуть ті, хто найбільш різко (позитивно або негативно) ставиться до програми. Може статися, що виступати будуть ті, хто обіймає в громаді керівні посади. Залежно від часу проведення зборів на їх відвідуваність може вплинути гендерний фактор. Активність жінок значною мірою залежить від культурних особливостей. Через згадані проблеми громадські збори не варто використовувати як основний метод збору даних для оцінювання.

Таблиця 8.3. Переваги й недоліки громадських зборів

Переваги	Дозволяють членам громади дізнатися про інтервенцію та обговорити проблемні питання
	Підвищують обізнаність та достовірність
	Не потребують великих витрат
	Дають фахівцю з оцінювання можливість отримати відомості з важливих питань
Проблемні сторони	Може мати місце нерепрезентативність, зумовлена факторами гендеру й влади

Джерело: Автори.

Картування

Створення нових або використання існуючих мап – процес, що його в контексті збору даних називають **картуванням**. Картування можна використовувати для обговорення місцевих пріоритетів розвитку та ступеня їх виконання. За його допомогою також можна перевіряти достовірність вторинних джерел інформації. Окрім того, картування також може враховувати зміни (об'єктивні або такі, що, як здається аудиторії, мали місце) протягом певного часу (наприклад, до й після інтервенції). Хоча процес картування часто застосовується при плануванні інтервенцій, він також може бути використаний під час оцінювання.

Картування:
Створення нових або використання існуючих мап в рамках збору даних

Картування може охоплювати окремих осіб або групи людей. Як груповий інструмент воно корисне для проведення спільних оцінювань за участі зацікавлених осіб, тому що надає можливість працювати разом. В той же час картування може підвищити розуміння процесу кожним членом громади. Це особливо важливо, якщо люди матимуть різне розуміння громади залежно від власного статусу й досвіду. Картування також є корисним інструментом для збору даних у групах з низьким рівнем писемності.

Існують, зокрема, наступні види картування:

- ресурсне картування;
- історичне картування;
- соціальне картування;
- картування за станом здоров'я;
- картування за статками;
- картування за землекористуванням;
- демографічне картування

Ресурсне картування часто використовується для збору й структурування інформації про розподіл, доступність і використання ресурсів усередині громади. Воно може застосовуватися до й після інтервенції з метою визначення змін, що відбулися.

Є багато способів створення мап. Приміром, для програми з водопостачання й санітарії фахівці з оцінювання можуть просто на папері малювати прості мапи, позначаючи відомі джерела води, відхожі місця, основні екологічні особливості та схеми розташування поселень. Потім фахівці з оцінювання можуть попросити селян за допомогою інших наданих їм матеріалів (таких як нитки, трубки, стрічки та мотузки) позначити, які джерела питної води чи відхожі місця використовують різні сім'ї. За допомогою кольорових олівців, чорнила та крейди можна малювати крапки, що вказують на розподіл спеціальних груп, таких як бідні жінки, багаті люди та лідери (Narayan 1996).

Ось деякі приклади використання картування в ході спільної з членами громади роботи:

- школярів попросили намалювати мапу розташування їхньої громади з позначенням власних домівок та домівок їхніх сусідів;
- групу чоловіків та жінок попросили за допомогою місцевих матеріалів – глини, трави, грязі, каміння, паличок – створити модель своєї громади й позначити всі водні (або інші) ресурси;
- під час зборів групи було започатковано обговорення, в ході якого членів громади залучили до розробки її мапи (Narayan 1996).

Соціальне картування: Створення концептуальної картини з урахуванням елементів, з яких складається громада, включаючи її ресурси й активи та їх взаємодію

Соціальне картування передбачає створення концептуальної картини з урахуванням елементів, з яких складається громада, включаючи її ресурси й активи та їх взаємодію. Соціальне картування може використовуватися для представлення інформації щодо схеми розташування селища, його інфраструктури, демографії, етномовних груп, громадських об'єктів, стану в сфері охорони здоров'я, розподілу багатства та інших аспектів. Цей підхід зводить докупи членів громади задля кращого розуміння як самої громади, так і ролі, яку відіграє для неї інтервенція (позитивної або негативної). Він також може використовуватися як частина будь-якого підходу, де соціальне картування може допомогти відповісти на питання оцінювання. Окрім того, даний вид картування може застосовуватися для порівняння стану до й після проведення інтервенції.

До інших інструментів картування відносяться аерофотозйомка, в тому числі з використанням засобів Google Earth (див. блок 8.5), топографічна зйомка, мапи, підготовлені професійними картографами, а також мапи, створені на місці працівниками, які перебувають на території громади й користуються допомогою важливих її членів. Отримана таким чином інформація може використовуватися в ході планування міжгромадських заходів (Narayan 1996).

Спостережні прогулянки

Прогулянки досліджуваною місцевістю значно відрізняються від розглядання фотографій, зроблених з повітря. Замість фотокамери дослідник використовує власні навички спостереження, щоб отримати уявлення про місцевість. Пройшовши стежками, якими щодня ходять члени громади, можна зрозуміти, чому жінки приносять воду лише раз на день або обирають інший шлях, коли починаються дощі. Прогулянка може розповісти про соціальне розмежування в громаді, наприклад про розподіл влади й кастовість. Вона також може дати уявлення про просторову структуру, архітектурні стилі, використання простору, екологічну санітарію, недостатність або надлишковість використання об'єктів благоустрою, а також дії, пов'язані з об'єктами водопостачання та санітарії (Narayan 1996).

Блок 8.5 «Використання програми Google Earth для картування»

Google Earth, комп'ютерна програма, яку можна завантажити з мережі Інтернет, містить мапи всієї планети. Вони створені шляхом розміщення зображень, отриманих із супутників, від аерофотозйомки та систем географічної інформації, на зображення тримірного глобусу. Знімки багатьох великих міст доступні в такій якості, що дозволяє побачити окремі будівлі, будинки та навіть автомобілі (див. фото). Ступінь деталізації певною мірою залежить від того, який інтерес викликає певна територія.

Зображення міста Дакка (Бангладеш) у програмі Google Earth

Програма Google Earth може бути корисною для збору базових даних та даних про тенденції. Фахівці з оцінювання можуть знайти потрібну місцевість в Google Earth, зберегти й роздрукувати її зображення. Отримані пізніше дані можна порівняти з вихідним зображенням, щоб продемонструвати зміни, приміром, у стані дорожньої мережі або в під'єднанні сільських громад до централізованого електропостачання.

Спостережна прогулянка – це прогулянка, яку може здійснити фахівець з оцінювання територією громади, щоб отримати репрезентативні уявлення про її членів, оточуючу місцевість та ресурси. Спостережні прогулянки – інструмент для збору просторових даних. Спостережна прогулянка може зайняти всього годину, а може розтягнутися й на цілий день.

Спостережна прогулянка: Прогулянка, яка має на меті спостереження за членами громади, оточуючою місцевістю та ресурсами

Спостережну прогулянку планують, проводячи «лінію спостереження» на мапі громади. Для того, щоб фахівець з оцінювання отримав репрезентативні враження про громаду, ця лінія має проходити через всі зони території громади.

Фахівець з оцінювання у супроводі кількох місцевих мешканців пішки проходить маршрут, позначений на мапі лінією спостереження. Він розмовляє з членами громади, спостерігаючи за умовами, людьми, проблемами та шукаючи можливості для покращення ситуації (Academy for Educational Development 2002).

Нижче наведені приклади різних аспектів становища громади, які можуть бути об'єктом спостереження під час спостережної прогулянки:

- житлові умови;
- наявність бездомних дітей;
- характер дитячої праці;
- нелегальна вулична торгівля та проституція;
- доступність громадського транспорту;
- види неурядових (НУО) та церковних організацій;
- види крамниць;
- види їжі, що продається на відкритих ринках;
- санітарні умови;
- взаємодія між чоловіками та жінками;
- наявність медичних установ;
- громадські об'єкти (Academy for Educational Development 2002).

Спостережні прогулянки дають фахівцю з оцінювання загальне уявлення про громаду. Вони допомагають виявити питання, що потребують додаткового розгляду.

Інструмент 2: Аналіз документів/ записів та вторинних даних

Іноді для відповіді на питання оцінювання можна використати вже зібрані дані. Використовуючи масиви даних, зібраних іншими, необхідно розуміти, навіщо ці дані збиралися, як визначалися змінні та як дані кодувалися й очищалися. Це також передбачає з'ясування того, які заходи вживалися щодо відсутніх даних і відповідей та низького відсотка наданих відповідей під час опитувань. До джерел вторинних даних відносяться не лише масиви даних, створені під час попередніх заходів, а й газетні статті, телевізійні програми, веб-сторінки, блоги та форуми в мережі Інтернет.

Використання документів/записів

Урядові установи, клініки, школи, асоціації та організації розвитку – це лише деякі з організацій, що створюють документи. Ці документи можуть стати основою оцінки.

Документи організацій – популярне джерело інформації для оцінювання. Більшість організацій вже мають зібрані й упорядковані дані від клієнтів та громад. Вони також, цілком імовірно, вже сформували висновки й звіти на основі цієї інформації – у формі звітів для внутрішнього керівництва, бюджетних документів, звітів для суспільства або фінансуючих установ або ж у вигляді звітів моніторингу й оцінювання.

МакКестон (McCaston 2005) представляє набір видів вторинних даних та відомостей, що часто використовуються для аналізу бідності:

- демографічні (кількість населення та темпи її зростання, рівень урбанізації, статеві показники, етнічні групи, міграційні тенденції);
- дискримінація (за ознакою статі, етнічної приналежності, віку);
- гендерна рівність (з урахуванням віку, етнічної приналежності);
- політичний клімат;
- економічні умови (зростання, коефіцієнт заборгованості, торгівельні умови);
- рівні бідності (відносний, абсолютний);
- зайнятість і рівень заробітної плати (офіційні й неофіційні показники; показники доступності);
- режими отримання засобів для існування (сільські, міські, сільськогосподарські, несільськогосподарські, неофіційні);
- сільськогосподарські показники й практики (опади, врожайність, види ґрунтів та їх використання, зрошення);
- здоров'я (недостатнє харчування, дитяча смертність, рівень імунізації, коефіцієнт фертильності, рівень поширеності використання контрацептивів);
- послуги з охорони здоров'я (кількість, рівень, показник кількості закладів на душу населення, доступність з урахуванням статі й етнічної приналежності);
- освіта (рівень грамотності серед дорослих, показники залучення дітей до школи, відсоток учнів, що переривають навчання у школі, співвідношення чоловіків та жінок, співвідношення кількості учнів за ознакою етнічної приналежності);
- школи (кількість і рівень, показник кількості шкіл на душу населення, доступність з урахуванням статі й етнічної приналежності);
- інфраструктура (дороги, постачання електроенергії, телекомунікації, вода, санітарія);
- екологічний стан і проблеми;
- шкідливі культурні практики.

Наведену вище інформацію можна знайти у:

- файлах і записах;
- комп'ютерних базах даних;
- галузевих звітах;
- урядових звітах;
- інших звітах, включаючи звіти про попередні оцінювання;
- даних перепису й квартирних опитувань;

- електронних розсилках та дискусійних групах;
- інших документах (бюджети, політики й процедури, організаційні діаграми, мапи);
- газетах і телерепортажах.

Для виокремлення даних з паперових документів фахівці з оцінювання розробляють інструмент збору даних, що точно вказує, які дані потрібно брати з файлу або запису та як їх кодувати. Метою є розробка простого і зрозумілого інструменту. Одразу після створення він має пройти попереднє тестування.

Розглянемо, наприклад, оцінювання, покликане встановити, чи діють підготовлені в рамках урядових навчальних програм медсестри-реаніматологи ефективніше за інших медсестер тієї ж спеціалізації. Інструмент для збору даних може бути використаний для систематичного отримування відповідної інформації з особистих справ медсестер (див. рис. 8.1). Фахівці з оцінювання можуть сформувати вибірку реанімаційних закладів, де працює одна або декілька медсестер, що пройшли державну програму підготовки, та проаналізувати усі їхні записи. Ці документи містять відомості про освіту, досвід роботи медсестрою та оцінки якості роботи.

Працюючи з документами, що описують поточні заходи або практики, фахівець з оцінювання має перевіряти, наскільки точно документи відображають реальний стан речей. Це можна робити за допомогою спостережень (якщо програма все ще триває) та інтерв'ю. Приміром, використовуючи документи, що описують програму підготовки, фахівець з оцінювання може перевірити, чи справді заняття проводилися п'ять днів на тиждень, чи доступні були матеріали, а також чи представляли учасниці різні етнічні групи, як про це свідчать документи. Ця перевірка може бути виконана за допомогою кількох інтерв'ю з людьми, яким відомо про програму, її історію та впровадження. Залежно від конкретного випадку може бути важливим знайти людей, що виконують різні ролі, в тому числі відповідальних за клієнтів та формування бюджету. Їхні погляди можуть пролити світло на отриману інформацію та розповісти про те, що залишилося поза документами. Коли ні спостереження, ні інтерв'ю не є доцільними, фахівець з оцінювання може перевіряти інформацію, виявляючи та аналізуючи інші допоміжні документи, що містять аналогічну інформацію.

Зазвичай необхідно на додачу читати й аналізувати офіційні документи. Наприклад, в ході оцінювання програми з метою покращення реагування урядовців на телефонні дзвінки громадян фахівці з оцінювання використовували б офіційні документи для отримання відповідей на наступні питання:

- Коли було санкціоновано програму?
- Які цілі й завдання програми було визначено керівним документом?
- Скільки державних службовців мало бути залучено?
- Які агенції та департаменти мали бути залучені?
- Протягом якого періоду мала бути реалізована програма?
- Які показники ефективності були вказані (прямо чи опосередковано)?
- Який був бюджет програми?
- Які заходи передбачалися?

Рис. 8.1. Приклад інструменту збору даних

Дата _____

Ід. номер: _____

1. Рівень освіти: _____
2. Дипломований медпрацівник? Так _____ Ні _____
3. Закінчені державні курси? Так _____ Ні _____
4. Якщо так, вкажіть рік навчання. _____
5. Стаж роботи медпрацівником в цій клініці? _____
6. Стаж роботи медпрацівником в інших місцях? _____
7. Оцінки роботи протягом останніх п'яти років:

   Рік: _____ Оцінка: _____

   Рік: _____ Оцінка: _____

   Рік: _____ Оцінка: _____

   Рік: _____ Оцінка: _____

   Рік: _____ Оцінка: _____

8. Нагороди за роботу протягом останніх п'яти років:

   _____ Так _____ Ні

9. Якщо так, вкажіть кількість нагород, отриманих за останні п'ять років:

Стать: _____ Чоловіча _____ Жіноча

10. Коментарі: _____

_____

_____

Джерело: Автори.

Використання існуючих масивів даних

Аналіз вторинних даних:
Аналіз даних, отриманих з іншого джерела, а не зібраних самими фахівцями з оцінювання

Електроні масиви даних, зібрані однією організацією, нерідко можуть бути отримані й повторно проаналізовані іншою організацією для відповіді на нові питання. Цей вид аналізу називається **аналізом вторинних даних**. Фахівці з оцінювання часто використовують великі комп'ютерні бази даних, що містять, наприклад, дані квартирного опитування або відомості про кредити, надані фінансовим посередником малим та середнім підприємствам. Під час аналізу вторинних даних фахівці з оцінювання мають концентруватися на меті й схемі оцінювання.

До ключових питань, які потрібно розглядати в ході визначення, чи варто отримувати й використовувати дані для вторинного аналізу, відносяться, зокрема, наступні:

- Чи є наявні дані валідними?
- Чи є наявні дані надійними?
- Чи є наявні дані точними?
- Якими є показники відповідей та відсутніх даних?

МакКестон (McCaston 2005) також пропонує перевіряти:

- відомості про професійний рівень організації, що володіє базою даних;
- методи й правила прийняття рішень, що використовувалися для очищення даних;
- давність інформації, що міститься в базі даних (актуальна чи застаріла?);
- чи узгоджуються ці дані з даними інших джерел.

Візьмемо для прикладу підрозділ оцінки, що розглядає можливість використання даних з великої інформаційної системи комп'ютеризованого управління. Ця система містить відомості про кредити на навчання, отримані в рамках масштабної державної програми. Одна з вимог цієї програми – щоб рівень доходів студента не перевищував певної межі. Фахівці з оцінювання, які розглядають можливість використання таких даних, повинні пересвідчитись, що, згідно з наявними даними, фінансова підтримка не надавалася студентам, котрі не мали на неї права. Якщо ж буде встановлено, що гроші надавалися тим, хто не мав їх отримувати, потрібно буде визначити, чи це пов'язано з помилково вказаними даними (неточність), чи мова йде про неналежне виконання вимог. Фахівці з оцінювання повинні також переконатися, що у фінансуванні не відмовляли тим, хто мав право його отримувати, а також що студентам не надавали більше одного кредиту. Спочатку вони мають запустити програмні перевірки даних, щоб визначити характер та масштаб відповідних проблем у базі даних. Перевірити дані можна, наприклад, зателефонувавши певній вибірці студентів, щоб переконатися, що вони отримали і гроші, і навчання.

Фахівці з оцінювання також могли б зв'язатися з навчальними закладами й поцікавитись, чи містять їхні документи відомості про те, що студенти проходили в них навчання та отримували фінансування за рахунок державної програми.

Під час використання вторинних даних необхідно дотримуватися наступних процедур:

- Дізнатися, що потрібно зробити для того, щоб отримати дані на свій комп'ютер. Іноді організація, які володіє базою даних, може самостійно виконати аналіз для фахівця з оцінювання і не погоджується на будь-які інші варіанти. (Якщо можливо, краще отримати файл з вихідними даними.)
- Перш ніж завантажувати файл на свій комп'ютер, слід перевірити його на наявність вірусів.
- Отримати структуру бази даних, словник даних та схеми кодування.
- Пересвідчитися в точності даних за допомогою тестування й перехресної перевірки.
- Передати дані у такий спосіб, що виключає утворення нових помилок (наприклад, не допускається передрукування). Вручну перевірити окремі фрагменти даних, щоб переконатися, що їх передача пройшла успішно.
- Належним чином розставити посилання на всі вторинні дані (дата публікації, імена авторів).

У табл. 8.4 в стислому вигляді вказано переваги й недоліки використання наявних даних.

Таблиця 8.4. Переваги й недоліки використання наявних даних

Переваги	Отримати наявні дані зазвичай дешевше й швидше, ніж збирати нові.
Недоліки	В наявності можуть бути не зовсім ті дані, які потрібні.
	Може бути складно забезпечити доступ до цих даних протягом тривалого часу.
	Необхідно перевіряти валідність і надійність даних, а також виправляти помилки кодування.

Джерело: Автори.

Інструмент 3: Спостереження

Спостереження може бути потужним інструментом збору даних. Наприклад, лише візуального спостереження достатньо для отримання інформації про закономірності організації дорожнього руху, землекористування, схему міського та сільського середовища, якість житла, стан доріг або контингент пацієнтів сільської лікарні.

Таблиця 8.5. Причини застосування та приклади спостереження

Причини застосування спостереження	Приклади
Для збору прямої інформації	Випадкові відвідування шкіл, домівок, фермерських господарств та інших місць з метою спостереження, а не просто опитування людей
	Операції й процедури зі спостереження в офісах, школах, лікарнях та в інших місцях замість обмеження джерел інформації одними лише звітами
	Ненав'язлива фіксація кількісних показників етнічних, гендерних або вікових груп учасників певної діяльності
Для розуміння поточної поведінки, процесу, ситуації в динаміці або певної події	Спостереження та опис того, що робиться на кожному етапі проекту або програми
	Спостереження за взаємодією дітей на дитячому майданчику, поводженням батьків з дітьми або викладачів з учнями, медичних працівників з пацієнтами, а також керівників з підлеглими
	Спостереження за тим, як керівники проводять засідання до й після навчальних програм
Для вивчення речових доказів, продуктів або результатів	Спостереження за харчовими продуктами та іншими товарами на ринку
	Періодичне спостереження за узбережжям озера, на якому проводиться програма очистки
	Залучення групи експертів до вивчення якості злакових і бобових рослин на пасовищі
	Інспектування садів, інформаційних бюлетенів, програмних документів проектів і т. ін.
Для надання альтернативи, коли письмові або інші процедури збору даних можуть бути економічно недоцільними або невідповідними потребам	Добровільне залучення кількох учасників до спостереження й звітування щодо практичної реалізації програми замість охоплення анкетуванням всіх її учасників
	Спостереження динаміки й проблемних питань, озвучених під час семінару для нових іммігрантів
	Залучення тренерів до спостереження занять одне одного; нотування динаміки, питань і рівня участі слухачів

Джерело: Взято з University of Wisconsin Cooperative Extension 1996.

Спостереження – корисний інструмент збору даних у випадку, якщо метою оцінювання є збір еталонних і описових даних або документування програмних заходів, процесів та результатів. Цей підхід працює у найрізноманітніших ситуаціях (табл. 8.5).

Спостереження можуть бути структурованими або напівструктурованими:

- Структуровані спостереження передбачають чітке визначення того, що має спостерігатися та протягом якого періоду. Спостерігачі зазвичай використовують контрольний список для підрахунку подій та повторень або ж секундомір для вимірювання тривалості заходів.
- Метод проведення неструктурованих спостережень обирають залежно від ситуації, без попереднього задуму або плану стосовно того, що і скільки спостерігати. Спостерігач веде спостереження й записи відповідно до того, як розвиваються події.

Напівструктуровані спостереження проводяться тоді, коли фахівець з оцінювання має загальне уявлення про те, що потрібно спостерігати, але не має конкретного плану. Фахівець з оцінювання може просто відзначати те, що вважає цікавим, типовим, незвичним або важливим. Він також може постійно вести записи про події, що відбуваються, або ж зосередитися на якихось окремих діях.

Структуроване спостереження

Застосовуючи структурований підхід до спостережень, Стейк (Stake 1995) розробив форму ведення спостережень з прив'язкою до наявних питань (див. рис. 8.2). Багато форм структурованих спостережень використовують скорочення й спеціальні позначки, що дозволяють заощадити місце для вписування нових даних. Форми можна змінювати з урахуванням потреб конкретних оцінок.

Рис. 8.2. Форма для спостережень з прив'язкою до існуючих питань

Спостерігач:	Навчальний заклад:	Дата:	Час спостереження: з до
Викладач: чол. жін.	Вік 25 35 50 65	Клас:	Час складання звіту:
Досвід викладача: 0 - - ... місяців	Пряме викладання: Мало - - Багато	Кількість учнів:	Тема:
Стисле викладення уроку, видів діяльності:		Коментарі щодо проблем наукової освіти: 1 реакція на урізання бюджету 2 центр влади 3 підготовка викладачів 4 інтерактивні матеріали	
Опис аудиторії	Педагогічна орієнтація	Мета викладача	Зроблено посилання на:
місце для навчання Н - - В	підручник Н - - В	дидактична Н - - В	науковий метод 0 - - хв.
місце для науки Н - - В	станд. тестування Н - - В	евристична Н - - В	технології 0 - - хв.
місце для змагання Н - - В	вирішення задач Н - - В	філетична Н - - В	етика, реліг. 0 - - хв.

На рис. 8.3 показано ще один приклад форми, що використовується для структурованого спостереження. Це інструмент для збору даних про комунікативні навички дітей.

Рис. 8.3. Зразок форми для структурованого спостереження

Показник раннього спілкування (ПРС)

Інститут дослідження раннього дитинства **з вимірювання зростання й розвитку**

Ім'я або номер дитини _____ Дата тесту _____ (ДД.ММ.РР)
Тривалість тесту _____ - _____
Хв. Сек.
Форма: Дім або Комора Зміна умов (див. список нижче): _____
Головний кодувальник: _____ Фахівець з оцінювання: _____
Розташування (обведіть один пункт): Дім Центр Інше
(поясніть в примітках) Мова проведення тесту: _____
Якщо кодується надійність, ім'я кодувальника: _____
Примітки:

	Жести	Вокалізації	Окремі слова	Групи слів
Початок 0:00 Сек.	Ж	В	С	Г
1:00 Сек.	Ж	В	С	Г
2:00 Сек.	Ж	В	С	Г
3:00 Сек.	Ж	В	С	Г
4:00 Сек.	Ж	В	С	Г
5:00 Сек.	Ж	В	С	Г
6 хв. Кінець Усього	Ж	В	С	Г

Список умов

інтерпретатор

набір інструментів мовної інтервенції

медична інтервенція (наприклад, трубки)

консультант з психічного здоров'я

навчання у середовищі

немає

дієтолог

інше

фізіотерапевт

постачальник первинних медичних послуг

дипломований медпрацівник

чутлива взаємодія

соціальний працівник

логопед

	Жести	Вокалізації	Окремі слова	Групи слів	Усього
Основний кодувальник					
Кодувальник надійності					
Згода					
Незгода					
Відсоток					

Загальний відсоток

Згода

З/З + Н

Визначення надійності:

1. Записати бали основного кодувальника в першому рядку.
2. Записати бали кодувальника надійності в другому рядку.
3. Записати номер, щодо якого у них була згода, в третьому рядку.
4. Записати номер, щодо якого у них була незгода, в четвертому рядку.
5. Розрахувати відсоток згод для кожної ключової категорії.
6. Розрахувати загальний відсоток згод за допомогою загальних балів.
7. Розрахувати середній відсоток згод за різними категоріями (скласти згоди й незгоди в різних категоріях (третій та четвертий рядки)).

Джерело: Early Childhood Research Institute on Measuring Growth and Development 2008.

Напівструктуроване спостереження

Напівструктуроване спостереження може не мати певного плану. Натомість воно може визначати те, що саме має спостерігатися (табл. 8.6).

Документування спостережень

Існує щонайменше три способи запису спостережень.

- *Робочий зошит для спостережень:* Друкований документ, в якому передбачено місце для запису спостережень (приклади робочих зошитів для спостережень можна знайти в наступних джерелах: Stake (1995), University of Wisconsin Cooperative Extension (1996), Yin (2003)).
- *Лист для записів або контрольний список:* Форма, що використовується для запису спостережень у форматі «так/ні» або за допомогою шкали оцінки. Листи для записів використовуються тоді, коли спостереження ведеться за конкретними, наочними об'єктами, особами або характеристиками.

Таблиця 8.6. Використання напівструктурованого спостереження

Спостереження за:	Приклади
характеристиками учасників (окремих осіб або груп)	Стать, вік, професія/рід занять, одяг, зовнішній вигляд, етнічна приналежність
	Ставлення до певної теми, до інших осіб, до себе
	Навички, звички та знання
взаємодіями	Рівень участі, зацікавленість
	Владні стосунки, прийняття рішень
	Загальний клімат для навчання, вирішення проблем
	Рівень підтримки, співпраця й конфлікти
невербальною поведінкою	Вирази обличчя
	Жести
	Пози
лідерами та фронтменами програм	Ясність передачі думок
	Лідерські навички, заохочування до участі
	Обізнаність про клімат у групі
	Гнучкість, здатність пристосовуватися
	Знання предметної області, наочне викладання, інші викладацькі/навчальні техніки
	Послідовність заходів
фізичним оточенням	Приміщення (простір, зручність, придатність)
	Додаткові зручності (напої і т. ін.)
	Розташування сидячих місць
тим, що було створено в рамках програми	Демонстрації, облаштоване місце, план
	Брошури, інструкції, інформаційні бюлетені

Джерело: Cloutier and others 1987.

Польові нотатки: Найменш структурований спосіб запису спостережень. Спостереження записуються у вільній описовій формі, коли спостерігач помічає щось важливе.

Польові нотатки містять «поточні дані, які збирає спостерігач. Вони складаються з опису власного досвіду й спостережень, слів людей, вражень і думок спостерігача щодо того, що він бачить, а також сформованих на місці інтуїтивних здогадок і тлумачень ситуації» (Patton 2002, p. 305).

Польові нотатки:
Опис власного досвіду й спостережень, слова людей, враження та думки спостерігача щодо того, що він бачить, а також сформовані на місці інтуїтивні здогадки і тлумачення

За словами Лофленда (Lofland 1971, p. 102), польові нотатки є «найважливішими визначальними факторами для проведення якісного аналізу. Польові нотатки передають саму сутність спостережень. Спостерігач, який не веде їх, з такиим же успіхом міг би залишатися вдома».

Петон (Patton 2002) розглядає методики ведення польових нотаток під час спостереження. Він вказує на те, наскільки важливо описувати все, що спостерігач вважає вартим уваги, не покладаючись на пам'ять. Спостерігач має якомога швидше записувати інформацію, що допомогла йому зрозуміти контекст, ситуацію та події. Польові нотатки мають містити описову інформацію, яка в ході аналізу дозволить фахівцю з оцінювання подумки повернутися до спостереження й власних вражень. Петон пропонує занотовувати основну інформацію – наприклад, де проходило спостереження, хто був присутній, як виглядало це місце, які соціальні взаємодії відбувалися, а також яка діяльність велася.

Він відзначає важливість використання конкретних термінів (на відміну від більш загальних): такі слова як «поганий», «злий» та «ніяково» є недостатньо описовими. «Такі неконкретні слова приховують те, що відбувалося насправді, замість того, щоб розкривати подробиці ситуації», – стверджує Петон.

Навчання й підготовка спостерігачів

Петон (Patton, 2002, p. 260) вказує, що «наявність слуху та зору сама по собі не робить людину кваліфікованим спостерігачем». Він говорить про важливість підготовки спостерігачів, визначаючи шість її компонентів:

- Вчитися бути уважним, бачити й чути все, що відбувається.
- Практикуватися у веденні описових нотаток.
- Виробляти дисципліну ведення польових нотаток.
- Вчитися відділяти корисні подробиці від незначущих фактів, аби останні не створювали інформаційного шуму.
- Ретельно виконувати валідацію та тріангуляцію спостережень.

Повідомляти про сильні сторони й обмеження власної точки зору – це вимагає одночасно знання себе та готовності поступитися своєю приватністю.

Таблиця 8.7. Переваги й недоліки збору даних шляхом спостереження

Переваги	Збір даних про фактичну поведінку замість добровільно наданих даних або суб'єктивних вражень.
	Збір даних в режимі реального часу, а не ретроспективно.
Недоліки	Формування вибірки, фіксація й тлумачення даних можуть бути складними.
	Збір даних може бути трудомістким.

Джерело: Автори.

Якщо є така можливість, слід використовувати більше одного спостерігача. Всі спостерігачі повинні бути навчені вести спостереження згідно з погодженими процедурами.

Процедура збору даних за допомогою спостережень має пройти попередню перевірку. Для цього щонайменше двоє спостерігачів повинні відправитися в одне й те саме місце й вести польові нотатки за допомогою форм кодування та листів оцінки. Потім ці документи порівнюються. Якщо виявляється, що між нотатками існують значні відмінності, потрібно провести додаткову підготовку та краще роз'яснити методику. Якщо ж відмінності незначні, процедуру можна запроваджувати.

У табл. 8.7 в стислому вигляді вказано переваги й недоліки збору даних за допомогою спостережень.

Інструмент 4: Опитування й інтерв'ю

Опитування – прекрасний інструмент для збору даних про враження, думки й ідеї людей. Він менш корисний, коли мова йде про вимірювання поведінки, адже те, що люди кажуть, може не збігатися з тим, що вони власне роблять.

Опитування може бути структуроване або напівструктуроване, його можна проводити особисто, по телефону або взагалі незалежно, коли респонденти заповнюють форму, яку вони отримали поштою або через веб-сайт. Опитування може охоплювати вибірку цільової аудиторії або ж усю цільову аудиторію (формування вибірки розглядається в главі 9).

Структуровані та напівструктуровані опитування

Структуроване опитування: Опитування, що включає певний набір варіантів відповідей, з яких респонденти можуть вибрати один або декілька

Структурованими називають опитування, що включають певний набір варіантів відповідей, з яких респонденти можуть вибрати один або декілька. Всім респондентам в абсолютно однаковий спосіб ставлять абсолютно однакові запитання, надаючи одні й ті самі варіанти відповідей.

Скільки варіантів слід пропонувати? Для питань, що стосуються такої інформації, як регіон країни, рід занять або вікова група респонден-

та, кількість варіантів визначається кількістю можливих відповідей. Якщо для відповіді використовується шкала, варіантів має бути непарне число (3, 5 або 7), щоб респондент міг легко обрати нейтральний варіант. (Іноді використовуються шкали з парним числом балів – якщо необхідно, щоб респондент зробив чіткий вибір між оцінками «задовільно» та «частково незадовільно».) Напівструктуровані опитування ставлять ті самі загальні питання, але на більшість або й на всі з них респонденти можуть давати відповідь у вільній формі.

Блок 8.6 «Питання структурованих та напівструктурованих опитувань»

Приклади структурованих питань:

1. Чи допоміг вам цей семінар навчитися оцінювати свою програму?
   - Ні або майже ні
   - Певною мірою
   - Допоміг
   - Дуже допоміг
   - Надзвичайно сильно допоміг
   - Не маю думки
   - Н/з
2. Чи всі люди у вашому селищі мають джерело чистої води в межах 500 метрів від дому?
   - Так
   - Ні

Приклади напівструктурованих питань:

1. Які знання, отримані під час семінару з оцінювання програм, ви застосували у своїй роботі?
2. Де знаходяться джерела чистої води, якими користуються селяни?

Напівструктуроване опитування: Опитування, на більшість питань якого респонденти дають відповіді у вільній формі

Напівструктуроване опитування – опитування, на більшість питань якого респонденти дають відповіді у вільній формі. Опитування цього типу особливо корисні, якщо фахівець з оцінювання прагне отримати поглиблене розуміння реакцій на досвід або зрозуміти причини того, чому респонденти дотримуються тих чи інших точок зору. Напівструктуровані опитування повинні мати чітко визначену мету. Якщо фахівець з оцінювання хоче отримати відомості про кроки у процесі, про ролі й сфери відповідальності різних членів громади чи команди або про те, як працює програма, доцільніше провести інтерв'ю з людьми, ніж намагатися розробити письмове опитування, що враховує всі можливі варіанти відповідей. Якщо з потенційними респонден-

тами вдається встановити хороший контакт, якщо їм ставити цікаві питання, вони можуть погодитися брати участь в інтерв'ю протягом години або й більше. При цьому мало хто захоче провести годину за заповненням анкети. В ідеалі кожне інтерв'ю мають проводити дві особи, щоб потім мати змогу порівняти їхні записи. Залучення одразу двох інтерв'юерів може допомогти розв'язати суперечки з приводу того, що саме було сказано.

У табл. 8.8 дається короткий перелік переваг структурованих та напівструктурованих опитувань.

Таблиця 8.8. Переваги структурованих та напівструктурованих опитувань

Структуровані опитування	Напівструктуровані опитування
Важче розробляти: опитування має враховувати всі можливі варіанти	Легше розробляти: опитування може містити відкриті питання, що дозволяє отримати будь-яку інформацію, не включену в структуровані розділи, завдяки чому зменшується ризик залишити щось поза увагою
Легше проводити: поставити відмітку в якомусь полі швидше, ніж записати описову відповідь	Складніше проводити: людям обтяжливо заповнювати анкети
Легше аналізувати	Важче аналізувати, проте такі опитування дають багатші дані; інтерпретація відповідей у вільній формі може бути суб'єктивною
Ефективніші при роботі з великою кількістю людей	

Джерело: Автори.

Способи проведення опитування

Найчастіше використовуються наступні засоби проведення опитувань:

- телефонні анкети;
- анкети для самостійного заповнення, що розповсюджуються за допомогою пошти, електронної пошти або веб-сайтів;
- анкети, що заповнюються спільно (розповсюджений варіант в контексті розвитку).

Телефонні опитування є корисними для розуміння досвіду й думок, а також для отримання опису процесу, складеного окремими людьми. Особисті та телефонні опитування схожі; відмінність між ними лише у формі.

Анкети для самостійного заповнення можуть доставлятися респондентам особисто або поштою, електронною поштою чи через веб-сайти. Респонденти відповідають на питання на бланку анкети, який потім повертають інтерв'юеру. Опитування, що передбачають самостійну відповідь респон-

дентів, мають бути короткими. Заповнення анкети повинно займати не більше 20 хвилин (а краще – ще менше). Результати досліджень свідчать, що люди найчастіше чесно відповідають на незручні питання саме в анонімних опитуваннях, коли вони самі заповнюють анкету.

Використання пошти, електронної пошти або веб-сайтів набагато менш витратне, ніж проведення опитування особисто. Форми опитувань можуть бути надіслані в будь-яку місцевість, де потенційні респонденти мають доступ до цих методів розповсюдження. Якщо доступність поштового й телефонного зв'язку, мережі Інтернет в певній місцевості є обмеженою, участь цієї місцевості також буде обмеженою, що стане джерелом помилкових висновків. Низький рівень грамотності респондентів робить ці підходи у багатьох контекстах невідповідними.

Типовими для контексту розвитку є опитування, коли анкета заповнюється співробітником програми оцінювання. Такі опитування можуть бути структурованими, напівструктурованими або змішаного типу. Спеціаліст зі збору даних зачитує вголос питання й записує відповіді. Загалом краще використовувати закриті питання, додаючи в кінці одне–два відкритих. Ці відкриті питання подобаються людям, тому що дозволяють висловити щось, що, на їхню думку, опитування не враховує, або ж висловити свою думку стосовно самого опитування. Одного або двох відкритих питань достатньо; відповідати на більшу кількість питань може бути обтяжливо. Аналіз цих відкритих питань може займати багато часу, але вони дають важливу інформацію і можливість цитувати респондентів для ілюстрування основних тем оцінювання.

Комп'ютеризована система телефонних опитувань (КСТО, анг. CATI) використовує інтерактивну комп'ютерну систему, що допомагає інтерв'юерам, які ставлять питання по телефону. Інтерв'юер починає інтерв'ю, а комп'ютерна програма контролює перехід між питаннями. Вибір наступних питань ґрунтується на відповідях на попередні, що дозволяє проводити більш індивідуалізовані, складніші інтерв'ю. Інтерв'юери вводять дані й прості коди безпосередньо в комп'ютерну систему. Більшість питань йдуть з варіантами відповідей, тож інтерв'юеру треба лише натиснути на

Комп'ютеризована система телефонних опитувань (КСТО):
Опитування з використанням інтерактивної комп'ютерної системи, яку застосовують інтерв'юери, ставлячи питання по телефону

Таблиця 8.9. Переваги й недоліки опитувань

Переваги	Респонденти викладають свої думки, переконання та враження.
Недоліки	Люди можуть бути неготові точно пригадати свої дії, або ж не хотіти розказувати про якусь свою поведінку, особливо якщо вона незаконна або вважається ганебною.
	Те, що люди, на їхню думку або за їхніми словами, роблять, не завжди збігається з тим, що вони роблять насправді.

Джерело: Автори.

відповідь; комп'ютерна система перетворює дані на код, який зберігає в базі даних (UNESCAP 1999).

У табл. 8.9 в стислому вигляді вказані переваги й недоліки використання опитувань.

Розробка питань для опитування

Розробка питань для опитування включає:

- вибір питань;
- формулювання питань і відповідей;
- визначення порядку питань;
- вибір схеми опитування;
- перегляд, переклад і попереднє тестування питань.

Вибір питань. Фахівці з оцінювання можуть використовувати відкриті або закриті питання. На відкриті питання неможливо дати коротку відповідь або вибрати один з варіантів. Респонденти часто не бажають (або ж їм незручно) давати письмові відповіді на відкриті питання.

На закриті питання можна відповідати, надаючи один простий елемент інформації. Питання «Ваша дата народження?» – закрите, тому що для відповіді для нього достатньо вказати дату народження, без додаткових подробиць. Другою формою закритого питання є альтернативне питання, тобто таке, що передбачає відповідь «так/ні» або «вірно/невірно». Третя форма закритого питання – питання з варіантами відповідей. Багато експертів радять використовувати переважно закриті питання, але ставити одне–два відкритих наприкінці опитування (Burgess 2001) чи залишати місце для коментарів після кожного питання.

Використання одразу великої кількості різних форм питань може збити респондента з пантелику. Джексон і Агревал (Jackson and Agrawal 2007) рекомендують використовувати в опитуваннях не більше трьох форм питань.

Згідно з Дослідженням критеріїв оцінки рівня життя (ДКОРЖ) (World Bank 1996), доцільно розробляти питання опитувань поетапно (див. табл. 8.10). Таким чином можна отримати добре збалансовану анкету.

Визначення типу питання змушує фахівця з оцінювання чітко визначитися з тим, про що він хоче спитати, а також допомагає респонденту надати належну відповідь. Петон (Patton 2002) описує шість видів питань:

- *Питання про досвід і поведінку:* Що робить або збирається робити людина? Такі питання ставляться для того, щоб дізнатися про поведінку, досвід, вчинки або діяльність.
- *Питання про думки й цінності:* Вони спрямовані на розуміння когнітивних та інтерпретативних процесів людей (а не вчинків чи поведінки).

- *Питання про відчуття:* Такі питання дають можливість дізнатися про емоції людей, пов'язані з враженнями й думками. Відповіді зазвичай даються у вигляді прикметників («тривожний», «щасливий», «наляканий», «пригнічений», «впевнений» і т. ін.). Питання має показати респонденту, що воно стосується думок, переконань і зважених суджень, а не почуттів.

Таблиця 8.10. Етапи розробки змісту анкети

Етап	Опис
Визначення загальних цілей та завдань	Визначення цілей (дослідити бідність; зрозуміти вплив державної політики на домашні господарства).
Балансування розділів	Визначення найважливіших питань (частота надання субсидій на їжу, вплив змін доступності або вартості державних послуг в сфері охорони здоров'я й освіти, вплив змін економічного клімату як результат структурних змін або переходу від центрального планування до ринкової економіки).
Балансування питань всередині розділів	У рамках, приміром, освітнього сектору визначається, які з наступних факторів є найважливішими для країни на даний момент: • рівні й визначальні фактори зарахування до навчальних закладів, поганої відвідуваності занять, навчання, а також відмінностей у показниках для чоловіків та жінок; • вплив кількості років навчання на заробіток в офіційному секторі та сільському господарстві, різниця доходів у цих секторах; • хто з дітей має підручники або отримує шкільні обіди чи стипендії; • скільки доводиться батькам платити за навчання своїх дітей.
Запис питань для вивчення конкретних проблем або програм	Наприклад, досліджуючи доступність підручників, фахівець з оцінювання має знати: • для скількох предметів наявні підручники; • підручники видаються кожній дитині чи планується їх спільне використання; • підручники видаються додому чи використовуються лише в класі; • скільки років мають використовуватися підручники – один чи кілька; • чи є підручники безкоштовними; • коли підручники мають бути в наявності; • що краще – підручники з магазину чи зі шкільної бібліотеки.

Джерело: World Bank 1996.

- *Питання про знання:* Питання про те, що саме відомо респонденту.
- *Сенсорні питання:* Питання про те, що можна побачити, почути, відчути на дотик, за допомогою нюху та смаку.
- *Питання про анкетні дані:* Вік, освіта, вид занять та інші характеристики особи, з якою проводиться інтерв'ю.

Формулювання питань і варіантів відповіді. Питання кожного з визначених Петоном шести типів можна поставити в теперішньому, минулому або майбутньому часі. Приміром, фахівець з оцінювання може спитати:

- Що ви знали про лікування ВІЛ/СНІД п'ять років тому? (минуле)
- Що ви дізналися із сьогоднішньої презентації про ВІЛ/СНІД? (теперішнє)
- Що б ви хотіли дізнатися про ВІЛ/СНІД? (майбутнє) (Patton 2002)

Блок 8.7 «Рекомендації щодо формулювання питань в опитуваннях»

Наступні поради можуть допомогти фахівцям з оцінювання формулювати питання таким чином, щоб респонденти добре розуміли, яку інформацію вони мають надати:

- Використовуйте прості слова, що мають однакове значення для всіх респондентів.
- Питання й відповіді мають бути прив'язаними до конкретної категорії.
- Уникайте двозначних питань і варіантів відповідей (уникайте питань або відповідей, в яких використовуються слова «та» чи «або»).
- Уникайте питань, що передбачають наявність у респондента певних знань. В разі необхідності надавайте відповідну інформацію на початку питання.
- Уникайте дублювання заперечення в комбінації «питання-відповідь» (наприклад, уникайте використання слова «не» в питанні, на яке треба відповісти «так» або «ні»).
- Робіть варіанти відповіді взаємовиключними.
- Варіанти відповіді мають бути збалансованими.
- Уникайте питань, що можуть викликати в респондента заперечення, звучати нав'язливо або зверхньо.

Джерело: TC Evaluation Center 2007. Інші корисні джерела: Foddy (1993), Fowler (2002), Sue and Ritter (2007).

Питання оцінки повинні бути зрозумілими (блок 8.7). В іншому випадку респонденту буде незручно, він відчуватиме свою неосвіченість, почуватиметься спантеличеним або ворожо налаштованим. Фахівці з оцінювання повинні вивчити терміни, які вживають люди з цільової аудиторії опитування, говорячи про проект, програму або політику, щоб згодом вживати ці ж слова при формулюванні опитування.

Фахівець з оцінювання має не допустити підштовхування респондента до бажаної відповіді. Візьмемо, наприклад, наступні питання:

«Чи збільшилися ваші прибутки після тренінгу, і якщо так, то наскільки?»

«Якщо ваші прибутки збільшилися (незначно, помірно, суттєво або дуже суттєво), чи пов'язуєте ви це з проходженням тренінгу?»

Очевидно, що ці питання «ведуть» респондента тим шляхом, який був би бажаний для фахівця з оцінювання. Краще було б запитати про тенденції зміни прибутків у довгостроковому періоді, а потім, якщо мало місце якесь підвищення, запитати про фактори, які його зумовили.

Незначні зміни у формулюванні питань можуть сильно вливати на відповіді респондентів.

> Кілька дослідників вивчали, які наслідки може давати зміна прикметників та прислівників. Такі слова, як «зазвичай», «часто», «іноді», «час від часу», «зрідка» та «нечасто», регулярно вживаються в анкетах, хоча зрозуміло, що для різних людей вони мають різне значення. Деякі прикметники є дуже варіативними, в той час як інші – ні. Наступні слова і фрази мають дуже варіативні значення, а тому використання їх в опитуваннях варто уникати: «чітка сфера повноважень», «переважно», «численний», «переважна більшість», «менша частина», «значна частина», «значна кількість», «багато», «суттєва кількість», «мало». Інші слова і фрази мають меншу варіативність і тлумачаться більш-менш однаково. Це, наприклад: «майже всі», «практично всі», «більшість», «одноголосно», «невелика кількість», «не дуже багато», «майже ніхто», «пара», «кілька» (StatPac 2007).

Фрері (Frary 1996) дає кілька рекомендацій щодо створення ефективних анкет. Зокрема, стосовно відповідей з використанням шкали він рекомендує наступне:

- Вказуйте варіанти відповіді в порядку зростання зліва направо:
  1) Ніколи 2) Рідко 3) Час від часу 4) Часто.
- Можна спробувати комбінувати категорії відповідей, якщо малоймовірно, що респонденти будуть відповідати «ніколи», та якщо «рідко» означає майже той самий рівень активності:
  1) Рідко або ніколи 2) Час від часу 3) Часто.
- Просіть респондентів оцінювати як позитивні, так і негативні стимули (питання з обох боків шкали). В цьому випадку респонденти повинні оцінювати кожну відповідь, а не загалом погоджуватися чи не погоджуватися з усіма варіантами.
- Наскільки можливо, скорочуйте кількість варіантів відповіді. Наприклад, якщо думки респондентів з великою імовірністю будуть зрозумілі, використовуйте простий набір: *1) Згоден 2) Не згоден*. Якщо думка багатьох респондентів є неоднозначною, пропонуйте більше варіантів:
  1) Згоден 2) Скоріше згоден 3) Скоріше не згоден 4) Не згоден.

Сформулювавши питання, фахівець з оцінювання має прочитати кожне з них з відповідями, щоб перевірити якість формулювання й логіки. Чи все добре з граматикою? Чи узгоджена відповідь з питанням з точки зору логіки?

Формулювати питання для опитування непросто, адже вони мають бути зрозумілими для кожного. Слова мають багато значень і конотацій. Якщо питання ставиться таким чином, що люди його не розуміють або ж розуміють по-різному, вони відповідатимуть зовсім на інше питання, а не на те, яке ви мали на увазі. Також існує ризик збору непотрібних даних. Наприклад, голова відомства хоче дізнатися, який обсяг навчання роботі з комп'ютерами отримали його підлеглі. Можна поставити наступні питання:

- Чи проходили ви якесь навчання протягом останніх трьох місяців?
- Чи проходили ви якесь навчання протягом останніх шести місяців?
- Чи проходили ви якесь навчання протягом останнього року?

Проблема з цим набором питань полягає у тому, що кожен, хто проходив навчання протягом останніх трьох місяців, ствердно відповість на всі три питання.

Більш продумане питання матиме такий вигляд: «Скільки тренінгів ви відвідали протягом кожного з наступних періодів?

a. Протягом попередніх 1 – 3 місяців: _____
b. Протягом попередніх 4 – 6 місяців: _____
c. Протягом попередніх 6 – 9 місяців: _____
d. Протягом попередніх 10 – 12 місяців: _____
e. Протягом попередніх 12 – 24 місяців: _____

Погано сформульовані питання можуть викликати роздратування респондентів, які відповідатимуть навмання або навіть відмовлятимуться від опитування. Оскільки шкоду від погано сформульованих питань компенсувати на етапі аналізу неможливо, поставитися до їх формулювання слід відповідально. Отже, фахівці з оцінювання повинні зарезервувати достатньо часу на перевірку й попереднє тестування опитування.

Необхідно забезпечити конфіденційність і анонімність. Конфіденційність – це етичний принцип, що захищає інформацію, відкриваючи доступ до неї лише уповноваженим особам. Анонімність означає, що особа та персональна інформація респондента залишаться закритими. Багато фахівців з оцінювання використовують наступну заяву: «Ваші відповіді будуть конфіденційними, а дані завжди представлятися таким чином, щоб жоден елемент опублікованих даних не можна було пов'язати з вашою особою» (Burgess 2001).

Визначення порядку питань. Визначення порядку питань наступним чином може підвищити шанси на те, що опитування буде пройдено респондентом до кінця:

- Почніть з цікавих для респондента питань.
- Питання мають бути упорядковані логічно; слід уникати складних відгалужень.

- Групуйте схожі або пов'язані питання і намагайтеся встановити логічну послідовність для них (Burgess 2001).
- Ведіть респондента від простих і цікавих питань до складних і незручних, від загальних – до конкретних (Jackson and Agrawal 2007).
- Розміщуйте питання в хронологічній послідовності (Jackson and Agrawal 2007).
- Ставте найважливіші питання протягом перших двох третин анкети, тому що респонденти можуть не пройти опитування до кінця (Jackson and Agrawal 2007).
- Розміщуйте особисті або конфіденційні питання в кінці опитування – якщо поставити їх на початку, деякі респонденти можуть відмовитися відповідати далі (Frary 1996).
- Визначте пріоритетність кожного питання й розгляньте можливість відкидання питань з низьким пріоритетом (Jackson and Agrawal 2007).

Джексон і Агравел (Jackson and Agrawal 2007) рекомендують, щоб більшість анкет для самостійного заповнення займали не більше однієї сторінки. Анкети для респондентів з університетською освітою можуть займати дві сторінки.

Вибір схеми опитування. Потрібно уважно поставитися до схеми друкованої анкети:

- Вказати назву опитування й дату редакції анкети.
- Анкета повинна бути лаконічною, без зайвих заголовків.
- На бланку має бути вказана контактна інформація й відомості про те, куди повертати заповнену анкету.
- Зробити коротку вступну заяву (на випадок, якщо респондент не отримує супроводжувального листа), в якій викладається мета опитування та пояснюється, як воно використовуватиметься.
- Питання й варіанти відповіді повинні мати єдиний формат. Кожне питання має бути пронумероване.
- Анкета має містити вказівки стосовно того, як треба відповідати на питання. Повинно бути зазначено, скільки варіантів відповідей можна вибрати (один або всі, що підходять), а також яким чином позначати вибір (поставити відмітку в клітинці, обвести відповідь, написати коротку відповідь) (Burgess 2001).

Джексон і Агравел (Jackson and Agrawal 2007) рекомендують надавати для відповіді порожні клітинки або лінії, між якими має бути достатньо місця, щоб фахівець з оцінювання міг чітко визначити, що саме вибрав респондент. Вони також наголошують на важливості професійної оцінки опитування.

До анкети необхідно додавати супроводжувальний лист. Такий лист визначає спонсора й надає учасникам відомості про мету опитування. Джексон і Агравел (Jackson and Agrawal 2007) пропонують включати в супроводжувальний лист наступні відомості:

- мета опитування (і чому респондент має заповнити анкету);
- як конкретна особа обиралася для участі в опитуванні;
- як будуть використані дані;
- чи буде забезпечено анонімність;
- скільки часу буде потрібно для заповнення анкети;
- заохочення;
- контактна інформація для тих, хто прагне дізнатися більше чи отримати роз'яснення.

Якщо опитування проводиться через пошту, необхідно додавати також підписаний конверт з марками або впровадити іншу процедуру для повернення заповненої анкети.

Перегляд, переклад і попереднє тестування питань. Процес розробки опитування – ітеративний. Після підготовки першої версії анкети зацікавлені сторони повинні уважно її переглянути. В ході перегляду вони мають робити нотатки, висловлювати зауваження й рекомендації. Потім в анкету вносяться зміни. Можливо, цей процес доведеться повторити кілька разів, перш ніж анкета буде готова для попереднього тестування.

Якщо для збору даних потрібно застосувати більш ніж одну мову, анкету доведеться перекласти. Особи, відповідальні за проведення опитування, повинні переконатися в правильності перекладу. Перший переклад анкети має виконувати підготовлений працівник, який знає обидві мови й обізнаний з темою опитування (див. Behling and Law 2000). Після перекладу анкети на місцеву мову інша особа має виконати зворотний переклад. Цей процес дозволяє виявити недоліки та непорозуміння стосовно перекладеного матеріалу. Це єдиний спосіб забезпечити точний переклад. Будь-які неузгодженості щодо анкети потрібно усунути до початку попереднього тестування.

В більшості випадків анкети друкуються лише офіційною мовою (або мовою) країни, а до збору інформації залучають групи інтерв'юерів, що володіють кількома місцевими мовами. В таких випадках кілька ключових питань або фраз перекладаються на місцеві мови й включаються до інструкції з проведення опитування. Для роботи з респондентами, які говорять іншими мовами, необхідно залучати місцевих перекладачів.

Питання обов'язково треба формулювати простими словами зрозумілою для всіх мовою. Не варто викладати їх академічною або офіційною мовою. Може існувати значна різниця між усною та письмовою мовою у певній

місцевості, тому що дуже складно віднайти баланс між простотою і точністю формулювань. Це особливо актуально для мов, які рідко використовуються в письмовому вигляді.

Після того, як анкету погоджено, вона має пройти попередню перевірку в польових умовах із залученням невеликої кількості респондентів. За результатами цієї перевірки, ймовірно, буде підготовлена нова редакція анкети. Якщо перша польова перевірка вказує на необхідність великої кількості змін, може знадобитися додаткова польова перевірка, що охопить всю анкету або її частину, особливо якщо мали місце кілька перекладів.

Якщо у вас так багато анкет та інтерв'ю, надзвичайно важливо якомога раніше провести польову або попередню перевірку інструмента для збору даних. Це означає необхідність опитування невеликої, але репрезентативної вибірки потенційних респондентів і виокремлення моментів, що потребують роз'яснення. Одна із ефективних стратегій – спостерігати за тим, як респонденти заповнюють анкету, і просити їх вголос пояснювати, чому вони дають ті чи інші відповіді. Ця вправа може дати уявлення про те, як потенційні респонденти тлумачать питання. Нерідко те, що вважається абсолютно зрозумілим, виявляється незрозумілим або двозначним. Якщо респонденти неправильно розуміють питання, фахівцям з оцінювання потрібно його переглянути й повторно перевірити, перш ніж впроваджувати цей інструмент.

Мета польової перевірки – переконатися, що анкета дає змогу збирати потрібну інформацію. Правильний попередній тест розглядає анкету на трьох рівнях:

- *В цілому:* Чи всі частини анкети логічно узгоджені? Чи не ставиться в різних її частинах одне й те саме питання?
- *Кожний розділ:* Якщо анкета складається більш ніж з одного розділу, чи забезпечує кожен з них збір необхідної інформації? Чи враховуються всі основні види діяльності? Чи є нерелевантні питання?
- *Окремі питання:* Чи зрозумілі використані формулювання? Чи передбачають питання двозначні відповіді? Чи можливі альтернативні тлумачення?

Під час проведення польової або попередньої перевірки важливо тестувати анкету на людях з різних місцевостей, що представляють всі основні мовні та соціально-економічні групи. До їх числа можуть входити сільські та міські жителі, особи, які працюють в офіційному та тіньовому секторі, а також представники всіх ключових мовних груп.

Якщо попереднє тестування проводиться лише однією мовою, його проведення зазвичай займає близько тижня. Якщо заключне опитування має проводитися більш ніж однією мовою, на його перевірку знадобиться більше часу, адже потрібно буде перевірити версію анкети кожною мовою.

Після попереднього тесту необхідно один-два тижні відвести для перегляду результатів. Тестувальники повинні зібратися разом і обговорити й узгодити потрібні зміни. В ході типового широкомасштабного опитування зазвичай проводять три цикли попередніх перевірок.

Блоки 8.8 і 8.9 містять вказівки щодо проведення опитувань. В першому з них акцент робиться на процесі адміністрування, в другому – на написанні хороших питань.

Картка оцінки для громадян

Картка оцінки для громадян: Різновид опитувальника, що дозволяє вимірювати результативність роботи державних або комунальних послуг

Картка оцінки для громадян – це різновид опитувальника, що дозволяє вимірювати ефективність державних або комунальних послуг. Її все частіше використовують в контексті розвитку. Наприклад, у місті Бангалор (Індія), група приватних осіб використовує такі картки для вирішення проблем з комунальними службами (Paul and Thampi 2008). Картки оцінки для громадян вже багато років використовуються в США для оцінки роботи державних установ.

Блок 8.8 «Загальні вказівки з проведення опитувань»

Нижче наведено загальні вказівки з проведення опитувань:

- Опитування має бути простим, чіткім, легким і коротким.
- Знайдіть інших людей, які проводили аналогічне оцінювання. Знайдіть та вивчіть опитування, схожі на те, яке ви хочете провести.
- Переконайтеся, що люди знають, чому ви просите їх взяти участь (але так, щоб це не вплинуло на їхні відповіді).
- Ставте питання, на які легко відповідати і які не заважають респондентам давати чітку відповідь.
- Не ставте питання, для відповіді на які респондентам треба шукати інформацію в якихось сторонніх джерелах. Якщо ж це необхідно, попередьте їх заздалегідь, щоб вони могли зібрати матеріал до того, як взятися за опитувальник.
- Поважайте приватність респондентів. Проводьте опитування конфіденційно, впроваджуйте процедури для забезпечення приватності. Ніколи не обіцяйте конфіденційності, якщо не можете її забезпечити.
- Поважайте час та інформацію респондентів.
- Повідомляйте респондентам, як їх було обрано та чому їхня участь важлива.
- Не завдавайте шкоди: зберігайте конфіденційність відповідей. У своєму звіті використовуйте узагальнені відповіді. Призначайте даним ідентифікаційний номер і знищуйте посилання на ім’я особи.
- Якщо можете надіслати респондентам копію заключного звіту, що використовує результати опитування, повідомте їх про це.

Джерело: Автори.

Блок 8.9 «Рекомендації щодо написання опитувальників»

Наступні рекомендації можуть допомогти фахівцям з оцінювання у написанні опитувальників:

1. Наскільки можливо, використовуйте в якості зразка існуючий опитувальник. Вносьте зміни з урахуванням своїх потреб. Змінювати легше, ніж створювати опитувальник з нуля.
2. Дотримуйтесь наступних базових вказівок:
   - Вживайте мову, зрозумілу респондентам.
   - Ставте одне питання за один раз, наприклад: «Наскільки матеріал є зрозумілим?», а не «Наскільки матеріал є зрозумілим та корисним?» (Якщо матеріал зрозумілий, але не корисний, респондент не зможе висловити свої враження.)
   - Ставте питання так, щоб всі респонденти відчували, що їхні відповіді прийнятні. Починайте своє питання, задаючи діапазон: «Наскільки…» або «Наскільки важливий…»
   - Пропонуйте в якості варіантів відповіді взаємовиключні категорії. Якщо мова йде про вікові групи, формуйте категорії на кшталт: 20 – 30, 31 – 40, 41 – 50, а не 20 – 30, 30 – 40, 40 – 50.
   - За можливості ставте питання так, щоб відповіді вибудовувалися в певному порядку – від негативних до позитивних:

«Якою мірою … було … Зовсім не корисне ← ... → Дуже корисне».

   - Уникайте відповідей «так/ні», якщо відповідь не можна сформулювати однозначно (позитивно або негативно). Натомість намагайтеся врахувати діапазон поглядів, запропонувавши респондентам шкалу. Це може бути, наприклад, шкала з п’яти позицій – від «ніяк або мало» до «дуже значною мірою».
   - Уникайте категоричних формулювань з будь-якого краю шкали (мало хто має категоричну думку стосовно будь-яких питань). Пом’якшуйте граничні формулювання, наприклад: «завжди або майже завжди» з одного краю шкали та «ніколи або майже ніколи» – з іншого.
   - Ставте питання про поточну ситуацію. Пам’ять з часом притупляється.
   - Залишайте вихід (використовуйте такі категорії, як «не маю даних, щоб винести судження» та «не маю якоїсь певної думки»). Якщо ви не залишите такого варіанту, респонденти можуть давати відповіді, які не мають сенсу.
   - Уникайте подвійних заперечень.
3. Прагніть того, щоб респондентам було легко заповнити опитувальник до кінця. Опитувальник повинен містити клітинки, в яких респондент відмічатиме свої відповіді. Надайте достатню інформацію, щоб респонденти знали, що саме вони мають робити: наприклад, «оберіть один варіант» або «оберіть все, що підходить».
4. Ставте спочатку загальні питання, далі – більш конкретні, а в кінці – одне або два відкритих питання, наприклад: «ви маєте зауваження чи хочете нам ще щось повідомити?»
5. Питайте лише про ті демографічні відомості, які збираєтесь використати. Майте на увазі, що деяких людей можна ідентифікувати за демографічними характеристиками. Розташовуйте демографічні питання ближче до кінця опитувальника.
6. Залучіть експертів до перевірки чернетки питань.
7. Якщо питання потрібно перекласти на іншу мову, виконайте переклад, а потім перекладіть їх в зворотному напрямку, щоб перевірити якість перекладу.
8. Якомога більше попередніх тестів! Замість того, щоб просто читати опитувальнник, дайте типовим респондентам заповнити його. Далі проаналізуйте з ними кожне питання, щоб почути їхню думку. Чи зрозуміле питання? Чи зрозуміло було, про що йдеться в питанні? Невідомі слова або незрозумілі фрази? Чи можна сформулювати кожне з питань якось краще?

Джерело: Автори.

Картки оцінки для громадян використовуються для:

- отримання зворотного зв'язку стосовно роботи державних і комунальних служб від користувачів їх послуг;
- оцінки роботи окремих постачальників послуг та порівняння їх показників;
- формування бази даних зворотного зв'язку стосовно послуг, які можна віднести до державного сектору (ADB and ADBI 2004).

В табл. 8.11 наведено приклад результатів заповнення громадянами карток оцінки якості наданих послуг.

Оцінка громадян слугує джерелом відомостей про наступні фактори:

- наявність послуг;
- доступність послуг;
- надійність послуг;
- загальна якість послуг;
- задоволення від послуг;
- реакція постачальника послуг на звернення громадян;
- приховані витрати;
- корупція й системи забезпечення;
- готовність платити (World Bank 2006).

Картки оцінки для громадян можуть використовуватися по-різному (див. табл. 8.12). Вони можуть також застосовуватися для різних рівнів урядування.

У табл. 8.13 в стислому вигляді наводяться переваги й недоліки карток оцінки для громадян в контексті збору даних.

Таблиця 8.11. Результати заповнення громадянами карток оцінки задоволення від інфраструктурних послуг

Послуга	Кількість користувачів серед респондентів (тис.)	Відсоток задоволених користувачів	Відсоток незадоволених користувачів
Енергопостачання	1024	43	15
Водопостачання	775	41	19
Телефон	203	87	12
Поліція	98	45	36

Джерело: Автори.

Інтерв'ю

Інтерв'ю – це діалог між кваліфікованим інтерв'юером та особою, з якою проводиться інтерв'ю. Мета – отримати інформацію, яку можна буде використати для відповіді на питання оцінки (блок 8.10). Якість отриманої інформації залежить переважно від уміння й особистості інтерв'юера (Lofland and Lofland 1995). Запорука успішного інтерв'ю – здатність добре слухати й добре ставити питання.

Таблиця 8.12. Використання карток оцінки для громадян

Вид дослідження або проекту	Приклад
Дослідження задоволення користувачів комунальними послугами	Сім індійських міст; Кампала, Уганда
Дослідження задоволення наданням послуг на рівні провінцій та національному рівні	Індія, Філіппіни, Уганда
Секторні дослідження	Державні лікарні в Бангалорі (Індія)
Оцінювання програм	Сільська продовольча безпека в Таміл-Наду, водопостачання й санітарія в сільській місцевості в Махараштрі (Індія)
Проекти реформування врядування	Бангладеш, Перу, Філіппіни, Шрі-Ланка, Україна, В'єтнам

Джерело: World Bank 2006.

Таблиця 8.13. Переваги й недоліки використання карток оцінки для громадян щодо збору даних

Переваги	• Отримання сумативного зворотного зв'язку щодо ефективності роботи
	• Структура карток забезпечує простоту комунікації
	• Картки сприяють зниженню кількості викривлень під час збору даних
Недоліки	• Для того, щоб зрозуміти цей вид оцінки, потрібна кваліфікація
	• Застосування карток обмежене порівнянням якості послуг
	• Потреба у великій вибірці різнорідної аудиторії для оцінки послуг, якими користується лише малий відсоток населення
	• Бракує прогнозованості реакції державних посадовців

Джерело: Автори.

Блок 8.10 «Рекомендації щодо проведення інтерв'ю»

Наступні рекомендації можуть стати в нагоді фахівцям з оцінювання, що проводять інтерв'ю:

1. Повідомляйте респондентам:
   - про мету і час проведення дослідження;
   - чому з ними проводять це інтерв'ю;
   - як їх було обрано;
   - як будуть використані дані;
   - чи є інтерв'ю конфіденційним;
   - скільки часу займе інтерв'ю;
   - про що йтиме мова (попередньо поінформуйте);
   - чи хочете ви поспілкуватися з ними ще раз;
   - чи отримають вони копію заключного звіту;
   - чи надасте ви їм стислий варіант своїх нотаток, якщо вони попросять.
2. Намагайтесь обирати час і місце, що дозволять провести інтерв'ю в спокійній обстановці та без відволікаючих факторів.
3. В ідеалі – залучіть помічника, що вестиме записи.
4. Розгляньте можливість запису інтерв'ю на диктофон. Якщо ви хочете скористатися диктофоном, попередньо отримайте дозвіл від свого співрозмовника.
5. Дотримуйтеся плану інтерв'ю. Якщо ви ставите закриті питання, озвучуйте їх точно так, як написано. Якщо ставите відкриті питання, «йдіть за течією», не намагайтеся її направляти.
6. Враховуйте культурні норми (чи припустимо дивитися людині в очі, ставити прямі запитання) і гендерні особливості.
7. Підтримуйте рівновагу: Якщо запитуєте респондентів про їхні думки, наприклад, щодо основних сприятливих факторів для проекту, спитайте також, що вони вважають основними перешкодами для проекту.
8. Намагайтеся уникати питань «чому?», якщо вони можуть бути сприйняті як агресивні або критичні.
9. Сприймайте все, що говорить респондент, зі співчуттям і без осудження.
10. Ведіть нотатки уважно, але не відволікаючись від бесіди.
    - Записуючи щось, підтримуйте зоровий контакт.
    - Записуйте ключові слова або фрази, а не дослівний виклад інтерв'ю.
    - Якщо респондент говорить щось, що б ви хотіли записати, попросіть його повторити цю фразу або допишіть її, перш ніж ставити наступне питання.
    - Якщо респондент говорить щось, що можна було б процитувати, спитайте, чи можна буде використати його слова.
11. Викладаючи інтерв'ю в письмовому вигляді:

- Пам'ятайте, що кожне слово й кожна думка можуть виявитися цінними.
- Розпишіть інтерв'ю на основі своїх нотаток якомога детальніше й точніше.
- Упорядкуйте (хоча б коротко) свої записи одразу після інтерв'ю (відповідно, резервуйте достатньо часу між інтерв'ю).
- Розпишіть інтерв'ю повністю не пізніше наступного дня.

Якщо мова йде про делікатні питання, можна показати респонденту записи, щоб перевірити їх точність і погодити остаточний варіант.

Джерело: Автори.

Кожній культурі властиві свої цінності й звичаї: питання або жест можуть бути образливими для представників однієї культури і нормальними для представників іншої. Тому культурна компетентність є обов'язковою вимогою для інтерв'юера, який має уважно стежити за реакціями респондента під час інтерв'ю, аби випадково його не образити (блок 8.11).

Під час роботи з представниками будь-яких культур інтерв'юери повинні пам'ятати наступне:

- Кожній людині подобається, якщо до неї ставляться з повагою.
- Навіть представники культур, відомих готовністю до самопожертви й схильністю до колективізму, усвідомлюють значення своєї особистості і хочуть, щоб до них ставилися відповідно.
- Кожній людині подобається відчувати, що її думка важлива для вас.

Блок 8.11. «Рекомендації щодо проведення інтерв'ю в міжкультурному контексті»

Наступні рекомендації можуть стати у нагоді фахівцям з оцінювання, які проводять інтерв'ю в країнах, що розвиваються:

1. Перед кожним інтерв'ю дізнайтеся більше про культуру респондента.
2. Обговоріть методику ведення інтерв'ю з кимось, добре обізнаним з цією культурою, або залучіть інтерв'юера з того ж культурного середовища.
3. Дізнайтеся якомога більше про наступні культурні характеристики:
   - комфортна відстань, якої інтерв'юер має дотримуватися щодо респондента;
   - належний рівень зорового контакту;
   - значення інтонацій, рухів головою та інших елементів мови тіла протягом бесіди;
   - відповідний одяг;
   - правильне вітання;
   - гендерні ролі.

Питання гендеру – особливо важливі. Наприклад, в деяких культурах неприпустимим є перебування чоловіка-інтерв'юера наодинці з жінкою-респондентом. Чоловіки з певних культурних середовищ можуть відмовитися розмовляти з інтерв'юером жіночої статі. В подібних ситуаціях може виявитися необхідним залучення інтерв'юерів тієї ж статі або присутність свідка іншої статі.

У табл. 8.14 стисло викладено переваги й недоліки інтерв'ю.

Таблиця 8.14. Переваги й недоліки інтерв'ю

Переваги	• Інтерв'ю можуть бути структуровані, напівструктуровані, неструктуровані або ж поєднувати характеристики цих типів. Звичайно вони напівструктуровані
	• Є можливість поглибленого розгляду складних питань
	• Помилки легко виправити: незрозумілі питання можна роз'яснити в ході інтерв'ю і надалі використовувати вже відредагований варіант
	• Можуть тривати годину або більше, залежно від важливості й зацікавленості
	• Завдяки інтерв'ю фахівець з оцінювання може сформувати для себе інтуїтивне відчуття ситуації
Недоліки	• Інтерв'ю можуть бути витратними, трудомісткими та забирати багато часу
	• Можуть не надати пояснень, чому різні люди мають різні точки зору
	• В результаті вибіркового слухання інтерв'юер може пропустити інформацію, що не відповідає його сталим переконанням
	• Інтерв'ю вимагають уваги до культурного контексту, зокрема щодо гендерних питань

Джерело: Автори.

Рівень відповідей

Рівень відповідей: Відсоток осіб, які беруть участь в опитуванні, від загальної кількості людей, до яких звернулися фахівці з оцінювання

Рівень відповідей (Response rate) – це відсоток осіб, які фактично беруть участь в опитуванні, від загальної кількості людей, до яких звернулися фахівці з оцінювання. У добре проведеному оцінюванні має бути вказана кількість осіб (або одиниць-респондентів – наприклад організацій), до яких звернулись з проханням взяти участь в опитуванні; кількість респондентів; відсоток фактичних респондентів; заходи, вжиті для збільшення відсотка фактичних респондентів (телефонні дзвінки, листи).

Низький рівень відповідей може свідчити про необ'єктивну вибірку (сформовану самими респондентами). Це проблема, адже люди, які прийняли рішення взяти участь у дослідженні, можуть сильно відрізнятися від тих, хто вирішив участі не брати. Уявімо, що організація проводить дослідження поглядів своїх працівників, але з усієї їх кількості відповідь надають лише 30 відсотків. Якщо найбільш незадоволені працівники, як правило, заповнювали анкети, а задоволені цього не робили, дані не відображатимуть погляди всіх працівників. Помилкою було б приймати рішення на основі цих результатів, не зібравши додаткової інформації про те, хто взяв участь в опитуванні, а хто – ні. Не слід звітувати про опитування, для яких відсоток фактичних респондентів становить менше 70%, доки фахівець з оцінювання не проаналізує викривлення відповідей та не доведе, що опитування є валідним (Jackson and Agrawal 2007).

Фахівець з оцінювання може розглянути демографічні характеристики респондентів, щоб встановити, чи відповідають вони показникам аудиторії у цілому в розрізі, наприклад, віку й статі. Якщо відповідність спостерігається, фахівець з оцінювання може переходити до наступного етапу, але обов'язково має згадати про цей фактор у звіті. В деяких випадках фахівець з оцінювання може одразу поспілкуватися з декількома особами, які відмовилися взяти участь в опитуванні, щоб спробувати виявити якісь закономірності. Якщо демографічні характеристики респондентів та нереспондентів відрізняються, результати опитування повинні це враховувати. Фахівці з оцінювання мають повідомляти про результати аналізу нереспондентів та їх вплив на висновки.

Результати опитування з низьким відсотком фактичних респондентів обов'язково мають містити згадку про кількість респондентів та визнання того, що ці результати можуть неточно відображати стан речей для групи в цілому. Таким чином, дані повинні бути вказані у форматі: «87 опитаних» або «75 відсотків (N = 60) респондентів вказали ...». Відсоткові показники не слід використовувати, якщо анкети загалом було розіслано менш ніж 50 особам. Показник «75 відсотків респондентів», якщо його застосувати до 45 осіб – зовсім не те саме, що 75 відсотків із 3000 опитаних.

Інструмент 5: Фокус-групи

Фокус-група: Методика якісного оцінювання: малі групи людей збираються і разом обговорюють конкретні теми під наглядом модератора

Фокус-група – це методика якісного оцінювання: малі групи людей збираються і разом обговорюють конкретні теми під наглядом модератора. Структура фокус-групи може здаватися неформальною, але насправді під час роботи з нею використовується сценарій, що містить набір відкритих питань. Підготовлений модератор веде обговорення й ставить зондуючі або додаткові питання залежно від ситуації.

Груповий процес зазвичай дає більше інформації, ніж індивідуальні інтерв'ю, тому що люди висловлюють різні погляди й ведуть діалог один з одним. Але фокус-група – це не просто групове інтерв'ю. Модератор підтримує діалог і питає в членів групи про їхні думки й відчуття. Розмова часто ведеться нелінійно; учасники можуть повсякчас згадувати нову інформацію або висловлювати різні точки зору.

Фокус-групи використовуються для збору даних щодо наступного:

- групова взаємодія;
- складність ресурсів;
- «як» і «чому», на відміну від «так/ні» та «скільки»;
- контекстуальні відповіді, а не позитивні або негативні;
- безпосередній зворотний зв'язок;
- складність поведінки та мотивів;

- інтенсивність висловлювання думок;
- погляди на делікатні питання (Billson and London 2004).

Фокус-група може бути використана для тріангуляції методів. Вона також буде корисною, якщо респондентам незручно працювати з паперами.

Не варто використовувати фокус-групи для збору статистичних даних. Вони також не підходять, якщо:

- постають непереборні мовні бар'єри;
- фахівець з оцінювання погано контролює групу;
- не вдається встановити довіру;
- не вдається забезпечити свободу висловлювань;
- конфіденційність є критичною (Billson 2004).

Фокус-група зазвичай складається з 6–12 осіб. Склад групи залежить від її призначення. Більшість фокус-груп однорідні, складаються з людей зі схожим статусом (викладачі в одній групі, студенти – в іншій; керівники – в одній групі, підлеглі – в іншій). Те, хто саме має увійти до певної фокус-групи, а хто не повинен входити до її складу, залежить від ситуації й культурного контексту. Підбір учасників не повинен жодним чином вплинути на достовірність результатів. Не можна пропустити жодну з категорій потенційних учасників, якщо їхні відповіді можуть змінити або розширити результати. За наявності потенційних додаткових впливових груп необхідно зібрати ще одну фокус-групу.

Фокус-групи мають забезпечувати учасникам:

- зручне й безпечне місце для зустрічі (обов'язково);
- закуски та напої (обов'язково);
- грошові заохочення (не обов'язково);
- організацію транспортування та догляду за дітьми (не обов'язково);
- вмілого модератора (обов'язково);
- помічника для ведення записів (не обов'язково).

Сесії фокус-груп записуються на диктофон. В ідеальному варіанті фахівці з оцінювання виконують повну розшифровку кожного запису. Сесія фокус-групи починається з чіткого пояснення:

- призначення фокус-групи;
- чому погляди учасників важливі;
- як було обрано учасників;
- що таке фокус-група;
- які правила діятимуть в процесі, в т. ч. головне правило: що було сказано, залишається в цій кімнаті.

Модератор керує процесом, забезпечує зосередженість групи на темі та слідкує за тим, щоб всі учасники мали можливість висловити свої погляди, а також щоб розмови не забивалися лише кількома особами. Модератор ставить кілька питань, виконуючи спеціально написану інструкцію. Всі питання – відкриті. Розмова, яка спочатку точиться навколо простих, легких питань, переходить до питань більш серйозних. Все закінчується підсумковими й завершальними питаннями, що дозволяють в разі необхідності скорегувати враження та записати додаткові коментарі й думки.

Фахівці з оцінювання відповідають за всі аспекти проведення фокус-групи, в тому числі за те:

- хто в ній має бути;
- скільки сесій має бути проведено;
- де проводяться сесії;
- коли проводяться сесії;
- про що запитувати учасників;
- як аналізувати й представляти дані.

Типовий проект фокус-групи вимагає від шести до восьми тижнів підготовки. Тривалість процесу залежить від факторів логістики, складності та терміновості проекту, доступності осіб, які приймають рішення, а також складності набору учасників.

Оцінювання з використанням фокус-групи дає найкращі результати, якщо виконувати її наступним чином (Billson 2004):

- Крок 1: Дайте пояснення ключових питань оцінки. Якщо клієнти й фахівці з оцінювання не мають чіткого уявлення про ключові питання, на які має відповісти фокус-група, весь процес буде марним.
- Крок 2: Сплануйте підхід оцінювання. Виходячи з мети й ключових питань оцінювання, визначте загальний підхід та порядок тем для відповіді на ключові питання.
- Крок 3: Складіть інструкцію для модератора, яка повинна описувати протокол (структуру або правила поведінки) для проведення оцінювання таким чином, щоб не вплинути на відповіді учасників, але підштовхнути групу до обговорення ключових питань.
- Крок 4: Наберіть учасників.
- Крок 5: Визначте методи модерації. Виберіть та застосуйте ефективні методи модерації для сесій фокус-групи.
- Крок 6: Вислухайте звіти спостерігачів, фахівців з оцінювання і клієнтів та зафіксуйте додаткову інформацію. Одразу після кожної фокус-групи, поки дані ще свіжі, обговоріть свої враження від сесії з клієнтами та іншими зацікавленими сторонами. Зафіксуйте додаткову інформацію, що

не стала предметом відкритого обговорення (враження, висновки й т. ін.), для використання під час наступного кроку.

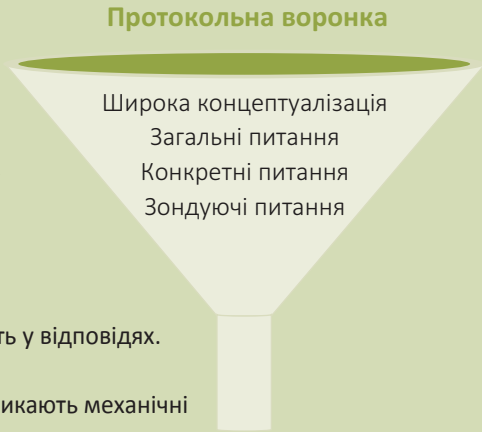
- Крок 7: Проаналізуйте дані. Якщо фокус-група спрацювала добре, ви отримаєте дуже багато ідей. Ці якісні дані вимагають застосування особливих методів аналізу, зокрема аналіз змісту (див. главу 10).
- Крок 8: Представте висновки. Викладіть висновки в осмислений і зрозумілий для інших (зокрема, для клієнта) спосіб. Використовуйте усний, письмовий або відеоформат, або ж комбінацію будь-яких з цих форматів.

Сесії фокус-груп зазвичай тривають від однієї до двох годин. У більшості випадків ста хвилин достатньо, щоб обговорити кожен тематичний напрям, не втомлюючи при цьому учасників і модератора та не викликаючи в них нудьги (Billson 2004). Графік проведення фокус-груп не повинен бути занадто насиченим. Модератор не витримає більш ніж дві або три сесії на день.

Фокус-групи потрібно збирати, за можливості, в нейтральній обстановці. Окрім того, місце має бути легкодоступним. Зазвичай учасники сидять за столом або на стільцях, виставлених у коло. Це сприяє їх спілкуванню. В кімнаті під час обговорення можуть бути присутні кілька спостерігачів. Їхньої необхідно представити на початку сесії, пояснивши причину їх присутності.

Блок 8.12 «Рекомендації щодо формування питань для фокус-груп»

Наступні поради допоможуть фахівцям з оцінювання отримати максимум користі від сесій з фокус-групами:



- Спочатку подумайте про те, як буде використана інформація, отримана від фокус-групи. Переконайтеся, що всі питання прямо чи опосередковано ведуть до корисної інформації.
- Уникайте нечітких, розпливчатих формулювань.
- Ставте одне питання за один раз.
- Ставте відкриті питання.
- Уникайте припущень, що можуть привести вас до неправильних висновків.
- Уникайте питань, що провокують необ’єктивність у відповідях.
- Не пропонуйте альтернативні відповіді.
- Робіть сесії цікавими. Механічні запитання викликають механічні відповіді.
- Використовуйте «протокольну воронку», що починається з питань про широкі концептуальні речі й завершується зондуючими питаннями.

Джерело: Billson 2004: 29.

Для якомога точнішої фіксації подій під час сесії дуже корисними є засоби аудіо- та відеозапису. Більшість учасників швидко забувають, що їх записують. Якщо запис із застосуванням технічних засобів не ведеться, двоє спостерігачів мають вести нотатки. Ведення нотаток за допомогою ноутбука може прискорити процес початкового аналізу даних та формування звіту. Модератор не повинен робити нотатки.

Таблиця 8.15. Чотири етапи фокус-групи та дії модератора

Етап	Дії
I. Преамбула або вступна заява	Модератор знімає напруження серед учасників, ставлячи їм прості початкові питання.
	Модератор пояснює призначення фокус-групи.
	Модератор розповідає про базові правила.
	Кожен з присутніх представляється.
II. Представлення й «розігрів»	Учасники діляться своїм досвідом і розказують про свій зв'язок з темою.
	Модератор стимулює роботу групи, спонукаючи учасників думати на обрану тему.
	Модератор починає з найлегших, найменш приватних питань.
III. Основна частина групового обговорення	Модератор переходить до складніших та делікатніших питань.
	Модератор витягує з групи поглиблені відповіді.
	Отримувані дані поєднуються, створюючи складну, інтегровану базу для аналізу.
	До обговорення залучаються всі учасники.
IV. Завершення	Модератор ставить завершальні питання.
	Модератор підбиває підсумки та уточнює ключові теми.
	Модератор представляє теорії, враження й передчуття членам групи, щоб вони висловили свою реакцію, підтвердили, роз'яснили або виправили почуте.
	Модератор запрошує учасників висловити ще один круг коментарів або здогадів («ключові висновки»), а також «останні думки», те, що не було висловлено раніше, та будь-які зауваження, які учасники прагнуть донести до групи оцінювання.
	Модератор дякує учасникам і винагороджує їх (якщо було обіцяно якесь матеріальне заохочення).

Джерело: Автори.

Немає однозначного правила стосовно того, скільки сесій фокус-групи треба проводити. Загальний принцип полягає у тому, що їх треба проводити до тих пір, поки не почнуть спливати одні й ті самі теми, або поки не перестане надходити нова інформація. Це зазвичай трапляється після трьох-шести сесій. Може бути корисно поставити набір основних питань кожній фокус-групі, а потім додавати різні питання чи натяки, якщо чітко визначиться коло тем. Фахівці з оцінювання мають вислуховувати один одного після кожної сесії, щоб можна було приймати рішення про внесення коректив у протокол наступних сесій.

Необхідно брати до уваги повсякденні справи учасників: фахівці з оцінювання повинні пам'ятати про час, який вони забирають (в тому числі час на поїздки); участь в одній сесії може зайняти півдня або й більше. Фокус-група, запланована на другу половину дня, ближче до вечора, може завадити приготуванню вечері, а це призведе до зменшення кількості людей, готових взяти участь в дослідженні.

Фокус-групи вимагають серйозної підготовки, що займає немало часу. В блоці 8.12 наведено рекомендації щодо написання питань для фокус-груп.

Сесію фокус-групи можна розбити на чотири окремих етапи. Ці етапи, а також кроки, що їм відповідають, наведено в табл. 8.15.

Питання для фокус-групи з блоку 8.13 мають ілюструвати види питань, які фахівець з оцінювання захоче поставити за допомогою цього інструмента.

Блок 8.13 «Приклад питань для фокус-групового обговорення»

Наступний сценарій може бути корисним для моделювання питань для фокус-групи.

Представлення: Давайте розпочнемо з того, що ви відрекомендуєтеся. Будь ласка, назвіть свою посаду й місце роботи, а також згадайте про свій досвід отримання послуг цієї лікарняної програми – в тому числі причину, чому ви до неї звернулися.

Про враження: Критики безкоштовних лікарняних програм говорять про низьку якість медичних послуг. Що ви можете розповісти зі свого досвіду отримання послуг у безкоштовних лікарнях?

Про врядування: Останні опитування показують, що поточні районні бюджети не доходять до місцевих лікарень. Як би ви оцінили цю проблему, виходячи з власного досвіду отримання послуг в лікарнях?

Джерело: Взято з Billson 2004.

Таблиця 8.16. Переваги й недоліки фокус-груп

Переваги	Фокус-групу можна провести швидко й легко.
	Потрібно менше персоналу, ніж для поглиблених індивідуальних інтерв'ю.
	Є можливість внесення змін в процедуру й питання.
	Можна вивчати різні точки зору.
	Проведення фокус-групи може бути цікавим для учасників.
Недоліки	Аналіз займає багато часу.
	Учасники можуть не бути репрезентативними представниками аудиторії, що, в свою чергу, може вплинути на достовірність даних.
	Група може зазнати впливу з боку модератора або домінуючих учасників.

Джерело: Автори.

У табл. 8.16 в стислому вигляді вказано переваги й недоліки фокус-груп.

Інструмент 6: Щоденники та опитувальні листи для самостійного заповнення

Інший метод збору даних полягає у використанні щоденників або опитувальних листів для самостійного заповнення.

Щоденники або журнали

Щоденник: Письмовий звіт, який заповнюється учасником, як правило, щодня

Щоденник (або журнал) – письмовий звіт, який заповнюється учасником, як правило, щодня. Щоденники використовуються для збору детальної інформації про події з повсякденного життя людей. Вони також можуть використовуватися для вивчення таких явищ, як соціальні мережі; здоров'я, захворювання та пов'язана з ними поведінка; дієта й харчування; сільськогосподарська робота; звички у сфері навчання; використання контрацептивів; виховання дітей. На додачу до збору даних про діяльність, що викликає зацікавленість, протягом певного часу, щоденники також можуть характеризувати таку діяльність. Наприклад, курці, які беруть участь у програмі відмови від паління, можуть вести щоденники, щоб записувати, коли в них виникає сильне бажання палити, де вони при цьому знаходяться та що роблять. Ідея полягає в наступному: якщо будуть виявлені певні закономірності, учасники зможуть їм протидіяти. Якщо бажання палити виникає одразу після їжі, закономірність можна подолати, швидко вставши з-за столу.

Щоденники використовуються як на загальнонаціональному рівні, так і в малих програмах (Stafford 2006). Це корисні інструменти для доповнення даних, отриманих з інших джерел (табл. 8.17). Ведення щоденника вимагає грамотності й бажання присвячувати цьому час. Однак в деяких випадках вони можуть бути представлені в форматі опитувального листа, що дозволяє виключити потребу в грамотності. Наприклад, піктограмний щоденник використовувався для збору даних про споживання й витрати домашніх господарств в Танзанії та Гамбії (Wiseman, Conteh, and Matovu 2005).

Протягом якого періоду має заповнюватися опитувальний лист? Для отримання достатньої вибірки поведінки потрібно зазвичай не менше тижня, при цьому період більше місяця вже буде задовгим. Після цього ведення щоденника стає обтяжливим, якість його заповнення погіршується.

Зазвичай учасникам видають буклети, що містять чіткий набір вказівок, а також приклад запису в щоденнику. Наводяться пояснення всіх незвичних термінів. На останній сторінці буклета розміщується запитання про те, охоплює щоденник типовий чи нетиповий період. Учасники також можуть додати від себе будь-які коментарі, що не увійшли до щоденних записів.

Таблиця 8.17. Рекомендації щодо використання щоденників або журналів

Дія	Подробиці
Особиста зустріч з метою залучення потенційних учасників	Треба заохочувати потенційних респондентів до участі, залучаючи до роботи з ними мотивованих інтерв'юерів.
	Апелюйте до альтруїзму й прагнення допомагати людям.
	Забезпечте конфіденційність.
	Забезпечте винагороду за заповнення щоденника.
Видача буклета кожному учаснику	Розпочніть з чітких вказівок щодо того, як потрібно заповнювати щоденник.
	Надайте приклад запису в щоденнику.
	Додайте короткі пам'ятки.
	Поясніть значення всіх термінів, таких як «подія» або «сесія».
	На останній сторінці буклета запитайте, був це типовий чи нетиповий період, залиште місце для коментарів та пояснень.
	Буклет повинен містити календар, де буде вказано, коли має бути зроблений запис.
Прийняття рішення щодо періоду збору даних	Якщо період задовгий, ведення щоденника може стати обтяжливим для учасників.
	Якщо ж період закороткий, ви можете не отримати дані про поведінку або подію.

Джерело: Автори.

Опитувальний лист для самостійного заповнення

Опитувальний лист для самостійного заповнення: Перелік заходів або подій, які респонденти відзначають у відповідних випадках

Опитувальний лист для самостійного заповнення – це поєднання опитувальника й щоденника. Учасників просять відстежувати певний набір вчинків або подій, список яких наводиться, щоб респонденти могли легко їх відзначати. Опитувальні листи можуть заповнюватися щодня, щотижня або кожного разу, коли має місце певна подія чи вчинок. Опитувальний лист можна використовувати, приміром, для відстеження прийому ліків від малярії або кількості й часу подорожей до джерела води.

Опитувальні листи для самостійного заповнення використовувалися в дослідженні акушерок в Індонезії (Abdallah 2002). Той проект базувався на теорії про те, що для зміни поведінки після тренінгу потрібно додаткове підкріплення навичок. Метою було підвищення якості консультацій, що їх акушерки надавали щодо планування сім'ї. Учасниці заповнювали анкету самоконтролю раз на тиждень, зосереджуючись на певних навичках міжособистісного спілкування, щоб краще зрозуміти, як вони спілкуються, проаналізувати власний досвід і покращити свої навички.

Опитувальний лист заповнювати й аналізувати легше, ніж щоденник. Однак розробка опитувального листка вимагає від фахівця з оцінювання глибокого розуміння ситуації.

Таблиця 8.18. Переваги й недоліки щоденників та опитувальних листів для самостійного заповнення

Переваги	Цей інструмент дає змогу отримувати поглиблені детальні дані («багаті дані»), які в іншому випадку будуть забуті.
	Є можливість зібрати дані про те, як люди використовують свій час.
	Можна збирати делікатну інформацію.
	Даний інструмент є допоміжним для інтерв'ю, надаючи багатші дані.
Недоліки	Вимагає певного рівня грамотності.
	Може викликати зміну поведінки через те, що люди знають, що за їхньою поведінкою ведеться спостереження.
	Вимагає від учасника цілеспрямованості та самодисципліни, точного й чесного запису інформації; учасник може недостатньо точно згадувати поведінку.
	Записані дані можуть бути неточними, або ж учасник може не одразу записати інформацію і неправильно згадати події.
	Проблема може виникнути через нерозбірливий почерк учасника.
	Фрази в щоденниках може бути важко зрозуміти.

Джерело: Автори.

У табл. 8.18 стисло викладено переваги й недоліки щоденників та опитувальних листків для самостійного заповнення.

Інструмент 7: Експертні судження

Іноді в якості джерела інформації або думки варто залучати експертів. Їхня роль нагадує роль літературного або кінокритика. Люди використовують свої судження для інформованого прийняття власних рішень або вибору. Урядові групи фахівців – різновид джерела експертних суджень.

Експертні судження:
Думка окремого експерта або експертної групи

Експертні судження можуть надаватися окремим експертом або експертною групою. Експертна група може виконувати перевірку лікарні в рамках процесу акредитації. Ця група відвідує лікарню для вивчення її програм, обладнання й персоналу. Окремий експерт (наприклад, шкільний інспектор) може відвідувати школи з метою оцінки шкіл, їх адміністрації та вчителів.

Експертні судження можуть базуватися на визначених професійних критеріях та процедурах або ж виноситися з урахуванням конкретного випадку. Якщо експертні судження ґрунтуються на визначених критеріях, ці критерії зазвичай являють собою опубліковані стандарти й інструменти, які гарантують, що експерти ставитимуть питання й проводитимуть оцінку в єдиному ключі. Наприклад, організація, що надає фінансування, може захотіти оцінити якість навчальної програми, для якої був затверджений план навчання з цілями, завданнями й довгостроковими результатами. Добре обізнаний з навчальним планом, цілями, завданнями й довгостроковими результатами експерт розглядає навчальну програму й проводить оцінку, орієнтуючись на документацію.

Ситуативні перевірки проводяться нерегулярно, залежно від обставин. Їх зазвичай проводять з конкретною метою. Такі перевірки звичайно не базуються на якихось визначених стандартах; експерти самостійно визначають стандарти для винесення суджень. Приклад ситуативної перевірки – відвідування директором програми водопідготовки місця, де ця програма реалізується, щоб отримати уявлення про хід її виконання.

При виборі експертів необхідно користуватися «тестом поміркованої особи»: чи вважатиме поміркована особа цього експерта або цю експертну групу гідними довіри? Експерти повинні відображати широкий спектр поглядів, досвіду та ролей.

Важливо встановити критерії для вибору експертів на основі не лише визнання людини як експерта, а й з урахуванням наступних параметрів:

- області компетенції;
- різні точки зору;
- різні політичні погляди;
- різноманітна технічна кваліфікація.

Критерії відбору всіх експертів мають зазначатися в звіті про оцінювання.

Міжнародна рада з безпеки й розвитку (International Council on Security and Development 2006) залучила експертну групу для перевірки ефективності роботи Міжнародного комітету з контролю за наркотиками. В даному випадку було особливо важливо не лише залучити визнаних експертів, а й переконатися, що вони представляють різні частини світу, зокрема ті, де проблема наркотиків є особливо гострою.

І хоча використання групи експертів не вважається ефективним підходом до виконання оцінювання, в деяких випадках це може бути найкращий варіант з урахуванням обмежень у часі й ресурсах. Групи експертів більше підходять для оцінки схеми або ранніх чи проміжних етапів реалізації програми чи проекту, ніж для оцінювання впливу. Цей підхід особливо ефективний для проведення швидкої оцінки.

Таблиця 8.19. Переваги й недоліки експертних суджень

Переваги	Швидкість і відносно низький рівень витрат.
Недоліки	Слабкість в оцінюванні впливу.
	Оцінка може ґрунтуватися переважно на суб'єктивних враженнях.
	Цінність зібраних даних залежить від того, наскільки гідною довіри видається група експертів.

Джерело: Автори.

У табл. 8.19 в стислому вигляді вказано переваги й недоліки використання експертних суджень.

Інструмент 8: Метод Дельфі

Метод Дельфі: Метод, що передбачає отримання інформації та суджень від учасників задля сприяння вирішенню проблем, плануванню та прийняттю рішень без особистої зустрічі учасників

Метод Дельфі дозволяє експертам, які мешкають у різних місцях, вести діалог і знаходити консенсус за допомогою ітеративної процедури. Експертам ставлять конкретні запитання; їх відповіді надсилаються до центрального вузла, підсумовуються та знову відправляються експертам. Тоді експерти коментують отримані підсумки. Вони можуть піддавати сумнівам конкретні точки зору або додавати нові аспекти розгляду проблеми, надаючи додаткову інформацію. Завдяки тому, що нікому не відомо, хто що сказав, вдається уникнути конфліктів.

Метод Дельфі передбачає отримання інформації та суджень від учасників задля сприяння вирішенню проблем, плануванню та прийняттю рішень без особистої участі учасників (Dunham 1996). Експерти обмінюються інформацією за допомогою пошти, факсів або електронної пошти. Цей метод

допомагає групам знаходити консенсус з питань, стратегій та пріоритетів. Він застосовувався для прогнозування туристичного потенціалу Південної Африки (Kaynak, Bloom, and Leibold 1994), знаходження консенсусу стосовно пріоритетів національної політики боротьби з наркотиками країн, що розвиваються (Rainhorn, Brudon-Jakobowicz, and Reic 1994), а також встановлення пріоритетів у боротьбі з шистосоміазом в Кенії (Kirigia 1997).

Метод Дельфі вимагає наявності координатора, який визначає експертів-учасників, організує запити інформації, підсумовує отриману інформацію та підтримує зв'язок з учасниками. Робота координатора може потребувати багато часу. Координування 20 учасників методу Дельфі засобами електронної пошти протягом трьох раундів обробки анкет може становити 30–40 годин роботи координатора.

Процедура методу Дельфі складається з наступних кроків:

1. Визначення й запрошення експертів до участі. Необхідно визначити критерії, яким мають відповідати експерти (як правило, група експертів з різними сферами компетенції). Потім треба запросити експертів до участі в програмі.
2. Визначайте проблему та просіть надавати ідеї. Наприклад, можна спитати: «Що можна зробити для підвищення швидкості реакції на запити пацієнтів між візитами?» Підготуйте й надішліть першу анкету, що спонукатиме учасника виконати індивідуальний брейн-стормінг і запропонувати якомога більше ідей для роботи з проблемою.
3. Учасники заповнюють перший опитувальник. Детальне викладення ідей не потрібне. Навпаки, краще, якщо кожна ідея буде викладена в одному реченні, без спроби оцінки або обґрунтування. Потім учасники повертають анкети координатору.
4. Створіть і розішліть другий опитувальник. Координатор готує і розсилає учасникам другий опитувальник, який містить всі ідеї, надіслані у відповідь на перший. (Тільки координатор знає, який експерт запропонував ті чи інші ідеї.) Учасники отримують можливість вдосконалити кожну ідею, висловитися стосовно сильних та слабких її сторін, а також запропонувати нові ідеї.
5. Зберіть відповіді на другий опитувальник. Учасники анонімно пишуть свої відповіді на другий опитувальник і повертають його координатору.
6. Створіть і розішліть третій опитувальник. Координатор створює й розсилає третій опитувальник, в якому підсумовуються пропозиції з попереднього кроку, та просить додаткових роз'яснень, вказівок щодо сильних і слабких сторін, а також пропозицій нових ідей.
7. Продовжуйте процес. За бажання координатор повторює ітерації цієї процедури до тих пір, поки не перестануть виникати нові ідеї і не будуть визначені всі сильні й слабкі сторони, озвучені всі думки.

8. Прийміть рішення. Існує два варіанти прийняття рішення:
   - Якщо експерти знаходять консенсус стосовно основних ідей, які оцінювалися, процедура завершується. Кінцевим її продуктом буде список ідей із зазначенням їхніх сильних та слабких сторін.
   - За відсутності консенсусу координатор проводить формальну оцінку думок групи стосовно переваг ідей. Існує чимало способів проведення формальної оцінки. Один підхід полягає в тому, що координатор готує анкету з переліком всіх ідей і пропонує учасникам оцінити кожну з них. При цьому можна застосувати семибальну шкалу від 0 (потенціал для усунення проблеми відсутній) до 7 (дуже високий потенціал усунення проблеми). В разі використання цього підходу учасники надсилають форми оцінки координатору, який об'єднує результати й розставляє ідеї за рангом на основі отриманих оцінок. Другий підхід полягає в тому, що фахівець з оцінювання просить кожного учасника визначити п'ять найкращих ідей і дати п'ять балів найбільш перспективній з них, чотири – наступній за перспективністю і т. д. Результати голосування повертаються координатору, який підсумовує їх і готує звіт. Звіт містить список ідей, впорядкованих на основі загальної кількості набраних ними балів. Також вказується, скільки людей проголосувало за кожну з ідей (Dunham 1996).

Таблиця 8.20. Переваги й недоліки методу Дельфі

Переваги	Метод дозволяє учасникам зберігати анонімність.
	Не потребує великих витрат.
	Відсутні соціальний тиск, персональний вплив та особистісне домінування.
	Метод сприяє незалежному мисленню та поступовому формулюванню ідей.
	Учасники мають змогу обмінюватися інформацією й аргументами.
Недоліки	Методу може бракувати репрезентативності.
	Має тенденцію до виключення крайніх точок зору та просування поміркованих рішень.
	Вимагає навичок письмового спілкування.
	Вимагає часу й цілеспрямованості учасників.

Джерело: Michigan State University Extension 1994.

У табл. 8.20 в стислому вигляді вказано переваги й недоліки використання методу Дельфі.

Інструмент 9: Інші інструменти вимірювання

Під час роботи з деякими проектами або програмами фахівці з оцінювання стикатимуться з іншими інструментами вимірювання, наприклад з електронними, хімічними й механічними пристроями. Для збору даних для певних проектів та програм можуть використовуватися такі інструменти, як тести на папері та оцінка навичок. Нижче наведено приклади інших інструментів збору даних.

Багато проектів і програм (особливо в сфері охорони здоров'я) вимагають вимірювання фізіологічного стану для визначення наслідків інтервенцій. Для збору даних про вагу піддослідних використовують ваги. Зріст людини, довжину й обхват її кінцівок можна вимірювати за допомогою рулетки або іншого засобу. Для вимірювання часу, що знадобився для виконання завдання, використовують секундомір. В деяких випадках для отримання додаткових фізіологічних даних про вміст холестерину, артеріальний тиск, вміст цукру в крові, склад тканин тіла, об'єм легенів, аеробну витривалість, м'язову силу, гнучкість суглобів та інші параметри використовують спеціалізовані інструменти.

Для визначення якості води й ґрунту можуть використовуватися хімічні тести колодязної води й ґрунтів. Інші тести дозволяють вимірювати температуру плавлення, тиск пари та коефіцієнти біонакопичення для визначення фізичних та хімічних властивостей кожної конкретної хімічної речовини. Подібні тести допомагають визначати фактори небезпеки для здоров'я та довкілля (U.S. EPA 2007). Використовуючи пристрої для вимірювання, важливо перевіряти калібрування приладів та, в разі необхідності, вносити корективи. Проекти та програми, що стосуються харчування, можуть використовувати оцінки хімічних або поживних компонентів їжі.

Кваліфікаційні тести та тести досягнень можуть застосовуватися для оцінки знань і фактів. Вони нерідко зустрічаються у програмах навчання й освіти, але можуть бути корисними і в багатьох інших сферах.

Раніше багато тестів проводилися за допомогою паперу й олівця. Тепер їх можна проводити в електронному вигляді, в тому числі за допомогою мережі Інтернет.

Перед тим, як вибрати тест, важливо зрозуміти, який саме тест найкращим чином виявить те, про що хоче дізнатися фахівець з оцінювання. Існує чотири підходи до тестування досягнень:

- тестування за нормами;
- тестування за критеріями;
- тестування за цілями;
- тестування за галуззю (Fitzpatrick, Sanders, and Worthen 2004).

Тестування за нормами застосовується переважно для порівняння показників різних студентів, що проходять один і той самий тест. Такий вид тестування часто використовується для оцінки прогресу. Багато освітніх адміністрацій використовують тестування за нормами для порівняння успішності своїх учнів з учнями інших освітніх адміністрацій у розрізі загальноприйнятих знань і навичок. Слабкість цього підходу полягає в складності визначення «загальноприйнятих знань і навичок». Різні культури, суспільства та середовища можуть ставити перед викладачами різні цілі (Fitzpatrick, Sanders, and Worthen 2004).

Тестування за критеріями використовується для порівняння показників з абсолютним критерієм. Країна може прийняти критерії оцінки роботи шкіл та освітніх адміністрацій (а не окремих учнів) і використовувати їх для порівняння. Критерії нерідко застосовуються для оцінки навчального плану або програми (Fitzpatrick, Sanders, and Worthen 2004).

Тестування за цілями базується на цілях та довгострокових результатах, що стоять перед освітніми та навчальними програмами. Об'єкти оцінки в таких тестах визначаються заданими цілями й довгостроковими результатами. Тестування за цілями використовується для отримання зворотного зв'язку формативної оцінки. Цей зв'язок має допомагати викладачам та тренерам визначати ті сфери, де програми проходять успішно, і ті, що потребують вдосконалення (Fitzpatrick, Sanders, and Worthen 2004).

Тестування за галуззю пов'язане з предметною областю, в якій проводиться вимірювання. Це може бути, наприклад, знання історії або математики. Розробка цих тестів обходиться дорого, але вони можуть запропонувати цінну інформацію, наприклад: «Наскільки добре наші учні знають географію світу?» або «Що нашим студентам відомо про гендер?» (Fitzpatrick, Sanders, and Worthen 2004).

Тестування за нормами та критеріями формує стандарти для суджень про успішність. Тестування за цілями та галуззю не дозволяє формувати стандарти. Воно надає описові дані про академічну успішність, не оцінюючи учнів або школи (Fitzpatrick, Sanders, and Worthen 2004).

В мережі Інтернет можна знайти багато стандартизованих тестів (див. посилання наприкінці цієї глави). Вони дають змогу оцінити особисті характеристики, професійну зацікавленість, досягнення, розвиток, інтелект, схильність до читання й уміння навчатися. Деяким зацікавленим особам може бути потрібно більше відомостей про інтервенцію, ніж просто бали учасників на стандартизованому тесті знань. Рівень навичок можна оцінювати за допомогою пристроїв моделювання, розробки портфоліо, усних презентацій та дебатів. Важливо підібрати відповідний тест для вміння, яке ви хочете оцінити.

Тести навичок можуть бути корисні під час оцінки інтервенції, спрямованої на формування нових навичок (практики санітарії для матерів, сільськогосподарські навички для фермерів, ведення документації для держслужбовців). Оцінки розвитку часто мають на меті встановити, чи впроваджене у практику нове вміння, але вони не здатні визначити, чи використовується воно успішно. Експерти, навчені спостерігати за роботою, вміють оцінювати практичні навички. Встановлюються критерії, за якими й оцінюється робота піддослідних. Для документування ефективності можуть використовуватися фото- та відеозйомка.

Підсумки

Немає єдиного найкращого способу збору даних. Вибір одного з багатьох описаних в цій главі методів залежить від потреб оцінювання та конкретних питань, відповіді на які треба знайти.

Під час оцінювання не слід обмежуватися лише одним методом збору даних. Рішення про вибір методу мають базуватися на тому:

- про що має дізнатися фахівець з оцінювання;
- де знаходяться дані;
- скільки ресурсів і часу є в розпорядженні фахівця з оцінювання;
- наскільки складні дані потрібно зібрати;
- якою є частота збору даних.

Завдання до Глави 8

Практична вправа 8.1: Збір інформації з документації

Вас просять за допомогою вступних документів визначити кваліфікацію й досвід слухачів семінару з садівництва. Розробіть коротку форму з п'яти питань, які можна застосувати для отримання даних зі вступних документів.

Практична вправа 8.2: Збір інформації за допомогою інтерв'ю

Вас попросили розробити питання для короткого інтерв'ю з метою оцінки вражень учасників щодо якості семінару, що проходив на конференції, яку ви недавно відвідали. Розробіть п'ять відкритих питань стосовно змісту, рівня й практичної реалізації семінару. Якщо можливо, знайдіть партнера, який відвідував той самий семінар або конференцію. Проведіть інтерв'ю зі своїм партнером та дозвольте йому провести інтерв'ю з вами, використовуючи створений вами інструмент. Потім напишіть поглиблений звіт за результатами інтерв'ю. Нехай ваш партнер оцінить його точність, легкість для читання та охоплення.

Практична вправа 8.3: Збір інформації за допомогою фокус-групового обговорення

Вас попросили сформувати фокус-групу для оцінки впливу низки семінарів та пакету фінансової допомоги, покликаних допомогти жінкам започаткувати й вести власний малий бізнес. Напишіть п'ять питань для фокус-групи, які можна було б поставити жінкам, що проходили програму шість місяців тому. Обов'язково дізнайтеся, які цілі програми було досягнуто та як вона вплинула на життя, друзів і родини учасниць.

Посилання та рекомендована література

Abdallah, H. 2002. *Cost-Effectiveness of Self-Assessment and Peer Review in Improving Family Planning Provider-Client Communication in Indonesia.* Quality Assurance Project Case Study. Study conducted for the U.S. Agency for International Development by the Quality Assurance Project, Bethesda, MD.

Academy for Educational Development, Population Communication Services. 2002. "Session 10 Transect Walks and Observation." *Empowering Communities: Participatory Techniques for Community-Based Programme Development,* vol. 2, *Participant Handbook.* Washington, DC: Academy for Educational Development. http://pcs.aed.org/manuals/cafs/handbook/sessions10-12.pdf.

ADB (Asian Development Bank), and ADBI (Asian Development Bank Institute). 2004. *Improving Local Governance and Service Delivery: Citizen Report Card Learning Tool Kit.*

http://www.citizenreportcard.com/index.html#.

BBC (British Broadcasting Corporation). 2008. "Data Validation and Verification." *GCSE Bitesize.* http://www.bbc.co.uk/schools/gcsebitesize/ict/databases/3datavalidationrev1.shtml.

Behling, Orlando, and Kenneth S. Law. 2000. *Translating Questionnaires and Other Research Instruments: Problems and Solutions.* Quantitative Applications in the Social Sciences, Series 07-133. Thousand Oaks, CA: Sage Publications.

Billson, Janet Mancini. 2002. *The Power of Focus Groups for Social and Policy Research.* Barrington, RI*:* Skywood Press.

———. 2004. *The Power of Focus Groups: A Training Manual for Social, Policy, and Market Research: Focus on International Development.* Barrington, RI: Skywood Press.

Billson, Janet, and N. T. London. 2004. "The Power of Focus Groups." International Program for Development Evaluation Training (IPDET) presentation, Ontario, July.

Burgess, Thomas F. 2001. *Guide to the Design of Questionnaires,* ed. 1.1. University of Leeds, United Kingdom. http://www.leeds.ac.uk/iss/documentation/top/top2/top2-5.html.

Cloutier, Dorothea, Bill Lilley, Devon Phillips, Bill Weber, and David Sanderson. 1987. *A Guide to Program Evaluation and Reporting.* University of Maine Cooperative Extension Service, Orono.

CNCSTE (Chinese National Centre for Science and Technology Evaluation), and IOB (Policy and Operations Evaluation Department). 2006. *Country-Led Joint Evaluation of the ORET/MILIEV Programme in China.* Amsterdam: Aksant Academic Publishers.

Cnossen, Christine. 1997. *Secondary Research.* Learning Paper 7, School of Public Administration and Law, Robert Gordon University, Aberdeen, United Kingdom.

Dawson, Susan, and Lenor Manderson. 1993. *Methods for Social Research in Disease: A Manual for the Use of Focus Groups.* International Nutrition Foundation for Developing Countries, Boston. http://www.unu.edu/Unupress/food2/UIN03E/uin03e00.htm.

Denzin, K. 1978. *The Research Act.* New York: McGraw-Hill. Dunham, Randall B. 1996. *The Delphi Technique.*

http://www.medsch.wisc.edu/adminmed/2002/orgbehav/delphi.pdf.

Early Childhood Research Institute on Measuring Growth and Development. 2008.

Early Communication Indicator (ECI). http://cehd.umn.edu/ceed/projects/ecri/.

EuropeAid Co-operation Office. 2005. *Evaluation Methods.*

ec.europa.eu/europeaid/evaluation/methodology/egeval/index_en.htm.

Fitzpatrick, Jody L., James R. Sanders, and Blaine R. Worthen. 2004. *Program Evaluation: Alternative Approaches and Practical Guidelines.* New York: Pearson.

Foddy, William. 1993. *Constructing Questions for Interviews and Questionnaires.*

New York: Cambridge University Press.

Fowler, Floyd J. Jr. 2002. *Applied Social Research Methods Series,* vol. 1 *Survey Research Methods.* 3rd. ed. Thousand Oaks, CA: Sage Publications.

Frary, Robert B. 1996. "Hints for Designing Effective Questionnaires." *Practical Assessment, Research & Evaluation* 5 (3). http://PAREonline.net/getvn.asp?v=5&n=3.

Hofstede, G. 2001. *Culture's Consequences.* 2nd ed. Thousand Oaks, CA: Sage Publications.

Jackson, Gregg, and Rashmi Agrawal. 2007. "Guidelines for Developing Survey Instruments as a Part of the Designing and Conducting Surveys Workshop." International Program for Development Evaluation Training (IPDET) presenta- tion, Ottawa, Ontario.

Kaynak, Erdener, Jonathan Bloom, and Marius Leibold. 1994. "Using the Delphi Technique to Predict Future Tourism Potential." *Marketing Intelligence and Planning* 12 (7): 18–29. http://www.emeraldinsight.com/Insight/viewContentItem.do;jsessionid=372 EF38765D6F8133CDF8DABFB2071B4?contentType=Article&hdAction=lnkpdf&contentId=854271.

Kirigia, J. M. 1997. *Economic Evaluation in Schistosomiasis Using the Delphi Technique to Assess Effectiveness.* http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed& list_uids=9107365&dopt=Abstract.

Krueger, R. A., and M. A. Casey. 2000. *Focus Groups.* 3rd ed. Thousand Oaks, CA: Sage Publications.

Kellogg Foundation. 2008. *"Tips on Using Tests and Assessments."* *Evaluation Toolkit.*

http://www.wkkf.org/Default.aspx?tabid=90&CID=281&ItemID=2810034&NID =2820034&LanguageID=0.

Lofland, John. 1971. *Analyzing Social Settings.* Belmont, CA: Wadsworth.

Lofland, John, and L. H. Lofland. 1995. *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis.* 3rd ed. Belmont, CA: Wadsworth Publication.

McCaston, M. Katherine. 2005. *Tips for Collecting, Reviewing, and Analyzing Secondary Data.* Partnership and Household Livelihood Security Unit. http://pqdl.care.org/pv_obj_cache/pv_obj_d_8F453F01C87B8BB24774628B 95B 42BBCBD020200.

McNamara, Carter. 2007. *General Guidelines for Conducting Interviews.*

http://www.managementhelp.org/evaluatn/intrview.htm#anchor615874

Michigan State University Extension. 1994. *Delphi Technique.*

http://web1.msue.msu.edu/msue/imp/modii/iii00006.html.

Miles, Matthew B., and A. Michael Huberman. 1994. *Qualitative Data Analysis an Expanded Sourcebook.* Thousand Oaks, CA: Sage Publications.

Minnesota Department of Health. 2007. *Community Engagement: Community Forums and Public Hearings.* http://www.health.state.mn.us/communityeng/needs/needs.html.

Narayan, Deepa. 1996. *Toward Participatory Research.* World Bank Technical Paper 307, Washington, DC.

http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/ IB/ 1996/04/01/000009265_3980625172923/Rendered/PDF/multi0page.pdf.

O'Brien, G. 2001. *Data Flow, Application of Data Quality Flags, and Data Validation Processes for CCAQS.* Preliminary draft. http://www.arb.ca.gov/airways/Documents/reference percent20tables/March percent2006, percent202001/DataFlow_PrelimDraft.pdf.

Patton, Michael Q. 1987. *How to Use Qualitative Methods in Evaluation.* Thousand Oaks, CA: Sage Publications.

———. 2002. *Qualitative Evaluation and Research Methods.* 3rd ed. Thousand Oaks, CA. Sage Publications.

Paul, Samuel, and Gopakumar K. Thampi. 2008. "Monitoring Public Service Delivery: Citizen Report Cards Score in India." *Capacity.org.* 35 (December). United Nations Development Programme, New York. http://www.capacity.org/ en/journal/tools_and_methods/citizen_report_cards_ score_in_india.

Phuyal, Kamal. 2006. "Sharing Some Practical Experiences on Participatory Appreciative Planning Approach (PAPA): An Appreciative Approach to

Working in Partnership with Community People." Paper presented at the international workshop on "Action Learning: Lessons from the Field," organized by the Faculty of Policy Studies, Chuo University, Tokyo, October. http://www.scn.org/cmp/modules/emp-papa.htm.

Porteous, Nancy L., B. J. Sheldrick, and P. J. Stewart. 1997. *Program Evaluation Tool Kit: A Blueprint for Public Health Management.* Ottawa: Ottawa-Carleton Health Department.

Rainhorn, J.-D. P. Brudon-Jakobowicz, and M. R. Reich. 1994. "Priorities for Pharmaceutical Policies in Developing Countries: Results of a Delphi Survey." Bulletin of the World Health Organization 72(2): 257–64. 72 (2): 257–64. http://74.125.47.132/search?q=cache:D1zbVUz7oG0J:whq libdoc.who.int/ bulletin/1994/Vol72-No2/bulletin_1994_72(2)_257-264. pdf+Use+of+the+ delphi+technique+in+developing+countries&cd=3&hl=en &ct=clnk&gl=us& client=firefox-a.

Sanders, J. R. 2000. *Evaluating School Programs.* 2nd ed. Thousand Oaks, CA: Sage Publications.

Stafford, Frank P. 2006. *Timeline Data Collection and Analysis: Time Diary and Event History Calendar Methods.* Department of Economics, University of Michigan, Ann Arbor. http://www.atususers.umd.edu/wip2/papers_i2007/Stafford_Diaries.pdf.

Stake, Robert E. 1995. *The Art of Case Study Research.* Thousand Oaks, CA: Sage Publications.

StatPac. 2007. *Survey Software, Question Wording.* http://www.statpac.com/surveys/index.htm#toc.

Sue, Valerie M., and Lois A. Ritter. 2007. *Conducting Online Surveys.* Thousand Oaks, CA: Sage Publications.

TC Evaluation Center. 2007. *Wording Questions: Some General Guidelines.* University of California, Davis. http://ucce.ucdavis.edu/files/filelibrary/5715/27621.pdf.

Trochim, William M. K. 2006. *Types of Data, Research Methods Knowledge Base.* http://www.socialresearchmethods.net/kb/datatype.php.

UNESCAP (United Nations Economic and Social Commission for Asia and the Pacific). 1999. *Guidelines on the Application of New Technology to Population Data Collection and Capture.* http://www.unescap.org/stat/pop-it/pop-guide/capture_ch04.pdf.

University of Wisconsin, Cooperative Extension. 1996. *Program Development and Evaluation, Collecting Evaluation Data: Direct Observation.* http://learningstore.uwex.edu/pdf/G3658-5.pdf.

21.S. EPA (U.S. Environmental Protection Agency). 2007. *CTSA-Chapter 2: Data Collection.* http://www.epa.gov/dfe/pubs/lithography/ctsa/ch02/ch02.html.

Wadsworth, Y. 1997a. *Do It Yourself Social Research.* 2nd ed. St. Leonards,

New South Wales, Australia: Allen and Unwin.
———. 1997b. *Everyday Evaluation on the Run.* St. Leonards, New South Wales, Australia: Allen and Unwin.
Wengraf, Tom. 2001. *Qualitative Research Interviewing: Biographic Narrative and Semi-Structured Methods.* Thousand Oaks, CA: Sage Publications.
Wiseman, V., L. Conteh, and F. Matovu. 2005. "Using Diaries to Collect Data in Resource-Poor Settings: Questions on Design and Implementation." Health Policy and Planning 20 (6): 394–404. http://heapol.oxfordjournals.org/cgi/reprint/20/6/394.
World Bank. 1996. "A Manual for Planning and Implementing the Living Standards Measurement Study Survey." Working Paper 126, Washington, DC.
———. 2006. "Citizen Report Cards: A Presentation on Methodology." Participation and Civic Engagement Group, Social Development Department, Washington, DC. http://info.worldbank.org/etools/docs/library/94360/Tanz_0603/Ta_0603/ CitizenReportCardPresentation.pdf.
———. 2007. "Community-Driven Development." Washington, DC. http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTSOCIALDEVELOPMENT/EXTCDD/0,,menuPK:430167~pagePK:149018~piPK:149093~theSitePK:430161,00.html.
Yin, Robert K. 2003. *Case Study Research: Design and Methods.* 3rd ed. Thousand Oaks, CA: Sage Publications.

Веб-сайти

Alphabetical List of Aptitude or Achievement Tests. http://www.yorku.ca/psycentr/ tests/alpha.html.
Evaluation Portal. http://www.evaluation.lars-balzer.name/.
International Council on Security and Development. 2006. "Think tank announces expert panel to review effectiveness of International Narcotics Control Board." http://www.icosgroup.net/modules/press_releases/narcotics_control_board.
Measurement Group. *Evaluation/Research Tools.* http://www.themeasurement group.com/evalbttn.htm.
Nielsen, J. 1997. *The Use and Misuse of Focus Groups.* http://www.useit.com/papers/ focusgroups.html.
Standardized Monitoring and Assessment of Relief and Transitions. *Instruments and Tools.* http://www.smartindicators.org/IT.php?page=80.

ГЛАВА 9

Вибір стратегії формування вибірки

У цій главі розповідається, як визначити, скільки саме даних потрібно зібрати. Також розглядаються способи вибору джерел даних, які відображають генеральну сукупність та допомагають знаходити відповіді на питання оцінки.

Ця глава складається з трьох основних частин:

- Вступ до формування вибірки
- Види вибірок: випадкові та невипадкові
- Визначення розміру вибірки

Введення до формування вибірки

Генеральна сукупність:
Загальний набір елементів, щодо якого фахівець з оцінювання прагне дійти логічних висновків

Перепис:
Отримання даних від усієї генеральної сукупності

Вибірка:
Підмножина генеральної сукупності, в межах якої виконується збір даних

Вирішивши збирати дані з певної **генеральної сукупності** (людей, лікарень, шкіл і т. д.), фахівці з оцінювання мають прийняти рішення, чи потрібні їм дані про кожен елемент цієї сукупності чи достатньо буде лише даних про частину цих ***елементів***. Збір даних, що охоплює всі елементи генеральної сукупності, називають переписом. Збір даних в межах підмножини елементів називають ***вибіркою***. Якщо генеральна сукупність невелика, а вартість збору даних є низькою, перевагу віддають перепису. Якщо ж генеральна сукупність велика або якщо вартість збору даних є високою, зазвичай необхідно використовувати вибірку. Вибірка іноді також використовується для мінімізації навантаження на респондентів.

Люди постійно користуються вибірками. Наприклад, якщо ви здаєте аналіз крові, лабораторія використовує вибірку – зразок, а не всю вашу кров. Тести проводяться з використанням саме цього зразка. При цьому вважається, що результати таких тестів демонструють стан всієї вашої крові.

Але використання вибірки може бути актуальне не лише для масштабних, кількісних досліджень. Навіть під час проведення якісного польового дослідження тривалістю в один тиждень, покликаного оцінити програму, що охоплює великий географічний регіон, фахівцям з оцінювання потрібно добре обміркувати, які саме частини цього регіону вони хочуть дослідити. Висновки можуть бути необ'єктивними, якщо офіційні особи програми, які прагнуть показати її з найкращого боку, самі обиратимуть місцевості й учасників для включення у вибірку. Уникнути такої ситуації дозволить використання випадково визначеної вибірки. Розуміння основних концепцій формування вибірки може покращити точність оцінки.

Види вибірок: випадкові та невипадкові

Випадкова вибірка:
Вибірка, до якої з однаковою ймовірністю може потрапити кожен елемент генеральної сукупності

Помилка вибірки:
Спотворення висновків, зумовлене способом збору даних

Якщо неможливо отримати дані з кожної країни, від кожної особи або з кожної ферми, фахівці з оцінювання досліджують підмножину генеральної сукупності – вибірку. Вибірки бувають випадковими та невипадковими.

Випадкові вибірки

Випадковою називають вибірку, до якої з однаковою імовірністю може потрапити кожен елемент генеральної сукупності. Прикладом випадкової вибірки є лотерея, адже кожне число з однаковою ймовірністю може стати виграшним.

Однією з переваг випадкових вибірок є виключення **помилок вибірки** або перекручення даних, зумовленого способом збору даних. Якщо випустити ризик помилки вибірки з поля зору, висновки можуть бути сформовані

на основі неправильних даних. Якщо ж кожен елемент має однакові шанси бути обраним, помилка вибірки виключається. Існують різні форми помилки вибірки. Поширеною є помилка, пов'язана з ініціативністю, яка також може проявлятися у декількох формах. Помилки, пов'язані з ініціативністю, трапляються в будь-якій програмі, де люди виступають добровольцями, а також у випадках, коли респонденти самі зголошуються взяти участь в опитуванні. Джерело помилки полягає в тому, що для добровольців і респондентів часто притаманні якості, не властиві іншим людям.

Випадкова вибірка має бути репрезентативною для генеральної сукупності в цілому. Така вибірка дозволяє фахівцям з оцінювання робити загальні висновки щодо генеральної сукупності, з елементів якої було сформовано вибірку.

Для створення випадкової вибірки потрібний повний перелік усіх елементів генеральної сукупності. Такий перелік називають ***основою вибірки***. Кожному елементу генеральної сукупності має бути привласнений унікальний ідентифікаційний номер. Потім з генеральної сукупності вибираються окремі елементи. Для цього використовують інструмент рандомізації – наприклад таблицю випадкових чисел. Це таблиця з набором цифр (0, 1, 2, 3…), наведених у довільному порядку. Посилання на веб-джерела таблиць випадкових чисел наведено в кінці даної глави. Багато людей використовують онлайн-генератори випадкових чисел. Ресурси на кшталт Stattrek дозволяють вказати, скільки випадкових чисел потрібно та якими мають бути мінімальне та максимальне значення. Після цього генеруються випадкові числа.

Основа вибірки:
Повний перелік елементів, на основі якого формується вибірка

Існує шість видів випадкових вибірок:

- прості випадкові вибірки;
- механічні вибірки з випадковим інтервалом;
- механічні випадкові вибірки;
- розшаровані випадкові вибірки;
- серійні випадкові вибірки;
- багатоступеневі випадкові вибірки.

Прості випадкові вибірки

Проста випадкова вибірка – найбільш розповсюджений та найпростіший вид випадкової вибірки. Він застосовується, якщо основною метою є логічні висновки про всю генеральну сукупність, а не про якісь її частини.

Прості випадкові вибірки – хороший метод, якщо треба сформувати вибірку розміром від 50 до 500 елементів з однорідної генеральної сукупності або більшу вибірку з більш неоднорідної сукупності. Формувати просту випадкову вибірку можна по-різному. Наприклад, фахівці з оцінювання, зацікавлені у спостереженні класних занять з метою вимірювання кількості часу, що його учні проводять за виконанням практичних завдань, можуть

випадково обрати аудиторії, час та день. Фахівці з оцінювання, які хочуть спостерігати за інтенсивністю дорожнього руху на шляху між селищем та великим містом, можуть довільно обирати час або день, пору року та точки спостереження при дорозі.

Процедура формування простої випадкової вибірки виглядає наступним чином (Jackson 2008):

1. Ретельно визначте генеральну сукупність, зазначаючи, що в неї входить, а що – ні.
2. Знайдіть або сформуйте основу вибірки, що містить перелік всіх елементів генеральної сукупності та привласнює кожному з них унікальний номер. Номери не обов'язково мають бути суміжні (з використанням кожного наступного номера), але формування вибірки піде швидше, якщо між привласненими номерами не буде великих проміжків.
3. Визначте, якого розміру буде вибірка.
4. Визначте розрядність найбільшого привласненого номера.
5. Візьміть таблицю випадкових чисел.
6. Вирішіть, в якому порядку ви будете читати номери (наприклад, зверху вниз, зліва направо, стовпець за стовпцем).
7. Випадково оберіть початкову позицію в таблиці.
8. Читайте номери в обраний спосіб, називаючи стільки розрядів, скільки було визначено в кроці 4.
9. Кожного разу, коли ви читаєте число, що відповідає номеру елемента з вашої основи вибірки, відзначайте цей елемент як обраний.
10. Продовжуйте, поки не виберете стільки елементів вибірки, скільки потрібно.
11. Натрапивши на випадковий номер, що збігається з уже використаним, переходьте до наступного числа. Один і той самий елемент не можна обирати двічі.

Або ж, приміром, група оцінювання вибирає 100 файлів з 500. Всі ці файли отримали послідовні номери від 001 до 500. Потім фахівці з оцінювання можуть застосувати генератор випадкових чисел, зазначивши, що потрібно 100 чисел, та вказавши найменшим число 001 і найбільшим – 500.

(Механічна) вибірка з випадковим інтервалом:
Вибірка, що формується за допомогою випадкових інтервалів між елементами

Вибірки з випадково обраним інтервалом та випадково оборaним стартом і фіксованим інтервалом

Вибірка з випадково обраним інтервалом може використовуватися, якщо існує упорядкована генеральна сукупність ненумерованих елементів, нумерація яких є складною або трудомісткою. Як і для простої випадкової вибірки, вам знадобиться таблиця випадкових чисел. Однак таблиця ви-

користовується не для того, щоб обирати елементи вибірки, а для вибору інтервалів (кроку між елементами), що їх розділяють.

За допомогою наступних кроків можна сформувати вибірку з випадковим інтервалом (Jackson 2008):

1. Оцініть розмір генеральної сукупності.
2. Визначте бажаний розмір вибірки.
3. Поділіть оціночну кількість елементів генеральної сукупності на бажаний розмір вибірки. Округліть результат до двох знаків після коми: це буде середній розмір випадкового інтервалу, необхідний для формування вибірки бажаного розміру з генеральної сукупності.
4. Помножте отриманий після кроку 3 результат на 1/5, добуток округліть до двох знаків після коми. Ви отримаєте «множник».
5. Випадково оберіть початкову позицію в генеральній сукупності.
6. За допомогою таблиці випадкових чисел оберіть випадкове однорозрядне число. Помножте його на множник, добуток округліть до найближчого цілого числа, відрахуйте отримане число позицій від початкової, внесіть цю позицію до вибірки й позначте її.
7. Візьміть наступне випадкове однорозрядне число, помножте його на множник, округліть результат до найближчого цілого числа й відрахуйте далі це число позицій.
8. Продовжуйте виконувати ці кроки для генеральної сукупності, поки не повернетеся в початкову точку.
9. УВАГА: Якщо важливо повернути елементи до їх оригінальних позицій, обов'язково позначте кожне місце, з якого ви берете запис, спеціальним ідентифікатором.

Механічну вибірку з випадковим інтервалом також можна формувати за допомогою комп'ютерних програм.

Іноді неможливо виконати справді випадковий добір. В такому випадку можна застосувати систематичний добір з випадковим початком: випадково обравши відправну точку, ви вносите до вибірки кожну n-ну позицію.

Механічні вибірки з випадковим стартом та фіксованим інтервалом іноді вважаються квазівипадковими, тому що початкова точка є випадковою, але інтервал – ні. Хоча цей метод формування вибірки дещо простіший, ніж механічна вибірка з випадковим інтервалом, в певних умовах він виявляє серйозний недолік. Розглянемо, приміром, оцінку, що вимагає вибірки зі щоденних документів ринку, який працює сім днів на тиждень. Генеральна сукупність включає 700 записів, а вам потрібно відібрати 100. В цьому випадку фіксований інтервал – кожен сьомий запис, а це означає, що до вибірки потраплять записи лише за один день тижня. Ця вибірка навряд чи буде репрезентативною для інших днів. Щоб уникнути подібної ситуації,

Механічна вибірка з випадковим стартом та фіксованим інтервалом:
Метод систематичного добору, за якого початкова точка обирається випадково, а потім до вибірки включається кожна n-на наступна позиція

за будь-якої підозри, що фіксований інтервал може дати нерепрезентативну вибірку, використовуйте механічну вибірку з випадковим інтервалом.

Механічна вибірка з фіксованим інтервалом формується наступним чином:

1. Оцініть розмір генеральної сукупності.
2. Визначте бажаний розмір вибірки.
3. Поділіть результат кроку (1) на результат кроку (2) – це довжина інтервалу.
4. Випадково оберіть початкову позицію в генеральній сукупності.
5. Відрахуйте вказаний інтервал та включіть відповідний елемент до вибірки.
6. Продовжуйте відраховувати той самий інтервал і обирати елементи, поки не повернетесь у початкову точку.

Стратифіковані випадкові вибірки

Нерідко буває важливо включити до вибірки окремі групи, що можуть бути недостатньо представлені в разі використання простої випадкової вибірки. Щоб гарантувати представленість таких груп, фахівці з оцінювання можуть розділити генеральну сукупність на шари (на основі певних істотних характеристик – таких як гендер, вік, етнічна приналежність і т. ін.), після чого сформувати вибірку. Цей різновид вибірки називається ***стратифікованою випадковою вибіркою***.

Стратифікована випадкова вибірка: Вибірка, основа якої ділиться на два або більше шари (підмножини генеральної сукупності), з яких випадково вибираються учасники

Припустимо, фахівці з оцінювання хочуть вивчити наслідки інтервенції для міських та сільських жителів. Якщо сільські жителі представляють лише невелику частину загального населення регіону, проста випадкова вибірка може не містити достатньої їх кількості для проведення значущого аналізу.

Розшарована випадкова вибірка формується шляхом розділення генеральної сукупності на незбіжні групи (шари) ($n1, n2, n3, \ldots ni$, такі, що $n1 + n2 + n3 + \ldots + ni, = n$), та утворення простої випадкової вибірки в кожному шарі. Кількість елементів кожного шару повинна бути пропорційною розміру шару відносно генеральної сукупності.

Випадкові кластерні вибірки

Випадкова кластерна вибірка: Відбір елементів ведеться з природно утворених серій цих елементів

Серійна випадкова вибірка – це вибірка, сформована з природно утворених кластерів (груп) елементів аналізованої сукупності. Домогосподарства в даному випадку – кластери людей, міста – кластери домашніх господарств. Кластери мають бути взаємовиключними та охоплювати всі елементи генеральної сукупності. Кластерні вибірки часто застосовуються в наступних ситуаціях:

- Не існує повного списку всіх одиниць генеральної сукупності, але є повний перелік кластерів, в яких вони зустрічаються.

- Є повний список, але його члени знаходяться так далеко один від одного, що відряджати спеціалістів зі збору даних для проведення простої випадкової вибірки було б задовго й занадто дорого.

Якщо застосовується кластерна випадкова вибірка, формується випадкова вибірка з кластеру, після чого збираються дані стосовно всіх відібраних елементів. Цей метод особливо ефективний, якщо відібрані елементи розташовані далеко один від одного.

Головний недолік такого різновиду вибірки полягає в менш точних оцінках параметру (характеристики) генеральної сукупності порівняно з простою випадковою вибіркою або стратифікованою випадковою вибіркою того ж розміру. Може бути, що лікарні, вибрані в блоці 9.1, обслуговують клієнтів іншого економічного або релігійного стану, аніж ті, що до вибірки не потрапили. В такому випадку результати вибірки дадуть недостовірну оцінку генеральної сукупності пацієнтів зі СНІДом, що обслуговуються в лікарнях регіону.

Блок 9.1 «Використання серійної випадкової вибірки для відбору людей зі СНІДом для проведення інтерв'ю»

Група оцінювання хоче провести інтерв'ю приблизно з 200 особами з ВІЛ-інфекцією, що проживають в даному регіоні. Списків ВІЛ-інфікованих осіб, які мешкають в регіоні, немає. Створення таких списків обійшлося б неприпустимо дорого. Крім того, в процесі могли б постати певні питання етичного характеру і проблеми з конфіденційністю. Однак в регіоні працюють 25 лікарень, і фахівцям з оцінювання відомо, що більшість цих лікарень обслуговують приблизно по 50 пацієнтів зі СНІДом. Тому фахівці з оцінювання вирішили створити випадкову вибірку лікарень і охопити їхніх пацієнтів зі СНІДом. В регіоні працюють 25 лікарень. Вони віддалені одна від одної. Дороги в регіоні погані. Фахівці з оцінювання знають, що більшість лікарень обслуговують приблизно по 50 пацієнтів зі СНІДом. Отже, вони випадково відбирають чотири з 25 лікарень і вивчають всіх пацієнтів зі СНІДом в цих лікарнях. Розмір вибірки – близько 200 пацієнтів.

Багатоступеневі випадкові вибірки

Багатоступенева випадкова вибірка:
Вибірка, в якій послідовно використовуються дві або більше процедури випадкового відбору

Багатоступенева випадкова вибірка поєднує дві або більше форми випадкового відбору. Процес зазвичай починається з серійного випадкового відбору. Потім утворюється проста випадкова вибірка або розшарована випадкова вибірка. У прикладі, наведеному в блоці 9.1, багатоступенева випадкова вибірка почалася б із серійного відбору восьми лікарень. Потім

для кожної лікарні утворилася б проста випадкова вибірка з 25 пацієнтів. Цей метод на виході дає вибірку з 200 пацієнтів, як і в наведеному вище прикладі, але вони представляють більше лікарень.

В умовах багатоступеневого відбору можна поєднувати випадковий та невипадковий відбір. Наприклад, вибірку лікарень, згаданих у блоці 9.1, можна було б сформувати у невипадковий спосіб, а вибірку з їх пацієнтів – у випадковий.

Недоліком багатоступеневих та серійних вибірок є те, що вони забезпечують дещо менш точні оцінки генеральної сукупності порівняно з простою випадковою вибіркою. У прикладі про СНІД-лікарні, приміром, лише вісім лікарень з 25 потрапляють до випадкової вибірки, і вони можуть бути недостатньо репрезентативними.

Невипадкові вибірки

Цільова вибірка: Вибірка, що формується на основі заздалегідь визначених критеріїв

Типова (медіанна) вибірка: Різновид цільової вибірки, елементи якої беруться з середньої зони розподілу

Вибірка максимальної варіації (гетерогенна): Цільова вибірка, елементи якої представляють весь діапазон значень параметру

Квотна вибірка: Цільова вибірка, для якої відбирається певна кількість різних типів елементів

Якщо утворення випадкових вибірок неможливе, необхідно застосувати інший підхід. Зазвичай виділяють три види невипадкових вибірок: цільова вибірка; вибірка за методом снігової кулі; вибірка зручності.

Методи формування невипадкових вибірок дають фахівцям з оцінювання обмежені можливості робити статистичні висновки щодо загальної сукупності. Тим не менш, такі вибірки здатні додати ваги результатам оцінки.

Цільові вибірки

Під час формування **цільової вибірки** відбір ведеться з урахуванням конкретної мети дослідження. Відбір ґрунтується на заздалегідь визначених критеріях, які, на думку фахівців з оцінювання, дадуть змогу отримати необхідні дані.

Ось найбільш розповсюджені форми цільових вибірок (Jackson 2008):

- ***Типова (медіанна) вибірка:*** Спеціально відбираються елементи з типовими або найбільш розповсюдженими серед генеральної сукупності характеристиками (це середня зона кривої нормального розподілу). Мета дослідження – уважно вивчити саме типові, а не атипові випадки.
- ***Вибірка максимальної варіації (гетерогенна):*** Елементи вибірки представляють весь діапазон характеристики, що викликає інтерес дослідників. Ці елементи відбираються з усіх частин кривої нормального розподілу.
- ***Квотна вибірка:*** Елементи відбираються таким чином, що кожен шар представлений рівноцінно (в кількісному або пропорційному вираженні). Наприклад, фахівець з оцінювання вибирає п’ять елементів з верхньої, п’ять – з середньої та ще п’ять – з нижньої третини розподілу.

- ***Вибірка екстремальних випадків:*** Елементи відбираються з крайніх зон розподілу. Їх беруть з лівого й правого краю кривої нормального розподілу. Вибірки екстремальних випадків дають змогу розглянути найнетиповіші випадки.
- ***Вибірка підтверджувальних та заперечних випадків:*** Елементи відбираються з випадків, що підтверджують або заперечують загальновизнану думку, принцип або теорію (наприклад, добре підготовлені проекти, що були успішними, та добре підготовлені проекти, що зазнали невдачі).

Вибірка екстремальних випадків:
Цільова вибірка, елементи якої беруться з лівого та правого країв розподілу

Вибірка підтверджувальних та заперечних випадків:
Цільова вибірка, елементи якої відбираються з випадків, що підтверджують або заперечують загальновизнану думку, принцип або теорію

Вибірки за методом снігової кулі

Вибірки за методом снігової кулі використовуються тоді, коли фахівці з оцінювання не знають, кого або що включити у вибірку. Вони застосовуються, якщо межі генеральної сукупності невідомі і основа вибірки відсутня.

Вибірки за методом снігової кулі часто використовуються для проведення інтерв'ю. Фахівці з оцінювання, які використовують цей різновид цільової вибірки, запитують в людей, що відповідають критеріям участі в дослідженні, з ким ще варто поговорити.

Вибірки за методом снігової кулі потенційно можуть дати серйозну похибку через свою залежність від персональних рекомендацій. Один з прикладів їх застосування описує Девакі Намбіар (Nambiar 2008). Використовувати такі вибірки слід обережно.

Вибірка за методом снігової кулі:
Вибірка, для формування якої учасників інтерв'ю запитують, кого ще можна опитати

Вибірки, сформовані для зручності

Коли мова йде про **вибірку зручності,** відбір елементів виконується так, як зручно фахівцю з оцінювання. Ось поширені приклади таких вибірок в контексті оцінювання розвитку:

- відвідування місць реалізації проекту, що розташовані найближче до аеропорту;
- проведення інтерв'ю з тими менеджерами проекту, з якими можна поспілкуватися в день візиту фахівця з оцінювання;
- проведення спостережень в тих місцях, які показують офіційні особи проекту;
- спілкування з тими представниками неурядових організацій (НУО) та міськими лідерами, з якими стикається фахівець з оцінювання (Jackson 2008).

Вибірка, сформована для зручності:
Вибірка, сформована у найзручніший для фахівця з оцінювання спосіб

Вибірки, сформовані для зручності – дуже слабкий спосіб формування логічних висновків через те, що неможливо встановити, наскільки вибірка відрізняється від генеральної сукупності.

Недоліки невипадкових вибірок

Хоча використання невипадкових вибірок для оцінювання розвитку буває виправданим, цей різновид вибірок має кілька недоліків:

- Особи, які формують політику, вважають їх менш достовірними порівняно з випадковими вибірками.
- Через їхню невідповідність припущенням статистики висновків, до невипадкових вибірок не можна застосовувати перевірки статистичної значущості та розрахунки довірчих інтервалів.
- Невипадкові вибірки схильні до різних видів помилок.

Без формування випадкової вибірки не можна робити висновків щодо генеральної сукупності. Однак дані можуть бути дуже корисними, і, за обмежених можливостей, такий підхід може бути найкращим варіантом.

Комбіновані вибірки

Методи формування випадкових і невипадкових вибірок можна поєднувати. Наприклад, група оцінювання, що збирає дані про школи, може відібрати дві школи з найбідніших громад та дві – з найбагатших. Потім з цих чотирьох шкіл випадково відбирають учнів для включення до вибіркової сукупності.

Визначення розміру вибірки

Навіть за умови використання випадкової вибірки зберігається певна ймовірність помилки; вибірка може відрізнятися від генеральної сукупності. Статистика використовується для оцінки ймовірності того, що результати вибірки є репрезентативними для генеральної сукупності. Спеціалісти зі статистики розробили теорії та формули для таких оцінок та визначення відповідного розміру вибірки. Даний розділ присвячений базовим концепціям статистичного аналізу та їх застосуванню в ході розробки оцінки.

Розмір вибірки:
Кількість спостережень у вибірці

Визначаючи **розмір вибірки**, фахівці з оцінювання повинні вирішити, наскільки впевненими їм треба бути в тому, що результати вибірки точно відображають генеральну сукупність. Звичайно довірчий рівень – 95%. Це означає, що в 95 випадках зі 100 результати вибірки точно відображатимуть генеральну сукупність. Якщо фахівці з оцінювання хочуть бути впевненими на 90%, вони можуть взяти меншу вибірку. Якщо на 99% – більшу.

Визначивши розмір вибірки, фахівці з оцінювання мають вирішити, наскільки точними мають бути їхні оцінки. На відміну від помилок, похибку вибірки можна розрахувати. Це середнє значення різниці між всіма оцінками для всіх можливих вибірок та значенням, яке було б отримане

у разі вивчення всієї генеральної сукупності. Це називається **похибкою вибірки**. Припустимо, що за даними опитування 48 відсотків населення підтримують підвищення податків, а 52 відсотки виступають проти нього. При цьому похибка вибірки складає +/– 3 відсотки. На практиці це означає, що, якби було опитано кожного громадянина, фактичні результати були б наступні: 45–51 відсоток (48+/–3) за підвищення і 49–55 відсотків (52+/–3) проти. Ці ±3 відсотки – ***довірчий інтервал*** (не треба плутати з довірчим рівнем). Зауважте, ці два інтервали частково перекриваються. Це свідчить про відсутність можливості з 95-відсотковою впевненістю стверджувати, що більшість населення виступає проти підвищення податків. Вибірка дала занадто близькі результати.

Розмір вибірки залежить від розміру генеральної сукупності, бажаного довірчого рівня, а також бажаного рівня точності. Існує два способи визначення потрібного рівня вибірки. Перший – застосувати формулу, другий – звернутися до таблиці, що демонструє розмір вибірки, необхідний для даного довірчого рівня (табл. 9.1).

Похибка вибірки:
Оціночна величина похибки, спричиненої тим, що замість цілого розглядається вибірка

Довірчий інтервал:
Діапазон, у межах якого із заданою довірчою ймовірністю можна очікувати значення оцінюваної випадкової величини

Таблиця 9.1. Мінімальний розмір вибірки для забезпечення довірчого рівня у 95% з 5-відсотковою похибкою

Розмір генеральної сукупності	Розмір вибірки	Розмір генеральної сукупності	Розмір вибірки
10	10	550	226
20	19	600	234
40	36	700	248
50	44	800	260
75	63	900	269
100	80	1000	278
150	108	1200	291
200	132	1300	297
250	152	1500	306
300	169	3000	341
350	184	6000	361
400	196	9000	368
450	207	50000	381
500	217	100 000+	385

Джерело: Krejcie and Morgan 1970.

Що менше розмір генеральної сукупності, то більшим відносно неї є розмір вибірки. Наприклад, для генеральної сукупності з 300 елементів довірчий рівень у 95% забезпечує вибірка зі 169 елементів (трохи більше половини). Для генеральної сукупності з 900 елементів достатньо буде вибірки з 269 елементів – менше третини. Якщо ж генеральна сукупність складається зі 100 000 елементів, знадобиться вибірка лише з 385 елементів, а це всього 0,385 % від розміру генеральної сукупності.

Під час формування вибірки елементів (такими елементами можуть бути райони, школи, вчителі, батьки, громадяни) для добровільної участі в оцінці немає гарантій, що всі обрані елементи вдасться знайти (навіть за умови використання найактуальнішої основи вибірки), що вони відгукнуться на запрошення взяти участь або ж погодяться стати учасниками оцінки. Через це багато фахівців з оцінювання формують вибірку більших розмірів, ніж необхідно насправді, щоб отримати в результаті потрібну кількість респондентів. Якщо відсоток фактичних респондентів відрізняється від запланованого показника на 10% або більше, існує висока ймовірність **викривлення відповідей** (response bias), що потребує аналізу (див. главу «Інструменти збору даних»).

Рівень відповідей: Відсоток елементів вибірки, від яких фактично отримуються дані

У табл. 9.2 наведено розміри вибірки для дуже великих генеральних сукупностей (від 1 млн). Багато національних досліджень використовують вибірки розміром близько 1100, що забезпечує похибку на рівні +/–3% з довірчим рівнем 95%.

В мережі Інтернет можна знайти інструменти, що визначають розмір вибірки для генеральної сукупності, який дозволить забезпечити бажані довірчий рівень та похибку. Посилання на сайти з цими інструментами наведено в кінці цієї глави.

Таблиця 9.2. Розмір вибірки для генеральної сукупності розміром від 1 млн елементів

Похибка (%)	Довірчий рівень		
	99%	95%	90%
± 1	16 576	9604	6765
± 2	4144	2401	1691
± 3	1848	1067	752
± 5	666	384	271

Джерело: Jackson 2007.

Підіб'ємо підсумки:

- Точність даних можна підвищити, збільшивши розмір вибірки. Щоб підвищити точність і зменшити похибку, збільшіть розмір вибірки.
- Стандартний довірчий рівень – 95% (похибка – +/–5%).
- Що більшою є похибка, то менш точними є результати.
- Що менше розмір генеральної сукупності, то більшим відносно неї є розмір вибірки.

Якщо оцінка вимагає застосування складної багатоетапної стратегії формування вибірки, фахівці з оцінювання можуть звернутися до когось по допомогу. В Американської статистичної асоціації є каталог статистичних консультантів. Спілка статистичних консультантів пропонує такі послуги, як управління даними, аналіз даних, консультації під час написання наукових робіт, а також навчання в сфері статистики. На HyperStat Online розміщено посилання на багато інших ресурсів з допоміжними матеріалами щодо статистики. Посилання на ці сайти наведено в кінці даної глави.

Підсумки

Зазвичай неможливо або нераціонально збирати дані з кожного існуючого джерела. Тому фахівці з оцінювання збирають їх з певної вибірки джерел.

Існує два підходи до формування вибірки – випадковий і невипадковий. Випадковою називають вибірку, до якої з однаковою ймовірністю може потрапити кожен елемент генеральної сукупності. За допомогою випадкової вибірки вивчають підмножину генеральної сукупності, після чого роблять висновки щодо генеральної сукупності в цілому.

Існує шість видів випадкових вибірок:

- прості випадкові вибірки;
- механічні вибірки з випадковим інтервалом;
- механічні випадкові вибірки;
- розшаровані випадкові вибірки;
- серійні випадкові вибірки;
- багатоступеневі випадкові вибірки.

Існує три види невипадкових вибірок:

- цільові вибірки;
- вибірки за методом снігової кулі;
- вибірки, сформовані для зручності.

Під час оцінки не обов'язково використовувати лише один підхід. Допускається поєднання різних методів.

Розмір вибірки залежить від розміру генеральної сукупності, бажаного довірчого рівня, а також бажаного рівня точності. Навіть при застосуванні випадкової вибірки залишається ймовірність помилки. Фахівці з оцінювання використовують різноманітні статистичні підходи для визначення довірчого рівня й довірчого інтервалу.

Завдання до Глави 9

Практична вправа 9.1: Формування вибірки

Населений пункт, який ви вивчаєте, має населення 300 осіб. Кожний мешканець зареєстрований в міській обліковій книзі під унікальним ідентифікаційним номером. У вас є час і ресурси для опитування 50 містян. Ви б хотіли зробити висновки з цієї вибірки для генеральної сукупності з 300 осіб? Чи можливо це? Який вид вибірки ви б обрали? Обґрунтуйте своє рішення.

Практична вправа 9.2: Використання таблиці випадкових чисел

У поселенні, яке ви вивчаєте, 90 домівок. У вас є пронумерований список домівок, з якого ви хочете зробити просту випадкову вибірку з 10 елементів. Як можна використати наведену нижче таблицю випадкових чисел для формування такої вибірки?

44	14	12	12	03	12	73	72	62	33	35	62	80	34	77
69	59	54	90	01	50	04	93	76	69	43	95	47	60	80
23	95	24	95	24	55	69	89	41	18	12	94	43	21	43
40	76	50	38	18	05	44	23	72	61	58	67	99	05	75
54	05	51	52	04	34	25	64	90	95	02	86	51	14	37
36	82	03	65	38	93	49	64	06	93	01	30	62	05	68
96	19	97	24	16	26	94	14	17	45	22	51	09	92	16
75	85	18	50	50	60	80	52	42	11	05	70	89	53	38
57	78	12	98	55	51	48	77	54	07	66	15	33	44	64
58	20	10	51	62	06	25	56	63	67	73	73	79	05	65
55	84	17	67	52	38	16	29	05	24	12	05	35	87	31
92	44	84	04	17	47	18	78	54	40	02	59	74	06	73
86	96	79	86	75	67	31	41	40	20	87	17	85	98	70
78	84	03	69	43	38	43	98	90	75	56	49	88	52	78
25	05	76	72	06	59	37	56	24	36	95	05	30	62	02
26	67	04	13	77	37	21	57	77	41	82	30	32	80	09

Практична вправа 9.3: Стратегія формування вибірки

Працюючи в малих групах, визначте відповідний захід або статистику для кожного з наступних питань оцінки. Потім вирішуйте, яку стратегію формування вибірки ви б застосували в кожній з наступних ситуацій, і поясніть свій вибір.

1. Якою є якість доріг в селищах на північному заході Камбоджі одразу після сезону дощів?
2. Яка частина дітей до 10 років у штаті Таміл-Наду хворіє на малярію щонайменше один раз?
3. Які демографічні характеристики властиві відвідувачам сільських лікарень Шрі-Ланки?

Посилання та рекомендована література

Easton, V. J., and J. H. McColl. 2007. *Statistics Glossary: Sampling.* http://www.stats.gla.ac.uk/steps/glossary/sampling.html.

Guba, E., and Y. S. Lincoln. 1989. *Fourth Generation Evaluation.* Thousand Oaks, CA: Sage Publications.

Henry, G. T. 1990. *Practical Sampling.* Thousand Oaks, CA: Sage Publications. Jackson, Gregg B. 2007. *Sampling for IEG Managers.* Presentation at George Washington University, Washington, DC, December 18.

———. 2008. *Sampling in Development Evaluations.* International Program for Development Evaluation Training (IPDET) presentation, Ottawa, June 30 and July 1.

Kish, L. 1995. *Survey Sampling.* New York: John Wiley & Sons.

Krejcie, R. V., and D. W. Morgan. 1970. "Determining Sample Size for Research Activities." *Educational and Psychological Measurement* 30: 607–10.

Kumar, R. 1999. *Research Methodology: A Step-by-Step Guide for Beginners.* Thou- sand Oaks, CA: Sage Publications.

Laws, S., with C. Harper and R. Marcus 2003. *Research for Development: A Practical Guide.* Thousand Oaks, CA: Sage Publications.

Levy, P., and S. Lemeshaw. 1999. *Sampling of Populations.* 3rd ed. New York: John Wiley & Sons.

Lipsey, M. W. 1990. *Design Sensitivity: Statistical Power for Experimental Research.*

Thousand Oaks, CA: Sage Publications.

Lohr, S. 1998. *Sampling: Design and Analysis.* Pacific Grove, CA: Duxbury Press. Nambiar, Devaki. 2008. "The Delhi Snowball: Sampling Escapades in Urban India."

Paper presented at the Annual Meeting of the International Communication Association, Montreal, May 22–26.

Neuman, W. Lawrence. 2006. *Social Research Methods: Qualitative and Quantitative Approaches.* 6th ed. Boston: Allyn and Bacon.

Patton, M. Q. 2002. *Qualitative Research and Evaluation Methods.* Thousand Oaks, CA: Sage Publications.

Scheyvens, R., and D. Storey, eds. 2003. *Development Fieldwork: A Practical Guide.* Thousand Oaks, CA: Sage Publications.

Stattrek. http://stattrek.com/tables/randon.asph.

Tryfos, P. 1996. *Sampling Methods for Applied Research.* New York: John Wiley & Sons.

Веб-сайти

Alliance of Statistics Consultants.
http://www.statisticstutors.com/#statistical- analysis.
American Statistical Association Directory.
http://www.amstat.org/consultant directory/index.cfm.
Dr. Drott's Random Sampler.
http://drott.cis.drexel.edu/sample/content.html.
HyperStat Online. "Chapter 11: Power."
http://davidmlane.com/hyperstat/power.html.
HyperStat Online: Help with Statistics: Statistical Consultants and Tutors.
http:// davidmlane.com/hyperstat/consultants.html.
Probability Sampling. http://www.socialresearchmethods.net/kb/sampprob.htm.
Power Analysis. http://www.statsoft.com/textbook/stpowan.html.
Research Randomizer. http://www.randomizer.org.
StatPages.net. Web Pages that Perform Statistical Calculations,
http://www.Stat Pages.net.
Survey Research Methods Section.
http://www.fas.harvard.edu/~stats/survey-soft/ surveysoft.html.
The Survey System: Sample Size Calculator.
http://www.surveysystem.com/sscalc.htm.
UCLA Statistics Calculator: http://calculators.stat.ucla.edu.

ГЛАВА 10

Планування і проведення аналізу даних

Після того, як дані було зібрано, фахівцям з оцінювання потрібно вивчити їх і зробити висновки. Цей процес починається з вибору стратегії аналізу даних. Якісні й кількісні дані вимагають різних стратегій і методик.

Ця глава складається з чотирьох основних частин:

- Стратегія аналізу даних
- Аналіз якісних даних
- Аналіз кількісних даних
- Поєднання кількісних і якісних даних

Стратегія аналізу даних

Аналіз якісних даних:
Процедура, що використовується для аналізу інформації, зібраної в нечисловій формі, наприклад, викладених письмово або в аудіоформаті відповідей респондентів напівструктурованих інтерв'ю або інших документів і записів. Мета аналізу – зрозуміти й інтерпретувати поведінку й ситуацію. Ця процедура передбачає також визначення тем

Розробка стратегії **аналізу даних** – важлива складова процесу планування. Під час планування оцінювання фахівцям з оцінювання слід пам'ятати про техніки аналізу даних та їхні слабкі й сильні сторони. Визначена в матриці плану оцінки мета повинна бути конкретною та вказувати, який аналіз і графічні матеріали мають бути створені за результатами збору інформації. Розповсюджена помилка – збір величезної кількості даних, які потім ніде не використовуються.

Незалежно від того, якісним чи кількісним даним приділяється особлива увага в плані оцінки, збір даних та їх аналіз будуть частково збігатися. На початку збору даних на їх аналіз витрачається невеликий обсяг часу, особливо якщо проводиться попередня перевірка. На наступних етапах більше часу приділяється аналізу даних і менше – їх збору.

Якісний аналіз має проводитися в ситуаціях, коли, приміром, для отримання поглибленого розуміння інтервенції застосовуються напівструктуровані інтерв'ю. Можуть аналізуватися відповіді на наступні питання:

- З якими складнощами зіткнулися фахівці з оцінювання?
- Як учасники пояснюють свій вихід з програми?
- Які враження отримують учасники?

Кількісний аналіз використовується для відповіді на питання, для роботи з якими застосовуються структуровані інструменти збору даних (наприклад, опитування). Аналізуються, зокрема, такі питання:

- Якими є середні показники різних груп учасників?
- Як учасники оцінюють актуальність інтервенції за шкалою від 1 до 5?
- Яка спостерігається варіативність у відповідях?
- Чи існують статистично значущі відмінності між двома групами?

Аналіз якісних даних

Аналіз якісних даних використовується для тлумачення нечислових даних, зібраних в ході оцінки. Аналіз напівструктурованих спостережень, інтерв'ю з відкритими питаннями, письмових документів та розшифровок сесій фокус-груп вимагає застосування якісних методів.

Аналіз якісних даних починається безпосередньо на місці їх збору, коли приходить розуміння стану речей. Частиною польової роботи є запис та відстеження аналітичних висновків, що з'являються під час збору даних. Збір і аналіз даних мають частково збігатися в часі, але фахівець з оцінювання не повинен допускати, щоб початкові інтерпретації обмежували подальшу аналітичну роботу (табл. 10.1).

Таблиця 10.1. Рекомендації щодо збору й аналізу якісних даних

Завдання	Рекомендації
Зібрати дані	• Старанно ведіть записи.
	• Записуйте інтерв'ю, враження й зауваження про фокус-групи одразу після збору даних.
	• Постійно робіть порівняння.
	• Регулярно зустрічайтеся зі своєю групою для порівняння нотаток, визначення тем і внесення коректив.
Підсумувати дані	• Пишіть резюме (на одну сторінку) одразу після кожного важливого інтерв'ю або фокус-групи.
	• Згадуйте всі основні питання.
	• Визначайте найбільш цікаві, пізнавальні або важливі питання чи отриману інформацію.
	• Зазначте, які нові питання потребують вивчення.
Відстежувати параметри за допомогою інструментів	• Створіть окремий файл для фіксування власних реакцій під час дослідження й заносьте до нього свої відчуття, інтуїтивні здогадки тощо.
	• Записуйте свої ідеї, щойно вони з'являються.
	• Ведіть запис цитат в ході збору даних: ви зможете використати їх згодом при написанні звіту.
Зберігати дані	• Вся ваша інформація має бути в одному місці.
	• Робіть копії всієї інформації й зберігайте оригінали в одному місці.
	• Якщо вам треба дописувати, вирізати й вставляти якісь фрагменти даних, використовуйте для цього копії.

Джерело: Автори.

Якщо збір даних ведеться за допомогою опитувань, стандартизованих тестів і експериментальних планів, збір та аналіз даних є чітко розмежованими. Але плинний і мінливий характер природничих досліджень робить межі між збором і аналізом даних розпливчатими. В ході польових робіт з'являтимуться ідеї щодо напрямів аналізу. Формуються закономірності. На думку приходять можливі теми. З'являються гіпотези, на яких будується подальша робота дослідників. Початкові етапи польової роботи зазвичай є активними і мінливими – дослідники розробляють напрями, що ґрунтуються безпосередньо на отриманих даних. Пізні етапи завершують розпочате, акцентуючи увагу на зборі підтверджувальних даних, тобто на поглибленні розуміння і підтвердження чи заперечення виявлених закономірностей.

Іноді аналіз виявляє прогалини або неясності. Якщо дозволяють час, кошти та інші ресурси, фахівець з оцінювання може повернутися до польової роботи для збору додаткових даних, щоб проаналізувати відповіді або зробити нові спостереження. Під час збору даних члени групи нерідко проводять щоденні або щотижневі зустрічі для обговорення нових тем та впровадження протоколів.

Правильне ведення записів

Під час збору якісних даних важливо точно записувати всі спостереження. Отже, правильне ведення записів обов'язкове. Це означає, що треба приділяти багато уваги тому, що говорять люди та як вони це говорять. Ведучи записи, фахівці з оцінювання не повинні тлумачити слова респондентів. Вони повинні записувати те, що спостерігають, в тому числі мову тіла і все потенційно релевантне, що відбувається під час збору даних (наприклад, треба фіксувати випадки, коли інтерв'ю кимось переривається). Фахівці з оцінювання мають записувати свої думки, реакції й інтерпретації в окремому розділі своїх нотаток.

Важливо, щоб у фахівців з оцінювання одразу після інтерв'ю, спостереження, картування або фокус-групи був час для перевірки, дописування й аналізу записів. Тоді вони зможуть згодом краще розібратися в написаному. Дивно, як важко іноді зрозуміти поспішні нотатки вже наступного дня.

Навіть якщо вівся аудіо- або відеозапис сесії, фахівцям з оцінювання варто виділити час на перегляд письмових записів, поки враження ще свіжі. Це допоможе потім заощадити години прослуховування або перегляду записів чи вичитування розшифровок.

Тріангуляція: Використання трьох (або більше) джерел, видів інформації або видів аналізу для перевірки та підкріплення оцінки шляхом перехресної звірки результатів

Тріангуляція – використання трьох (або більше) джерел, видів інформації або видів аналізу для перевірки та підкріплення оцінки шляхом перехресної звірки результатів. Тріангуляція – корисний інструмент для якісного аналізу даних. Розглянемо, приміром, наступні приклади змішаних джерел даних:

- інтерв'ю, фокус-групи, анкети;
- анкети, наявні дані, експертні групи;
- спостереження, документи програми, картування;
- інтерв'ю, щоденники, наявні дані.

Комбінація висновків з цих трьох джерел свідчить про наявність закономірності.

Упорядкування якісних даних для аналізу

Після збору якісних даних фахівець з оцінювання матиме багато сторінок записів і розшифровок спостережень, інтерв'ю та інших джерел даних. Упорядкувати цю інформацію й розібратися в ній нелегко.

Існують інструкції з упорядкування даних. Документування цього процесу важливе для демонстрації валідності результатів (IDRC 2008).
Перш ніж перейти до упорядкування, фахівці з оцінювання повинні:

- переконатися у відсутності неповних даних;
- зробити кілька копій усіх даних;
- розсортувати дані по різних файлах (IDRC 2008).

Фахівець з оцінювання упорядковує записи по-різному, використовуючи паперові копії або електронні файли, частина з яких може представляти собою скановані документи. Деякі фахівці з оцінювання створюють чотири файли: в першому дані зберігаються в хронологічному порядку, в другому зберігаються аналітичні або щоденникові записи, в третьому – релевантні записи щодо методології дослідження, а в четвертому – копія всіх записів (IDRC 2008).

Петон (Patton 2002, зазначає про інші варіанти упорядкування якісних даних і складання відповідних звітів. Він нагадує фахівцям з оцінювання, що, виконуючи аналіз даних, варто пам'ятати про питання оцінювання. Обираючи спосіб упорядкування даних, треба прагнути відповісти на питання оцінювання.

- Розповідні підходи передбачають представлення даних або хронологічно (історія розповідається від початку до кінця), або ретроспективно (рухаючись з кінця до початку, розкриваючи послідовність подій).
- Підходи, пов'язані з вивченням практичних прикладів, дають інформацію про людей або групи; критичні ситуації зазвичай представляються в порядку виникнення.
- Аналітичні підходи включають опис процесів, висвітлення ключових питань (що часто збігаються з основними питаннями оцінювання), упорядкування питань, а також обговорення ключових концепцій – наприклад, порівняння лідерів з послідовниками.

Читання й кодування якісних даних

Визначення та використання категорій для упорядкування інформації – це лише початкові кроки. Наступний крок – прочитання даних. Після кількох читань фахівець з оцінювання має визначити потенційні теми. Петон (Patton 2002) відзначає, що визначення центральних тем подібне до створення пред-

метного покажчика для книжки або ярликів для реєстраційної системи. Він рекомендує фахівцям з оцінювання переглядати інформацію і привласнювати кожній темі назву або ярлик. Після впорядкування даних за темами, їх треба кодувати (блок 10.1). «Коди – ефективний засіб маркування й пошуку даних. Вони прискорюють аналіз і надають йому нові можливості» (Miles and Huberman 1994, p. 65).

Кодування – ітеративний процес. Хоча створення списку кодів до початку польової роботи – корисний крок, фахівцям з оцінювання доведеться перевіряти, переглядати, змінювати, нарощувати, а іноді й скасовувати коди в ході оцінювання, коли польові нотатки диктуватимуть більш практичне маркування.

Блок 10.1 «Приклад кодування»

Нижче наведено приклади кодів, встановлених для оцінювання освітньої програми. P означає учасників, S – персонал.

Код: Ps Re Prog (реакції учасників на програму)

Код: Ps Re Ps (реакції учасників на інших учасників)

Код: Ob PP (спостереження за взаємодіями учасників)

Код: Ob SS (спостереження за взаємодіями персоналу)

Код: Ob SP (спостереження за взаємодіями персоналу і учасників)

Код: Phil (заяви про філософію програми)

Код: Prc (приклади програмних процесів)

Код: P/outs (впливи програми на учасників, довгострокові результати)

Код: S-G (формування підгруп)

Код: GPrc (груповий процес)

Код: C! (конфлікти)

Код: C-PP (конфлікти між учасниками)

Код: C-SP (конфлікти між персоналом і учасниками)

Код: C-SS (конфлікти всередині персоналу)

Якщо кодування проводиться вручну, абревіатури наносяться на поля прямо поряд з відповідними абзацами або цитатами. Повні розшифровки кодів в дужках позначають окремі файли, що містять аналогічно закодовані уривки.

Проведення аналізу змісту

Аналіз якісних даних називають **аналізом змісту (контент-аналізом)**. Він передбачає визначення й кодування наявності в тексті, мові тощо певних слів, фраз і понять. Це систематичний підхід, що визначає й підсумовує приховані в даних ідеї.

Аналіз змісту (контент-аналіз): Різновид системного аналізу якісних даних, що, використовуючи коди, визначає і відзначає наявність у тексті, мові тощо певних слів, фраз і понять

Говорячи про аналіз змісту, мають на увазі аналіз книжок, брошур, розшифровок, новин, інших письмових документів, промов і відеозаписів. Цей підхід застосовується до розповідей – наприклад щоденників і журналів – а також до якісних відповідей на відкриті питання опитувань, інтерв'ю і фокус-груп. Аналіз змісту можна використовувати, наприклад, для вивчення шкільних підручників, щоб встановити, чи охоплюють вони навчальний матеріал, необхідний для вивчення конкретного предмета, чи викладають матеріал доступною мовою та представляють його з урахуванням контексту, в якому живуть і навчаються школярі. Поглиблений аналіз міг би бути спрямований на те, щоб встановити, чи не транслюють підручники якісь політичні догми чи викривлене трактування історії.

Блок 10.2 «Використання аналізу змісту для визначення сприйняття позитивних наслідків використання ручних насосів»

Щоб дізнатися, як селяни сприймають використання ручних насосів для води, фахівці з оцінювання опитали 100 селян. Вони проаналізували результати за наступною схемою:

1. Прочитати всі відповіді й записати всі їх варіанти.
2. Вибрати відповіді, що зустрічаються найчастіше, і коротко їх викласти. Згрупувати всі основні відповіді в категорії. Якщо, приміром, є відповіді «ручними насосами користуватися набагато легше» та «вода завжди чиста, і насос ні разу ще не ламався», фахівець з оцінювання може визначити наступні категорії: простота використання, постачання чистої води, надійність.
3. Переконатися, що всі категорії взаємовиключні і кодувальники будуть в змозі визначити, до якої категорії відноситься та чи інша відповідь.
4. Завершити кодування.

Звести у таблиці кількість згадувань кожної відповіді.

Джерело: Narayan 1996.

Аналіз змісту зазвичай починається з кодування даних. Цей процес спирається на припущення, що найчастіше вживані слова й фрази відображають важливі теми. Таким чином, аналіз змісту починається з визначення кількості згадування слів, вимірювання простору (сантиметрів колонки, якщо мова йде про газети), часу (для теле- та радіомовлення) та частоти вживання ключових слів.

Аналіз змісту виходить за межі простого підрахунку слів. З метою аналізу слова кодуються й розподіляються за категоріями. Ці нові кодовані категорії перевіряються на предмет частоти відповідей та взаємовідносин (блок 10.2).

Можна виділити два види аналізу змісту – кількісний і якісний. **Кількісний аналіз змісту** – це аналіз, що розглядає кількість згадувань вибраних термінів у тексті. **Якісний аналіз змісту** виходить за межі підрахунку згадувань і розглядає взаємовідносини між визначеними поняттями (Busch and others 2005).

Кількісний аналіз змісту:
Аналіз, що розглядає кількість згадувань вибраних термінів у тексті

Якісний аналіз змісту:
Аналіз, що виходить за межі підрахунку згадувань і розглядає взаємовідносини між визначеними поняттями

Наприклад, кількісний аналіз змісту міг би визначати категорії за наступним списком:

- спільні вислови щодо теми (що сприймалося як очевидне, а що вимагало додаткових пояснень з боку інших учасників);
- переконання та міфи стосовно теми, що є спільними, очевидними та спірними;
- аргументи, що їх наводять учасники, якщо їхні погляди оспорюються;
- джерела інформації, на які посилаються учасники для обґрунтування своїх поглядів і вражень, а також реакцій інших людей;
- аргументи, джерела й види інформації, що стимулюють зміну думки та нову інтерпретацію досвіду;
- стиль мовлення, мова тіла й рівень емоціональної участі, коли учасники обговорюють один з одним дану тему (Catterall and Maclaran 1997).

Кріпендорф (Kripendorff 2004) стверджує, що аналіз змісту має відповідати на шість питань:

- Які дані аналізуються?
- Як визначаються дані?
- Що представляє собою генеральна сукупність, з якої отримуються дані?
- У якому контексті аналізуються дані?
- Де проходить межа аналізу?
- Якою є мета логічних висновків?

Після відповіді на ці питання можна робити висновок щодо релевантних і нерелевантних даних.

Нойєндорф (Neuendorf 2006) представляє блок-схему, що демонструє процес аналізу змісту (рис. 10.1). Спочатку розглядається теорія і підстави для проведення аналізу.

Добре виконати аналіз змісту непросто. Фахівцям з оцінювання варто пам'ятати про потенційні проблеми, пов'язані з його застосуванням (табл. 10.2).

Комп'ютеризований контент-аналіз

Якщо для аналізу змісту передбачається використання комп'ютерів, вся інформація має міститися в файлах, які здатна прочитати комп'ютерна програма. Фахівцям з оцінювання може знадобитися друкування, сканування або відтворення файлів даних.

Існує багато програмних пакетів для упорядкування даних, отриманих від спостережень, інтерв'ю та фокус-груп. До них відносяться системи управління текстово-орієнтованих баз даних, текстові редактори, а також програми автоматичної індексації. Ці програми були розроблені спеціально для роботи з текстовими додатками. Деякі програми для аналізу змісту можна використовувати і з іншими носіями.

Рис. 10.1. Блок-схема дослідження з використанням контент-аналізу

1. Теорія й обґрунтування: Який зміст і чому? Чи є у вас питання дослідження? Гіпотези?

2. Рішення щодо концептуалізації: Які змінні будуть використовуватися під час дослідження? Як ви визначаєте їх теоретично?

3. Операційні засоби вимірювання: Повинні відповідати концептуалізації (внутрішня валідність). Яку одиницю збору даних ви використовуватимете? Таких одиниць може бути кілька. Чи проводиться вимірювання показників на високому рівні, чи є категорії взаємовиключними, чи охоплюють вони всю генеральну сукупність? Необхідно попередньо створити схему кодування, що описуватиме всі засоби вимірювання. На цьому етапі можна оцінювати як зовнішню, так і змістовну валідність.

Кодування вручну

4a. Схеми кодування:
Треба створити наступне:
(1) Кодова книга (з повним переліком одиниць вимірювання всіх змінних)
(2) Форма кодування

Комп'ютерне кодування

4b. Схеми кодування: Навіть на використання комп'ютеризованого аналізу змісту текстів вам знадобиться кодова книга – повне пояснення ваших словників та методу їх застосування. Ви можете використовувати словники для визначення ключових слів. Під час створення вихідних словників спочатку обов'язково розробіть список частоти згадувань на основі вибірки тексту і виокреміть ключові слова й фрази.

5. Формування вибірки: Чи можна виконати перепис змісту? (Якщо так, перейдіть до п. 6). Як ви сформуєте випадкову вибірку підмножини змісту (за періодом часу, питанням, сторінкою, каналом чи іншими засобами)?

6. Навчання й початкова надійність: Протягом навчальної сесії, в ході якої кодувальники працюють разом, з'ясуйте, чи можуть вони погодити код змінних. Потім, під час незалежного тесту кодування, відзначте надійність кожної змінної. На кожному етапі в разі необхідності редагуйте кодову книгу/форму кодування.

7a. Кодування (вручну): Залучайте щонайменше двох кодувальників, щоб забезпечити надійність кодування. Кодування має виконуватися незалежно.

7b. Кодування (комп'ютерне): Застосуйте словники під час тесту вибірки, щоб отримати дані про частоту згадувань на одиницю для кожного словника. Забезпечте валідизацію за допомогою вибіркової перевірки.

8. Кінцева надійність: Розрахуйте показник надійності для кожної змінної.дослідження. Гіпотези.

9. Створення таблиць і звітів: Може виконуватися у різний спосіб. Цифри й статистика можуть представлятися окремо або за допомогою багатокоординатної таблиці. Також можна використовувати діаграми.

Таблиця 10.2. Потенційні переваги й недоліки контент-аналізу

Потенційні переваги	Потенційні недоліки
Аналіз змісту розглядає безпосередньо спілкування, використовуючи для цього тексти або розшифровки бесід, завдяки чому дістається самої суті соціальної взаємодії	Може бути надзвичайно трудомістким
Можна виконувати як кількісні, так і якісні операції	Можливі похибки, особливо якщо для підвищення рівня інтерпретації даних застосовується якісний аналіз
Аналіз текстів може надавати цінні історичні/культурні відомості в динаміці	Аналізу часто бракує теоретичної бази. Нерідко трапляються занадто вільні спроби робити значущі логічні висновки про відносини і впливи
Може використовуватися для інтерпретації текстів для розробки, наприклад, експертних систем (тому що знання й правила можна кодувати у вигляді чітких заяв про зв'язки між поняттями)	Існує схильність до спрощення, особливо якщо мова йде про обробку складних текстів
Є ненав'язливим засобом аналізу взаємодій	Занадто часто складається просто з підрахунку кількості слів
Дає уявлення про комплексні моделі людської думки й використання мови	Часто випускає з поля зору контекст, в якому створювався текст, а також стан речей після оформлення тексту
	Комп'ютеризація процесу може бути складною

Джерело: Busch and others 2005.

Програмне забезпечення для комп'ютеризованого аналізу якісних даних (CAQDAS): Застосування комп'ютера для пошуку, впорядкування, категоризації й анотації даних

Програмне забезпечення для комп'ютеризованого аналізу якісних даних (CAQDAS), також відоме як ПЗ для аналізу якісних даних (QDAS або QDS), дозволяє відшукувати, упорядковувати, категоризувати й анотувати текстові та візуальні дані (блок 10.3). Програми цього типу дозволяють фахівцям з оцінювання візуалізувати відносини й теоретичні конструкції та будувати теорії. До таких програм відносяться Ethnograph, Qualpro, Hyperqual, Atlas-ti, QSR's N6 (раніше відомий як NUD*IST) та NVivo 8, AnSWR, HyperRESEARCH, Qualrus тощо.

На веб-сайті Американської асоціації оцінювання є публічний розділ, присвячений ресурсам. Однією з категорій таких ресурсів є «програмне забезпечення для аналізу якісних даних». В цій категорії дається стислий опис доступного програмного забезпечення, часто включно з вартістю, обсягом і видами носіїв інформації, а також посилання на сайти постачальників ПЗ. Описи містять дані про можливість безкоштовного завантаження в повному або випробувальному режимі.

Аналіз якісних даних вручну

Портос, Шелдрік і Стюарт (Porteous, Sheldrick, and Stewart 1997) наводять цінні рекомендації щодо аналізу якісних даних вручну. Звичайно, якщо було зібрано значну кількість даних, краще застосовувати комп'ютерні інструменти. Для початку підготуйте наступні матеріали:

- кілька маркерів (різного кольору для кожного питання оцінки);
- окремий робочий лист для кожного питання оцінки (рис. 10.2);

Блок 10.3 «Стратегії електронного кодування даних»

Наступні рекомендації стосуються використання CAQDAS:

- Пам'ятки – найпростіший спосіб анотації даних. Режим електронних наліпок дозволяє користувачам прикріпляти пам'ятки до даних.
- Вільне кодування дозволяє фахівцям з оцінювання маркувати розділи даних і призначати для них коди.
- Процедури автоматичного кодування працюють по-різному. Найбільш поширений спосіб – автоматичне кодування комп'ютерною програмою результатів пошуку. Дані також можуть перекодовуватися автоматично на основі попередньо використаних запитів.
- Рекомендовані програмою коди – це нова функція Qualrus: алгоритм пропонує коди, орієнтуючись на раніше привласнені.
- Мультимедійне кодування – функція, що її пропонують N6, HyperRESEARCH і Qualrus. Ці програми дозволяють кодувати послідовності аудіо- та відеофайлів, а також частин зображень. Низка інших програм CAQDAS дозволяють використовувати посилання на зовнішні мультимедійні файли.

Джерело: Loughborough University Department of Social Sciences 2007.

Рис. 10.2. Бланк робочого листа для аналізу якісних даних

Тема/кількість згадувань	Цитати	Висновки/Результати

Джерело: Porteous, Sheldrick, and Stewart 1997.

- дані, в тому числі нотатки, розшифровки та аудіозаписи інтерв'ю і фокус-груп;
- інструменти збору самостійно заповнених анкет, реєстраційних форм, спостережень та аналізів амбулаторних карток.

Використовуйте щонайменше один робочий лист для кожного питання оцінки. Вписуйте питання оцінки в призначене для цього місце згори на кожному робочому листі. При роботі з кожним питанням оберіть код для ідентифікації даних. Це може бути колір ручки, олівця чи маркера, або ж якийсь символ. Вкажіть цей колір чи символ в другому проміжку згори на кожному робочому листі.

Заповнюючи робочі листи, переглядайте зібрані записи й матеріали та кодуйте інформацію. Слід дотримуватися наступної процедури:

- Прочитайте всі заповнені анкети, нотатки й розшифровки за один раз.
- Позначте різнокольоровими маркерами ті місця, де йдеться про певне питання оцінки.
- Тепер уважно прочитайте всі дані, що відносяться до першого питання оцінки.
- У стовпець «Тема» робочого листа впишіть кожну думку, ідею або враження, що стосуються очікувань щодо цього питання оцінки.
- Залишайте місце після кожної теми: там ви будете відмічати кожну згадку про цю тему.
- Ведіть підрахунок згадувань думки, ідеї або враження.

Опрацюйте решту робочого листа наступним чином (рис. 10.3):

- Виберіть і впишіть з наявних записів цитати, що найкраще ілюструють кожну тему.
- Сформулюйте попередні висновки щодо конкретних моментів і запишіть їх до стовпця «Висновки».
- Впорядкуйте висновки за типом або категорією.
- За допомогою числа відповідей (N = x) додайте точності та відчуття масштабу.

Якщо для аналізу якісних даних використовують картки для записів, метою є узагальнення побаченого або почутого простими словами, фразами, темами або закономірностями. Пізніше можуть з'явитися нові теми, і попередні матеріали треба буде перечитати, щоб переконатися, чи не була ця тема присутня в них з самого початку і лише пропущена через те, що її важливість була сумнівною.

Визначаючи слова, питання, теми або закономірності, відмічайте їх розташування, щоб пізніше знайти їх, якщо треба буде перевірити цитату чи контекст. Перший раз це може бути виснажливо, але з досвідом приходить вміння швидше знаходити потенційно важливу інформацію.

Рис. 10.3. Заповнений робочий лист для аналізу якісних даних

Тема/кількість згадувань	Цитати	Висновки
Батьки ~~////~~ ~~////~~ ~~////~~ ~~////~~ ~~////~~ ~~////~~ ~~////~~ ////	Пряма мова: «Я думаю, процедура прийняття рішень була б дуже доречною»	Склалося стійке враження, що батьків слід активніше долучати до вибору тем
Охоплювати кілька тем за одну сесію ~~////~~ ~~////~~ ~~////~~ ///	«Іноді ми починали одну тему, а потім наставав час іти або переходити до наступної теми»	Багато учасників (38 з 52 опитаних) вважали, що треба відводити більше часу на обговорення
Кожній темі приділяється недостатньо часу ~~////~~ ~~////~~ ~~////~~ ////	«Потрібно більше часу на обговорення»	

Джерело: Porteous, Sheldrick, and Stewart 1997.

Часом думка меншості є важливою, і про неї необхідно повідомляти. Використовуйте власне судження, але завжди уточнюйте, що цю думку висловив лише один респондент (або кілька респондентів). Як зазначається на веб-сайті Університету Західної Англії:

> В житті рідко все буває розкладено по поличках. Завжди трапляються глухі кути, теми, що загасають або суперечать одна одній. Під час якісного дослідження існує спокуса ігнорувати категорії, що не вписуються в нову теорію. Такі категорії – наче непарні шкарпетки, які можна іноді знайти у шухляді. В усіх проектах якісного дослідження зустрічаються дивні випадки, що вибиваються із загальної канви. Але їх не треба відкидати, вони – частина цілого.

Інтерпретація якісних даних

Фахівці з оцінювання описують дані й інтерпретують їх. Перш ніж якісні дані можна буде проаналізувати, їх потрібно чітко описати. Інтерпретація даних означає знаходження можливих причинно-наслідкових зв'язків, формулювання логічних висновків, тлумачення даних та роботу з випадками, що суперечать аналізу.

Багато людей бояться використовувати статистику. Відповідно, поширена думка, що застосовувати якісні методи легше. Насправді ефективний якісний аналіз висуває більше вимог, ніж здається на перший погляд. Аналіз якісних даних може бути трудомістким і забирати чимало часу, але він здатний давати уявлення про поведінку або процеси, що не знаходять відображення в кількісних даних. Фахівець з оцінювання має одразу відвести певний час на проведення аналізу, щоб отримати хороші результати.

Якісні методи – ефективний інструмент для вивчення причинно-наслідкових зв'язків. (Чудову покрокову інструкцію з проведення системного якісного аналізу даних наводять Майлз і Г'юберман (Miles and Huberman 1994)) Петон (Patton 2002) описує два різновиди якісного аналізу: індуктивний і дедуктивний. **Індуктивний аналіз** передбачає виявлення закономірностей, тем і категорій даних. **Дедуктивний аналіз** – це аналіз даних з використанням існуючої системи. Зазвичай якісні дані піддаються індуктивному аналізу на ранніх етапах (визначення категорій, закономірностей і тем). Після визначення цих категорій, закономірностей і тем можна переходити до дедуктивного аналізу. Дедуктивний етап включає в себе тестування й підтвердження автентичності та відповідності індуктивного аналізу.

Індуктивний аналіз:
Аналіз даних, що передбачає виявлення закономірностей, тем і категорій

Дедуктивний аналіз:
Аналіз даних з використанням існуючої системи

Існує певний ризик помилки при роботі з якісними даними (якщо не використовувати програмне забезпечення для аналізу змісту); люди часто бачать те, що хочуть бачити, і не помічають речі, що не відповідають їхнім очікуванням. Рішенням (яке, однак, не завжди дає стовідсотковий результат) є запросити іншу особу для аналізу даних. Порівнявши результати двох аналізів, фахівці з оцінювання можуть виявити нові теми або інші способи тлумачення даних.

Звіти про якісні дані не завжди включають кількісні або відсоткові дані про те, хто що зробив чи сказав, або ж це може бути не завжди доцільно. Через те, що респондентам не ставилося одне й те саме питання, важко зробити висновки про їх думку стосовно цього питання. Інший спосіб усунення помилок – залучення двох кодувальників до перевірки одних і тих самих документів та їх кодування за темами. Якщо фахівці з оцінювання добре підготовлені і операційні визначення та системи оцінки зрозумілі й заздалегідь погоджені, оцінки матеріалів обома фахівцями з оцінювання будуть значною мірою збігатися. Високий рівень такої збіжності – індикатор достовірності. Низький рівень збіжності вказуватиме на потребу переглянути операційні визначеня та/або системи оцінки.

У табл. 10.3 наведено рекомендації щодо упорядкування й інтерпретації якісних даних.

Звітування про якісні дані

У багатьох оцінках використовуються як якісні, так і кількісні дані. Якщо ж фахівці з оцінювання застосовують змішаний метод (який поєднує кількісні та якісні підходи), їм потрібно буде знайти коментарі щодо кількісних даних. Якщо, приміром, 55 відсотків респондентів були незадоволені рівнем доступності інтервенції, корисно включити репрезентативний набір коментарів, що ілюструє джерела незадоволення.

Таблиця 10.3. Упорядкування й інтерпретація якісних даних

Завдання	Рекомендація
Розробити категорії	• Використовуйте часто повторювані теми, ідеї, слова й фрази.
	• Використовуйте категорії, достатньо великі для охоплення певного спектру поглядів, але не настільки великі, щоб бути позбавленими змісту.
	• Створіть категорії, що чітко відрізняються одна від одної.
Виконати кодування категорій	• Розробіть схему кодування.
	• Розробіть правила прийняття рішень щодо кодування; вони повинні бути чіткі та охоплювати всі випадки.
	• Навчіть кодувальників користуватися схемою кодування.
Перевірити надійність, якщо залучається більш ніж один спостерігач	• Проведіть попередню перевірку на невеличкій вибірці якісних даних.
	• Перевірте збіжність оцінок різних фахівців з оцінювання – чи однакові оцінки виставляють спеціалісти, які оцінюють одну й ту саму річ в один і той самий спосіб?
	• Якщо існують проблеми, усуньте їх. Потім повторіть попередню перевірку.
Проаналізувати дані	• Упорядкуйте дані.
	• Ви можете записати дані на картки.
	• Ви можете записати дані в електронні таблиці.
	• Ви можете скористатися комп'ютером для проведення аналізу даних.
	• Відсортуйте дані, щоб виявити закономірності й теми.
Інтерпретувати дані	• Якщо можливо, залучіть групи по двоє осіб до перегляду й категоризації даних, щоб порівняти їх висновки. Якщо висновки будуть різні, перегляньте та скорегуйте їх.
	• Намагайтеся знайти сенс і значення даних.
	• Встановіть зв'язок між темами й категоріями з одного боку і процесами та/або довгостроковими результатами програми з іншого. Чи піднімаються якісь теми частіше за інших, коли респонденти обговорюють процесові питання? Чи видаються якісь теми більш релевантними за інші, коли респонденти обговорюють питання довгострокових результатів? Шукайте альтернативні пояснення та інші способи трактування даних.
Переглянути інформацію й розповсюдити її	• Розповсюджуйте інформацію швидко й часто, надаючи її ключовим інформантам.
	• Залучайте інших людей до перевірки ранніх проектів звітів, щоб отримати інформацію, питання, інші погляди на тлумачення даних, а також відомості про інші можливі джерела даних.
Написати звіт	• Опишіть основні теми (тематичний підхід) або представте матеріали таким чином, щоб відобразити динаміку подій (хронологічний підхід).
	• Висвітліть цікаві точки зору, навіть якщо їх висловили лише одна–дві особи.
	• Тримайте концентрацію; дуже легко загубитися у великих обсягах даних.
	• Включайте лише важливу інформацію. Запитайте себе, чи дає інформація відповіді на питання оцінки та чи буде вона корисною для зацікавлених осіб.

Джерело: Взято з Porteous 2005.

Фахівцю з оцінювання знадобляться «нормальні цитати». Необхідно вибрати заяви учасників, адже вони чітко представляють тему або важливий аспект, який варто підкреслити. В цих словах є сила: той, хто читатиме звіт, запам'ятає цитату, а не сторінку описового тексту. Для уникнення помилок фахівці з оцінювання мають включати питання, що демонструють спектр проблем і перспектив з певної теми.

Аналіз кількісних даних

В ході аналізу кількісних даних підсумовується числова інформація, зібрана в ході оцінювання. Фахівці з оцінювання вводять дані в комп'ютерний файл з метою упорядкування даних або ж для застосування програм, що одразу аналізують дані.

Кодування даних

Для аналізу кількісних даних їх спочатку треба кодувати, але робити це простіше, ніж проводити аналіз змісту якісних даних. Кодування кількісних даних застосовується, якщо дані потрібно перекласти в числову відповідь. Кодування дозволяє здійснювати осмислену обробку даних. Дані щодо таких характеристик, як зріст, вага, вік та кількість пропущених днів навчання, не потребують кодування, адже вони вже числові. Інші види даних – наприклад, відомості про те, чи має респондент банківський рахунок – для проведення аналізу потребують числових кодів. Фахівець з оцінювання може кодувати відповіді, наприклад, позначаючи відповідь «так» одиницею, а «ні» – двійкою.

Інші дані можуть збиратися в діапазонах. Припустимо, що є питання: «До якої вікової групи ви належите?» Кожній віковій групі можна привласнити код (до 18 = 1, 18–25 = 2, 26–35 = 3 і т. д.). Ми наводимо рекомендації з роботи з кількісними даними (блок 10.4).

Блок 10.4 «Рекомендації щодо роботи з кількісними даними»

- Інформаційні змінні обов'язково мають бути промарковані.
- Переконайтеся, що до словника даних вносяться усі зміни ярликів даних та кодів відповідей. Критично важливо добре вести документацію!
- Створіть резервну копію масиву даних на випадок аварійної ситуації. Створюйте тимчасові та постійні масиви даних помірковано. Подумайте про те, що потрібно буде зробити у разі втрати даних.
- Завжди зберігайте копію оригінального масиву даних.

Джерело: Ініціатива щодо вимірювання здоров'я дітей та підлітків (CAHMI) , 2009.

Кожну особу або запис називають «випадком». Файли даних складаються зі змінних та їх значень. Змінними називають всі об'єкти аналізу. Надзвичайно важливо, щоб фахівці з оцінювання вказували, як кодується кожна змінна та до якої категорії вона відноситься. Коди необхідно заносити до кодової книги (яку іноді називають словником даних). Кодувальники повинні мати доступ до кодової книги та вміти користуватися нею.

Очистка даних

Очистка даних:
Процес усунення помилок і невідповідностей з метою покращення якості даних

Очистка даних – це процес усунення помилок і невідповідностей з метою покращення якості даних (Rahm and Do 2000). Дані з помилками й невідповідностями часто називають «брудними даними». За оцінками аналітиків, до половини часу, що витрачається на аналіз, йде на очистку даних. Зазвичай це недооцінка. Після очистки масиву даних аналіз стає простішим (P.A.N.D.A. 2000).

Ось деякі з найбільш розповсюджених джерел помилок в базі даних:

- відсутність даних;
- «незастосовні» або пусті відповіді;
- помилки при вводі даних;
- неправильний формат даних;
- зсув стовпця (дані, що належать до одного стовпця змінних, введені в сусідній стовпець);
- сфабриковані дані;
- помилки кодування;
- помилки вимірювань та інтерв'ю;
- застарілі дані (P.A.N.D.A. 2000).

В деяких випадках респонденти не знають відповіді на питання або відмовляються її дати. В інших випадках респондент може ненавмисно пропустити питання. Багато фахівців з оцінювання використовують такі коди для відповідей:

- не знаю = 8;
- відмова відповідати або відсутні дані = 9.

Розглянемо, наприклад, випадок, коли дані вводилися в базу кількома різними людьми. Один кодувальник вводив всі дані, використовуючи великі літери, інший починав прізвища з великої літери, а решту літер набирав маленькими. Один кодувальник вводив кожну адресу як один запис, інший окремо вводив вулицю й дім, місто та країну. На додачу дані були об'єднані з даними оцінки, що проводилася 10 років тому. Новий інструмент опитування ставив більше питань, ніж попередні, тому попередні дані не містили відповідей на багато питань.

Ще одна проблема з даними полягає у тому, що багато респондентів були збиті з пантелику деякими питаннями. Люди, що записували відповіді, докладали усіх зусиль, але всі вони використовували різні правила кодування. В кількох випадках респонденти, яким пропонували вибір за шкалою від 1 до 5, називали в якості відповіді десяткове число (наприклад, 2,5). Дехто з кодувальників округляв такі відповіді до найближчого цілого, інші округляли до 2, а дехто записував таку відповідь як «дані відсутні».

Всі ці проблеми викликані людським фактором: помилками у відповідях або введенні даних. Чимало з цих проблем можна уникнути, якщо встановити та суворо дотримуватися правил вводу й кодування даних. Однак помилки під час вводу даних будуть траплятися завжди. Саме тому введені дані необхідно перевіряти. Фахівці з оцінювання мають встановити правила кодування відповідей та обліку оригіналів анкет (до яких можна звернутися в разі виявлення помилок). Фахівці з оцінювання мають переглянути дані й спробувати виявити «брудні».

Введені дані потрібно перевірити й очистити. Тільки після цього можна переходити до їх аналізу. Розглянемо, наприклад, дані, зібрані за допомогою шкільних документів. Для питання про стать можливі лише дві відповіді: 1 – чоловіча, 2 – жіноча. Інші відповіді мають вважатися помилковими. Для питання про проходження медогляду можливі лише такі варіанти: 1 – так, 2 – ні, 8 – не знаю, 9 – відмова відповідати або відсутні дані. Будь-які інші відповіді слід вважати помилкою кодування. Для виявлення можливих помилок в даних щодо зросту фахівець з оцінювання може розглянути дані, що є значно вищими або значно нижчими за очікувані для учня такого віку показники (O'Rourke 2000b). Якщо існують сумніви в точності даних, фахівець з оцінювання може звернутися до оригіналів робочих листів і перевірити точність кодованих даних. (Чепмен (Chapman 2005) рекомендує корисні матеріали щодо принципів та методів очистки даних. Він посилається на кілька онлайн-ресурсів і доступне програмне забезпечення. Хоча він говорить про проблеми біорізноманіття, значна частина інформації є універсальною. Демонструючи техніки очистки даних, Коді (немає даних) пропонує вправу. Посилання на неї наведено в кінці цієї глави).

Існують комп'ютерні програми для очистки даних. Ці програми перевіряють, чи немає даних поза дозволеним діапазоном значень. Одним з прикладів є WinPure. Надзвичайно важливо реєструвати прийняті рішення щодо очистки даних.

Використання статистики

Кількісні дані аналізуються з використанням статистики. В цьому розділі представлені деякі з найважливіших статистичних понять, з якими мають бути обізнані люди, які проводять або читають звіт про оцінювання розвитку.

Спеціалісти поділяють статистику на дві великі категорії:

Описова статистика:
Статистика, що використовується для опису й підсумування кількісних даних

Статистика висновків (інференційна):
Статистика, що використовується для формулювання висновків щодо генеральної сукупності на основі вивчення вибірки

- **описова статистика**, яка (згідно з найвужчим визначенням) описує та підсумовує кількісні дані;
- **статистика висновків (інференційна)**, яка зазвичай використовується для аналізу даних випадкової вибірки шляхом прогнозування діапазону значень генеральної сукупності для кількісної або якісної змінної на основі інформації про цю змінну в межах випадкової вибірки. Частина прогнозу включає твердження про надійність, в якому вказується ймовірність того, що істинне значення для генеральної сукупності знаходиться в межах вказаного інтервалу значень.

Описова статистика

Зазвичай дані підсумовуються за допомогою двох різновидів описової статистики:

Міри центральної тенденції:
Статистичні показники, що демонструють, наскільки схожими є характеристики

Показники дисперсії:
Статистичні показники, що демонструють, наскільки різними є характеристики

- **міри центральної тенденції:** спосіб описання групи даних з метою визначення центральної точки;
- **показники дисперсії:** спосіб описання групи даних з метою визначення її розсіювання.

Міри центральної тенденції. Існує три таких міри/показники.

Група даних нерідко представляється у вигляді графіка щільності ймовірності з метою представлення частотності кожної змінної в групі. Після цього визначається центральне значення, для чого застосовується один або кілька з трьох показників середнього значення: середнє, медіана й мода.

Мода: Відповідь, що дається найчастіше.
Медіана: Серединна точка або середнє значення розподілу; половина всіх значень розподілу більше значення медіани, а половина – менше. В масивах даних з парною кількістю елементів медіана визначається як середнє від двох середніх випадків.
Середнє: Сума всіх зібраних значень, поділена на їх число (розмір вибірки).

Найчастіше використовуються саме середнє та медіана.

У табл. 10.4 представлено дані щодо пропорції міського населення 16 країн. Уявімо, що ви хочете підсумувати цю інформацію й повідомити про середній рівень урбанізації для цієї групи країн. Середнє – сума рівнів урбанізації, розділена на кількість країн. (90 + 64 + 84 . . . + 93)/16 = 1141/16 = 71,3). Два середні випадки, нижче й вище яких знаходяться 50% всіх випадків – 71 і 73. Відповідно, медіана – 72 (73+71 = 144/2 = 72), мода – 47,6. Зауважте, що в цьому випадку розбіжність між середнім та медіаною невелика, але мода від них значно відрізняється.

Таблиця 10.4. Відсоток міського населення в країнах Латинської та Центральної Америки

Країна	Відсоток населення, що проживає в місті станом на 2007 р.	Країна	Відсоток населення, що проживає в місті станом на 2007 р.
Аргентина	90	Гондурас	47
Болівія	64	Мексика	76
Бразилія	84	Нікарагуа	59
Чилі	88	Панама	71
Колумбія	73	Парагвай	59
Коста-Ріка	62	Перу	73
Еквадор	63	Уругвай	92
Гватемала	47	Венесуела	93

Джерело: World Bank. World Development Indicators 2008, 162–64.

Вибір показника середнього значення залежить від типу даних: номінальні, порядкові або інтервально-пропорційні (див. табл. 10.5).

- **Номінальні дані**, які іноді ще називають *категорійними*, – це дані, що відносяться до однієї з незбіжних категорій – таких як стать, релігія або країна походження.
- **Порядкові дані** – це дані, які можна розмістити на впорядкованій шкалі, але «відстань» між послідовними відповідями не обов'язково буде однаковою. Шкали зі значеннями від «найважливіший» до «найменш важливий» або від «абсолютно згоден» до «зовсім не згоден» ілюструють порядкові дані. В порядкових даних немає нульової точки.
- **Інтервально-пропорційні дані** – дійсні числа. Це дані з нульовою точкою й фіксованими інтервалами, як лінійка, і можуть бути розділені й порівняні з іншими відносними одиницями.

Номінальні дані:
Дані, що відносяться до однієї з незбіжних категорій – таких як стать, релігія або країна походження

Порядкові дані:
Дані, які можна розмістити на впорядкованій шкалі

У випадку з інтервально-пропорційними даними вибір також залежатиме від розподілу. Якщо мова йде про розподіл, наближений до нормального, середнє, медіана й мода мають бути розташовані дуже близько одне від одного. В цьому випадку середнє буде найкращим показником середнього значення. При цьому, якщо серед оцінок є кілька дуже високих або дуже низьких, середнє перестане бути наближеним до центру. В цьому випадку найкращим показником середнього значення буде медіана.

Таблиця 10.5. Бажані ступені центральної тенденції для різних типів даних

Вид даних	Найкращий показник середнього значення
Номінальні	Мода
Порядкові	Мода або медіана
Інтервально-пропорційні	Мода, медіана або середнє

Джерело: Автори.

Інтервально-пропорційні дані:
Дані з нульовою точкою й фіксованими інтервалами, як лінійка, що можуть бути розділені й порівняні з іншими відносними одиницями

Діапазон:
Різниця між найвищим та найнижчим значеннями змінної

Нормальний розподіл:
Розподіл ймовірностей випадкової величини, що описує дані, згруповані навколо середнього

Показники дисперсії. Традиційно для вимірювання розсіювання кількісних змінних використовуються два показники: діапазон та стандартне відхилення. **Діапазон** – це різниця між найвищим та найнижчим значеннями змінної. Для даних, наведених в табл. 10.4, діапазон для показника урбанізації складає 93 – 47 – 46.

Діапазон мало що виявляє, тому що визначається він лише за двома спостереженнями, решта – ігноруються. Якщо два кінцеві значення рознесені максимально далеко, діапазон не дає жодного уявлення про те, що відбувається з усіма іншими значеннями.

Найбільш розповсюджений показник дисперсії для інтервальних або пропорційних даних – стандартне відхилення. Стандартне відхилення – це показник розсіювання значень з обох боків середнього. Що більше значення відрізняються від середнього, то більшим буде стандартне відхилення.

Для того, щоб краще зрозуміти поняття стандартного відхилення, важливо зрозуміти сутність **нормального розподілу** (рис. 10.4), графік якого формою нагадує дзвін. За нормального розподілу більшість даних припадає на його середину. Що ближче до краю розподілу, то менше даних.

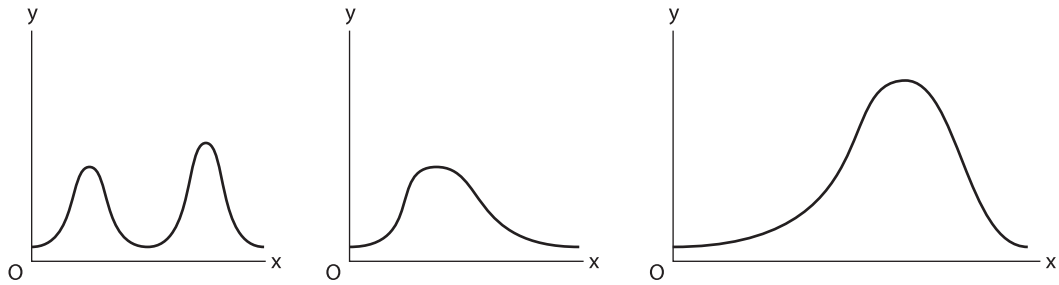
Дані не завжди розподіляються нормально. У деяких випадках криві більш пологі, в деяких – більш круті. Існують також криві, підняті з того чи іншого боку (див. рис. 10.5).

Стандартне відхилення показує, наскільки тісно дані розташовані навколо середнього. Воно вимірює відстань від середнього. За нормального розподілу одне стандартне відхилення від середнього в обох напрямках на горизонтальній осі охоплює близько 68 відсотків даних (див. рис. 10.6). Два стандартні відхилення від середнього охоплюють близько 95 відсотків даних. Три стандартні відхилення від середнього охоплюють близько 98 відсотків.

Рис. 10.4. Нормальний розподіл

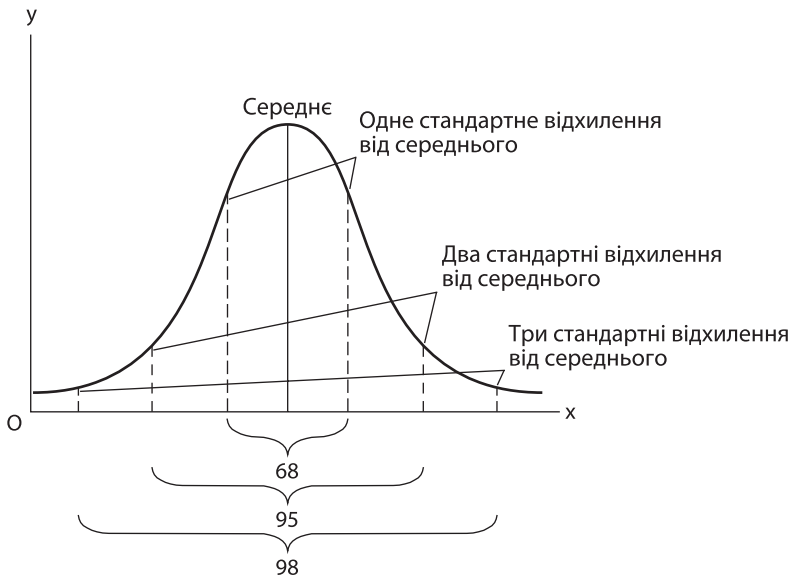
Джерело: Автори.

Рис. 10.5. Ненормальний розподіл



Джерело: Автори.

Рис. 10.6. Стандартні відхилення за нормального розподілу



Джерело: Автори.

Якби крива для масиву даних була більш пологою, стандартне відхилення було б більшим. Значення стандартного відхилення вказує на розсіювання даних навколо середнього (блок 10.5).

Якщо кожен на екзамені набере 75 балів, середнє буде дорівнювати 75, а стандартне відхилення – 0. Якщо кожен набере від 70 до 80 балів, середнє також буде дорівнювати 75, а стандартне відхилення буде меншим, ніж у випадку, коли діапазон балів – від 40 до 95 (і середнім – 75).

Блок 10.5 «Розрахунок стандартного відхилення»

Стандартне відхилення розподілу вираховується наступним чином:

1. Розрахуйте середнє значення даних.
2. Відніміть середнє від кожної точки даних, щоб знайти відхилення.
3. Піднесіть до квадрата кожне значення відхилення.
4. Складіть всі квадрати відхилень.
5. Відніміть 1 від числа точок даних.
6. Розділіть суму всіх квадратів відхилень на результат, отриманий на кроці 5 (кількість елементів списку мінус 1).
7. Добудьте квадратний корінь з результату, отриманого на кроці 6.

Формула для розрахунку стандартного відхилення має наступний вигляд:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})}{N - 1}},$$

де σ = стандартне відхилення, $\sum$ = сума, $\bar{x}$ = середнє.

Розрахунок стандартного відхилення навіть для невеликої вибірки займає багато часу. На щастя, більшість статистичних програм, включаючи Excel та SPSS для Windows, здатні виконати необхідні розрахунки.

Інакше кажучи:

Невелике стандартне відхилення = невелика дисперсія
Велике стандартне відхилення = значна дисперсія

Стандартне відхилення важливіше за діапазон, адже на його значення впливає кожний випадок.

Часто використовувані описові міри/показники. Іноді питання стосується конкретних цифр («Скільки у вас кіз? Як часто ви користуєтесь послугами банку?») Відповіді на такі питання даються в абсолютних числах та у відсотках.

В інших випадках респондентів просять скористатися шкалою. Наприклад, фахівець з оцінювання може попросити респондентів відповісти, чи застосовували вони отримані навички, за допомогою п’ятибальної шкали від «зовсім ні» до «багато разів». Аналізуючи цей різновид даних, треба встановити правило прийняття рішень: зосередьтеся на відсотку тих, хто дав відповідь, використовуючи граничні оцінки, тих, хто дав відповідь з будь-якого боку середньої категорії, або ж на середній відповіді. Хоча твердих правил тут не існує, деякі рекомендації можуть стати у нагоді (блок 10.6).

Блок 10.6 «Вказівки щодо аналізу даних кількісного дослідження»

Наступні вказівки можуть стати у нагоді фахівцям з оцінювання, які аналізують дані кількісного дослідження:

1. Оберіть стандартний спосіб аналізу даних і послідовно його застосовуйте.
2. Не поєднуйте середню категорію з категоріями з будь-якого кінця шкали.
3. Не давайте інформації щодо категорії «згоден» або «не згоден» у відриві від інформації щодо категорій «повністю згоден» або «зовсім не згоден» (якщо такі використовуються).
4. Аналізуйте й повідомляйте у звіті як відсотки, так і числа.
5. В якості орієнтиру вказуйте кількість респондентів.
6. Якщо результати мало відрізняються, зверніться до порівняння: як виглядають результати для питань, стосовно яких більшість респондентів відповіла «дуже задоволений» або «зовсім не згоден»?
7. Пам'ятайте, що аналіз даних – це одночасно мистецтво й уміння; тренування і практика значно полегшать процес.

Джерело: Автори.

Розглянемо дослідження клієнтів медичного центру (табл. 10.6). Один зі способів аналізу цих даних – повідомити, що половина респондентів згодні або абсолютно згодні з тим, що отримують якісні медичні послуги, а 55 відсотків згодні або абсолютно згодні з тим, що працівники лікарні готові відповідати на питання. Однак 60 відсотків згодні або абсолютно згодні з твердженням про те, що їм доводиться довго чекати прийому. В даному випадку фахівець з оцінювання вирішив повідомити про об'єднані відсотки категорій «згоден» і «абсолютно згоден».

Якби дані були іншими, фахівець з оцінювання міг би застосувати іншу стратегію.

Візьмемо, наприклад, результати, представлені в табл. 10.7.

Аналіз в цьому випадку міг би показати, що 80 відсотків респондентів згодні або абсолютно згодні з тим, що отримують якісні медичні послуги, а 70 відсотків згодні або абсолютно згодні з твердженням, що їм доводиться довго чекати прийому. Найбільш позитивним є показник готовності працівників закладу відповідати на питання: 95 відсотків респондентів згодні або абсолютно згодні.

Опис двох змінних в один момент часу. Іноді фахівці з оцінювання хочуть описати одночасно дві змінні. Уявімо, приміром, що фахівець з оцінювання хоче описати склад учнів на практичних заняттях і на лекціях. Для кожного заняття потрібно знати відсоток хлопчиків і дівчат. Аналіз даних показує, що на практичних заняттях хлопчиків – 55 відсотків, а дівчат – 45. При цьому на традиційних лекціях дівчат – 55 відсотків, хлопчиків – 45.

Таблиця 10.6. Думки клієнтів щодо медичних послуг у місцевій лікарні (відсоток респондентів)

1. Якщо говорити про ваші враження від відвідування місцевої лікарні, чи погоджуєтесь ви з наступними твердженнями?

Твердження	Зовсім не згоден	Не згоден	Важко відповісти	Згоден	Абсолютно згоден
Мені довго доводиться чекати прийому	10	20	10	35	25
Працівники лікарні готові відповідати на мої питання	5	10	30	30	25
Лікарня надає мені якісні медичні послуги	15	25	10	25	25

Джерело: Автори.

Примітка: N = 36.

Таблиця 10.7. Думки клієнтів щодо медичних послуг у місцевій лікарні (відсоток респондентів)

1. Якщо говорити про ваші враження від відвідування місцевої лікарні, чи погоджуєтесь ви з наступними твердженнями?

Твердження	Зовсім не згоден	Не згоден	Важко відповісти	Згоден	Абсолютно згоден
Мені довго доводиться чекати прийому	50	20	10	15	5
Працівники лікарні готові відповідати на мої питання	0	5	0	30	65
Лікарня надає мені якісні медичні послуги	0	20	0	55	25

Джерело: Автори.

Примітка: N = 36.

Крос-табуляція демонструє спільний розподіл двох або більше показників, зазвичай у вигляді матриці. Тоді як густина ймовірності свідчить про розподіл однієї змінної, таблиця спряженості описує розподіл двох або більше змінних одночасно. Кожна комірка показує відсоток і кількість респондентів, що дає конкретну комбінацію відповідей.

Крос-табуляція: Спільний розподіл двох або більше показників, зазвичай представлений у вигляді таблиці

Дані щодо відвідування занять можна інтерпретувати наступним чином: в даній вибірці хлопчики з дещо більшою ймовірністю (55 відсотків), ніж дівчата (45 відсотків) відвідуватимуть практичні заняття. Цей висновок свідчить про наявність зв'язку між статтю й відвідуваністю занять. Але наскільки сильний цей зв'язок? Розглядаючи показники зв'язку, важливо розуміти поняття незалежних та залежних змінних.

Незалежна змінна пояснює зміну значення іншої (залежної) змінної. Наприклад, під час оцінки курсу навчання до незалежних змінних можуть відноситися досвід викладачів, попередня підготовка учасників, навчальна програма, тривалість курсу, методика викладання тощо.

Незалежна змінна: Змінна, що пояснює зміну значення іншої (залежної) змінної

Залежні змінні потребують пояснень. У випадку з курсом навчання залежними можуть бути такі змінні, як бали, набрані під час тестування знань, рівні на матриці плану оцінки або ж покращені плани оцінки.

Залежна змінна: Змінна, яку треба пояснювати

Фахівців з оцінювання нерідко цікавить, чи існує різниця між середніми значеннями кількісної змінної для пари вибірок. Наприклад, їх може цікавити дослідження таких питань:

- Чи є середня врожайність після реалізації проекту зрошення вищою, ніж раніше?
- Чи виявляють опитування пацієнтів старої лікарні та лікарні, побудованої в рамках проекту розвитку, різницю у відсотку задоволених медичним обслуговуванням?

Фахівцям з оцінювання треба визначити, чи відображають наявні відмінності фактичну зміну середніх значень для сукупності, або ж мова йде про випадкові варіації у вибірках з двох генеральних сукупностей. Під час статистичного тесту зазвичай вважається, що різниці між двома середніми генеральної сукупності (або пропорціями) немає. Це питання розглядається далі в розділі, присвяченому статистиці висновків.

Міри зв'язку (або залежності) вказують, наскільки сильно пов'язані змінні. Зв'язок сам по собі не доводить причинності. Він здатен лише в загальному сенсі вказувати на можливу наявність причинно-наслідкових зв'язків, якщо показник зв'язку високий.

Міри зв'язку: Міри, що показують, наскільки сильно пов'язані змінні

Міри зв'язку зазвичай представляються за допомогою шкали від –1 до +1. Позитивне значення міри означає, що змінні рухаються в одному напрямку: зростають або знижуються. Це називають **прямою залежністю**. Ідеальній прямій залежності відповідає значення +1.

Пряма залежність: Залежність, за якої змінні рухаються в однаковому напрямку (зростають або знижуються)

Зворотна залежність:
Залежність, за якої змінні рухаються в протилежних напрямках

Негативне значення міри вказує на **зворотну залежність** між змінними. Це означає, що вони рухаються в протилежних напрямках (наприклад, зі збільшенням віку погіршується стан здоров'я). Ідеальній зворотній залежності відповідає значення –1. Що ближче цей показник до 0, то слабшою є залежність (0 вказує на відсутність будь-якої залежності); що ближче показник до +1 або –1, то залежність сильніша.

Статистика висновків (інференційна статистика)

Статистика висновків дозволяє фахівцям з оцінювання прогнозувати стан генеральної сукупності на основі випадкової вибірки з неї. Основна проблема з використанням випадкової вибірки полягає в тому, що результати відображають певну примхливість вибірки, а отже, не дають точного уявлення про генеральну сукупність. Якби фахівці з оцінювання сформували іншу вибірку, чи були б результати аналогічними або відмінними?

Перевірки статистичної значущості вимірюють імовірність отримання таких самих результатів у випадку повної однорідності всієї генеральної сукупності. Фахівці з оцінювання називають це нульовою гіпотезою; вона завжди базується на нульовій різниці всередині генеральної сукупності.

Уявімо, що опитування з використанням випадкової вибірки населення Пакистану показало, що середній річний дохід чоловіків та жінок відрізняється на 5000 рупій. Питання для перевірки можна було б сформулювати наступним чином: якби в середині генеральної сукупності не існувало жодних відмінностей, якою була б імовірність не знайти різницю в 5000 рупій у випадковій вибірці? Якщо ймовірність складає 5% (0,05) або менше, ми робимо висновок, що результати вибірки точно відображають стан генеральної сукупності, різниця у 5000 рупій має місце, і ця різниця статистично значуща.

5-відсоткова ймовірність помилки означає, що фахівець з оцінювання на 95% впевнений, що результати вибірки не випадкові. Це означає, що результати статистично значущі на рівні 0,05.

Всі перевірки статистичної значущості частково базуються на розмірі вибірки. Якщо вибірка дуже велика, невеликі відмінності, ймовірно, будуть статистично значущі. Фахівцям з оцінювання треба вирішити, чи важливими є відмінності з урахуванням природи дослідження. Три популярні види статистичної перевірки – *хі*-квадрат тест, *t*-тест і критерій дисперсійного аналізу. Далі наводиться стислий опис кожного з них.

***Хі*-квадрат тест:**
Статистичний тест, що використовується для визначення, чи існують явні відмінності між спостереженими частотами та очікуваними випадковими

Хі-квадрат тест. Хоча ***Хі*-квадрат тест** – не найефективніший показник зв'язку, це один з найпопулярніших статистичних тестів, тому що його легко розраховувати й інтерпретувати. Мета *хі*-квадрат тесту – визначити, чи існують явні відмінності між спостереженими частотами та очікуваними випадковими. Цей тест використовується для порівняння двох номі-

нальних величин (наприклад, шлюбного стану й віросповідання). Він також використовується для порівняння двох порядкових показників (відповідей з використанням шкали) або поєднання номінальних і порядкових змінних.

Статистика *хі*-квадрат – це сума внесків з кожної окремої комірки таблиці даних. Кожна комірка таблиці додає щось до загальної статистики *хі*-квадрат. Якщо конкретна комірка значно відрізняється від очікуваної частоти, її внесок буде великим, якщо наближена до очікуваної для неї частоти – її внесок буде незначним. Велике значення статистики *хі*-квадрат вказує на те, що – десь у таблиці – спостережені частоти помітно відрізняються від очікуваних. Тест не вказує на те, яка комірка (або комірки) зумовлює високе значення статистики *хі*-квадрат, а лише на те, що така комірка існує. *Хі*-квадрат вимірює, чи незалежні дві змінні одна від одної, базуючись на спостережених даних.

Хі-квадрат тест вимірює значущість крос-табуляцій. Значення хі-квадрат не слід розраховувати для відсотків. Для проведення тесту крос-табуляції мають конвертуватися в абсолютні показники (числа). Хі-квадрат тест також складно проводити, якщо в будь-якій комірці значення частоти нижче п'яти (поглиблене обговорення цієї проблеми див. у Fienberg 1980).

t-тест. Розглядаючи різницю між показниками двох груп, фахівці з оцінювання повинні оцінити різницю між їх середніми в контексті розсіювання або варіативності цих показників. Саме це й робить ***t*-тест**. Він використовується для визначення того, чи є одна з груп числових показників статистично вищою або нижчою за іншу. Цей аналіз варто проводити під час порівняння середніх для двох груп. Фахівці з оцінювання використовують *t*-статистику для порівняння середніх балів групи, що зазнала впливу проекту, з середніми балами контрольної групи.

***t*-тест:**
Статистичний тест, що використовується для визначення того, чи є одна з груп числових показників статистично вищою або нижчою за іншу

Дисперсійний аналіз (ANOVA). *t*-тест стає дуже незручним, якщо мова йде про порівняння трьох або більшої кількості груп. Якщо для оцінки потрібно одночасно порівняти середні для кількох різних груп, краще застосовувати **дисперсійний аналіз (ANOVA)**.

Дисперсійний аналіз (ANOVA):
Статистична техніка для оцінки різниці між масивами даних

ANOVA – це статистична техніка для оцінки різниці між масивами даних. Її можна застосовувати за допомогою програми EXCEL. ANOVA використовується для одночасного порівняння двох або більшої кількості середніх з метою встановлення, чи є спостережена зміна випадковою чи вона пов'язана з фактором або комбінацією факторів. Таким чином, ANOVA оцінює те, як номінальні незалежні змінні впливають на постійну залежну змінну. Припускається, що генеральні сукупності для всіх порівнюваних груп мають однакові значення стандартного відхилення (припущення про однорідність дисперсії), а також що вибірки з генеральної сукупності

сформовані довільно. Важливо переконатися в дійсності таких припущень, перш ніж застосовувати ANOVA. Тести в ANOVA базуються на використанні *F*-коефіцієнта – дисперсії, пов'язаної з експериментальною дією або ефектом, розділеної на дисперсію, зумовлену помилкою експерименту. Нульова гіпотеза полягає в тому, що цей коефіцієнт дорівнює 1,0 (тобто активний вплив дорівнює експериментальній помилці). Ця гіпотеза відкидається, якщо *F*-коефіцієнт достатньо великий, щоб імовірність того, що він дорівнює 1,0, була меншою за певний встановлений показник, приміром 0,05 (1 з 20).

Пов'язування кількісних і якісних даних

Майлз і Г'юберман (Miles and Huberman 1994) розглядають як можна пов'язати кількісні та якісні дані. Вони починають з цитування Фреда Керлінгера, авторитетного спеціаліста з кількісних досліджень: «Якісних даних не існує. Все – або 1, або 0». Тоді вони пропонують протилежну точку зору: всі дані, в принципі, є якісними.

Суперечки стосовно кількісних та якісних даних тривають в США вже багато років, однак в контексті розвитку ця проблема не піднімалася. Оцінювання розвитку використовує як кількісні, так і якісні методи, щоб зрозуміти оточуючий світ: «Кількість якостей, а також виміряна якість – це все просто кількісно виражений розмір величини» (Miles and Huberman 1994). Якісні методи дають більше контексту; кількісні підходи дозволяють переносити висновки на інші ситуації.

Майлз і Г'юберман відзначають, що пов'язування в плані дослідження кількісних та якісних даних дозволяє фахівцям з оцінювання:

- підтвердити або підкріпити кожен різновид даних за допомогою тріангуляції;
- доробити або уточнити аналіз для отримання більш детальної інформації;
- ініціювати нові напрями думок, звертаючи увагу на несподіванки чи парадокси, «вивертаючи» ідеї та використовуючи нові здогадки.

Грін, Караселлі та Грем (Greene, Caracelli, and Graham 1997) описують гносеологічну й політичну цінність поєднання методів у оцінці:

- гносеологічна: ми краще щось пізнаємо, якщо використовуємо кілька шляхів пізнання;
- політична: всі шляхи пізнання часткові; отже, численні, різноманітні шляхи пізнання потрібно цінувати й поважати.

Відзначається, що «хороша оцінка із застосуванням змішаних методів активно сприяє різноманітності способів мислення та спонукає цінувати

спільну роботу заради досягнення кращого розуміння», а «різні методи краще підходять до вивчення різних типів явищ».

Гокінз (Hawkins 2005) наводить перелік наступних переваг від використання інтегрованого підходу оцінювання, що застосовує змішані методи:

- Можна вбудувати перевірки узгодженості на основі процедур тріангуляції. Вони дозволять робити дві або більше незалежних оцінок ключових змінних.
- Можна отримати різні точки зору. Наприклад, фахівці з оцінювання вважають рівень доходів або споживання ключовими показниками достатку домашніх господарств, але при цьому вивчення практичних прикладів може виявити, що жінок більше турбують питання уразливості, безпомічності або незахищеності від насильства.
- Аналіз може проводитися на різних рівнях. Опитування можуть давати хороші оцінки добробуту на рівні окремої особи, домашнього господарства або громади, але вони набагато менш ефективні, якщо мова йде про аналіз соціальних процесів (наприклад, соціального конфлікту) або проведення інституційного аналізу (наприклад, визначення ефективності надання громадських послуг – на думку громади або фактичної ефективності). Чимало якісних методів призначено для аналізу таких питань, як соціальні процеси, інституціональна поведінка, соціальна структура, конфлікт.
- Можна надавати можливість зворотного зв'язку, що допоможе фахівцям з оцінювання інтерпретувати висновки. Звіти за результатами опитувань часто включають посилання на очевидні невідповідності у знайдених фактах або цікаві відмінності між групами, які не вдається пояснити шляхом аналізу даних. В більшості кількісних оцінок, щойно завершився етап збору даних, повернутися на місце й отримати додаткові дані вже неможливо.
- Фахівці з оцінювання, які проводять опитування, часто застосовують якісні методи для перевірки аномальних значень. В багатьох випадках аналітик має прийняти вольове рішення щодо того, чи слід виключати респондента, який повідомляє про умови, що значно відрізняються від норми.
- Вигоди від інтегрованих підходів залежать від підготовки фахівця з оцінювання. З точки зору спеціаліста з кількісних оцінок, оціночний компонент допоможе визначити ключові проблеми, які має розглянути оцінка, доопрацювати питання у відповідності з поглядами респондентів, а також надати інформацію про соціальний, економічний і політичний контекст, в якому проводиться оцінка. Також можна повернутися до польової роботи, щоб більше дізнатися про факти, які викликали зацікавленість.

- Спеціалісти з якісного оцінювання можуть отримати вигоду від використання кількісних методів. Можна застосувати методи формування вибірки, щоб забезпечити можливість перенесення висновків на генеральну сукупність. Процес формування вибірки можна узгодити з поточною або попередньою опитувальною роботою. Це дозволить порівнювати висновки якісного оцінювання з висновками опитування. Статистичний аналіз може знайти застосування для контролю характеристик домашніх господарств, а також соціально-економічних умов у різних районах, де ведеться дослідження. Це, в свою чергу, дозволить усунути альтернативні пояснення спостережених змін.

Гокінз (Hawkins 2005) розглядає, коли треба та коли не треба застосовувати змішаний підхід. За його словами, змішані методи слід застосовувати у наступних випадках:

- Потрібне поглиблене розуміння інтервенції та її контексту.
- Існують значні бюджетні й часові обмеження. Тріангуляція може допомогти з валідацією інформації, отриманої з різних джерел за допомогою різних методів і малих вибірок.

Змішані методи не варто застосовувати в наступних випадках:

- Для відповіді на питання достатньо одного методу.
- Потрібна можливість узагальнення висновків, і показники очевидні.
- Фахівці з оцінювання з досвідом застосування обраних методів не мають можливості брати участь в дослідженні протягом усього часу його проведення (краще користуватися тільки тими методами, спеціалісти з яких доступні).
- Ключові зацікавлені особи є настільки активними прибічниками певної парадигми, що відкидають всі інші та не довіряють змішаним підходам, як би добре не проводилася оцінка.
- На етап аналізу й інтерпретації відведено замало часу.

Підсумки

Якісний аналіз даних використовується для роботи з нечисловими даними. Якісні дані можна збирати, приміром, за допомогою неструктурованих спостережень, інтерв'ю з відкритими питаннями, аналізу письмових документів та розшифровок фокус-груп. Нотатки, що їх робить дослідник під час збору якісних даних, дуже важливі й повинні бути детальними.

Аналіз змісту – процес для аналізу якісних даних. Аналіз якісних даних трудомісткий і займає багато часу, але може виявити цінну інформацію.

Зібравши якісні дані, фахівці з оцінювання мають їх упорядкувати. Дані можна відсортувати таким чином, щоб виявити закономірності й спільності. Після сортування (вручну або електронним способом) дані можна кодувати, а потім – інтерпретувати.

Кількісні дані аналізують із застосуванням описової статистики та статистики висновків. Описова статистика підсумовує дані й описує центральне значення за допомогою середнього, медіани й моди. Для демонстрації дисперсії зазвичай використовують такі показники, як діапазон і стандартне відхилення. До засобів описової статистики також відносяться щільність ймовірності, відсотки, коефіцієнти, пропорції та темпи змін.

Статистика висновків дозволяє фахівцям з оцінювання прогнозувати стан генеральної сукупності на основі випадкової вибірки з неї. Стандартні інструменти статистики – *xi*-квадрат тест, *t*-тест і ANOVA.

Фахівці з оцінювання користуються як якісними, так і кількісними методами. Використання одночасно кількох методів у багатьох випадках є дуже корисним. У разі, якщо ставиться всього кілька легких питань, рекомендується застосовувати один підхід.

Завдання до Глави 10

Практична вправа 10.1: Кодування й аналіз якісних даних

Зберіть кілька великих газетних статей про проблеми розвитку. За допомогою цих статей підготуйте таблицю зі стовпцями «стаття» й «витяг». Потім внесіть до неї витяги зі статей, що здаються вам важливими. Визначивши за допомогою статей теми, додайте їх у якості заголовків наступних стовпців та відзначте комірку з відповідною темою для кожного витягу. Напишіть описовий текст з підсумками, які ви зробили зі статей.

Практична вправа 10.2: Уникнення поширених помилок інтерпретації кількісних даних

1. Вісімдесят відсотків респондентів зазначають, що програма була для них корисною. Як краще повідомити про результати опитування: «Програма корисна» або «Учасники вважають програму корисною»? Респондентів опитування попросили назвати фактори, що перешкоджають і допомагають програмі. У чому полягає проблема звітування результатів у вигляді позитивних і негативних моментів?
2. Дослідники попросили студентів оцінити різні компоненти курсу. Більшість студентів оцінили кожен з компонентів позитивно. Чому не варто звітувати, що більшість (70 відсотків) студентів вважають курс успішним?
3. Сорок відсотків жінок і 30 відсотків чоловіків схвалюють зміни у плані навчання. Чи правильно буде повідомити, що більшість жінок схвалює зміни у плані навчання?
4. П'ятдесят один відсоток респондентів схвалює зміни у плані навчання. Чи правильно буде повідомити, що більше половини респондентів підтримали зміни у плані навчання?
5. Було опитано п'ять із 20 інструкторів курсу. Всі п'ять повідомили, що готові викладати курс. Чи правильно буде сказати, що всі інструктори добре підготовлені? Чи правильно буде сказати, що 25 відсотків інструкторів добре підготовлені?
6. Кількість жінок-депутатів у законодавчому органі, до якого входить 50 членів, збільшилася з 2 до 4. Чи правильно буде говорити про збільшення на 100 відсотків?
7. Учасники програми навчання заробили на 20 відсотків більше, ніж ті, хто у програмі участі не брав. Чи правильно буде повідомити, що програма забезпечила зростання заробітної плати на 20 відсотків?

Практична вправа 10.3: Аналіз результатів анкетування

Заповніть анкету. Нехай такі ж анкети заповнять щонайменше двоє ваших колег. Зберіть анкети й підрахуйте результати. Працюючи самостійно або разом з іншими, підбийте підсумки в описовій формі. Зробіть висновки щодо загальних результатів.

1. Наскільки ви готові сказати, що здатні виконати наступні завдання?

Навичка	Ні або майже ні	Певною мірою	Готовий	Значною мірою	Повністю готовий
a. Спланувати оцінювання.					
b. Проаналізувати дані.					
c. Розробити анкету.					
d. Провести фокус-групу.					
e. Координувати зустріч зацікавлених осіб.					
f. Написати звіт оцінювання.					
g. Підготувати стислий усний звіт.					

2. Зараз, на цьому етапі даної навчальної програми, чи погоджуєтеся ви з кожним із наступних тверджень:

Твердження	Зовсім не згоден	Не згоден	Важко відповісти	Згоден	Абсолютно згоден
a. Матеріал новий для мене.					
b. Матеріал цікавий.					
c. Для лекцій виділено достатньо часу.					
d. Для обговорення в класі виділено достатньо часу.					
e. Вправи є корисними.					
f. Я вивчаю матеріал, який зможу використати у своїй роботі.					

3. Якщо ви хочете якось іще прокоментувати курс, зробіть це тут.

Посилання та рекомендована література

Babbie, E., F. Halley, and J. Zaino. 2000. *Adventures in Social Research.* Thousand Oaks, CA: Pine Forge Press.

Busch, Carol, Paul S. De Maret, Teresa Flynn, Rachel Kellum, Sheri Le, Brad Meyers, Matt Saunders, Robert White, and Mike Palmquist. 2005. *Content Analysis.* Writing@CSU, Colorado State University, Department of English, Fort Collins, CO. http://writing.colostate.edu/guides/research/content/.

Child and Adolescent Health Measurement Initiative (CAHMI). 2006. "Promoting Healthy Development Survey: Implementation Guidelines." Portland, OR: Oregon Health and Science University, Department of Pediatrics.

———. 2009. "Step 4: Monitor Survey Administration and Prepare for Analysis." http://www.cahmi.org.

Catterall, M., and P. Maclaran. 1997. "Focus Group Data and Qualitative Analysis Programs: Coding the Moving Picture as Well as the Snapshots." *Sociological Research Online* 2 (1). http://www.socresonline.org.uk/socresonline/2/1/6.html.

Chapman, Arthur D. 2005. *Principles and Methods of Data Cleaning: Primary Species-Occurrence Data.* Report for the Global Biodiversity Information Facility, Version 1.0, Copenhagen. http://www2.gbif.org/DataCleaning.pdf.

Cody, Ronald. n.d. "Data Cleaning 101." Robert Wood Johnson Medical School, Piscatawy, NJ. http://www.ats.ucla.edu/stat/sas/library/nesug99/ss123.pdf.

Constable, Rolly, Marla Cowell, Sarita Zornek Crawford, David Golden, Jake Hartvigsen, Kathryn Morgan, Anne Mudgett, Kris Parrish, Laura Thomas, Erika Yolanda Thompson, Rosie Turner, and Mike Palmquist. 2005. *Ethnography, Observational Research, and Narrative Inquiry.* Writing@CSU, Colorado State University, Department of English, Fort Collins, CO. http://writing.colostate.edu/guides/research/observe/.

Dehejia, Rajeev H., and Sadek Wahba. 2002. "Propensity Score-Matching Methods for Nonexperimenal Causal Studies." *Review of Economics and Statistics* 84 (1): 151–61.
http://www.nber.org/~rdehejia/papers/matching.pdf.

Denzin, N., and Y. Lincoln, eds. 2000. *Handbook of Qualitative Research.* 2nd ed.
Thousand Oaks, CA: Sage Publications.

Fienberg, S. E. 1980. The *Analysis of Cross-Classified Categorical Data.* 2nd ed. Cambridge, MA: MIT Press.

Firestone, W. 1987. "Meaning in Method: The Rhetoric of Quantitative and Qualitative Research." *Educational Researcher 16* (7): 16–21. http://edr.sagepub.com/cgi/content/abstract/16/7/16.

Gao, Jie, Michael Langberg, and Lenard J. Schulman. 2006. *Analysis of Incomplete Data and an Intrinsic-Dimension Helly Theorem.* http://www.cs.sunysb.edu/~jgao/paper/clustering_lines.pdf.

Glass, G., and K. Hopkins. 1996. *Statistical Methods in Education and Psychology.* 3rd ed. Boston: Allyn and Bacon.

Greene, J. C., and V. J Caracelli. 1997. *Advances in Mixed Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms.* New Directions for Evaluation No. 74. San Francisco: Jossey-Bass.

Greene, J. C., V. J. Caracelli, and W. F. Graham. 1989. "Toward a Conceptual Framework for Mixed-Method Evaluation Designs." *Educational Evaluation and Policy Analysis* 11 (3) 255–74.

Hawkins, Penny. 2005. "Thinking about Mixed Method Evaluation." International Program for Development Evaluation Training (IPDET) presentation, Ottawa, July.

IDRC (International Development Research Centre). 2008. *Qualitative Research for Tobacco Control, Module 6: Qualitative Data Analysis.* Ottawa, Canada. http://www.idrc.ca/en/ev-106563-201-1-DO_TOPIC.html.

Jaeger, R. M. 1990. *Statistics: A Spectator Sport.* 2nd ed. Thousand Oaks, CA: Sage Publications.

Krippendorf, Klaus. 2004. *Content Analysis: An Introduction to Its Methodology.* 2nd ed. Thousand Oaks, CA: Sage Publications.

Loughborough University Department of Social Sciences. 2007. *New Methods for the Analysis of Media Content. CAQDAS: A Primer.* Leicestershire, United Kingdom. http://www.lboro.ac.uk/research/mmethods/research/software/caqdas_ primer.html#what.

Miles, Mathew. B., and A. Michael Huberman. 1994. *Qualitative Data Analysis an Expanded Sourcebook.* 2nd ed. Thousand Oaks, CA: Sage Publications.

Morse, Janice M., and Lyn Richards. 2002. "The Integrity of Qualitative Research." In *Read Me First for a User's Guide to Qualitative Methods,* ed. J. M. Morse and

L. Richards, 25–46. Thousand Oaks, CA: Sage Publications.

Narayan, Deepa. 1996. "Toward Participatory Research." World Bank Technical Paper 307, Washington, DC.

http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/ IB/1996/04/01/000009265_3980625172923/Rendered/PDF/multi0page.pdf.

NCSTE (Chinese National Centre for Science and Technology Evaluation), and IOB (Policy and Operations Evaluation Department. 2006. *Country-Led Joint Evaluation of the ORET/MILIEV Programme in China.* Amsterdam: Aksant Academic Publishers.

Neuendorf, Kimberly A. 2006. *The Content Analysis Guidebook Online.* http://academic.csuohio.edu/kneuendorf/content/index.htm.

O'Rourke, Thomas W. 2000a. "Data Analysis: The Art and Science of Coding and Entering Data." *American Journal of Health Studies* 16 (3): 164–66. http://Findarticles.com/p/articles/mi_m0CTG/is_3_16/ai_72731731.

———. 2000b. "Techniques for Screening and Cleaning Data for Analysis." *American Journal of Health Studies* 16: 217–19. http://Findarticles.com/p/articles/mi_m0CTG/is_4_16/ai_83076574.

P.A.N.D.A. (Practical Analysis of Nutritional Data). 2000. "Chapter 2: Data Cleaning." http://www.tulane.edu/~panda2/Analysis2/datclean/dataclean.htm.

Patton, Michael. Q. 2002. *Qualitative Research and Evaluation Methods.* 3rd ed. Thousand Oaks, CA: Sage Publications.

Porteous, Nancy. 2005. "A Blueprint for Public Health Management: The ToolKit's Approach to Program Evaluation." Presentation at the International Program for Development Evaluation Training, June–July 2005. Ottawa, Ontario.

Porteous, Nancy L., B. J. Sheldrick, and P. J. Stewart. 1997. *Program Evaluation Tool Kit: A Blueprint for Public Health Management.* Ottawa-Carleton Health Department, Ottawa.

http://www.phac-aspc.gc.ca/php-psp/tookit.html.

Rahm, Erhard, and Hong Hai Do. 2000. "Data Cleaning: Problems and Current Approaches." University of Leipzig, Germany. http://homepages.inf.ed.ac.uk/wenfei/tdd/reading/cleaning.pdf.

Rossman, G. B., and B. L. Wilson. 1994. "Numbers and Words Revisited: Being 'Shamelessly Methodologically Eclectic.'" *Quality and Quantity* 28: 315–27.

Sieber, S. D. 1973. "The Integration of Fieldwork and Survey Methods." *American Journal of Sociology* 78 (6):1335–59.

Smith, J. A. 1991. *The Idea Brokers: Think Tanks and the Rise of the New Policy Elite.* New York: Free Press.

StatSoft. 2008. *Electronic Textbook.* http://www.statsoft.com/textbook/glosp.html. Stemler, Steve 2001. "An Overview of Content Analysis." *Practical Assessment, Research & Evaluation* 7 (17). http://PAREonline.net/getvn.asp?v=7&n=17.

Swimmer, Gene. 2006. "Qualitative Data Analysis, Part I." *IPDET Handbook 2005.* International Program for Development Evaluation Training (IPDET) presentation, Ottawa, July.

University of the West of England. 2006. *Analysis of Textual Data.* Bristol. http://hsc.uwe.ac.uk/dataanalysis/qualTextData.asp.

U.S. GAO (General Accounting Office). 1996. *Content Analysis: A Methodology for Structuring and Analyzing Written Material.* GAO/PEMD-10.3.1. Washington, DC. (Available free of charge.)

Washington State University. 2000. A *Field Guide to Experimental Designs: What Is an ANOVA*? Tree Fruit Research and Extension Center. Wenatchee, WA. http://www.tfrec.wsu.edu/ANOVA/basic.html.

Weber, Robert Philip 1990. *Basic Content Analysis.* 2nd ed. Thousand Oaks, CA: Sage Publications.

Weiss, Christopher, and Kristen Sosulski. 2003. *Quantitative Methods in Social Sci- ence: QMSS E-lessons.* Columbia Center for New Media Teaching and Learning, Columbia University, New York. http://www.columbia.edu/ccnmtl/projects/qmss/anova_about.html.

Wolcott, H. F. 1990. "On Seeking–and Rejecting–Validity in Qualitative Research." In *Qualitative Inquiry in Education: The Continuing Debate,* eds. E. W. Eisner and A.Peshkin, 121–52. New York: Teachers College Press.

Веб-сайти

Навчальні посібники онлайн

CAQDAS (Computer-Assisted Qualitative Data Analysis Software). 2008. http:// caqdas.soc.surrey.ac.uk/.

Lane, D. M. Hyperstat Online Textbook. http://davidmlane.com/hyperstat/index.html.

Statistics at Square One. http://bmj.bmjjournals.com/collections/statsbk/index.shtml.

Stat Primer. http://www2.sjsu.edu/faculty/gerstman/Stat Primer.

Комп'ютерне програмне забезпечення для якісних даних

AEA (American Evaluation Association). http://www.eval.org/Resources/QDA.htm.

AnSWR. http://www.cdc.gov/hiv/topics/surveillance/resources/software/answr/index.htm (developer site), http://www.cdc.gov/hiv/software/answr/ver3d.htm (free download site).

Atlas-ti. http://www.atlasti.com/.

CDC EZ-Text. http://www.cdc.gov/hiv/topics/surveillance/resources/software/ez-text/index.htm (developer site), http://www.cdc.gov/hiv/software/ez-text.htm (free trial download site)

Ethnograph. http://www.qualisresearch.com/default.htm.

Friese, Susanne. 2004. Overview/Summary of Qualitative Software Programs. http:// www.quarc.de/software_overview_table.pdf.

Hyperqual. http://home.satx.rr.com/hyperqual/.

QSR Software. NVivo 8. http://www.qsrinternational.com/. Qualpro. http://www.qualproinc.com/.

Комп'ютерне програмне забезпечення

SPSS. http://www.spss.com.
OpenStat. Version 4. http://www.statpages.org/miller/openstat/. Tutorial for SPSS v. 11.5. http://www.datastep.com/SPSSTraining.html/.
Getting Started with SPSS for Windows. http://www.indiana.edu/~statmath/stat/ spss/win/.
WinPure. Data Cleaning Software. http://www.winpure.com.

Приклади аналізу даних

Carleton University. *Canadian Foreign Policy Journal.* http://www.carleton.ca/ npsia/cfpj.
IISD (International Institute for Sustainable Development). http://www.iisd.org/ measure/default.htm.
IMF (International Monetary Fund). http://www.imf.org/external/pubs/res/index.htm.
———. http://www.imf.org/external/np/sta/index.htm.
North-South Institute. http://www.nsi-ins.ca/ensi/research/index.html.
OECD (Organisation for Economic Cooperation and Development). http://www.oecd.org/dac/.
UNDP (United Nations Development Programme). *Human Development Report 2002.* http://www.undp.org/hdr2002.
UNEP (United Nations Environmental Programme). http://www.grid.unep.ch. UNESCO (United Nations Educational, Scientific, and Cultural Organization). http://www.uis.unesco.org/en/stats/stats0.htm.
UNHCR (United Nations High Commission for Refugees). http://www.unhcr.ch/ cgi-bin/texis/vtx/home.
UNSD (United Nations Statistics Division). http://unstats.un.org/unsd/databases.htm.
USAID (U.S. Agency for International Development.) http://www.usaid.gov/educ_ training/ged.html.
———. http://www.dec.org/partners/eval.cfm.
U.S. Census Bureau. http://www.census.gov/ipc/www/idbnew. html. WHO (World Health Organization). http://www.who.int/health-systems-performance.
———. Statistical Information System (WHOSIS). http://www.who.int/topics/ statistics/en/.
World Bank. http://www.worldbank.org/data.
———. *World Development Indicators 2005.* http://worldbank.org/data/wdi2005.

ПОДОЛАННЯ ВИКЛИКІВ

«Докладаючи всіх можливих зусиль, ви здатні творити незбагненні дива зі своїм життям та життям інших людей»

—ГЕЛЕН КЕЛЛЕР

Глава 11: Оцінювання комплексних інтервенцій

- Загальна картина через призму оцінювання розвитку
- Спільні оцінювання
- Оцінювання національних програм
- Оцінювання секторальних програм
- Тематичні оцінювання
- Оцінювання програм глобального й регіонального партнерства

ГЛАВА 11

Оцінювання комплексних інтервенцій

Інтервенції стають все складнішими для оцінювання. З'являються нові виклики, сподівання, парадигми боротьби з бідністю, методи, клієнти, фахівці з оцінювання. В цьому розділі розглядається оцінювання на об'єднаному, національному, тематичному, секторальному та глобальному рівнях.

Ця глава складається з шести основних частин:

- Загальна картина через призму оцінювання розвитку
- Спільні оцінювання
- Оцінювання національних програм
- Оцінювання секторальних програм
- Тематичні оцінювання
- Оцінювання програм глобального й регіонального партнерства

Загальна картина через призму оцінювання розвитку

Іноді оцінювання має давати уявлення про загальну картину. Це означає, що треба виходити за межі окремого проекту, програми або політики й оцінювати також супутні проекти, програми та політики.

Загальна картина може охоплювати узагальнений досвід й вплив інтервенцій розвитку в межах сектора (наприклад, охорони здоров'я, освіти, транспорту). Приміром, міністерство хоче визначити загальний вплив і зробити висновки з досвіду інтервенцій, спрямованих на покращення добробуту дітей або жінок; донор прагне вивчити ефективність своєї стратегії в освітньому секторі.

Все частіше, у відповідь на Паризьку декларацію й Аккрську угоду, донорські програми базуються на секторальних підходах. Складні економічні, політичні та соціальні фактори впливають на заходи розвитку й оцінки. Розвиток і оцінювання розвитку все частіше набувають багатостороннього характеру, до них залучаються неурядові організації – приватний сектор, неприбуткові й правозахисні агенції, а також громадянське суспільство в цілому.

Інтервенції розвитку не є лінійними. Хоча за ними не стоїть логіка типу «якщо… то…», багато факторів – економічних, політичних, кліматичних та соціальних – поширюють свою дію на програму й можуть вплинути на її довгострокові результати. Фахівці з оцінювання мають визначати ці фактори й планувати оцінювання таким чином, щоб можна було виміряти їхній вплив.

Складність оцінювання зростає, і взаємодія цих факторів теж стає складнішою. Країни та їх партнери прагнуть визначити кумулятивні наслідки змін у секторі, країні, або ж крос-тематичних проблем – таких як зміна клімату.

Фахівцям з оцінювання доводиться проводити оцінювання на фоні цієї складності. Але складність оцінювання зростає не просто через розширення масштабів. Зростає обізнаність про те, якою важкою ношею є оцінювання для країн, де відповідні структури лише з'являються. У зв'язку з цим Паризька декларація та Аккрська угода підкреслили необхідність партнерських оцінювань якості допомоги країнам, а також результатів такої допомоги.

Розвиток рухається в напрямку більш комплексних програм, охоплюючи політичні реформи, нарощування потенціалу країн і глобальні питання. Оцінювання, в свою чергу, також розширило свою дію за рахунок:

- переорієнтації з рівня окремого проекту, програми або діяльності на національний, секторальний, тематичний, регіональний або глобальний рівень;
- визначення найкращих шляхів узагальнення довгострокових результатів інтервенцій на рівні діяльності та країни з метою оцінки загальнопрограмних або глобальних результатів;
- знаходження шляхів оцінки впливу плану програми, партнерського підходу та управління на загальні результати;

- прагнення до відтворюваності на вищому рівні та застосовності на рівні систем (Heath, Grasso, and Johnson 2005).

Національні оцінювання – один зі способів отримання загального уявлення про те, що відбувається, та розуміння загального ефекту й досвіду розвитку в країні. Але не завжди зрозуміло, як саме проводити такі оцінювання. Секторальні чи тематичні оцінювання – в масштабах однієї або кількох країн – також здатні показати загальну картину. Вони, скоріше за все, потребуватимуть використання кількох методів, в тому числі поєднання наявних даних, інтерв'ю, польових візитів, опитувань та фокус-груп.

У даній главі розглядаються п'ять видів досліджень загальної картини:

- Об'єднані оцінювання
- Оцінювання національних програм
- Оцінювання секторальних програм
- Тематичні оцінювання
- Оцінювання програм глобального й регіонального партнерства

Спільні оцінювання

Спільне оцінювання: Оцінювання, участь в якому бере більш ніж одна організація

Спільні оцінювання – оцінювання, участь в яких бере більш ніж одна організація.

> Є різні рівні «спільності», що визначаються мірою співпраці партнерів у процесі оцінки, об'єднанням їхніх ресурсів та поєднанням звітів. Спільні оцінювання допомагають вирішити проблеми атрибуції в оцінці дієвості програм і стратегій, компліментарність зусиль різних партнерів, якість координації допомоги і т. д. (OECD 2002, p. 26).

Об'єднані оцінювання проводяться з початку 1990-х років. Їх кількість і якість зросли після Паризької декларації 2005 р. та Аккрської угоди 2008 р. Однак ідея, що полягала у використанні національних систем для отримання даних, орієнтованих на результат, усе ще слабка. Оцінка Паризької декларації вказує на те, що кількість спільних оцінювань зростає, вони стають все більш якісними та корисними. Тим не менш, донори, що не мають змоги розраховувати на надійні національні системи даних, усе ще стикаються з необхідністю звітування про результати з використанням своїх власних систем. Існують наступні схеми організації об'єднаного оцінювання:

- донор + донор;
- донор + країна-партнер;
- мультидонор + мультипартнер;
- партнер + партнер (Chen and Slot 2007).

Брайєр (Breier 2005) визначає типологію таких оцінювань на основі режиму роботи. Згідно з класичною мультипартнерською схемою, участь (на рівних умовах) можуть брати всі зацікавлені партнери, які активно співпрацюють. За мультипартнерською схемою з обмеженнями участь відкрита лише для тих, хто відповідає певним вимогам – наприклад, належить до певної групи (приміром, Європейського Союзу) або має сильну зацікавленість в питанні, що оцінюється.

Планування спільних оцінювань – необхідна умова його успіху. Брайєр зазначає:

> Один з майже універсальних висновків з недавнього досвіду об'єднаних оцінювань – надзвичайно важливо на початку процесу виділити достатньо часу для розробки й узгодження спільного розуміння мети, цілей, основних напрямів та рамок запропонованої оцінки. Це вимагає наполегливості й терпіння. Якщо їх не буде – як це траплялося з багатьма проаналізованими оцінками – існує висока ймовірність проблем на подальших етапах (2005, p. 40).

Якщо у спільному оцінюванні беруть участь лише кілька агенцій, структура управління може бути простою:

- Фахівці з оцінювання регулярно зустрічаються й обмінюються інформацією про всі управлінські рішення.
- Фахівці з оцінювання однаковою мірою залучають всі агенції до управління, але призначають одну або кілька з них керівними.
- Фахівці з оцінювання делегують керівні повноваження одній агенції, дозволяючи решті агенцій переглядати ключові довгострокові результати (OECD 2006).

Якщо ж мова йде про масштабніші спільні оцінювання, найпоширенішою є дворівнева система, що складається із:

> (a) керівного комітету з широким представництвом;
>
> (b) меншої за розмірами адміністративної групи, яка займається повсякденними справами оцінки. В межах цієї структури існує значна свобода щодо прийняття рішень: чи будуть якісь агенції брати участь в якості пасивних партнерів, на якому рівні деталізації буде керівний комітет залучатися до прийняття рішень, а також скільки партнерів мають бути представлені в адміністративній групі та скільки відповідальності потрібно делегувати (OECD 2006, p. 20).

Інший варіант – гнучкий, або децентралізований підхід. Кожна агенція може керувати окремими субкомпонентами оцінювання – це схоже на складання пазлу. Також можна запровадити змішаний підхід, за якого певні

компоненти процесу оцінювання проводяться разом, в той час як інші реалізуються окремими партнерами (OECD 2006).

І централізовані, і децентралізовані структури мають сильні й слабкі сторони. Децентралізація управління оцінюванням полегшує делегування або розподіл повноважень, що забезпечує вищу ефективність процесу. Однак децентралізовані структури можуть також створювати дублювання роботи або випускати з поля зору важливі питання. Використання централізованої структури управління дозволяє кожному з партнерів докластися до всіх компонентів процесу оцінювання й впливати на них. Проте за цієї системи партнерам бракує відповідальності й наполегливості, які забезпечує децентралізована стратегія. Партнерам потрібно розглянути плюси й мінуси кожної структури, а потім прийняти рішення, яка з них найкраще підходить для потреб даного оцінювання.

КСР визначає наступні ключові сфери, в яких партнери з оцінювання повинні досягти згоди:

- базові правила управління оцінюванням;
- технічне завдання оцінювання;
- методи збору й аналізу даних;
- засоби вибору групи оцінювання (проведення тендеру й укладання договору);
- бюджетування, розрахунок витрат, фінансування;
- звітування про результати, включаючи правила розголошення даних (OECD 2006).

Брайєр (Breier 2005) розглядає шляхи вирішення проблем юридичного характеру для проведення спільного оцінювання. До таких проблем, зокрема, відносяться:

- договірні потреби різних партнерів;
- погоджені процедури, що зазвичай відображають юридичну систему, вимоги та прийняту практику агенції, яка керує від імені групи;
- вибір між акордними договорами та договорами без проведення тендера, включно з умовами розривання контрактів (у разі незадовільного виконання зобов'язань);
- умова щодо подачі звітів про хід виконання робіт, які мають демонструвати цільове використання коштів.

Після впровадження офіційної структури й базових правил лідери об'єднаного оцінювання мають звернутися до операційних та адміністративних питань. Фрімен (Freeman 2007) описує 16 правил організації та управління зовнішньою командою оцінки (табл. 11.1).

Таблиця 11.1. Вибір, організація й управління зовнішньою групою оцінювання та звітування про результати

Питання	Правило
Вибір зовнішньої групи оцінювання	• Переконайтеся, що група має необхідний досвід проведення складних, масштабних оцінок співпраці з розвитку.
	• Максимально спрощуйте організацію. Якщо можливо, працюйте з тими суб'єктами, з якими ви вже працювали.
	• Інтегруйте національних консультантів у міжнародний тендерний процес, підключіть їх до вибору й планування методології.
	• Якщо дослідження охоплює більш ніж одну країну, зробіть так, щоб кожна польова група користувалася ресурсами інших організацій. Уникайте географічної або інституціональної спеціалізації.
Управління зовнішньою групою оцінювання	• На рівні керівної ради головної зовнішньої організації оцінки чітко заявіть про свій намір активно використовувати оцінку в своїй роботі.
	• Проводьте семінари для групи оцінювання, щоб виробити спільний підхід до вимірювань і звітування.
	• Завжди, якщо це можливо, створюйте адміністративну групу, що буде звітувати перед загальним керівним комітетом оцінювання.
	• Відведіть достатньо часу на засіданнях керівного комітету для повного обговорення й вироблення спільної позиції.
	• Під час об'єднаних оцінювань переконайтеся, що зовнішні фахівці з оцінювання працюють відкрито й прозоро та прямо пов'язують методи із зібраними даними, висновками і рекомендаціями.
	• Якщо знадобилися додаткові ресурси і ви ведете відповідні переговори, спочатку адміністративна група та група оцінювання повинні погодити розділ праці. Окремо розглядаються завдання, що мають бути виконані за першим договором (з використанням вихідного пакету ресурсів), з одного боку, і такі, що є результатом нових проблем та інтересів або пов'язані з непередбачуваними обставинами – з іншого. Для цього група оцінювання мусить підготувати детальні плани для нових робіт із зазначенням грошових і часових потреб.
Організація зовнішньої адміністративної групи	• Переконайтеся, що провідна організація в консорціумі серйозно ставиться до оцінки й має необхідний досвід міжнародної співпраці у сфері розвитку. Проект має повністю відповідати її спеціалізації та ринковій орієнтації.
	• Переконайтеся, що зацікавлені особи протягом всього часу зберігають позитивний настрій.
	• Переконайтеся, що зовнішня група оцінювання, яка працює з усіма членами керівного комітету, уникає інституційної упередженості й однаково поводиться з усіма.
	• Відведіть істотну частину бюджету на поширення результатів і додаткові заходи після закінчення оцінювання.
Звітування про результати зовнішньої групи оцінювання	• Перш ніж представляти сам проект звіту, представте попередні висновки.
	• Нехай фахівці з оцінювання й члени керівного комітету обговорять проекти документів з готовністю покращувати їх. При цьому фахівці з оцінювання повинні мати змогу й бажання зберігати відповідальність за результати та висновки оцінки.

Джерело: За Freeman 2007.

Ще один цінний ресурс – «Керівні вказівки з управління об'єднаним оцінюванням» КСР (посилання на відповідний веб-сайт наведено в кінці цієї глави).

Оцінювання національних програм

У центрі уваги в загальній картині часто опиняється допомога країні. **Оцінювання національної програми** (яку іноді ще називають оцінкою допомоги країні) охоплює всю програму допомоги країні.

Оцінювання національної програми:
Оцінка всієї програми допомоги країні

Оцінювання національної програми – переважно нормативне дослідження, яке порівнює те, що вдається зробити, з тим, що було заплановано. Воно може бути спрямоване на:

- оцінку стратегічної релевантності програми допомоги країні в контексті її потреб;
- перевірку реалізації довгострокових результатів агенцій, аби встановити, чи були досягнуті цільові довгострокові результати;
- визначення успіхів і невдач у різних секторах або у випадках застосування різних підходів, а також визначення факторів, що вплинули на результати;
- визначення дієвості донорської допомоги даній країні (OECD 1999).

Оцінювання допомоги країні зазвичай концентрується на встановлених КСР критеріях релевантності, ефективності, впливу й сталості. Розглядатися можуть показники роботи донорів та/або країни (блок 11.1).

Під час оцінювання програми можуть виникати істотні проблеми:

- Загальній допомозі країні може бракувати чітких цілей та добре визначених бажаних довгострокових результатів; замість цього використовується кон'юнктурний підхід.
- Фінансування схожих інтервенцій розвитку може надходити з кількох джерел, що ускладнює атрибуцію результатів. (Примітка: співфінансування – це не те саме, що фінансування донорами схожих інтервенцій.)
- Зазвичай картування допомоги в країні не ведеться, через що важко встановити, чим займаються інші організації в районі проведення інтервенції.
- Як і в будь-якій оцінці, «на кону репутації, а страх наслідків може стати загрозою для морального духу й активності співробітників програми та партнерів. Оцінювання національної програми, як і будь-якої іншої, повинно проводитися делікатно» (OECD, 1999, p. 18).

Мережа з оцінювань розвитку КСР дає наступні рекомендації:

Блок 11.1 «Приклад методології оцінювання національної програми»

Починаючи з 1995 року Незалежна група оцінювання (НГО) Світового банку провела більше 70 оцінок національних програм і розробила чітко сформульовану методологію (World Bank 2008). Методологія оцінювання НГО передбачає рух знизу вгору і згори вниз. Під час оцінювання основних цілей визначаються:

- актуальність цілі;
- відповідність стратегії Банку в контексті досягнення цілі, в тому числі рівновага між кредитними та некредитними інструментами;
- ефективність реалізації стратегії;
- досягнуті результати.

Оцінювання проводиться у два кроки. Перший – перевірка, спрямована згори донизу. Встановлюється, чи вдалося програмі виконати конкретну мету Банку або досягти запланованого довгострокового результату й істотно вплинути на розвиток країни. Другий крок – перевірка знизу вгору. Перевіряються продукти й послуги Банку (кредитування, аналітично-консультаційні послуги, координація допомоги), їх використання для досягнення цілей. Разом ці два кроки забезпечують перевірку достовірності висновків щодо продуктів і послуг, а також масштабів впливу в контексті розвитку. Відповідно, оцінюється відносний внесок Банку, інших донорів, уряду й зовнішніх факторів у отримані результати.

Якщо НГО оцінює очікуваний вплив програми допомоги на розвиток, вона вимірює, наскільки актуальними були основні стратегічні цілі, та наскільки їх вдалося досягти. Зазвичай цілі програм виражаються через завдання вищого порядку – такі як Цілі розвитку тисячоліття (ЦРТ) та скорочення бідності. Стратегія допомоги країні також може передбачати безпосередні цілі, наприклад посилення цільового характеру соціальних послуг або сприяння інтегрованому розвитку села. Також може вказуватися, яким чином програми повинні сприяти досягненню цілей вищого порядку.

Оцінювання покликане перевірити, чи забезпечили проміжні цілі задовільні чисті вигоди та чи валідним виявився ланцюжок результатів, передбачений Стратегією допомоги країні. Якщо причинно-наслідкові зв’язки не були вказані, фахівець з оцінювання має відтворити їх, виходячи з наявних даних. Фахівцю з оцінювання також потрібно оцінити актуальність, ефективність та довгострокові результати проміжних цілей і цілей вищого порядку.

Фахівці з оцінювання також оцінюють те, наскільки відповідально ставляться клієнти до міжнародних пріоритетів розвитку. До таких пріоритетів відносяться ЦРТ, заходи із захисту прав людини, демократизації, забезпечення гендерної рівності й охорони довкілля. В ідеальному випадку ці питання визначаються й регулюються Стратегією допомоги країні, що дозволяє фахівцю з оцінювання зосередитися на тому, чи відповідними були запроваджені для такої допомоги підходи. Стратегія могла згладити певні конфлікти або уникнути розгляду проблем розвитку ключових клієнтів. Це могло призвести до зниження релевантності програми й почуття відповідальності у клієнта, а також до збільшення небажаних побічних ефектів, таких як втрата дієвості запобіжних засобів. Всі ці фактори варто брати до уваги, розглядаючи результати програми. Важливим є те, що, навіть якщо проекти Світового банку спрацювали добре, невдале виконання плану самою країною може призвести до того, що допомога вважатиметься незадовільною (і навпаки).

- Більша частина оцінювань має проводитися разом, за повної та активної участі реципієнтів допомоги та інших партнерів.
- Країни, що розвиваються, мають брати на себе більше ініціативи щодо планування, координування й складання графіків оцінювань.
- Країни, що розвиваються, мають отримувати допомогу в нарощуванні інституційних можливостей для ініціювання об'єднаних оцінювань і відігравання в них головної ролі.
- Потрібно забезпечити кращу координацію й обмін знаннями серед різних партнерів всередині країн, що отримують допомогу. Необхідно створити й розширити національні мережі й фахові об'єднання моніторингу та оцінювання.
- Якщо проводиться масштабне об'єднане оцінювання за участі кількох країн, що розвиваються, потрібно забезпечити умови для представників цих країн з метою обговорення й координування поглядів та внесків (OECD 2005, p. 7).

Група співпраці з оцінювання Міжнародного банку розвитку нещодавно опублікувала Стандарти добросовісної практики національних оцінювань. Ці стандарти покликані сприяти подальшій гармонізації критеріїв, процесів і процедур оцінки, а також забезпечити більшу порівнянність результатів оцінювань. Посилання на веб-сайт, де розміщена редакція цих стандартів від 2003 р., наведено в кінці цієї глави (Evaluation Cooperation Group 2008).

Оцінювання національної програми треба починати з чітко визначеного технічного завдання, розробленого спільно країною та організацією розвитку, для точного визначення очікувань зацікавлених осіб. Необхідно зазначити період, який має бути охоплено. Технічне завдання повинно:

- чітко вказувати мету оцінювання, його критерії та шляхи використання результатів;
- вказувати вихідні пріоритети організації стосовно національної програми (наприклад, скорочення бідності, підвищення врожайності);
- містити процедури звітування, розповсюдження даних і подальшої роботи; найкращий варіант – повне розкриття даних оцінювання.

Важливою складовою оцінювання національної програми є зіставний аналіз. Для порівняння звичайно використовуються схожі країни, зазвичай з того ж регіону. Що більше оцінок національних програм буде оприлюднено, то більше порівнянь між організаціями можна буде провести. Як зазначено в Стандартах добросовісної практики, складність національних оцінювань зростає. Наприклад, у зв'язку з Паризькою декларацією та Аккрською угодою виникла потреба в мультипартнерських оцінюваннях допомоги країнам, окрім тих оцінок, що їх проводить МБР. Потрібно охоплювати всі джерела зовнішньої підтримки, а це робить оцінювання значно проблематичнішим.

Оцінювання секторальних програм

Оцінювання секторальної програми:
Оцінка великого програмного сектору, наприклад освіти, охорони здоров'я, житлового будівництва або транспорту

Оцінювання секторальної програми – це оцінка великого програмного сектору, наприклад освіти, охорони здоров'я, житлового будівництва або транспорту. Керівні вказівки Міжнародної організації з міграції (МОМ) визначають секторальне оцінювання як «оцінку різноманітних заходів допомоги, зосереджених в одному секторі однієї або кількох країн. До сектору входить конкретна сфера діяльності – охорона здоров'я, промисловість, освіта, транспорт, сільське господарство» (IOM 2006, p. 30).

Через те, що оцінювання секторальної програми розглядає одночасно багато проектів з різними цілями й різними донорами, воно є складнішим за оцінювання проекту; може бути таким же складним, як і оцінювання національної програми (або ж іще складнішим, якщо вивчається стан сектору в кількох країнах). Як і у випадку з оцінюванням національних програм, секторальні оцінювання зазвичай нормативні.

Корисно порівнювати оцінювання національних і секторних програм НГО. На національному рівні НГО зазвичай проводить аналіз усього портфеля проектів Групи Світового банку за відповідний період в даній країні. Під час секторальних оцінювань НГО проводить аналіз портфеля всіх проектів у секторі за відповідний період у визначених країнах. Однак під час секторального аналізу нерідко виникають проблеми з визначеннями. Наприклад, якщо проект мультисекторальний, чи має НГО аналізувати лише ті проекти, для яких даний сектор є основним, або також ті, для яких він на другому, а то й третьому плані? Якщо НГО обирає ширший підхід, чи має вона зважувати результати за відсотком ресурсів, що виділяються саме на даний сектор? Таким чином, хоча і національні, і секторальні оцінювання передбачають аналіз портфеля, у випадку з секторальним аналізом можуть виникати складніші проблеми.

Аналогічні проблеми виникають і в питаннях планування та вибору методів оцінювань. Група Світового банку, так само як (усе частіше) й інші кредитори та донори, тепер використовує більш стандартизовані методи національного оцінювання. Секторальні ж оцінювання, навпаки, мають більш індивідуальний характер, а їх план і методи визначаються конкретними питаннями оцінки й наявними для проведення дослідження ресурсами. Хоча в більшості випадків вивчаються приклади з практики певної країни, це не є обов'язковою умовою. Якщо ж практичні приклади вивчаються, виникають важливі питання щодо вибору цих прикладів та методології.

Зацікавленість і партнерська відповідальність – ключові аспекти оцінювання секторальної програми (Danish Ministry of Foreign Affairs 1999). Як організації розвитку, так і партнерські установи прагнуть покращити надан-

ня допомоги, відповідальність і сталість у секторі. Стислий опис об'єднаних оцінювань сектору охорони здоров'я Танзанії ілюструє ці ключові моменти (блок 11.2).

Блок 11.2 «Об'єднане зовнішнє оцінювання сектору охорони здоров'я Танзанії, 1999–2006»

Протягом 1990-х років сектор охорони здоров'я Танзанії знаходився у стагнації. Місцеві медичні послуги потерпали від нестачі найпростіших медикаментів, обладнання й приладдя. Інші проблеми – занедбана інфраструктура, погане управління, брак керівництва, низька мотивація персоналу. Якість лікарняних послуг у кращому випадку залишалася на низькому рівні. Державний сектор, релігійні організації та приватні медичні заклади співпрацювали мало. На охорону здоров'я виділялося вкрай недостатньо коштів: витрати державного сектору складали 3,46 дол. США на душу населення. Партнери з розвитку також мало координували свої зусилля з підтримки медичного сектору.

Уряд Танзанії та його партнери з розвитку (Бельгія, Канада, Данія, Німеччина, Нідерланди та Швейцарія) відреагували на цю ситуацію, запустивши процес, що розпочався з місії спільного планування, яку уряд провів в середині десятиліття. Цей процес дозволив до 1999 року сформувати перший масштабний стратегічний план для сектору охорони здоров'я – Програму робіт у секторі охорони здоров'я (ПРСОЗ), а також згоду щодо того, що підтримка галузі має проводитися на основі секторального підходу. ПРСОЗ і складений пізніше Стратегічний план для сектору охорони здоров'я №2 (СПСОЗ-2) визначили процес реформування сектору, спрямований на усунення наявних проблем і досягнення конкретних цілей в сфері охорони здоров'я, узгоджених з Цілями розвитку тисячоліття (ЦРТ) та Національною стратегією скорочення бідності й розвитку (НССБР/MKUKUTA).

ПРСОЗ і СПСОЗ-2 визначили кілька пріоритетних областей стратегічних інтервенцій:

- покращення медичних послуг на районному рівні, реформування і вдосконалення лікарень регіонального та загальнодержавного рівнів;
- перетворення Міністерства охорони здоров'я та соціального добробуту на організацію, що забезпечує політичне сприяння розвитку сектору;
- вдосконалення систем централізованої підтримки (в т. ч. інфраструктури, медичних управлінсько-інформаційних систем, постачання медикаментів, транспорту, комунікацій та інформаційних технологій);

(Продовження блоку 11.2)

- прийняття й впровадження національної стратегії боротьби з ВІЛ/СНІД як чітко вираженого елементу СПСОЗ-2;
- покращення стосунків уряду й партнерів з розвитку для забезпечення гармонізації та узгодженості застосування зовнішніх та власне танзанійських ресурсів, що мало підвищити дієвість партнерства.

Оцінювання було зосереджене на чотирьох аспектах:

- релевантність стратегічних та практичних планів медичного сектору в контексті досягнення ЦРТ та виконання НССБР в сфері охорони здоров'я, а також відповідність і актуальність зовнішньої підтримки;
- прогрес та досягнення в розрізі кожного з дев'яти стратегічних пріоритетів процесу реформування сектору охорони здоров'я;
- конкретні покращення доступності, якості послуг та стану здоров'я населення протягом періоду проведення оцінки;
- зміни у партнерстві протягом періоду проведення оцінки, в т. ч. еволюція зусиль з гармонізації, узгодження й використання різних видів допомоги.

Оцінювання проводилося з грудня 2006 по вересень 2007 року групою з восьми іноземних консультантів з питань охорони здоров'я й оцінювання, троє з яких були громадянами Уганди й Малаві. Застосовувалися, зокрема, наступні основні методики:

- розширений аналіз документів;
- інтерв'ю ключових інформантів на національному рівні;
- самостійні оцінки, проведені персоналом Міністерства охорони здоров'я та соціального добробуту в 16 районах Танзанії;
- поглиблене вивчення практичних прикладів у шести районах (включаючи їх обговорення з членами громад), проведене групою оцінювання;
- аналіз потоків фінансових та інших ресурсів, спрямованих до сектору охорони здоров'я на рівні країни, регіонів, рад;
- аналіз даних про стан здоров'я населення.

Інформація оцінювання, отримана за допомогою всіх методологій, проходила перевірку й тріангуляцію з метою виявлення аналогій та посилення валідності результатів і висновків.

Джерело: Danish Ministry of Foreign Affairs 2007.

Тематичні оцінювання

Тематичне оцінювання:
Оцінка обраних аспектів або тем в низці видів діяльності, пов'язаних з розвитком

Тематичні оцінювання розглядають «обрані аспекти або теми в низці видів діяльності, пов'язаних з розвитком» (Danish Ministry of Foreign Affairs 1999, p. 30). Ці теми виникають з політичних заяв. Організація з розвитку може вирішити, приміром, що всі проекти або програми будуть спрямовані на вирішення конкретної проблеми (або проблем) – гендерної рівності, захисту довкілля, соціального захисту, полегшення стану бідних. Кожне з цих питань має розглядатися на всіх етапах проекту або програми в контексті всіх видів допомоги.

Як і у випадку з національними та секторальними оцінюваннями, тематичні оцінювання проводяться знизу вгору й згори вниз. Теми оцінюються проект за проектом; інформація з цих оцінок слугує чудовою базою для тематичного оцінювання. Окрім того, тематичні оцінювання виходять за межі окремих проектів, охоплюючи всі проекти одночасно; тематичні оцінювання використовують також для поглибленого вивчення певних країн (вивчення практичних прикладів). Тематичне оцінювання розглядає багато різних видів інформації (блок 11.3). Потім воно узагальнює інформацію з відповідних джерел.

Блок 11.3 «Приклад тематичного оцінювання, присвяченого дитячій праці на смітниках в Африці, Азії та Європі»

WASTE – організація, яка працює в напрямку сталого покращення умов життя міської бідноти та міського середовища в цілому. Міжнародна організація праці (МОП) залучила її для проведення тематичного оцінювання дитячої праці на смітниках. Метою оцінки було надання МОП рекомендацій щодо того, як краще боротися з експлуатацією дітей в цьому секторі.

Тематичне оцінювання виявило й критично оцінило це явище та різні підходи до вирішення проблеми дитячої праці на смітниках. Інформація була отримана з різних проектів, вже реалізованих в цьому секторі силами Міжнародної програми з викорінення дитячої праці, а також аналогічними зусиллями інших агенцій, установ та урядів. Результатом роботи стала стратегічна оцінка з рекомендаціями для МОП. (Звіт «Про експлуатацію дітей на смітниках: тематичне дослідження з дитячої праці» можна завантажити з сайту МОП.)

Польові дослідники WASTE в Єгипті, Індії, Румунії, Танзанії й Таїланді написали додаткові звіти, базуючись на своїх ключових висновках. В Танзанії, наприклад, звіт включав аналіз відповідних документів, інтерв'ю з керівництвом і персоналом проекту, іншими зацікавленими особами, польовий візит до двох транзитних пунктів в місті та головного сміттєзвалища в Дар-ес-Саламі, а також інтерв'ю з самими збирачами сміття та представниками їх об'єднання.

Джерело: Duursma 2004; WASTE 2005.

Одним з ключових висновків тематичного оцінювання, присвяченого дитячій праці на смітниках, була роль гендеру: більшість цих дітей складали дівчата, а серед дорослих майже всі були жінки. Це дослідження продемонструвало надзвичайну делікатність щодо гендерних питань.

Ось характерні риси підходу, що делікатно враховує гендерні нюанси:

- концептуальна система, що визнає гендерну природу розвитку й важливість гендерної рівності для економічного та соціального розвитку;
- створення гендерної бази даних на національному, секторальному або місцевому рівні, що об'єднувала б дані, які мають відношення до гендеру, та визначала б ключові гендерні питання, що мають бути враховані при плануванні й тематичному оцінюванні проектів.

Брак даних, в яких враховується гендерний фактор, часто використовується як виправдання недостатньої уваги, що приділяється гендерним питанням. Якщо гендерний фактор в даних не враховано, важливо визначити стратегії для розробки відповідних баз даних, що уможливили б проведення кращого гендерного аналізу в майбутньому. Такі стратегії включають:

- використання методів збору даних, що дають інформацію стосовно чоловіків та жінок зі включенням ключових гендерних питань в план дослідження (таких як гендерний розподіл праці, аналіз використання часу, контроль ресурсів, а також прийняття рішень на рівні домашніх господарств і громад);
- збір інформації про різних членів домогосподарств шляхом їх опитування, а також щоб «глава сім'ї» (яким зазвичай називають чоловіка) не був єдиним джерелом інформації;
- доповнення традиційних методів збору даних методами, що дозволяють враховувати гендер (наскільки можливо);
- забезпечення дослідниками гендерного балансу складу учасників;
- проведення консультацій із зацікавленими особами під час планування, аналізу та поширення інформації про висновки оцінки; при цьому до консультацій повинні залучатися групи, в яких представлені як чоловіки, так і жінки (World Bank 2007e).

Для планування та оцінювання розвитку часто використовують «гендерно-нейтральні» підходи, згідно з якими припускається, що чоловіки та жінки мають однакові потреби в сфері розвитку. В більшості суспільств чоловіки займають домінуюче положення в громаді та сім'ї, якщо мова йде про прийняття рішень, тому «гендерно-нейтральні» підходи орієнтуються переважно на пріоритети чоловіків. Ігнорування потреб і можливостей жінок істотно знижує ефективність та справедливість політик і програм. Багато інструментів планування й оцінювання розвитку застосовуються без урахування ген-

дерної специфіки. Через це навіть тоді, коли консультації з громадою або опитування сімей покликані зібрати погляди всіх членів громади, думки жінок залишаються поза увагою.

Оцінювання програм глобального й регіонального партнерства

Програми глобального й регіонального партнерства (ПГРП) – важлива форма спрямування й надання допомоги у розвитку для подолання нагальних глобальних чи регіональних проблем. Більшість ПГРП стосуються певного сектору або теми (це сільське господарство, довкілля, охорона здоров'я, фінанси, міжнародна торгівля).

ПГРП – це програмне партнерство, в рамках якого:

- партнери вносять і спільно використовують ресурси (фінансові, технічні, кадрові та репутаційні) задля досягнення погоджених цілей;
- програми носять глобальний, регіональний або міжнаціональний характер;
- партнери створюють нову організацію з управлінською структурою й керівництвом, яка покликана виконати поставлені цілі (World Bank 2007f).

Програма глобального й регіонального партнерства (ПГРП):

Програмне партнерство, за якого партнери вносять і спільно використовують ресурси задля досягнення погоджених цілей глобального, регіонального або міжнаціонального масштабу

Нова ПГРП зазвичай започатковується після того, як група глобальних або регіональних суб'єктів визначає проблему або можливість, що потребує реакції, але виходить за межі однієї держави. Цими суб'єктами можуть бути донори чи реципієнти з державного або приватного сектору.

Розглянемо для прикладу управління водними ресурсами в басейнах річок Меконг або Ніл. Ефективну роботу з проблемами або можливостями неможливо забезпечити на рівні однієї країни, тому що з водними ресурсами річок працюють одночасно багато країн: у басейні Меконгу їх шість, у басейні Нілу – 19. Крім того, колективні дії на глобальному та регіональному рівнях забезпечують істотну економію (мова йде про створення нових технологій або передових практик, актуальних для розвитку багатьох країн).

ПГРП часто зосереджуються на виробництві або збереженні глобальних суспільних благ (приклад: дослідження, що створює вдосконалене насіння або підвищує біорізноманітність) або подоланні глобальних суспільних проблем (наприклад, ВІЛ/СНІД або забруднення довкілля). Глобальні суспільні блага – це блага, які є за своєю природою неконкурентними (надання їх одній особі або країні не зменшує постачання благ іншим) та невиключними (після надання одній особі або країні всі отримують користь); користь від них виходить за межі однієї країни.

Партнери ПГРП зазвичай мають різні точки зору й керуються різною мотивацією. Організується формальна управлінська структура, що встановлює

формальні цілі, критерії членства або участі, процедури прийняття рішень, а іноді – критерії розподілу ресурсів.

Було засновано близько 150 мультипартнерських ПГРП. Рівень фінансування – від кількох мільйонів до більш ніж мільярда доларів на рік. Діяльність цих програм може обмежуватися обміном знаннями та налагоджуванням зв'язків, але може охоплювати й технічну допомогу, інвестиції, ринкову/торгову діяльність. До прикладів ПГРП відносяться: Консультативна група з міжнародних сільськогосподарських досліджень, Західноафриканська програма з ВІЛ/СНІД і транспорту, Глобальне водне партнерство, Інтегрована система з торгівлі, підприємство «Ліки від малярії».

ПГРП складно оцінювати через наступні їх риси:

- Партнерська природа програм має два наслідки. По-перше, можливо, потрібно буде враховувати різницю точок зору різних партнерів. По-друге, потрібно оцінити ефективність самої управлінської структури – як за внутрішніми критеріями добросовісної практики партнерства, так і в контексті її внеску до результатів програми (позитивного чи негативного).
- На відміну від проектів, ПГРП не має фіксованих часових рамок для досягнення цілей. І справді, цілі й стратегії часто змінюються через збільшення (або зменшення) фінансування, змінюються донори та партнери, а також зовнішні умови.
- Витрати й вигоди на різних рівнях: місцевому, національному, глобальному. Оцінка ефективності потребує врахування вкладу партнерів на різних рівнях та взаємозалежності цих рівнів; це робить формування ланцюжка результатів складним та створює проблему узагальнення результатів.
- Глобальний та регіональний вимір суспільних благ в контексті ПГРП означає розходження між витратами й благами, отриманими на національному та глобальному рівнях. Через це оцінка економічної ефективності витрат є складною, а оцінка дієвості координації та зв'язків між глобальним і національним рівнями – необхідною.
- Через подовжений термін реалізації фахівців з оцінювання часто просять розглянути рух в напрямку складних середньострокових стратегічних цілей – таких як мобілізація ресурсів для збільшення масштабу, передача виконання місцевим кадрам, незалежність від головної компанії, закриття й вихід з програми, зміна схеми співпраці з новими глобальними програмами.

З урахуванням активізації участі Групи Світового банку в ПГРП та все більш важливої ролі, яку відіграє Група в досягненні результатів глобального сталого розвитку, в 2006 році Незалежна група оцінювання (НГО) Світового банку започаткувала дві ініціативи для підтримання постійного вивчення й вдосконалення роботи. По-перше, у відповідь на запит від Мере-

жі оцінювання ОЕСР/КСР (яку підтримує Група оцінювання ООН та Група співпраці з оцінювання Міжнародного банку розвитку) НГО очолила роботу зі створення спільних принципів та стандартів оцінювання ПГРП. Першим результатом цієї роботи стала «Збірка матеріалів з оцінювання програм глобального і регіонального партнерства» (World Bank 2007f). Триває робота над дослідженням, яке має завершитися публікацією звіту, що міститиме більш детальні вказівки й приклади.

По-друге, НГО почала аналізувати незалежні оцінювання ПГРП, в яких Світовий банк бере участь, прагнучи визначити якість оцінки та надати зворотний зв'язок, незалежну оцінку висновків і рекомендацій та оцінити роботу Світового банку як партнера. Сім з цих аналізів глобальних програм було оприлюднено, з наведеної інформації робляться висновки.

Приклад такого аналізу – оцінка й аналіз глобальної програми підприємства «Ліки від малярії». Ця організація, заснована в 1999 році, прагне зменшити тягар малярії, що його несуть країни, територія яких охоплена цією хворобою. Для цього підприємство винаходить, розробляє й постачає нові доступні протималярійні препарати за допомогою ефективних партнерств між державним та приватним сектором. Сукупні внески за 2006 рік склали 151 млн доларів США; Секретаріат програми, організованої як незалежна неприбуткова організація згідно із законодавством Швейцарії, знаходиться в Женеві. Ще один офіс – у Нью-Делі. Рада директорів підприємства «Ліки від малярії» складається з 11 визначних представників промисловості, академічних кіл, Всесвітньої організації охорони здоров'я, а також Фонду Біла та Мелінди Гейтс, що надає 60 відсотків коштів.

Донори організації в 2005 році ініціювали зовнішнє оцінювання. Рада співпрацювала з фахівцями з оцінювання, хоча й не замовляла їхні послуги. Оцінювання було проведено групою з чотирьох спеціалістів під керівництвом видатного африканського професора-медика.

Технічне завдання передбачало використання стандартних критеріїв оцінювання; питання оцінювання стосувалися деталізації аспектів, властивих ПГРП в цілому та природі діяльності, що її фінансує підприємство «Ліки від малярії» зокрема. Далі розглядається, як саме ці критерії застосовуються в складних оцінках ПГРП.

Релевантність. Ставилися питання стосовно релевантності цілей програми: чи було досягнуто міжнародного консенсусу щодо проблеми, яку програма має розв'язати, та потреби, яку вона має задовольнити? Оцінювання свідчило про релевантність програми в контексті досягнення ЦРТ з профілактики поширення інфекційних хвороб, що також є глобальним суспільним благом. Під час оцінки релевантності також застосовувався такий критерій, як принцип субсидіарності, відповідно до якого програми повинні

реалізовуватися на рівні, максимально наближеному до місцевого. В ході оцінювання ставилося питання, чи надає програма послуги, що не можуть бути надані на національному або місцевому рівні. Через те, що мова йшла про глобальне суспільне благо, фахівці з оцінювання вирішили, що програма дотримується принципу субсидіарності.

Фахівці з оцінювання розглянули створювану програмою додаткову цінність в контексті інших джерел постачання й надання послуг (враховуючи як інші протималярійні програми, так і приватні програми розробки препаратів). Також була оцінена відповідність цілей програми потребам бенефіціарів. За цим критерієм програма була визнана відповідною, адже вона працює не лише в сфері розробки препаратів, а й займається питаннями забезпечення та доступності.

Фахівці з оцінювання також ставили наступні запитання: Чи релевантними є стратегія й план програми? Чи був вибір інтервенцій релевантним і відповідним? Якими були масштаб і географічне охоплення? Як застосовувалося партнерство між державним і приватним сектором? Чи була б більш релевантною альтернативна стратегія програми?

Дієвість (досягнення цілей). Під час оцінки вивчалися отримані короткострокові й довгострокові результати з урахуванням очікувань, а також фактори, що сприяли досягненню цілей або заважали йому. Цей аспект оцінки значною мірою залежав від якості системи моніторингу й оцінки (МіО) та збору даних. Оцінювалися впливи наступних факторів:

- участь зацікавлених осіб;
- дієвість управління портфелем;
- дієвість управління партнерством між державним та приватним секторами;
- застосування наукового підходу;
- дієвість основних підходів до продукту;
- побічні довгострокові результати.

Ефективність. Хоча цей аспект не згадувався в технічному завданні як окремий критерій, фахівці з оцінювання визнали важливість вивчення альтернативної вартості більш ефективних засобів надання послуг. Вони проаналізували вартість управління й адміністрування порівняно з програмними витратами та динамічними тенденціями, а також порівняли їх з еталонними зразками – такими як вартість розробки медичних препаратів для лікування інших хвороб, а також вартість типового процесу розробки продукту в великій фармацевтичній компанії.

Управління. Легітимність управління підприємства «Ліки від малярії» було оцінено шляхом вивчення представленості в керівному органі країн, де поширена ця хвороба, використання наукового досвіду (наприклад, до-

свіду дослідників зі Всесвітньої організації охорони здоров'я та з уражених країн), свідоцтв упевненості партнерів в успішності кампаній з мобілізації ресурсів, а також реакцій на оцінки. Оцінка також визначала дієвість управління для виконання наступних важливих функцій:

- вибір стратегії (наскільки добре керівна структура забезпечила інформованість учасників дискусії щодо нового напряму роботи щодо забезпечення доступності препаратів на національному рівні?);
- співпраця з іншими агенціями з аналогічними функціями (фахівці з оцінювання рекомендували приділити більше уваги цьому аспекту);
- мобілізація ресурсів (фахівці з оцінювання рекомендували використовувати більш упереджувальний підхід до отримання необхідних ресурсів та диверсифікувати донорську підтримку).

Сталість/середньострокові стратегічні питання. В ході оцінювання вивчалося, чи ймовірним було збереження переваг програми в поточних умовах, причому сталість розглядалася з точки зору фінансових потреб, збереження зацікавленості й підтримки цілей. Також визначалося, чи можна було внести зміни таким чином, щоб умови більше сприяли сталості. (Чи варто змінити масштаб програми? Чи можна було б підвищити ефективність стосунків з іншими організаціями? Чи потрібно розглянути альтернативну організаційну або управлінську структуру?) Фахівці з оцінювання також працювали над покращенням МіО. Технічні завдання вимагали проведення тестування запропонованого донорами шаблону для визначення ефективності партнерства між організаціями державного та приватного секторів. Крім того, фахівці з оцінювання внесли рекомендації щодо покращення моніторингу для підтримки ініціатив МіО в майбутньому. Методологія оцінювання передбачала аналіз програмних документів, розширені інтерв'ю із зацікавленими особами на рівні програми та виду діяльності, спостереження за зборами науково-консультаційного комітету, а також виїзди на місця всередині країни.

Аналіз глобальної програми визначав незалежність і якість оцінювання. Щоб оцінити рівень незалежності, розглядалося, хто замовив оцінку та як здійснювалося керування нею. Також вивчалося, чи охопила оцінка всі аспекти роботи, передбачені технічним завданням. Увага також приділялася якості оцінювання та її кінцевому впливу (якою мірою керівники програми впроваджували рекомендації). За результатами аналізу було винесено незалежне судження щодо висновків оцінювання. Аналіз підтвердив, що воно забезпечило надійну основу для наступних дій керівництва програми й Світового банку.

Під час аналізу глобальної програми була оцінена ефективність партнерської діяльності Світового банку. Розглянуто використання ним своїх можливостей з мобілізації ресурсів з метою впровадження програми, його пізнішу фінансову підтримку, що спочатку була важливою, але з часом зменшила

свою вагу порівняно з підтримкою інших донорів. Ще одним об'єктом вивчення стала участь банку в нарадах керівних органів та комітетів. Було визначено потенційні проблемні питання, пов'язані з роллю банку в програмі. Аналіз підсумував практичні висновки з оцінювання й аналітичних заходів, включно з тими, що стосуються самої програми (потреба у впровадженні відповідної системи моніторингу й оцінювання), інших ПГРП (потреба у дієвій координації та консультаціях з іншими ключовими гравцями на глобальному й національному рівнях), а також інших інструментів допомоги (висновки щодо проектів, що надають послуги за рахунок партнерства між організаціями державного й приватного секторів).

Підсумки

Останнім часом фахівці з оцінювання намагаються вивчати наслідки багатьох інтервенцій, зосереджуючись на секторі, країні або темі. Ці підходи спрямовані на визначення того, яким чином інтервенції впливають на загальну картину. Складніші підходи використовуються для оцінки одночасно кількох проектів, програм та політик з метою виявлення взаємодій та шляхів спільного використання ресурсів. До таких оцінок відносять спільні оцінки, оцінки національних програм, оцінки секторальних програм, тематичні оцінювання, а також оцінювання глобального або регіонального партнерства.

Посилання та рекомендована література

ADB (Asian Development Bank). 2006a. *Indonesia: Capacity Building Project in the Water Resources Sector.* Project Performance Evaluation Report for Indonesia. http://www.adb.org/Documents/PPERs/INO/26190-INO-PPER.pdf.

———. 2006b. *Private Sector Development and Operations: Harnessing Synergies with the Public Sector.*

http://www.oecd.org/dataoecd/6/59/39519572.pdf.

AfDB (African Development Bank). 2005. *Morocco: Evaluation of Bank Assistance to the Education Sector.*

http://www.oecd.org/dataoecd/36/19/37968449.pdf.

Bamberger, Michael. 2005. *IPDET Handbook for Evaluating Gender Impacts of Development Policies and Programs.* International Program for Development Evaluation Training, Ottawa.

Bamberger, Michael, Mark Blackden, Lucia Fort, and Violeta Manoukian. 2001. "Integrating Gender into Poverty Reduction Strategies." In *The PRSP Sourcebook.* 335–74. Washington, DC: World Bank. http://povlibrary.worldbank.org/files/4221_chap10.pdf.

Breier, Horst. 2005. *Joint Evaluations: Recent Experiences, Lessons Learned and Options for the Future.* Draft report to the Development Assistance Committee Network on Development Evaluation, Organisation for Economic Cooperation and Development, Paris.

Chen, Jhaoying, and Hans Slot. 2007. "Country-Led Joint Evaluation: Dutch ORET/ MILIEV Programme in China." Presentation at the sixth meeting of the Development Assistance Committee Network on Development Evaluation, Paris, June. http://www.oecd.org/dataoecd/63/28/38851957.ppt.

Compton, Donald W., M. Baizerman, and S. H. Stockdill, eds. 2002. *The Art, Craft, and Science of Evaluation Capacity Building.* New Directions for Evaluation 93 (Spring) (publication of the American Evaluation Association).

Danish Ministry of Foreign Affairs. 1999. *Evaluation Guidelines.* 2nd ed. Copenhagen. http://www.um.dk/NR/rdonlyres/4C9ECE88-D0DA-4999-9893-371CB351C 04F/0/Evaluation_Guidelines_1999_revised.pdf.

———. 2006. *Evaluation Guidelines.* Copenhagen.

———. 2007. *Joint External Evaluation: The Health Sector in Tanzania 1999–2006.*

http://www.oecd.org/dataoecd/53/46/39837890.pdf.

Duursma, M. 2004. *Child Labour in Scavenging, Tanzania, Dar es Salaam.* Updated August 16, 2005. http://www.waste.nl/page/724.

"Evaluation of the Implementation of the Paris Declaration. Synthesis Report.: 2008, 26.

Freeman, Ted. 2007. “Joint Evaluations.” International Program for Development Evaluation Training (IPDET) presentation, Ottawa, June–July.
Fullan, M. 1993. *Change Forces.* London: Falmer Press.
GTZ (Gesellschaft für Technische Zusammenarbeit). 2004. *National Monitoring of Sustainable Poverty Reduction Strategy Papers PRSPs.* Eschborn, Germany.
http://siteresources.worldbank.org/INTISPMA/Resources/Training-Events-and-Materials/summary_MainReport.pdf.
Heath, John, Patrick Grasso, and John Johnson. 2005. *World Bank Country, Sector, and Project Evaluation Approaches.* International Program for Development Evaluation Training (IPDET) presentation, Ottawa, July.
House, Ernest R. *Guiding Principles for Evaluators.* 1995. New Directions for Program Evaluation No. 66. San Francisco: Jossey-Bass.
IOM (International Organization for Migration). 2006. *IOM Evaluation Guidelines.*
Office of the Inspector General. http://www.iom.int/jahia/webdav/site/myjahiasite/shared/shared/mainsite/about_iom/eva_techref/Evaluation_Guidelines_2006_1.pdf.
Johnson, John. 2007. “Confronting the Challenges of Country Assistance Evaluation.” International Program for Development Evaluation Training (IPDET) presentation, Carleton University, Ottawa, June 26–27.
Kusek, Jody Zall, and Ray C. Rist. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System.* World Bank, Washington, DC. http://www.oecd.org/dataoecd/23/27/35281194.pdf.
Mackay, Keith. 1999. *Evaluation Capacity Development: a Diagnostic Guide and Action Framework.* ECD Working Paper 6 (January), World Washington, DC. http://lnweb18.worldbank.org/oed/oeddoclib.nsf/a4dd58e444f7c611852568080 06a0008/7f2c924e183380c5852567fc00556470?OpenDocument.
———. 2006. *Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management.* World Bank, Independent Evaluation Group, Washington, DC. http://siteresources.worldbank.org/INTISPMA/Resources/ecd_15.pdf.
———. 2007. *How to Build M&E Systems to Support Better Government.* World Bank, Washington, DC. http://www.worldbank.org/ieg/ecd/docs/How_to_build_ME_gov.pdf.
Mertens, D. M. 1994. “Training Evaluators: Unique Skills and Knowledge.” *New Directions for Program Evaluation* 62: 17–27.
MLB (Multilateral Development Bank), EC (Evaluation Cooperative Group), and WGEC (Working Group on Evaluation Criteria and Ratings for Public Sector Evaluation). 2003. *Good Practice Standards for Evaluation of MDB*

Supported Public Sector Operations. http://www.ifc.org/ifcext/oeg.nsf/AttachmentsByTitle/MDB-ECG/US$FILE/ MDB-ECG_Good-Practice.pdf.
OECD (Organisation for Economic Cooperation and Development). 1999. "Evalu- ating Country Programmes," Development Assistance Committee, Vienna Workshop, March 11–12.
http://www.oecd.org/dataoecd/41/58/35340748.pdf.
———. 2002. *OECD Glossary of Key Terms in Evaluation and Results Based Manage- ment.* Development Assistance Committee, Paris.
———. 2005. Workshop on "Challenging the Conventional Wisdom: The View from Developing Country Partners," Network on Development Evaluations, Development Assistance Committee, Nairobi, April 20–21. http://www.oecd.org/dataoecd/20/44/34981186.pdf.
———. 2006. *DAC Evaluation Series: Guidance for Managing Joint Evaluations.*
Development Assistance Committee, Paris. http://www.oecd.org/dataoecd/29/28/37512030.pdf.
Porteous, Nancy L., Barbara J. Sheldrick, and Paula J. Stewart. 1999. "En- hancing Managers' Evaluation Capacity: A Case Study for Ontario Public Heath." *Cana- dian Journal of Program Evaluation.* Special Issue: 137–54.
http://www.phac-aspc.gc.ca/php-psp/pdf/toolkit/enhancing_managers_evaluation%20_capacity%20_CJPE_1999.pdf.
UNFPA (United Nations Population Fund). 2005. *State of World Population 2005.*
http://www.unfpa.org/swp/2005/english/notes/index.htm.
WASTE. 2005. *Thematic Evaluation on Child Labour in Scavenging Africa, Asia, and Europe Assessment.* Gouda, the Netherlands. http://www.waste.nl/page/720.
World Bank. 2001a. *Design Paper for a Multipartner Evaluation of the Compre- hensive Development Framework.* Comprehensive Development Framework Secretariat, Washington, DC. http://www.worldbank.org/evaluation/cdf/cdf_evaluation_design_paper.pdf.
———. 2001b. *Engendering Development–through Gender Equality in Rights, Resources and Voice.* Washington, DC: World Bank.
———. 2001c. *The PRSP Sourcebook.* http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/ EXTPRS/0,,menuPK:384207~pagePK:149018~piPK:149093~theSitePK:384201,00.html.
———. 2005. *Evaluation of World Bank's Assistance to Primary Education.* Operations Evaluation Department, Washington, DC. http://www.worldbank.org/oed/education/evaluation_design.html.
———. 2007a. *Country Assistance Evaluation: CAE Retrospective.* Independent Evaluation Group, Washington, DC. http://www.worldbank.org/ieg/countries/cae/featured/cae_retrospective.html.

———.2007b. *Global Program Review: Medicines for Malaria Venture.* Independent Evaluation Group, Washington, DC. http://lnweb18.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c8525680800 6a0046/d591aea3bbb897de852573130077a0cb?OpenDocument.
———. 2007c. *Impact Evaluations: Bangladesh Maternal and Child Health.* Independent Evaluation Group, Washington, DC. http://www.worldbank.org/oed/ie/bangladesh_ie.html.
———. 2007d. *Poverty Monitoring Systems.* http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTPAME/0,,contentMDK:20203848~menuPK:435494~pagePK:148956~piPK: 216618~theSitePK:384263,00.html.
———. 2007e. *PRSP Sourcebook.* Washington, DC. http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/ EXTPRS/0,,contentMDK:20175742~pagePK:210058~piPK:210062~theSitePK: 384201,00.html.
———. 2007f. *Sourcebook for Evaluating Global and Regional Partnership Programs: Indicative Principles and Standards.* Independent Evaluation Group, Washington, DC. http://siteresources.worldbank.org/EXTGLOREGPARPRO/Resources/sourcebook.pdf.
———. 2008. *CAE Methodology.* Independent Evaluation Group, Washington, DC. http://web.worldbank.org/WBSITE/EXTERNAL/EXTOED/EXTCOUASSEVA/ 0,,contentMDK:21107046~menuPK:3098030~pagePK:64168445~piPK:64168309~theSitePK:2832047,00.html.

Веб-сайти

International Labour Organization. http://www.ilo.org.
Stufflebeam, D. L. 1999. *Evaluation Plans and Operations Checklist.* http://www.wmich.edu/evalctr/checklists/plans_operations.htm.
———. 2001. *Guiding Principles Checklist.* http://www.wmich.edu/evalctr/checklists/ guiding_principles.pdf.
World Bank. *ADePT: Stata Software Platform for Automated Economic Analysis.* http://econ.worldbank.org/programs/poverty/adept.
———. *Core Welfare Indicators Questionnaire (CWIQ).* http://www4.worldbank.org/ afr/stats/cwiq.cfm.
———. *Global and Regional Partnership Programs.* http://www.worldbank.org/ieg/ grpp.

УПРАВЛІННЯ

«Всі дії повинні приводити до мудрості»

—БХАГАВАД ГІТА

Глава 12: Управління оцінюванням

- Робота з матрицею плану
- Проведення оцінювання на підрядній основі
- Функції й обов’язки різних гравців
- Управління персоналом, завданнями й бюджетами

Глава 13: Представлення результатів

- Основи спілкування
- Написання звіту про оцінювання для вашої аудиторії
- Використання візуальної інформації
- Проведення усної презентації

Глава 14: Керівництво для фахівця з оцінювання: етика, політика, стандарти та керівні принципи оцінювання

- Етична поведінка
- Політика й оцінювання
- Стандарти й керівні принципи оцінювання

ГЛАВА 12

Управління оцінюванням

Оцінювання буває дуже складним: мотивація персоналу, своєчасне виконання зобов'язань, забезпечення високої якості робіт – непрості завдання. В цій главі розповідається про те, як фахівці з оцінювання можуть планувати оцінювання й управляти ним, дотримуватись стандартів якості та поширювати висновки, щоб результати проведеного оцінювання використовувалися особами, які формують політику, для впровадження змін.

Ця глава складається з чотирьох основних частин:

- Робота з матрицею плану
- Проведення оцінювання на підрядній основі
- Функції й обов'язки різних гравців
- Управління персоналом, завданнями й бюджетами

Робота з матрицею плану

Запорукою успішного оціювання й розвитку є планування. Якщо воно погано сплановане, ніякий подальший аналіз – яким би складним і потужним він не був – його не врятує. Китайське прислів'я говорить: «Навіть шлях у тисячу лі починається з першого кроку». Ймовірність дістатися цілі буде значно вищою, якщо перший крок – так само, як і наступні – мандрівник зробить у правильному напрямку. Блукання без чіткого розуміння цілі або напрямку забирає час, енергію й ресурси не тільки в мандрівника, а й у фахівця з оцінювання. Воно також зменшує шанси взагалі колись дістатися мети. Таким чином, розумно було б підготуватися до подорожі, зібравши необхідні мапи, вивчивши альтернативні маршрути та оцінивши час, витрати й небезпеки, що можуть чекати попереду.

Як уже говорилося в попередніх главах, матриця плану оцінювання – це спосіб візуалізації, мапа оцінювання. В центрі уваги опиняються усі основні компоненти його планування Але шаблон матриці не витесаний з мармуру, його можна змінювати відповідно до потреб оцінювання. Як і будь-який план, матриця, скоріше за все, проходитиме ітеративний процес оновлень та переглядів. Після надходження нової інформації деякі ідеї та підходи доведеться переглянути.

Але самої лише хорошої матриці плану для проведення вдалого оцінювання недостатньо. Навіть найкращим чином сплановане оцінювання може виявитися неякісним, якщо, наприклад, реалізувати план спробують люди, які не мають належної кваліфікації, або якщо оцінювання буде проведено занадто пізно для того, щоб воно могло вплинути на критичне рішення, або ж якщо кошти закінчаться вже на стадії збору даних.

Проведення оцінювання на підрядній основі

Керівник оцінювання відповідає за виявлення й усунення проблем нестачі кваліфікованих кадрів для проведення оцінювання. Кваліфікованих фахівців з оцінювання мало, а попит на їх послуги високий. Через це для проведення оцінювання розвитку нерідко на підрядній основі залучаються фірми або окремі консультанти. Деякі організації – наприклад Європейський Союз – доручають проведення всіх або майже всіх своїх оцінок стороннім спеціалістам. Умова забезпечення об'єктивності консультантів – відсутність будь-яких попередніх зв'язків з програмою.

Залучення підрядників має свої переваги й недоліки. До переваг відноситься той факт, що підрядник може мати глибокі знання щодо сектору й

виду програми, що оцінюються, говорити місцевою мовою та знати місцеві умови, а також, якщо його залучено за результатами тендеру, пропонувати найбільш економічно ефективне рішення. Недоліками є висока вартість (проведення тендеру може вийти дорожче, ніж сам підряд) й відсутність нарощування власних можливостей організації-замовника.

Запити пропозицій / Конкурс пропозицій

Запит пропозицій / конкурс пропозицій:
Запрошення надіслати тендерну/конкурсну пропозицію щодо участі в проекті

Процес найму консалтингової компанії або індивідуальних консультантів зазвичай розпочинається з розробки **запиту пропозицій** (ЗП). В залежності від організації та правил і процедур залучення підрядників договір може укладатися на безальтернативній основі, за результатами запрошення до торгів кількох підрядників, а також за підсумками відкритого конкурсу. Найкраща пропозиція обирається керівником оцінювання або призначеним для цього комітетом, який робить рекомендацію на основі визначених заздалегідь критеріїв.

Гокінз (Hawkins 2005) рекомендує включати в ЗП наступні пункти:

- цілі оцінювання;
- загальні відомості й контекст дослідження;
- ключові вимоги до інформації;
- завдання оцінювання;
- очікувані результати;
- часові рамки;
- критерії вибору найкращої пропозиції;
- відомості про контракт для менеджера проекту;
- кінцевий термін прийому пропозицій;
- бюджет та інші ресурси.

Вибір найкращої пропозиції має проводитися комітетом, до складу якого входять спеціалісти зі знаннями й досвідом в сфері оцінювання, обізнані щодо програмної сфери та культурного контексту, які зацікавлені в результатах та їх застосуванні. Цей комітет має вибрати пропозицію, керуючись наведеними в ЗП критеріями. Необхідно вести детальні записи під час процесу вибору.

Гокінз радить використовувати наступні критерії для вибору консультанта:

- Чи відповідає пропозиція вимогам, викладеним в ЗП?
- Чи пояснюється детально, як саме буде проводитися оцінювання?
- Які стратегії комунікацій і звітування пропонуються?
- Чи надається підтвердження компетентності? Який послужний список має консультант?
- Які витрати? Чи наводиться деталізація витрат?

Після укладення угоди з підрядником керівник оцінювання все ще матиме певні обов'язки стосовно підрядника й оцінки. Зокрема:

- чітке визначення цілей і завдань;
- загальне керівництво дослідженням;
- моніторинг роботи й своєчасне надання зворотного зв'язку;
- своєчасне прийняття рішень;
- відкритість до переговорів з підрядником у випадку, якщо в договір потрібно буде внести якісь зміни.

Технічне завдання

Технічне завдання (ТЗ): Документ, в якому викладено контекст, завдання й мету оцінювання; окремі обов'язки та сфери відповідальності членів групи оцінювання; графіки представлення результатів роботи

Технічне завдання (ТЗ) – документ, в якому викладено контекст, завдання й мету оцінювання; окремі обов'язки та сфери відповідальності членів групи з оцінювання; графіки представлення результатів роботи. ТЗ зазвичай дає загальний опис оцінювання та визначає початкові домовленості. Може йтися про домовленість між клієнтом і керівником оцінювання, керівником оцінювання й підрядником, або ж керівником оцінювання та персоналом з оцінювання. Всі члени групи оцінювання повинні мати чітке, описане розуміння своїх обов'язків щодо проведення оцінювання. Перелік цих обов'язків звичайно наводиться в ТЗ.

ТЗ мають реалістично визначати, що потрібно виконати в заданий проміжок часу.

ТЗ можуть включати матрицю плану. Інший варіант – розробку матриці плану вказано як один з перших очікуваних результатів. ТЗ також може описувати домовленості щодо етики й стандартів проведення оцінювання та звітування про неї (етика й стандарти обговорюються в главі 14).

У процесі розробки загальних ТЗ може бути корисно залучити зацікавлених осіб до обговорення й прийняття рішень про питання оцінювання. ТЗ визначає основні керівні вказівки, щоб кожен причетний розумів, що очікується від оцінювання (блок 12.1).

Згідно з Глосарієм ключових термінів оцінювання й управління з орієнтацією на результат (OECD 2002), ТЗ для оцінювання – це письмовий документ, що представляє:

- мету й обсяг оцінювання;
- методи, що мають використовуватися;
- стандарт для визначення якості роботи або проведення аналізу;
- доступні ресурси й час;
- вимоги до звітності.

Блок 12.1 «Рекомендації щодо написання технічного завдання для оцінювання»

Наступні поради можуть стати у нагоді під час написання ТЗ для оцінювання:

- Чітко вказуйте завдання оцінювання, визначаючи:
  - загальні теми й попередні питання оцінювання, на які треба знайти відповіді;
  - основних зацікавлених осіб та те, як вони будуть використовувати результати оцінювання;
  - загальний підхід оцінювання, який треба застосовувати;
  - очікувані продукти оцінювання, коли кожен з них має бути представлений, та як вони будуть використовуватися;
  - кваліфікація, яку повинні мати члени групи оцінювання;
  - логістичні заходи.
- Не просто вказуйте технічні або процесові завдання. Чітко зазначайте, як оцінювання має допомогти організації.
- Зосередьтеся на попередніх питаннях, відповіді на які треба знайти.
- Уникайте вибору завеликої кількості питань. Краще, щоб фахівці з оцінювання поглиблено вивчили кілька питань, ніж поверхово розглянули широкий спектр тем.

Джерело: Взято з UNDP 2006.

ТЗ зазвичай містять наступну інформацію:

- короткий і змістовний заголовок;
- опис проекту або програми;
- причини проведення оцінювання й очікування від неї;
- визначення обсягу й фокусу оцінювання (теми, які треба розглянути, питання, на які треба відповісти);
- визначення зацікавлених осіб (кого буде залучено, хто та що робитиме, процес контролю й обліку);
- опис процесу оцінювання (що буде зроблено);
- перелік очікуваних результатів (робочий план оцінювання, проміжний звіт, заключний звіт, презентації);
- визначення необхідних кваліфікацій (вимоги щодо освіти, досвіду, навичок, умінь);
- прогноз витрат з урахуванням видів діяльності, часу, кількості персоналу, оплати праці, витрат на відрядження тощо.

ТЗ складаються на додачу до матриці плану оцінювання, а не замість неї. ТЗ зосереджуються на обов’язках. Матриця плану – план проведення оцінювання.

Функції й обов'язки різних гравців

До оцінювання залучається багато людей. Вони займають різні посади, виконують різні функції й обов'язки. Основні ролі – клієнт (замовник), зацікавлені особи, керівник оцінювання та фахівець з оцінювання (сторонній або штатний). Обмежені функції виконують аналітик даних, радники, незалежні експерти та спеціалісти зі збору даних. Функції й обов'язки кожного учасника оцінювання мають бути чітко визначені й узгоджені.

Головний клієнт

В оцінюванні може бути багато зацікавлених осіб, але зазвичай лише один головний клієнт, який замовляє та, як правило, фінансує його. Структура оцінювання в цілому залежить від конкретних питань, вивчення яких потребує клієнт, того, як він планує користуватися результатами оцінювання, а також від відповідних часових обмежень.

Добре, якщо фахівці з оцінювання вже на ранніх етапах зустрічаються з клієнтом, щоб обговорити наступні моменти:

- природа й контекст запиту;
- потреба, проблема або інтерес, що спричинив запит на проведення оцінювання;
- критичні часові обмеження для надання висновків оцінювання;
- важливі питання, які мають бути розглянуті у процесі оцінювання;
- графік і частота контактів.

В «Основному контрольному списку оцінювання» Скрайвена (Scriven 2007) йдеться про те, як важливо зустрітися з клієнтом до початку оцінювання, щоб деталізувати його запит (як він його бачить) або спонукати клієнта пояснити свої погляди щодо різних нюансів. Скрайвен рекомендує ставити наступні питання:

- Що саме, конкретно, потрібно оцінити?
- Скільки контексту слід охопити?
- Чи має фахівець з оцінювання оцінити наслідки програми в цілому чи вклад кожного з її компонентів? Можливо, клієнт бажає перевірити свою теорію стосовно роботи компонентів?
- Чи повинен фахівець з оцінювання розглядати вплив в усіх можливих розрізах чи лише у деяких з них?
- Яким має бути оцінювання: формативним, сумативним та/або описовим?

Як уже зазначалося в главі 4, група оцінювання підтримує стосунки з головним клієнтом, зокрема, за допомогою плану комунікативної співпраці.

Цей план допомагає фахівцям з оцінювання працювати з клієнтом, основними зацікавленими особами, ЗМІ та громадськістю. Він відповідає на всі питання стосовно надання й отримання інформації. Він також встановлює план контактів із залученням різних ЗМІ для різних аудиторій. Запровадження плану комунікативної співпраці особливо важливе, якщо оцінка стосується інформації делікатного характеру.

Для багатьох фахівців з оцінювання комунікація – це те, що трапляється в кінці процесу, коли формується звіт. Насправді комунікація й розвиток комунікаційної стратегії мають пророблятися заздалегідь і вдосконалюватися в ході процесу. Комунікація повинна використовуватися для підтримки і вдосконалення оцінювання та його методології, формування позитивного образу дослідження, а також сприяння практичному застосуванню висновків.

Зацікавлені сторони

Як було визначено в главі 4, зацікавлені сторони – це люди й організації (окрім клієнта), які зацікавлені в інтервенції. Зазвичай це ті, на кого поширюється вплив інтервенції – під час її проведення або в наступні роки.

Зацікавлені особи можуть бути всередині або ззовні організації. В контексті оцінювання вони можуть виконувати широкий спектр функцій – від визначення тем і питань для оцінки до надання даних.

Під час спільних оцінювань зацікавлені особи можуть виконувати обов'язки на різних рівнях процесу. В таких оцінках окремі обрані зацікавлені особи можуть входити до складу групи оцінювання, брати участь у формулюванні питань, зборі даних, аналізі. Зацікавлені особи з більшою ймовірністю підтримуватимуть оцінювання та застосовуватимуть його результати й рекомендації, якщо самі беруть участь в цьому процесі (ми розглянули спільні оцінювання в главі 5).

Може бути корисно залучати до оцінювання критиків проекту, програми або політики. Критики визначають важливі проблеми, що мають бути розв'язані в ході оцінювання (і нерозв'язання яких може дискредитувати оцінювання), тому їх залучення сприяє зміцненню процесу оцінювання (CDC 2001).

Видана Світовим банком «Збірка матеріалів щодо учасників оцінювання» містить корисні поради із залучення зацікавлених осіб, особливо актуальні для спільного оцінювання (World Bank 1996). Визначивши коло зацікавлених осіб, треба переходити до їх залучення. Фахівці з оцінювання використовують для цього широкий спектр підходів. Нерідко потрібні особливі заходи для того, щоб групи, які зазвичай виключаються з процесу прийняття рішень, отримали право голосу. Для цього фахівцям з оціню-

вання спочатку доводиться організувати «мовчазні» групи, забезпечити їх представництво, проводити спільні сесії та використовувати техніки, що дозволяють зацікавленим особам на всіх рівнях бути почутими.

Багатьом зацікавленим особам стороння особа, яка пропонує «оцінювання із залученням зацікавлених сторін», може видатися підозрілою. Іноді досвід спілкування з державними установами, держслужбовцями та донорськими організаціями формує негативні враження, які треба подолати. Довіру можна завоювати, ділячись інформацією, працюючи через посередників, починаючи роботу якомога раніше і з широким підходом. Один зі шляхів побудови довірчих відносин полягає в наданні інформації про цілі, які переслідує оцінювання. Фахівці з оцінювання можуть робити це під час зустрічей з окремими особами або великими групами (наприклад, під час міських зборів). Під час цих зустрічей вони можуть надавати інформацію про методи й причини проведення оцінювання. Учасники зустрічей мають можливість висловити власні сподівання й стурбованість. Якщо оцінювання спільне, після встановлення довіри учасникам можна запропонувати створити власні комітети для роботи над оцінкою.

Публікація відкритих для коментування концептуальних документів у мережі Інтернет – один зі способів започаткувати спілкування з науково-дослідницьким співтовариством, неурядовими організаціями та зацікавленими громадянами. Звичайно, це необхідно робити тоді, коли пропозиції, що надходять, ще можна розглядати та використовувати для внесення змін. Цей підхід найбільш дієвий у тих випадках, коли фахівці з оцінювання публікують відповіді на коментарі та повідомляють про те, яких заходів було вжито (або про причини, чому рекомендації було відхилено).

В деяких випадках недовіра є настільки великою, що для подолання відстані між фахівцями з оцінювання й зацікавленими особами можуть знадобитися зусилля посередників. У таких випадках сторона, яка користується повагою зацікавлених осіб, може використати свій вплив для зближення фахівців з оцінювання із зацікавленими особами.

Фахівці з оцінювання також мають відвідувати зацікавлених осіб, особливо тих, хто зазнає безпосереднього впливу інтервенції. Вони можуть поцікавитися, чи чули зацікавлені особи про оцінювання та що про нього думають. Цей неформальний зворотний зв’язок можна порівняти з тими даними, що надходять через інші канали. Він дозволяє перевірити узгодженість та відсутність помилок.

Надання зацікавленим особам можливості брати участь у власній колективній оцінці підвищує шанси на подолання протиріч і формування консенсусу з дискусійних питань. Отже, залучення зацікавлених осіб на ранніх стадіях може запобігти виникненню конфліктів. Для тих, хто вва-

жає план оцінювання невигідним для себе, єдиним можливим виходом може бути відкрита протидія; що більша втрата, то сильнішою буде протидія. Після початку протидії розв'язати проблему буде важко, якщо взагалі можливо.

Керівник оцінювання

Керівник оцінювання – це особа, яка керує плануванням, підготовкою, реалізацією, аналізом, звітуванням, поширенням результатів та подальшим супроводженням оцінки. Він може керувати кількома оцінками. В деяких випадках фахівець з оцінювання одночасно і керує оцінюванням, і проводить його.

Керівник оцінювання: Особа, що керує плануванням, підготовкою, реалізацією, аналізом, звітуванням, поширенням результатів та подальший супровід оцінювання

Обов'язки керівника можна згрупувати за трьома етапами: підготовка, реалізація, подальший супровід (UNFPA 2007). Підготовка включає попередню роботу з визначення оцінювання й розробки ЗП, ТЗ та матриці плану оцінювання, а також вирішення логістичних питань. Вона передбачає наступні обов'язки керівника:

- Визначити мету оцінювання та користувачів її результатів.
- Визначити, кого треба залучити до процесу оцінювання.
- Зустрітися з основними зацікавленими особами.
- Визначити обсяг і підхід оцінювання, а також сформувати матрицю плану включно з питаннями.
- Розробити попереднє ТЗ оцінювання, вказати його часові рамки.
- Визначити, яке поєднання вмінь та досвіду потрібне для проведення оцінювання. Вести нагляд за розробкою та попереднім тестуванням інструментів збору даних, а також збором існуючої інформації. Використовувати селективний підхід і переконатися в надійності існуючих джерел інформації та в тому, що їх якість достатньо висока, щоб можна було отримати значущі результати оцінювання. Зібрана інформація має бути придатною для роботи.
- Вибрати, найняти та проінструктувати фахівців з оцінювання стосовно мети оцінювання, матриці, робочого плану. Провести необхідне навчання.
- Переконатися, що зібрану вихідну документацію/вихідні матеріали подано у достатньо ранній термін, щоб група оцінювання змогла скористатися отриманою інформацією.
- Контролювати план візитів на місця.
- Забезпечити доступність коштів для проведення оцінювання.

Реалізація – фактичне проведення оцінювання, включно з аналізом, написанням документів і звітуванням. Вона передбачає наступні обов'язки керівника:

- Забезпечити фахівцям з оцінювання повний доступ до записів, звітів, публікацій та всієї іншої актуальної інформації.
- Відстежувати хід оцінювання; надавати фахівцям з оцінювання зворотний зв'язок і вказівки протягом усього етапу реалізації.
- Оцінювати якість звітів, обговорювати сильні сторони й обмеження з фахівцями з оцінювання, щоб забезпечити відповідність проекту звіту ТЗ, обґрунтованість висновків і реалістичність рекомендацій.
- Організувати зустріч з фахівцями з оцінювання та ключовими зацікавленими особами для обговорення проекту звіту й отримання коментарів щодо нього.
- Затвердити кінцевий продукт оцінювання й організувати представлення його результатів зацікавленим особам.

Заключний етап, подальший супровід, передбачає оцінку ефективності, поширення результатів, відстеження рекомендацій та проведення аналізу оцінки постфактум. На цьому етапі керівник виконує такі обов'язки:

- Оцінювати роботу фахівців з оцінювання і вести відповідні записи.
- Поширити результати оцінювання серед основних зацікавлених осіб та інших цільових груп.
- Сприяти реалізації рекомендацій та використовувати результати оцінювання для формування програми зараз і в майбутньому, здійснювати регулярний моніторинг, щоб переконатися, що рекомендації виконуються.
- Очолити команду для вивчення вдалих аспектів, що заслуговують повторення, а також виявлення того, що варто було б зробити інакше.

На всіх трьох етапах керівник оцінювання може виступати в якості фасилітатора на зібраннях групи оцінювання, надаючи можливість всім учасникам висловлювати свої погляди й думки. В цій якості він відповідає за:

- формування порядку денного;
- дотримання групою встановленого порядку денного;
- надання можливості висловлення всіх точок зору;
- контроль за процесом прийняття рішень (консенсус або голосування).

Фахівець з оцінювання

Фахівці з оцінювання виконують основну частину роботи в ході оцінювання. Кількість фахівців залежить від масштабу оцінювання, бюджету й кількості вільних працівників.

Програма розвитку ООН (UNDP 2006) визначає наступні риси хорошого фахівця з оцінювання:

- досвід у конкретній предметній області;
- знання ключових проблем розвитку, особливо стосовно основних цілей, або ж здатність бачити «загальну картину»;
- обізнаність з діяльністю організації та її підходами;
- вміння планувати оцінювання, збирати дані, аналізувати їх та готувати звіти;
- вміння користуватися інформаційними технологіями.

Фонд народонаселення ООН (UNFPA 2007) визначає кілька потенційних обов'язків фахівців з оцінювання:

- Надавати дані стосовно плану оцінювання, вносити уточнення й конкретизувати завдання й питання оцінювання.
- Проводити оцінювання.
- Аналізувати надану інформацію/документацію.
- Розробляти/вдосконалювати інструменти збору додаткової інформації в міру необхідності, проводити або координувати додатковий збір інформації.
- Здійснювати візити на місця, проводити інтерв'ю.
- В ході спільних оцінок сприяти участі зацікавлених осіб.
- Надавати керівнику оцінювання регулярні звіти щодо процесу його проведення.
- Аналізувати й синтезувати інформацію, інтерпретувати отримані дані, розробляти й обговорювати висновки та рекомендації, робити висновки з отриманого досвіду.
- Брати участь в обговоренні проекту звіту оцінювання, виправляти будь-які фактичні помилки або хибні тлумачення.
- Брати участь у дискусіях на семінарах, представляючи результати оцінювання.
- Оформляти звіт оцінювання та готувати представлення його результатів.

Чимало організацій намагаються запровадити критерії професійної компетентності для осіб, що займаються оцінюванням. Також ведеться обговорення можливості сертифікації фахівців з оцінювання. (Питання компетентності та сертифікації фахівців з оцінювання розглядаються в главі 15.)

Управління персоналом, завданнями й бюджетами

Описуючи управління проектом, нерідко використовують порівняння. Хтось порівнює управління проектом з жонглюванням, тому що керівник повинен одночасно тримати в полі зору багато речей (люди, завдання, час, бюджет, якість). Інші порівнюють управління проектом з роботою диригента, адже керівник стоїть на віддалі й направляє зусилля багатьох компетентних людей з різними навичками.

Управління персоналом

Для того, щоб забезпечити добру спільну роботу групи оцінювання, потрібні різні вміння. В цьому розділі розповідається про навички розв'язання конфліктів та згуртування команди, які має опанувати керівник.

Керівникам також доводиться працювати зі страхами «господарів» програми. Хтось може непокоїтися через те, що оцінювання зосередиться на негативі, тому що саме негативні моменти привертають увагу. Це викликає стурбованість, адже негативні висновки можуть зруйнувати в цілому непогану програму. Можливі також страхи через те, що брак спостереженого впливу програми (який важко піддається вимірюванню) може призвести до хибного висновку про те, що програма недієва, а це приведе до скорочення фінансування.

Один зі шляхів подолання фахівцями з оцінювання страхів організаторів програм стосовно оцінювання – залучення їх до визначення проблем, які має охопити оцінювання. Їм також треба надавати можливість ознайомлюватися з робочими планами, висновками й рекомендаціями здійснюваного оцінювання.

Навички розв'язання конфліктів

Між людьми, які працюють у групах, нерідко виникають конфлікти. Для розв'язання конфлікту частіше за все потрібні вміння спілкуватися й слухати. Якщо ми говоримо про спілкування, то важливим є вміння говорити за себе, не переходячи на особистість співрозмовника. Вміння слухати передбачає звичку перевіряти, чи правильно ви розумієте співрозмовника («Тобто ви кажете, що... Я правильно зрозумів?»).

Блок 12.2 «Рекомендації щодо розв'язання конфліктів»

Керівники можуть допомагати розв'язанню конфліктів у різний спосіб:

- Організуйте зустріч конфліктуючих сторін. Дозвольте кожній з них стисло висловити свою точку зору, не перебиваючи. Якщо одна особа не дає другій висловитися чи починає її критикувати, зупиніть її.
- Дозвольте людям висловити свої почуття. Це сприяє зниженню градуса конфлікту.
- Попросіть обидві сторони описати дії, які б вони хотіли побачити від іншої сторони.
- Вислухайте обидві сторони. Запитайте себе, чи не спричинений цей конфлікт умовами роботи. Якщо це так, шукайте можливість змінити ці умови, щоб врегулювати конфлікт.
- Не ставайте на бік однієї із сторін. Нагадайте учасникам про ціль або завдання оцінки й спробуйте знайти спосіб допомогти обом сторонам досягти мети.
- Вимагайте, щоб учасники доклали зусиль для розв'язання протиріч. Дайте їм можливість зустрічатися для вирішення проблеми. Встановіть часові обмеження та відслідковуйте успішність дій з розв'язання конфлікту.

Не у всіх конфліктах має бути переможець і переможений. Найбільш конструктивні конфлікти завершуються «перемогою» обох сторін (блок 12.2).

Навички згуртування команди

Робота в команді вимагає набору різноманітних навичок. Зокрема:

- Вміння слухати: Члени команди повинні виробити вміння активно слухати. Таке вміння може бути найціннішим ресурсом команди.
- Вміння ставити питання: Члени команди мають ставити питання, з'ясовуючи й деталізуючи слова інших.
- Вміння переконувати: Членам команди може бути необхідно обмінюватися ідеями, деталізувати, захищати й переосмислювати власні ідеї.
- Вміння поважати: Члени команди мають поважати погляди інших, заохочувати й підтримувати їхні ідеї та зусилля.
- Вміння допомагати: Члени команди повинні допомагати один одному, пропонуючи підтримку, якщо вона потрібна.

Управління командою вимагає деяких додаткових навичок (див. блок 12.3).

Блок 12.3 «Рекомендації з покращення командної роботи»

Наступні поради допоможуть керівникам покращити спільну роботу членів команди:

- Повідомляйте про мету й важливість роботи.
- Слухайте активно, допомагайте іншим озвучувати їхні думки, перефразовуючи їхні слова.
- Співпрацюйте з колегами та іншими людьми, використовуючи техніку доброзичливого розпитування, цінуючи унікальний внесок і точку зору кожної людини.
- Відмежуйтеся від особистої упередженості заради пошуку відповідей на важкі питання на благо людей, заради яких працює проект, програма або оцінювання.
- Чесно й відкрито визнавайте результати оцінювання.

Джерело: NEIR TEC 2004.

Для максимального використання потенціалу команд можуть застосовуватися різні техніки. **Мозковий штурм** – техніка, що використовується для збору ідей групи людей за короткий проміжок часу. Кожна особа вносить свою ідею стосовно питання оцінювання, написаного на дошці. Одна особа пропонує одну ідею, потім інша особа пропонує іншу ідею. Координатор рухається по колу до тих пір, доки більше ніхто не зможе запропонувати нічого нового. Основне правило: кожна ідея записується на дошці – поганих ідей не буває. Ідеї не обговорюються. Завдяки цьому всі ідеї враховуються. Потім група разом визначає спільні ідеї, створюючи новий список.

Мозковий штурм: Техніка, що використовується для збору ідей групи людей за короткий проміжок часу

Картування концепцій:

Координований груповий процес, що використовується для отримання ідей від усіх учасників та їх групування й ранжування за важливістю

Картування концепцій – це координований груповий процес, що використовується для отримання ідей від усіх учасників та їх групування й ранжування за важливістю. Він починається з генерації ідей за допомогою мозкового штурму або діаграми групування. Кожна ідея записується на окремій картці й прикріпляється до стіни.

Коли нові ідеї перестають надходити, учасників просять подивитися на всі ідеї та сформувати з карток групи, що представляють однакову концепцію або тему. Після цього учасники вибирають відповідний ярлик для кожної концепції/теми. Далі координатор веде обговорення про важливість кожної з них та чому вони важливі (або ні). Деякі відносно неважливі концепції або теми на цьому етапі можуть виключатися, віднесені до них ідеї також виключаються або переносяться до інших рубрик. Розглядаючи групи карток на стіні, учасники можуть також можуть просити координатора перенести інші ідеї до інших категорій.

Якщо учасники будуть задоволені концепціями/темами та тим, які ідеї зібрано, координатор починає їх ранжування. Кожен учасник отримує певну кількість наліпок залежно від того, скільки важливих концепцій/тем координатор хоче бачити обраними. Якщо п’ять – кожен учасник отримує п’ять наліпок. Учасників просять прикріпити наліпку поряд з кожною з п’яти концепцій/тем, яку вони вважають найважливішою. Потім координатор визначає п’ять концепцій/тем з найбільшою кількістю наліпок і впорядковує їх за рангом (кількістю отриманих наліпок). Цей процес, за бажання, можна повторити також з ідеями, віднесеними до кожної концепції/теми.

Управління завданнями

Управління завданнями іноді здається простішим, ніж управління персоналом, але і цей процес має свої складнощі. Запорука успіху – постійна зосередженість на цілях оцінки й найважливіших завданнях.

Мапа завдань:

Таблиця, що показує доручення, відповідальну особу, дати початку й установленого завершення робіт

Діаграма Ґанта:

Діаграма, що використовується для візуалізації графіка проекту

Мапа завдань дозволяє звести докупи доручення, роздані всім учасникам, а також дати початку й завершення виконання цих доручень (див. табл. 12.1). Також можна використовувати **діаграму Ґанта** (див. рис. 12.1). Діаграма Ґанта – ефективний інструмент моніторингу. Смужку, що відповідає кожному завданню, можна заповнювати, відображаючи фактичний рух до виконання запланованого. Ця діаграма також може показувати залежність одного завдання від виконання іншого та вузькі місця, проходження яких вимагає виконання занадто багатьох завдань. Чимало програмних пакетів пропонують шаблони діаграми Ґанта.

Необхідно вести моніторинг діяльності, щоб забезпечити своєчасне виконання доручень. Якщо очікуваного руху вперед немає, фахівцю з оціню-

вання потрібно виявити перешкоди та вирішити, як їх усунути. Важливо, щоб члени команди не боялися повідомляти про проблеми, адже нерідко швидко виявлену проблему вирішити легше.

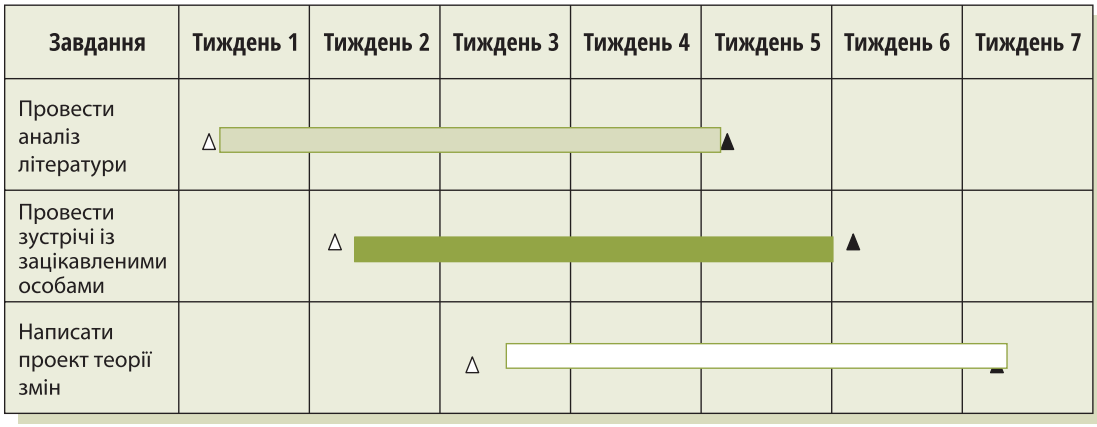
Хоча мати план важливо, важливо також зберігати гнучкість під час усунення перешкод. Корективи можна вносити, виділяючи більше часу чи ресурсів або скорочуючи кількість завдань.

Таблиця 12.1. Зразок мапи завдань

Завдання	Відповідальний(-і)	Початок	Закінчення
Провести аналіз зовнішньої літератури, виявити актуальні внутрішні й зовнішні теми	Ганна, Михайло й Тарас	01.07	04.07
Проаналізувати документи програми/проекту (наприклад, протокол ради, резолюцію, контрольні звіти), що стосуються інтервенції	Тарас	05.07	23.07
Запланувати й провести зустріч з клієнтом	Ганна	15.07	31.07
Визначити коло основних зацікавлених осіб і запланувати зустрічі	Тарас і Михайло	15.07	17.07
В письмовому вигляді стисло викласти результати зустрічі з клієнтом та прийняті рішення	Ганна	01.08	03.08
Провести зустрічі із зацікавленими особами й підсумувати питання	Ганна й Михайло	05.08	15.08
Написати проект плану оцінювання	Ганна	01.07	31.08

Джерело: Автори.

Рис. 12.1. Приклад діаграми Ґанта



Джерело: Автори.

Управління бюджетами

За Сандерсом (Sanders 1983), бюджет оцінювання зазвичай включає 10 категорій:

- заробітна плата й соціальний пакет персоналу оцінювання;
- консультанти;
- транспортні витрати й добові (для персоналу та консультантів);
- зв'язок (пошта, телефон тощо);
- друк і копіювання;
- обробка даних;
- друковані матеріали;
- приладдя й обладнання;
- субпідряди;
- накладні витрати (утримання приміщень, комунальні послуги).

Горн (Horn 2001) наводить детальний контрольний лист для формування бюджету оцінювання. Цей контрольний лист містить розбивку елементів витрат за наведеними вище категоріями. Керівник оцінювання може розрахувати витрати за кожною з категорій, щоб оцінити загальну вартість оцінювання.

Якщо витратна частина бюджету виявляється завеликою, існує необхідність взяти до уваги заходи економії, запропоновані Фітцпатріком, Сандерсом і Уортеном (Fitzpatrick, Sanders, and Worthen 2004):

- Залучайте добровольців або низькооплачуваних працівників.
- Користуйтеся послугами місцевих спеціалістів для збору даних, щоб не оплачувати відрядження.
- Навчіть низькооплачуваний персонал виконувати деякі необхідні завдання.
- Позичайте обладнання, кадри, матеріали та приладдя.
- Просіть внески в натуральній формі в організації, де працює зовнішній фахівець з оцінювання.
- Скорочуйте обсяг оцінювання (можливо, відкладаючи якісь його компоненти на майбутнє).
- Використовуйте існуючі показники, дані та звіти.
- Використовуйте недорогі методи збору даних, якщо точністю можна пожертвувати без серйозних наслідків.
- Використовуйте для поширення результатів громадські ЗМІ.
- Добийтеся вищої ефективності за рахунок правильного управління.

Те, яка частина бюджету програми або проекту буде виділена на оцінювання, залежить від багатьох факторів, у тому числі від помітності проекту або програми, запланованого масштабу оцінювання та репутаційного ризику, якщо оцінка охопить не все, що потрібно. Виданий Фондом Келлога «Ін-

струментарій оцінювання» (1998) повідомляє, що бюджет оцінювання має дорівнювати 5–10 відсоткам загального бюджету проекту. Інші спеціалісти з оцінювання рекомендують виділяти на оцінку 1–3 відсотки вартості програми або проекту (для більших проектів відсоток менший).

Управління проектами

Управління проектами означає одночасне управління всіма аспектами проекту – строками, обсягом, витратами, наявними ресурсами. Управління проектами – це процес, що складається з різних етапів. Існує багато моделей управління проектами й кілька програм сертифікації. Майкл Ґрір (2001), авторитетний спеціаліст з управління проектами, розробив модель, що акцентує на діях і результатах. Його модель розбиває управління проектами на п’ять етапів:

- започаткування;
- планування;
- виконання;
- контроль;
- завершення.

Кожний етап складається з кроків, які Ґрір називає діями. Він визначає 20 дій, що мають бути виконані за п'ять етапів (табл. 12.2).

Таблиця 12.2. 20 дій керівника проекту за Ґріром

Етап/дія	Результати успішного виконання
Етап започаткування	
1. Продемонструвати потребу в проекті і його доцільність	• Документи, що підтверджують потребу в очікуваних результатах проекту та описують, в загальному вигляді, ці очікувані результати, засоби їх створення, витрати на отримання й впровадження результатів, а також вигоди, що будуть отримані від впровадження очікуваних результатів
2. Отримати санкцію на виконання проекту	• Ухвала спонсора
	• Призначається куратор проекту, який офіційно визнає проект і видає керівнику проекту санкцію на використання ресурсів. Створюється статут проекту, який (a) офіційно визнає проект, (b) видається керівником, не задіяним у проекті, займає достатньо високу посаду в організації, щоб виділяти ресурси на потреби проекту, і (c) дає керівнику проекту право використовувати ресурси для виконання заходів проекту
3. Отримати санкцію на етап	• Спонсор дозволяє керівникові проекту використовувати ресурси організації для виконання певного етапу проекту

Етап/дія	Результати успішного виконання
	• Письмова ухвала для етапу, що (a) офіційно визнає існування етапу та (b) видається керівником, не задіяним у проекті, займає достатньо високу посаду в організації, щоб виділяти ресурси на потреби проекту
Етап планування	
4. Описати обсяг проекту	• Декларація обсягу проекту
	• План управління проектом
	• Деталізація робіт
5. Визначити й описати послідовність дій в рамках проекту	• Список дій, що мають бути виконані
	• Оновлення деталізації робіт
	• Мережевий графік проекту
6. Оцінити тривалість дій та потрібні ресурси	• Оцінка часу, необхідного для виконання кожної дії, а також припущення, зроблені для винесення кожної оцінки
	• Заява про потрібні ресурси
	• Оновлення списку дій
7. Розробити графік проекту	• Допоміжна інформація: використання ресурсів у динаміці, прогнозування грошових витрат, графіки замовлень/поставок тощо
8. Оцінити витрати	• Оцінка витрат, пов'язаних з виконанням кожної дії
	• Допоміжна інформація, в тому числі припущення й обмеження
	• План управління витратами, що описує рішення на випадок зміни обсягу витрат
9. Сформувати бюджет і план витрат	• Мінімальний обсяг витрат або бюджет в динаміці для вимірювання/моніторингу витрат
	• План витрат, в якому зазначається, скільки буде витрачено, на які ресурси та коли
10. Створити офіційний план якості (не обов'язково)	• План управління якістю, включно з операційними визначеннями
	• Контрольні листи для перевірки якості
11. Створити офіційний план комунікацій (не обов'язково)	• План управління комунікаціями, включаючи структури збору й розповсюдження, опис інформації, що має поширюватися, графіки, де вказується, коли буде створюватися інформація, метод оновлення плану комунікацій
12. Організувати й отримати в розпорядження персонал	• Розподілення функцій та обов'язків

Етап/дія	Результати успішного виконання
	• Штатний розклад • Організаційна схема з належною деталізацією • Персонал проекту • Список учасників команди проекту
13. Визначити ризики й запланувати заходи щодо мінімізації наслідків (не обов'язково)	• Документ з описом потенційних ризиків, включно з джерелами, властивостями й шляхами їх запобігання
14. Запланувати потребу й отримати сторонні ресурси (не обов'язково)	• План управління постачанням з описом того, як будуть залучатися підрядники
	• Технічне завдання або вимоги з описом предмета закупівлі (продукту або послуги) • Тендерні документи • Критерії оцінювання (для оцінки пропозицій підрядників) • Договір з одним чи кількома постачальниками товарів або послуг
15. Організувати план проекту	• Комплексний план проекту, що поєднує всі результати попередніх його планів
16. Завершити етап планування проекту	• Спонсор письмово затвердив план проекту («зелене світло» для початку реалізації)
17. Переглянути план проекту, якщо необхідно – внести зміни	• Впевненість у тому, що детальні плани з виконання конкретного етапу все ще актуальні й дадуть змогу отримати заплановані результати
Етап реалізації	
18. Виконати передбачені проектом заходи	• Результати роботи
	• Запити на внесення змін (у випадку розширення або звуження проекту) • Періодичні звіти про хід виконання проекту • Оцінка роботи команди та її покращення (в разі необхідності) • Запит тендерних пропозицій на очікувані результати, вибір підрядників (постачальників) та повторне укладання договорів • Адміністрування підрядів для отримання бажаних результатів роботи

Етап/дія	Результати успішного виконання
Етап контролю	
19. Контролювати заходи проекту	• Рішення щодо прийняття оглянутих результатів роботи • Корективні заходи, такі як коригування результатів роботи, внесення змін в робочий процес і т. д. • Оновлення плану й обсягу проекту • Список висновків з отриманого досвіду • Покращена якість • Заповнені контрольні листи оцінки (якщо є)
Етап завершення	
20. Завершити заходи проекту	• Офіційне, письмово задокументоване прийняття спонсором продукту цього етапу або заходу • Офіційне прийняття результатів роботи підрядника й оновлення його записів • Оновлена проектна документація, підготовлена для передання в архів • План подальшого супроводження та/або передачі результатів роботи

Джерело: Greer 2001. Передруковано з дозволу автора.

Підсумки

Планування надзвичайно важливе для забезпечення досягнення цілей оцінювання. Кожен, хто працює з оцінюванням, повинен розуміти свої функції й обов'язки відносно своїх колег. Технічне завдання – доступний друкований документ, в якому наводиться перелік цих обов'язків.

Вирішуючи питання управління персоналом, часом, завданнями й витратами, керівники повинні пам'ятати про мету оцінювання. В основі кожного рішення має бути прагнення досягти цієї мети.

Завдання до Глави 12

Практична вправа 12.1: Технічне завдання

Прочитайте наступне технічне завдання й виконайте вказівки, наведені після нього.

За пропозицією Світової організації торгівлі (СОТ) міжвідомча група, що займається програмою Інтегрованої системи, звернулася до Світового банку з проханням очолити проведення аналізу реалізації програми протягом останніх двох років. Спільний проект кількох організацій, Інтегрована система, покликана допомагати найменш розвиненим країнам скористатися можливостями, що їх надає міжнародна торгівельна система. Вона надає допомогу в сфері торгівлі: від семінарів з правил СОТ до покращення стану портів. Окремим країнам допомагають визначити їхні потреби, а потім збирають «круглий стіл» зі спонсорами, яких просять підтримати відповідну програму.

Аналіз має розглядати наступні теми:

- сприйняття завдань Інтегрованої системи (потрібно вивчити погляди зацікавлених сторін);
- аналіз допомоги в торгівельній сфері (інституційне будівництво, нарощування людських і організаційних ресурсів та інфраструктури);
- політичні міркування (розширення Інтегрованої системи, політичне середовище в контексті торгівлі та макроекономіки);
- адміністрування Інтегрованої системи;
- рекомендації на майбутнє.

Розглядаючи ці теми, консультант має оцінювати відповідність операцій Інтегрованої системи її цілям. Також необхідно оцінити ефективність витрат Системи для виконання завдань, а також дієвість координації дій основних відомств, що здійснюють нагляд за Системою, «круглими столами» та іншими заходами.

Очікується, що консультант вивчить документацію з реалізації програми, проведе інтерв'ю з операційним персоналом всіх залучених відомств, а також врахує думки представників бізнесу й уряду щонайменше двох з найменш розвинених країн, що отримували допомогу від Інтегрованої системи, причому одна з цих країн має бути африканською (пропонуються Ботсвана та Уганда). Також необхідно проконсультуватися з представниками основного донора. Звіт має займати близько 20 сторінок, не враховуючи додатків.

Працюючи, за можливості, в парах, проаналізуйте й висловіть критику технічного завдання та дайте відповідь на наступні питання:

1. Чи містить це ТЗ всі необхідні елементи?
2. Які елементи представлені повною мірою?
3. Які елементи потребують вдосконалення?

Практична вправа 12.2: Чи готові ви бути керівником?

Прочитайте цей список характеристик керівника (Reh 2007). Визначте, які з навичок у вас є, а які потребують вдосконалення.

Особисте

- Ви впевнені в собі та своїх здібностях. Ви задоволені тим, ким ви є, але продовжуєте навчатися і самовдосконалюватися.
- Ви, скоріше, екстраверт. Ви – не король вечірок, але й відсиджуватися в кутку нізащо не будете. Управління потребує роботи з людьми. Це заняття не для тих, хто не любить спілкування.
- Ви чесний і прямий. Ваші успіхи значною мірою залежать від довіри з боку інших.
- Ви не шукаєте причин, ви шукаєте можливості. Ви залучаєте людей до своєї роботи. Ви не відмовляєтесь від них через те, що їм чогось бракує.
- У вас є харизма. Керівники повинні бути лідерами. Ефективні лідери вміють змусити людей відчути свою присутність.

Професійне

- Ви послідовний, але сприймаєте зміни; на вас можна покластися, але ви здатні змінювати свою думку. Ви приймаєте рішення, але дослухаєтесь до порад.
- Ви – трохи ненормальний. Ви мислите нестандартно. Пробуєте нові підходи, і якщо вони не приносять результату, визнаєте свою помилку, але не просите вибачення за спробу.
- Вас не лякає необхідність працювати з цифрами. Ви розробляєте плани й графіки, а потім виконуєте їх.
- Ви готові швидко змінювати план, але вас не можна назвати легковажним.
- Ви розглядаєте інформацію як інструмент, яким треба користуватися, а не як джерело влади, яке треба тримати при собі.

Посилання та рекомендована література

Bemelmans Videc, M. L., R. C. Rist, and E. Vedung. 1997. *Sticks, Carrots, and Sermons: Policy Instruments and Their Evaluation.* Piscataway, NJ: Transaction Publishers.

Billson, Janet Mancini. 2004. *The Power of Focus Groups: A Training Manual for Social, Policy, and Market Research: Focus on International Development.* Burlington, RI: Skywood Press.

CDC (U.S. Centers for Disease Control and Prevention). 2001. *Introduction to Program Evaluation.* http://www.cdc.gov/tobacco/tobacco_control_ programs/surveillance_ evaluation/evaluation_manual/00_pdfs/Chapter1.pdf.

Chelimsky, E. 1987. "The Politics of Program Evaluation." *Social Science and Modern Society* 25: 24–32.

de Regt, Jacomina. 1996. "Mozambique: Country Implementation Review." In *World Bank Participation Sourcebook,* 83–88. Washington, DC: World Bank. http://www.worldbank.org/wbi/sourcebook/sb0211.pdf.

ESRC (Economic and Social Research Council). 2007. "Welcome to ESRC Today. Top Ten Tips." http://www.esrc.ac.uk/ESRCInfoCentre/Support/ Communications_toolkit/ communications_strategy/index.aspx.

Feuerstein, M. T. 1986. *Partners in Evaluation: Evaluating Development and Community Programs with Participants.* London: MacMillan, in association with Teaching Aids at Low Cost.

Fitzpatrick, Jody L., James R. Sanders, and Blaine R. Worthen. 2004. *Program Evaluation: Alternative Approaches and Practical Guidelines.* New York: Pearson Education.

Greer, Michael. 2001. "20 Key Project Manager Actions and Results." In *The Project Manager's Partner.* Amherst, MA: HRD Press. http://www.michaelgreer.com/20-actns.htm.

———. 2008. *Michael Greer's Project Management.* http://www.michaelgreer. com. Hawkins, Penny. 2005. "Contracting Evaluation." International Program for Development Evaluation Training (IPDET) presentation, Ottawa, June 30–July 1.

Horn, Jerry 2001. *A Checklist for Developing and Evaluating Evaluation Budgets.* http://www.wmich.edu/evalctr/checklists/evaluationbudgets.pdf.

Kellogg Foundation. 1998. *Evaluation Handbook.* Battle Creek, MI. http://www.wkkf.org/Pubs/tools/Evaluation/Pub770.pdf.

King, Jean A., Laurie Stevahn, Gail Ghere, and Jane Minnema. 2001. "Toward a Taxonomy of Essential Evaluator Competencies." *American Journal of Evaluation* 222: 229–47.

http://www.nbowmanconsulting.com/Establishing percent20Essentialpercent 20Programpercent20Evaluator percent20Competencies.pdf.

Kirkhart, K. E. 2000. “Reconceptualizing Evaluation Use: An Integrated Theory of Influence.” In *The Expanding Scope of Evaluation Use,* ed. V.J. Caracelli and H. Preskill, 5–24. New Directions for Evaluation No. 88. San Francisco: Jossey-Bass.

Kusek, Jody Zall, and Ray C. Rist. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System.* Washington, DC: World Bank.

Lawrence, J. 1989. “Engaging Recipients in Development Evaluation: The ‘Stake- holder’ Approach.” *Evaluation Review* 13 (3): 243–56.

Leeuw, Frans. 1991. “Policy Theories, Knowledge Utilization and Evaluation.” *OECD World Forum on Statistics: Knowledge and Policy,* vol. 4, 73–91. Organisation for Economic Co-operation and Development, Paris.

McNamara, C. 1999. *Checklist for Program Evaluation Planning.*
http://www.managementhelp.org/evaluatn/chklist.htm.

Muir, Edward. 1999. “They Blinded Me with Political Science: On the Use of Non– Peer Reviewed Research in Education Policy.” *Political Science and Politics* 32 (4): 762–64.

NCSTE (Chinese National Centre for Science and Technology Evaluation), and IOB (Policy and Operations Evaluation Department [the Netherlands]). 2006. *Country-Led Joint Evaluation of the ORET/MILIEV Programme in China.*
Amsterdam: Aksant Academic Publishers.

NEIR TEC (Northeast and the Islands Regional Technology in Education Consortium). 2004. *Gathering Together and Planning: Exploring Useful Skills for Educators to Develop Through Collaborative Evaluation*.
http://www.neirtec.org/evaluation/PDFs/Gathertogether3.pdf.

OECD (Organisation for Economic Cooperation and Development). n.d. *DAC Evaluation Quality Standards for Test Phase Application.* Development Assis- tance Committee, Paris.
http://www.oecd.org/dataoecd/30/62/36596604.pdf.

———. 2002. *OECD Glossary of Key Terms in Evaluation and Results-Based Management.* Development Assistance Committee, Paris.

Patton, Michael Q. 1977. “In Search of Impact: An Analysis of the Utilization of Federal Health Evaluation Research.” In *Using Social Research in Public Policy Making,* ed. C. H. Weiss, 141–64. Lexington, MA: Lexington Books.

———. 1997. *Utilization-Focused Evaluation.* 3rd ed. Thousand Oaks, CA: Sage Publications.

———. 2005. “Qualitative Methods and Analysis for Development Evaluation.” International Program for Development Evaluation Training (IPDET) presentation, Ottawa, June 27–29.

Reh, F. John. n.d. “How to Be a Better Manager.”
http://management.about.com/cs/midcareermanager/a/htbebettermgr.htm.

———. 2007. “Management Tips.” http://management.about.com/cs/generalmanagement/a/mgt_tips03.htm.
Rist, Ray C., and N. Stame. 2006. *From Studies to Streams: Managing Evaluative Systems.* Piscataway, NJ: Transaction Publishers.
Rossi, Peter Henry, Howard E. Freeman, and Mark W. Lipsey. 1999. *Evaluation: A Systematic Approach.* Thousand Oaks, CA: Sage Publications.
Rutman, L. 1980. *Planning Useful Evaluations: Evaluability Assessment.* Thousand Oaks, CA: Sage Publications.
Sanders, James R. 1983. “Cost Implications of the Standards.” In *The Cost of Evaluation,* eds. M. C. Alkin and L. C. Solman, 101–17. Thousand Oaks, CA: Sage Publication.
Schwartz, R. 1998. “The Politics of Evaluation Reconsidered: A Comparative Study of Israeli Programs.” *Evaluation* 4: 294–309.
Schwartz, R., and J. Mayne. 2004. *Quality Matters: Seeking Confidence in Evaluating, Auditing, and Performance Reporting.* Piscataway, NJ: Transaction Publishers.
Scriven, Michael. 2007. “Key Evaluation Checklist.” http://www.wmich.edu/evalctr/checklists/kec_feb07.pdf/.
Stufflebeam, Daniel L. 1999. “Evaluations Plans and Operations Checklist.” http://www.wmich.edu/evalctr/checklists/plans_operations.pdf.
Tilley, Nick. 2004. *Applying Theory-Driven Evaluation to the British Crime Reduction Programme: The Theories of the Programme and of Its Evaluations.* Thousand Oaks, CA: Sage Publications.
UNFPA (United Nations Population Fund). 2007. “Tool No. 5: Planning and Managing an Evaluation.” *Programme Manager’s Planning, Monitoring and Evaluation Toolkit.* New York. http://www.unfpa.org/monitoring/toolkit/5managing.pdf.
Weiss, Carol H. 1973. “Where Politics and Evaluation Research Meet.” *Evaluation* 1: 37–45.
———. 2004. *Identifying the Intended Uses of an Evaluation.* International Program for Development Evaluation Training (IPDET) presentation, Ottawa, July. http://www.idrc.ca/ev_en.php?ID=58213_201&ID2=DO_TOPIC.
Wholey, J. S. 1979. *Evaluation: Promise and Performance.* Washington, DC: Urban Institute.
———. 1994. “Assessing the Feasibility and Likely Usefulness of Evaluation.” In *Handbook of Practical Program Evaluation,* eds. Joseph S. Wholey, Harry P. Hatry, and Kathryn E. Newcomer, 15–39. San Francisco: Jossey-Bass.
Widavsky, A. 1972. “The Self-Evaluating Organization.” *Public Administration Review* 32: 509–20.
World Bank. 2006. *The World Bank Participation Sourcebook.* Washington, DC: World Bank. http://www.worldbank.org/wbi/sourcebook/sbhome.htm.

Веб-сайти

Canadian Evaluation Society. *Professional Designation Project.* http://www.evaluationcanada.ca/site.cgi?s=5&ss=6&_lang=EN.

Conflict Resolution Information Source. http://www.crinfo.org/index.jsp. Conflict Resolution Network. http://www.crnhq.org/.

EIB (European Investment Bank Group). 2008. "What Is the EIB Group?" http:// www.eib.org/attachments/general/whatis_eibgroup_2008_en.pdf.

ESRC (Economic and Social Research Council). "ESRC Society Today: Communication Strategy." http://www.esrc.ac.uk/ESRCInfoCentre/Support/Communications_toolkit/communications_strategy/index.aspx.

Europe Aid Cooperation Office. 2005. *Evaluation Methods.* http://ec.europa.eu/ europeaid/evaluation/methodology/index_en.htm.

———. 2006. *Methodological Bases.* http://ec.europa.eu/europeaid/evaluation/ methodology/methods/mth_en.htm.

European Commission. 2008. *Evaluation Network.* http://ec.europa.eu/regional_ policy/sources/docgener/evaluation/tech_en.htm.

Evalnet. 2000. http://www.evalnet.co.za/services/.

IBSTPI (International Board of Standards for Training, Performance, and Instruction). *Competencies for Internal Staff and External Consultants Conducting Evaluations in Organizational Settings.* http://www.ibstpi.org/Competencies/ evaluatorcompetencies.htm.

IDRC (International Development Research Centre). 2004. *Evaluation Planning in Program Initiatives.* http://web.idrc.ca/uploads/user-S/108549984812guideline-web.pdf.

MSH (Management Sciences for Health), and UNICEF (United Nations Children's Fund). "Quality Guide: Stakeholder Analysis." In *Guide to Managing for Quality.* http://bsstudents.uce.ac.uk/sdrive/Martin percent20Beaver/Week percent 202/Quality percent20Guide percent20-percent20Stakeholder percent20 Analysis.htm.

Treasury of Board of Canada Secretariat. 2005. *Improving the Professionalism of Evaluation.* http://www.tbssct.gc.ca/eval/dev/Professionalism/profession_e.asp.

UNDP (United Nations Development Programme). 2006. *Planning and Managing an Evaluation Website.* http://www.undp.org/eo/evaluation_tips/evaluation_ tips.html.

UNFPA (United Nations Population Fund). 2004. *Programme Manager's Plan- ning, Monitoring and Evaluation Toolkit.* http://www.unfpa.org/monitoring/ toolkit/5managing.pdf.

Weiss, Carol. *Identifying the Intended Use(s) of an Evaluation.* International Development Research Centre. http://www.idrc.ca/ev_en.php?ID=58213_201&ID2= DO_TOPIC.
Western Michigan University Evaluation Center. *Evaluation Checklists.* http:// www.wmich.edu/evalctr/checklists/checklistmenu.htm.
———. *The Checklist Project.* http://evaluation.wmich.edu/checklists.

Приклади технічних завдань

ADB (Asian Development Bank). 2008. *Model Terms of Reference Diagnostic City Water Assessment* (Links). http://www.adb.org/Water/tools/City-Water- Assessments.asp.
CIDA (Canadian International Development Agency). *Model for Evaluation Terms of Reference.* http://www.acdicida.gc.ca/INET/IMAGES.NSF/vLUImages/ Performancereview4/US$file/tor_sample_text.pdf.
FAO (Food and Agriculture Organization of the United Nations). *FAO Model Terms of Reference for a Joint Evaluation Mission.* http://www.fao.org/pbe/pbee/ common/ecg/233/en/torseng.doc.
UNDP (United Nations Development Programme). 2002. *Handbook on Monitoring and Evaluating for Results.* http://www.undp.org/eo/documents/HandBook/ ME-HandBook.pdf.

ГЛАВА 13

Представлення результатів

Після того, як збір та аналіз даних в основному завершено, настає час ознайомити інших зацікавлених осіб з попередніми результатами дослідження та розробити стратегію оприлюднення остаточних результатів. Обмін здобутими знаннями є одним із найважливіших компонентів оцінювання – необхідною передумова ефективних змін. Результати можна представляти у письмовій формі (у вигляді службових записок та звітів) або усно (під час брифінгів та презентацій).

Ця глава складається з чотирьох основних частин:

- Розробка комунікаційної стратегії
- Написання звіту з оцінювання
- Візуалізація інформації
- Проведення усної презентації

Розробка комунікаційної стратегії

Оцінювання, яке не надає інформації, необхідної для прийняття рішень, має незначну цінність. При розробці оцінки необхідно пам'ятати, що її метою є надання партнерам корисної інформації, яка дозволить приймати рішення. Оцінювання не ставить за мету отримання знань заради самих знань. У цьому полягає ключова відмінність між оцінюванням та дослідженням. Необхідно повідомляти результати оцінювання чітко, точно і в такій спосіб, який дозволить користувачам якомога краще скористатися цією інформацією.

Комунікаційна стратегія є необхідним компонентом оцінювання розвитку. В цій книзі ми наголошували на тому, що гарна комунікація починається з найбільш ранніх етапів та триває протягом всього процесу оцінювання. Це не та діяльність, яку можна здійснювати тільки наприкінці оцінювання. Клієнт і основні партнери мають бути залучені не лише до планування оцінювання, а й до розробки процесу та способів комунікації результатів і запрошення до надання зворотного зв'язку.

Комунікаційна стратегія повинна розроблятися на початковому етапі. У ній мають бути визначені наступні питання:

- хто має одержати інформацію щодо оцінювання;
- яка інформація потрібна;
- у якому форматі повинна бути ця інформація;
- коли має надаватися інформація;
- хто відповідає за її надання.

Якісна комунікаційна стратегія охоплює всі фази та продукти оцінювання. Вона використовує численні форми комунікації, які адаптують до потреб конкретних партнерів.

Комунікаційна стратегія для оцінювання може починатися з оглядового обговорення відповідних питань із замовником оцінки, після чого проводитимуться дискусії з групами місцевих зацікавлених осіб (табл. 13.1) Після обговорення із замовником можна провести формальний брифінг щодо попереднього дизайну оцінювання. Групам місцевих зацікавлених осіб можна надіслати повідомлення електронною поштою про те, що попередній дизайн розміщений на веб-сайті для коментування. Неформальна комунікація у процесі оцінювання є корисною. Клієнти та ключові зацікавлені особи, наприклад, можуть одержувати інформацію щодо проходження оцінювання за допомогою телефонних дзвінків або електронних листів.

Остаточні результати можна розсилати різним сторонам у різних формах. Донору, наприклад, можна забезпечити поглиблений брифінг щодо фактів, виявлених у процесі оцінювання, після чого надається формальний остаточний звіт. Місцевий програмний персонал може одержати загальний звіт та брифінг.

Таблиця 13.1. Зразок контрольного списку для попередньої фази оцінювання та фази планування

Аудиторія	Дія	Форма комунікації	Хто відповідальний	Термін виконання Дата здачі
Попередня робота				
Клієнт	Обговорення програмних потреб та графіка	Зустріч	Керівник групи	6/1
Національні та місцеві неурядові організації	Обговорення програмних питань	Зустрічі	Учасник групи B	6/5
Програмний персонал	Обговорення програмних питань	Зустрічі	Учасник групи C	6/11
Місцеві чиновники	Обговорення програмних питань	Зустрічі	Учасник групи B	6/10
Дорадчий комітет	Визначення кандидатур та розсилка запрошень	Ел. пошта	Учасник групи A	6/14
	Провести попередню зустріч щодо обговорюваних питань	Засідання дорадчого комітету	Учасник групи B	6/25
Спільнота, в якій проводиться оцінювання розвитку	Надати коментарі з обговорюваних питань	Повідомлення ел. поштою про відкритий веб-сайті щодо оцінювання	Учасник групи C	6/25
Дизайн/планування				
Дорадчий комітет	Перегляд та обговорення попереднього плану	Засідання дорадчого комітету	Учасник групи A	7/15
	Надання остаточного плану	E-mail	Учасник групи A	7/30
Клієнт	Надання остаточного плану	Усний брифінг	Керівник групи	7/22
Спільнота, в якій проводиться оцінювання розвитку	Перевірка проекту плану	Веб-сайт (моніторинг коментарів)	Учасник групи B	7/31

Джерело: Автори.

Що стосується фінального звіту, спеціалісти з оцінювання можуть використовувати брифінги, презентації та письмові звіти. Інформацію можна розповсюджувати для ширшої аудиторії за допомогою прес-релізів. Якщо планується прес-реліз чи прес-конференція, її часові рамки та метод проведення потрібно обговорити з основними зацікавленими особами.

Важливо включити процес отримання зворотного зв'язку, який поєднає спеціалістів з оцінювання та зацікавлених осіб і дозволить обговорити виявлені результати, висновки, альтернативні дії та наступні кроки. Якщо будуть проводитися обговорення у великих групах, спеціалісти з оцінювання мають визначити, чи існують будь-які перешкоди для того, щоб повідомляти результати одночасно різним групам зацікавлених осіб.

Навіть автори звітів з технічного оцінювання погоджуються, що це, можливо, не є найефективнішим способом повідомлення результатів оцінювання – Торрес, Прескілл та Пйонтек (Torres, Preskill, та Piontek 1997). Спеціалістам з оцінювання, можливо, доведеться використовувати альтернативні інструменти комунікації, такі як брошури, відео, дослідницькі звіти та короткі резюме з таблицями і графіками, щоб донести свої повідомлення (блок 13.1).

Написання звіту з оцінювання

Метою звіту є комунікація з читачами. Для забезпечення цього варто скористатися наступними порадами:

- Під час написання звіту не забувайте про його мету та цільову аудиторію. Дізнайтеся якомога більше про аудиторію, пишіть звіт таким чином, щоб він був якомога більш пристосованим до її особливостей.
- Використовуйте слова, що є
  - простими;
  - активними;
  - позитивними;
  - знайомими;
  - культурно чутливими.
- Якомога менше користуйтеся скороченнями та акронімами.
- Обмежуйте обсяг загальної інформації настільки, щоб лише представити звіт та зробити його контекст зрозумілим. Додатковий контекст можна, за потреби, включити у додаток.
- Надавайте достатньо інформації про планування та методи оцінювання, щоб читачі могли вважати звіт достовірним, але розуміти, що він охоплює обмежене коло питань. Попередьте читачів щодо уникнення інтерпретації результатів у неналежний спосіб. Знову ж таки, деталі можна вказати в додатку.

Блок 13.1. «Використання інноваційних комунікаційних технологій для підвищення інтересу до результатів оцінювання»

Для донесення результатів до зацікавлених осіб Лоренц, Галліксон і Тоул (Lawrenz, Gullickson, Toal) (2007) використовували різноманітні інноваційні техніки.

1. Створили опис випадку з відповідними висновками, який поширювали різними каналами, включаючи Інтернет, з метою задоволення потреб аудиторії. В описі випадку містилися описи виїзних візитів групи спеціалістів з оцінювання до 13 проектів. У ньому враховувалися не тільки потреби клієнта, а й потреби інших сторін, які могли б одержати користь від висновків.
2. Розподілили результати оцінювання на дев'ять дослідницьких звітів (стосовно співпраці, поширення, розробки матеріалів, підвищення кваліфікації, вдосконалення програм, прийняття на роботу та утримання на посаді, стабільності, дорадчих комітетів та оцінюваня) і розмістили їх на своєму веб-сайті. В цих дослідницьких звітах були поєднані звіти про виїзні візити, дані опитувань та попередні дослідження з кожного питання.
3. Для посилення інтересу до цих дослідницьких звітів була створена брошура, що мала привертати увагу. Брошура викликала настільки сильний інтерес до дослідницьких звітів, що весь комплект цих звітів був опублікований в альманасі, а результати дослідження розповсюджені серед ширшої аудиторії, ніж планувалося.
4. Розмістили онлайн посібник, що містив процедури, які використовувалися для проведення виїзних візитів. Після його появи на веб-сайті дослідників виявилося, що він цікавий для багатьох організацій та інших дослідників.
5. Розробили брошуру, в якій були викладені ключові кроки для забезпечення стабільності програми, яку вони досліджували, що була опублікована на веб-сайті у форматі PDF. Брошура була схвально сприйнята аудиторією, оскільки вона була простою в користуванні та надавала корисну інформацію стосовно вдосконалення проектів.
6. Провели відеоконференцію, на якій поглиблено обговорювали деякі зі своїх ідей.
7. Розмістили гіперпосилання на запис відеоконференції на своєму веб-сайті. Серед інформації на сайті були також посилання на інформацію про авторів дослідження, ключових спікерів на відеоконференції, допоміжні документи та відеоматеріали. Вся ця інформація була розміщена на веб-сайті, а також доступна на компакт-дисках.

- Напишіть виконавче резюме (описано далі).
- Розподіліть матеріал в основній частині звіту по розділах, які стосуються основних тем чи відповідають на ключові питання оцінювання.
- Помістіть основні питання на початку кожного розділу, а другорядні – ближче до кінця. Починайте кожен абзац, формулюючи питання, яке розглядається в ньому.
- Підкріплюйте висновки та рекомендації доказами.
- Розміщуйте технічну інформацію, включаючи матрицю планування та будь-які інструменти опитування, в додатку.
- Залиште достатньо часу для перевірки, перевірки і ще раз перевірки!
- Знайдіть людину, яка не бачила жодних матеріалів, і попросити її вичитати чернетку, визначити всі упущені чи незрозумілі моменти.
- Якщо можливо, запросіть незалежного редактора з досвідом у відповідній сфері та знанням методології оцінювання перевірити остаточний проект та запропонувати необхідні зміни до документа. Якщо перегляд силами колег недоцільний, попросіть переглянути звіт спеціаліста, який не займається оцінюванням.

Виконавче резюме

Виконавче резюме:
Коротке резюме звіту, що визначає розглянуті питання з оцінювання; описує використану методологію; підсумовує результати, висновки та рекомендації

Виконавче резюме звіту визначає питання, розглянуті в рамках оцінювання; описує використану методологію; підсумовує звіт.

Виконавче резюме надає читачеві можливість швидко зрозуміти основні повідомлення звіту. Це не просто стислий варіант висновків, що містяться у звіті. Воно не є «заманухою», яка обіцяє розкрити інформацію далі. Це має бути самостійний документ для читачів, які надто зайняті, щоб прочитати основну частину звіту.

За Скрівеном (Scriven 2007, ст. 1), метою виконавчого резюме є підсумок результатів, а не тільки процесу.

> Протягом всього процесу оцінювання ставте собі питання: як виглядатиме загальне резюме, враховуючи, що ви дізналися до цього, наскільки воно відповідає потребам клієнтів, аудиторії та зацікавлених осіб; це допоможе вам сконцентруватися на ключових моментах і не загубити нічого важливого.

Виконавче резюме має бути коротким: ідеальний обсяг – дві сторінки. Не варто робити його більшим за чотири сторінки. Дозвольте знову наголосити: в ідеальному варіанті – це дві сторінки з одинарним інтервалом; більше ніж чотири сторінки з одинарним інтервалом буде занадто!

Виконавче резюме може включати наступні компоненти:

- Короткий огляд чи вступний абзац, де вказана мета дослідження та проблема, що розглядається, який написаний так, щоб зацікавити читача.

- Опис оцінювання, де вказано основні питання, що розглядаються, а також стисла заява щодо обсягу та методики дослідження.
- Достатньо фонової інформації, щоб окреслити належний контекст дослідження.
- Визначте на власний розсуд, які результати можуть бути найбільш важливими для аудиторії.
- Певний спосіб зробити посилання на відповідні сторінки для читачів.
- Основні висновки та рекомендації.

Основна частина звіту

Основна частина звіту з оцінювання має містити наступні компоненти, що зазвичай поділені на глави (якщо звіт не дуже великий – розділи):

- вступ;
- опис підходу до оцінювання;
- результати;
- висновки;
- рекомендації.

У вступі до звіту вказуються наступні компоненти:

- мета оцінювання;
- загальна інформація, передумови оцінювання;
- цілі та мета програми, описані за допомогою моделі змін;
- питання оцінювання.

Короткий опис оцінювання включає наступні компоненти:

- мета;
- обсяг;
- питання;
- методологія;
- обмеження;
- залучені особи (дорадчий комітет, консалтингова фірма).

Результати потрібно викласти після опису оцінювання. При написанні цього розділу спеціалістам з оцінювання варто зробити наступне:

- представити результати у такий спосіб, який буде чітко зрозумілим для аудиторії;
- включати лише найважливіші результати;
- групувати результати відповідно до питань дослідження, основних тем чи проблем;
- використовувати діаграми, таблиці та інші графічні елементи, щоб ілюструвати основні моменти.

В кінці звіту зазвичай розміщують висновки та рекомендації, хоча читачі часто спочатку переглядають їх. Спеціалістам з оцінювання іноді важко відрізнити результати від висновків. Результати показують, що було виявлено під час оцінювання. Вони можуть вказувати на відповідність чи невідповідність критеріям. Результати потрібно підкріплювати доказами.

Висновки ґрунтуються на професійній оцінці результатів. Їх потрібно робити по кожній із другорядних цілей оцінювання, а також стосовно загальної мети проекту, програми, політики. У розділі «Висновки» не потрібно наводити нову інформацію.

Рекомендації спонукають до дій. Вони вказують, що варто було б зробити клієнту чи іншим зацікавленим особам за результатами звіту. Часто розробка рекомендацій виявляється складною. Вони не повинні бути надто деталізованими, щоб залишати за керівництвом прерогативу визначення певних рішень проблеми. Наприклад, оцінка може рекомендувати розробку політики ціноутворення для діяльності із технічної підтримки. Вона не надаватиме проекту такої політики у готовому вигляді та не визначатиме конкретні деталі. Проте вона може вказати основні компоненти, які мають включатися у політику ціноутворення.

Водночас рекомендації не повинні мати настільки узагальнений характер, щоб бути непридатними на практиці. Рекомендації мають бути достатньо чіткими та конкретизованими, щоб всі розуміли, яких заходів треба вжити для їх виконання, яка організація чи підрозділ має виконати певні дії і коли це варто зробити.

Звіти не повинні містити довжелезні списки рекомендацій. Варто обмежити кількість основних рекомендацій до трьох-чотирьох або згрупувати їх у зручну кількість груп (три-чотири) із допоміжними компонентами, за потреби. Треба зважати на тон рекомендацій. Пам'ятайте, що звіти не приймають рішень, це роблять люди.

Від рекомендацій мало користі, якщо їх не дотримуються. Одним із способів контролю виконання рекомендацій є створення системи відстежування дій за рекомендаціями. Така система дозволяє зацікавленим особам моніторити виконання рекомендацій за результатами оцінювання. Вона відстежує кожну рекомендацію, а також прогрес, якого було досягнуто внаслідок впровадження цієї рекомендації, в тому числі:

- дату рекомендації;
- хто відповідальний за вжиття заходів;
- реакцію/прогрес.

У таблиці 13.2 наведена проста матриця, що може використовуватися для системи відстежування рекомендацій. У цій матриці спеціалісти з оцінювання можуть заповнити перші дві колонки, а менеджери – відстежувати прогрес та забезпечувати виконання рекомендацій особами, яким вони були надані.

Як ще один приклад можна навести міжнародну фінансову корпорацію (МФК), структурний підрозділ Світового банку, що працює з приватним сектором. МФК має власну систему відстежування виконання рекомендацій. Звіти з оцінки, які готує Незалежна група оцінювання (НГО) МФК, включають рекомендації для Керівництва МФК та засоби реагування для Керівництва МФК. Ці рекомендації обговорюються на засіданні Комітету МФК з ефективності розвитку (КЕР). КЕР вимагає періодичних звітів щодо прогресу своїх рекомендацій, включаючи рівень прийняття та стан виконання. НГО разом із Керівництвом МФК розробила Звіт про відстеження управлінських дій (Management Action Tracking Record MATR).

Звіт MATR призначений для підтримання достовірності процесу звітування, оскільки в ньому ні НГО, ні МФК не можуть змінити фіналізовані рейтинги. Графік 13.1 ілюструє два етапи MATR. На першому етапі НГО та МФК узгоджують індикатори з оцінки кожної нової рекомендації, на другому статус та рівень прийняття кожної рекомендації, що впроваджується, періодично оновлюються та звітуються, не обов'язково повинні збігатися.

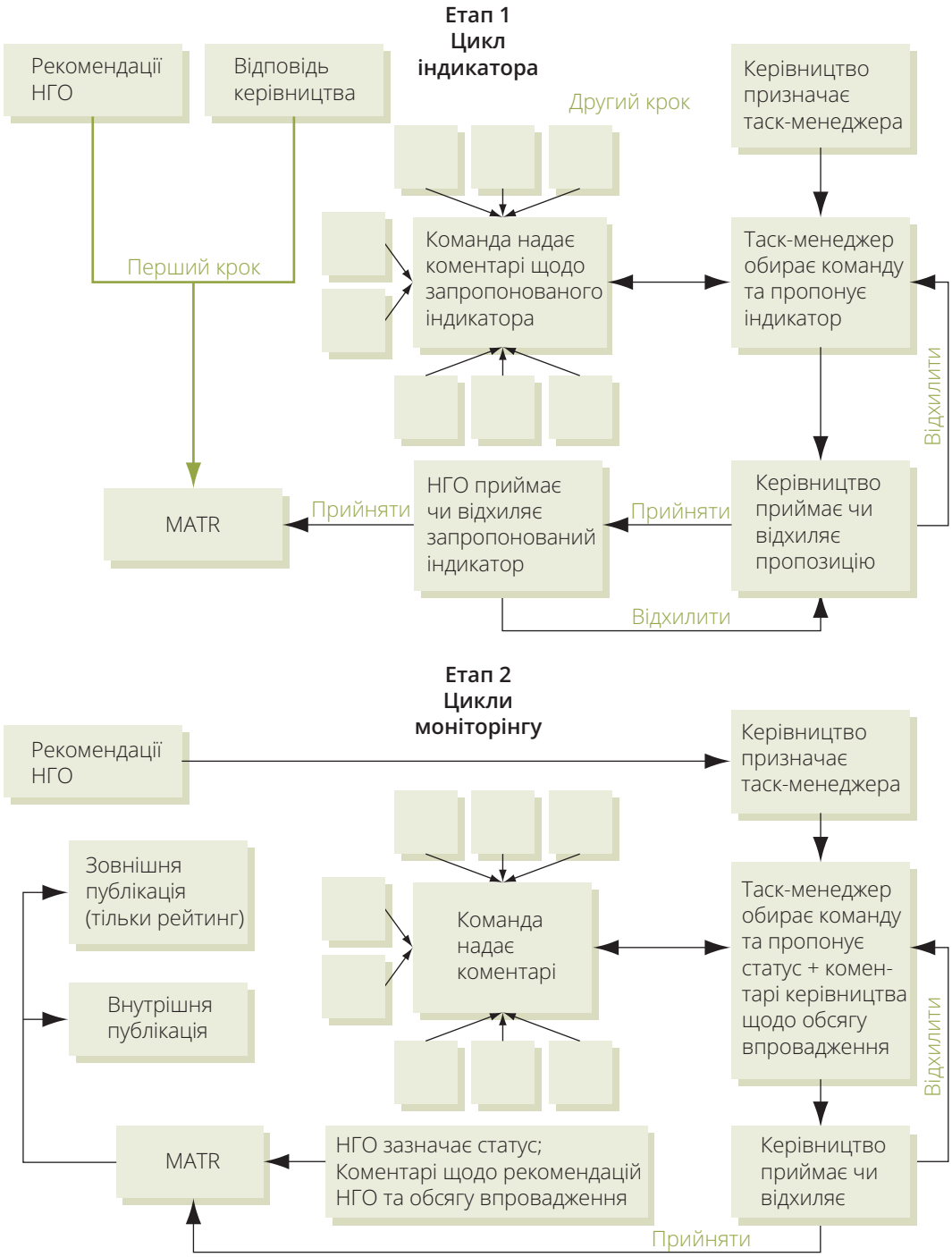
Таблиця 13.2. Система відстеження рекомендацій

Рекомендація	Джерело	Дата	Автор	Відповідальна особа/прогрес
1.				
2.				
3.				
4.				

Джерело: Автори.

Якщо рекомендації вже впроваджені, замінені чи більше не є релевантними, їх роблять неактивними. Рекомендації НГО, що не приймаються керівництвом МФК, не підлягають відстеженню.

Рис. 13.1. Двоетапний звіт щодо відстеження управлінських дій НГО MATR



Джерело: Незалежна група з оцінки Міжнародної фінансової корпорації 2008.

Візуалізація інформації

Візуальна інформація може зробити звіт цікавим, більш чітко доносити інформацію, ніж текст, а також привертати увагу читачів до певних пунктів. Отже, ефективне використання графіки може посилити вплив звіту з оцінювання.

Ілюстрації

Ілюстрації допомагають наочно підкріпити певні пункти звіту. Їх потрібно завжди використовувати з певною метою, а не лише в декоративних цілях. Ілюстрації на обкладинці звіту можуть допомогти візуалізувати загальну тематику звіту; для розміщення ілюстрацій в тексті звіту має бути конкретна та релевантна причина.

Всі ілюстрації варто згадувати в тексті, щоб читач розумів, що саме він може побачити на ілюстрації. Приклади ілюстрацій, що можуть використовуватися в оцінці, включають карти, ескізи, креслення, фотографії. За потреби можна одержати дозвіл на використання певних зображень.

Мапи

Мапи у звітах можуть використовуватися з такими цілями:

- вказати місце розташування програми;
- забезпечити контекст;
- вказати географічне охоплення чи поширення програми;
- показати базу вибірки;
- вказати коефіцієнти чи рівні розподілу по території за допомогою ізоліній (Cummings 2003).

Найбільш ефективне використання мап вимагає дотримання наступних вимог:

- забезпечення простоти читання та інтерпретації мап (заливка чи штрихування різних частин не повинні зливатися);
- використання найновішої версії;
- визначення джерела інформації;
- додавання стрілки, що вказує масштаб, за потреби.

Ескізи та креслення

В рамках збору даних спеціалісти з оцінювання інколи просять учасників зробити ескізи (ілюстрація 13.1). Ці ескізи можна використовувати з наступною метою:

- додати звіту цікавості;
- персоналізувати звіт;
- відобразити методологічний підхід;
- замінити більш вишукані ілюстрації, якщо технічні потужності не дозволяють створити їх;
- додати гумору;
- безпосередньо продемонструвати враження учасників.

Лінійні малюнки також можуть використовуватися як ілюстрація принципу роботи певних систем чи взаємозв'язків між елементами. Вони дуже корисні, якщо процес простіше схематично намалювати, ніж описати.

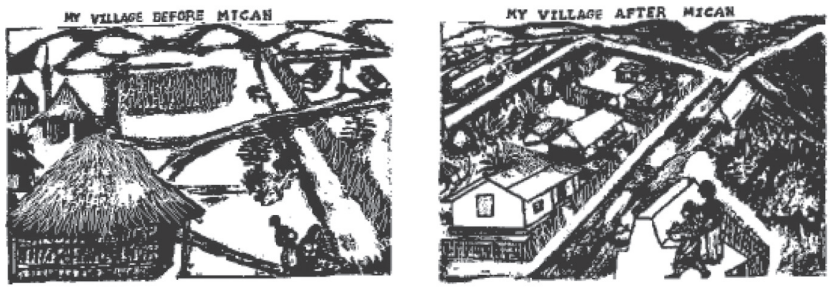
Фотографії

Фотографії є корисним додатком до звіту. Вони можуть:

- надавати контекст;
- показувати обсяг роботи на місцях (прогрес);
- фіксувати те, що безпосередньо спостерігається (наприклад, типи будинків, скупченість у населеному районі);
- ознайомлювати аудиторію із ситуацією на місцях;
- надавати докази (Cummings 2003).

Як й інші типи ілюстрацій, фотографії ніколи не можна використовувати виключно як оздоблення. Levin, Anglin, та Carney (1987) вивчили і підсумували інформацію щодо використання фотографій та малюнків у матеріалах. Вони дійшли двох висновків щодо того, яким чином ілюстрації можуть сприяти чи заважати вивченню інформації з тексту:

Ілюстрація 13.1. Малюнок дитини, що відображає її враження щодо стану села до та після проекту



Джерело: Cummings 2003.

- Якщо ілюстрації відповідають змісту, можна очікувати, що результативність вивчення інформації буде помірною чи високою.
- Якщо ілюстрації не відповідають змісту, чи навіть гірше, суперечать йому, читач не вивчить нову інформацію або навіть дійде неправильних висновків.

Отже, фотографії та інші картинки, що використовуються у звіті, потрібно додавати лише з чіткою та конкретною метою.

Графіки та діаграми

Графіки та діаграми є візуальним представленням даних. В ідеалі вони самі по собі є вичерпним зображенням даних, яке не потребує допоміжного тексту.

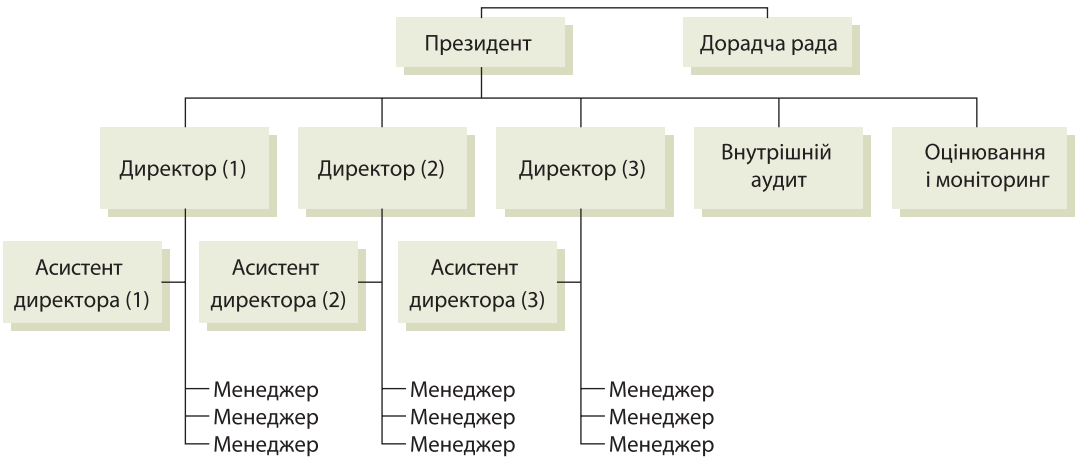
Організаційна діаграма

Організаційна діаграма ілюструє ієрархію в рамках організації (рис. 13.2). Такі схеми стисло і чітко демонструють обов'язки та структуру звітування.

Організаційна діаграма: Діаграма, що ілюструє ієрархію та структуру підзвітності в організації

Організаційні діаграми часто включають до звітів з оцінювання, оскільки розуміння структури установи є першим кроком до розуміння її діяльності. Більшість текстових редакторів мають функції побудови таких діаграм.

Рис. 13.2. Приклад організаційної діаграми



Джерело: Автори.

Діаграми Гантта

Діаграма Гантта: Діаграма, що ілюструє графік впровадження проекту

Діаграми Гантта використовуються для планування (рис. 13.3). Вони особливо корисні для управління проектами, зокрема їх планування.

Графіки

Графік: Візуальне представлення відносин між двома чи більше масивами даних

Графік повинен передбачати чітке повідомлення, що не потребує багато зусиль для інтерпретації.

Кожний графік повинен містити рядок назви, рядок цифр та посилання на джерело. У назві потрібно зазначати рік, до якого відносяться дані (якщо застосовується) (табл. 13.3).

Для представлення даних можуть бути корисними принаймні чотири типи графіків чи діаграм. Кожна з них слугує певній меті (табл. 13.4).

Лінійний графік: Графік, який створюється шляхом поєднання лінією кількох точок, що представляють дані; зазвичай використовується для демонстрації змін з плином часу.

Лінійний графік. **Лінійний графік** зазвичай показує, яким чином змінна (або змінні) змінюються з часом (рис. 13.4 та 13.5). Наприклад, спеціалісти з оцінювання можуть використовувати лінійний графік, щоб показати зростання чи зниження вартості харчових продуктів протягом року, зміни чисельності населення упродовж років чи оцінки студентів за шість тижнів.

Лінійні графіки можуть показувати зміну одного чи кількох елементів з плином часу. Лінійні графіки – гарний спосіб показати неперервні дані, тобто такі, що залежать від певних інтервалів чи коефіцієнтів. Інтервальні дані – це дані, поділені на діапазони, для яких відстань між діапазонами має значення. Дані коефіцієнтів – інтервальні дані, які можуть мати значення, що дорівнює абсолютному нулю. Прибуток є коефіцієнтом, оскільки нуль доларів – це дійсно «відсутність прибутку».

Діаграма Гантта: Діаграма, що ілюструє графік впровадження проекту

Рис. 13.3. Приклад діаграми Гантта

Завдання	Тиждень 1	Тиждень 2	Тиждень 3	Тиждень 4	Тиждень 5	Тиждень 6	Тиждень 7
Здійснити літ. редакцію	△				▲		
Провести засідання партнерів		△				▲	
Проект теорії змін			△				▲

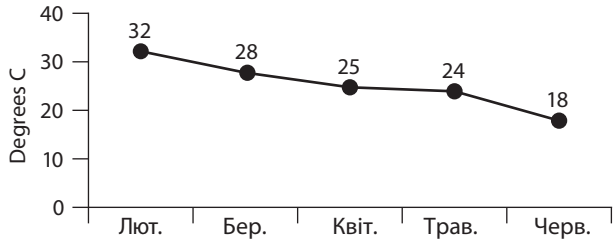
Джерело: Автори

Таблиця 13.3. Елементи графіків

Назва елемента	Опис
Заголовок	Всі графіки та діаграми повинні мати заголовок, щоб аудиторія могла відразу зрозуміти, про що в них йдеться.
Горизонтальна вісь, або вісь *x*	Горизонтальна вісь, або вісь *x*, є горизонтальною координатною віссю графіка чи діаграми, яка представляє певну змінну (наприклад, час)
Вертикальна вісь, або вісь *y*	Вертикальна вісь, або вісь *y*, є горизонтальною координатною віссю графіка чи діаграми, яка представляє другу змінну (наприклад, вартість)
Вихідна точка	Вихідна точка – точка перетину вертикальної та горизонтальної осей
Координатна сітка	Багато графіків мають координатну сітку, яка допомагає порівнювати дані, чітко показуючи рівні. Однак потрібно використовувати обмежену кількість ділень, щоб не нагромаджувати елементи на графіку
Підписи осей	Підписи осей *x* та *y* дуже важливі. Вони визначають, що саме вимірюється, та одиниці виміру (роки, метри, фунти, квадратні милі, кубічні тонни, долари, градуси). Наприклад: вартість (у дол. США), відстань (у км)
Шкала осей	Повинна бути шкала для осей *x* та *y*, що вказуватиме значення змінних. Ретельно обирайте шкалу, щоб вона охоплювала весь діапазон даних. Обирайте пропорцію між шкалами осей так, щоб якомога краще проілюструвати відношення між змінними
Фактичні значення	Багато графіків та діаграм також включають фактичні значення введених даних, що показуються у формі додаткового тексту в полі графіка. Ці примітки допомагають читачу краще зрозуміти фактичну ситуацію
Координати	Координата – точка на шкалі, де перетинаються значення по осі *x* та осі *y*. Координати можуть мати різний вигляд, залежно від обраного типу графіка (точка, пік, вершина смужки).

Джерело: Автори.

Рис. 13.4. Середнє зниження температури у провінції А, лютий–червень 2008



Джерело: Автори.

Таблиця 13.4. Типи та призначення різних видів графіків

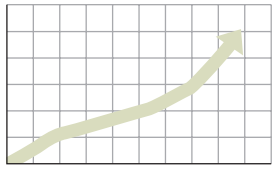
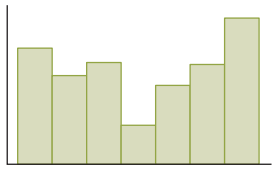
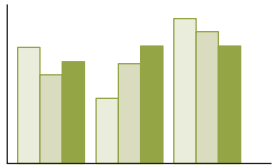
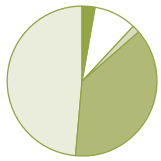
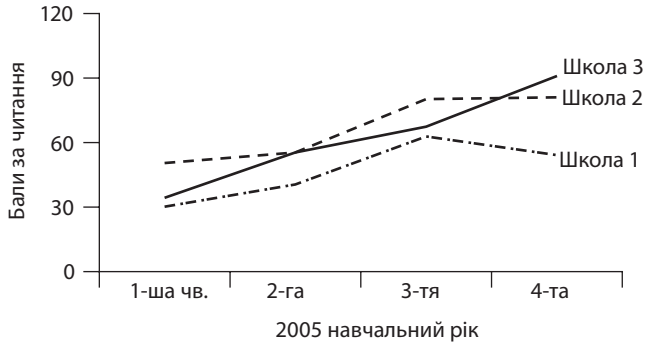
Тип графіка	Приклад	Мета
Лінійний графік		Показувати зміни ситуації з плином часу
Лінійчата діаграма		Порівнювати лінійні чи одновимірні характеристики
Гістограма		Порівнювати дві чи більше характеристики із значеннями звичайних змінних
Кругова діаграма		Показувати частини цілого
Діаграма розсіювання		Показувати тенденції чи відносини

Рис. 13.5. Відмінності між середніми балами за читання у трьох школах району, 2004/05



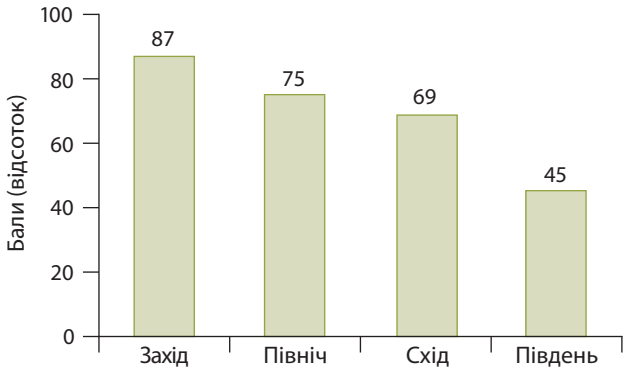
Стовпчаста діаграма. **Стовпчасті діаграми** використовують стовпчики (видовжені прямокутники) для того, щоб представити кількість та дозволити нам порівняти цифри. Треба бути уважними, називаючи їх, щоб читач розумів, який фактор вони представляють.

Стовпчаста діаграма: Тип діаграми, який використовує горизонтальні чи вертикальні прямокутники, щоб представити порівняння кількості

Існує два типи стовпчастих діаграм: лінійчаті, які показують інформацію про одну змінну, та гістограми, що показують інформацію про декілька змінних. Стовпчики можна форматувати вертикально чи горизонтально. Використовуйте гістограму, щоб порівняти кілька груп даних по одній змінній. Наприклад, метою оцінки є порівняння коефіцієнта рекультивації земель після гірничовидобувних робіт у трьох регіонах країни. Для представлення інформації можна використати гістограму. В іншому прикладі: діаграми з двома стовпчиками можна використовувати для порівняння відповідей хлопчиків з відповідями дівчат на певні запитання опитування.

Стовпчасті діаграми часто використовуються для представлення номінальних чи категорійних даних. Номінальні чи категорійні дані не є порядковими, і призначення номерів категоріям робиться на власний розсуд (1 = Схід, 2 = Північ, 3 = Південь), категорії необхідно чітко пояснити у легенді.

Рис. 13.6. Бали за тестом на знання щодо виховання дітей по регіонах



Джерело: Автори.

Нижче наведено приклади стовпчастих діаграм. На рис. 13.6 показано бали за тестом щодо догляду за дітьми, який проходили чотири вагітні жінки. **Лінійчаста діаграма** показує дані щодо конкретних змінних для однієї групи (рис. 13.6). **Гістограма** забезпечує дані щодо конкретних змінних для більш ніж однієї групи (рис. 13.7).

Кругова діаграма. Кругова діаграма може бути корисною при обмеженій кількості секцій (рис. 13.8). Її не варто використовувати, якщо кількість секцій перевищує вісім чи є меншою за три.

Кругова діаграма: Кругла діаграма, поділена на секції, кожна з яких представляє відносну величину відповідної змінної

Рис. 13.7. Різниця між відповідями на питання асистентів та директорів

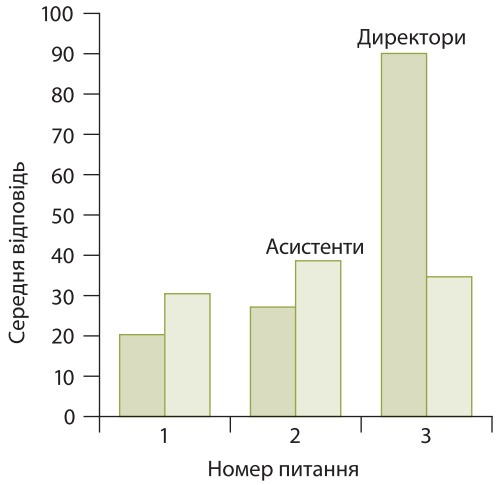
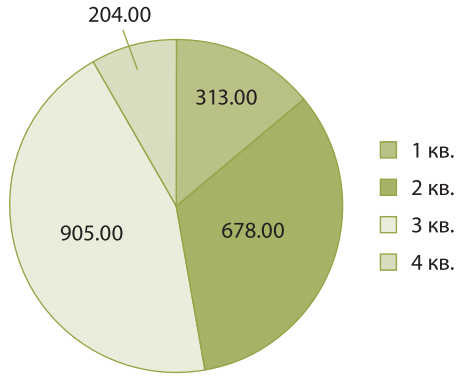


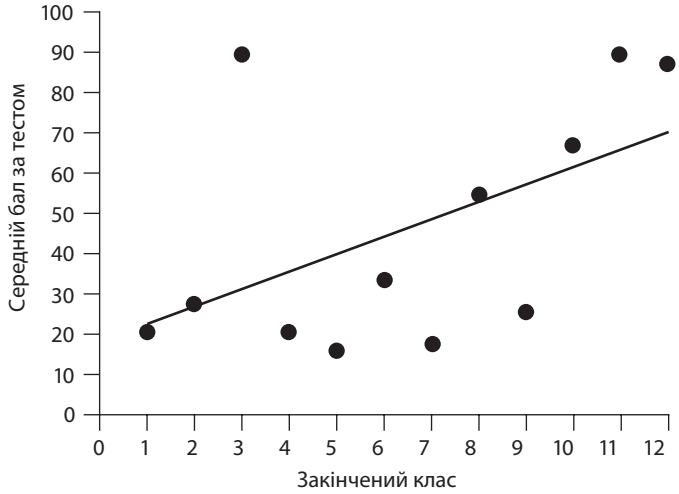
Рис. 13.8. Найвищі тарифи на електроенергію у канадських доларах, 3-й квартал



Джерело: Автори.

Діаграма розсіювання. **Діаграма розсіювання** – представляє взаємозв'язок набору даних, розміщуючи їх відносно горизонтальної та вертикальної осі (рис. 13.9). Якщо змінні корелюють, лінія чи крива будуть очевидними. Чим краща кореляція, тим ближче на лінії чи кривій знаходитимуться точки. Діаграма розсіювання, на якій не видно якоїсь конкретної фігури, означає, що між величинами немає безпосереднього зв'язку.

Рис. 13.9. Середні бали. Порівняння балів за тестом по роках навчання



Джерело: Автори, фіктивні дані

Таблиці

Таблиці можуть використовуватися для організованого представлення інформації. У звіті використовують два типи таблиць: таблиці даних і таблиці класифікації (матриці).

Таблиці даних

Зручно представити числову інформацію в таблицях, так званих таблицях даних, тільки якщо набори даних не дуже об'ємні (Tufte 1983). Більші таблиці даних зазвичай є підставою для обрання іншого формату представлення даних, такого як лінійні діаграми чи гістограми; зазвичай вони наводяться у додатках.

Як і у випадку картинок та ілюстрацій, аудиторія не зможе автоматично зрозуміти, про що йдеться в таблиці. Було б добре, якби у заголовку таблиці описувалося, які дані вона представляє і до чого вони відносяться. Короткий опис того, що представлено в таблиці, можна також включити у наративну частину звіту. Якщо ви представляєте дані у вигляді таблиці, завжди вказуйте рік та джерело даних.

Eherenburg (1977) підсумовує керівні принципи розробки таблиць для представлення інформації.

- Округляйте дані до не більш як двох знаків після коми. Це полегшує порівняння [*Примітка:* ми пропонуємо округляти до цілих, аудиторія не завжди налаштована порівнювати до десятих і сотих.]
- Вказуйте суми та середнє значення рядків і стовпчиків (за необхідності), щоб дати аудиторії змогу порівнювати значення окремих клітинок з цими значеннями.
- Розміщуйте найважливіші дані у колонки, оскільки це дозволить читачу робити зручне порівняння.

Приймаючи рішення щодо формату таблиці, не забувайте, що забагато ліній (які розділяють клітинки таблиці) ускладнюють читання. Це показано у наступних двох таблицях, причому табл. 13.5 є прикладом даних у таблиці, що містять забагато ліній, табл. 13.6 показує ті ж самі дані у таблиці з меншою кількістю рядків. Зверніть увагу, що у третій таблиці саме дані, а не лінії впадають в око. Також подивіться, що останній рядок таблиці показує середні значення стовпчиків з даними.

Класифікаційні таблиці (матриці)

Класифікаційна таблиця, або матриця, має шаблон, у якому показано, як різні елементи розподіляються за різними факторами (табл. 13.7). Принаймні два фактори сортування вказують на подібність чи розбіжність між

елементами, що класифікуються. Класифікаційні таблиці також допомагають проілюструвати комплексну інформацію. Матриця планування є класифікаційною таблицею.

Таблиця 13.5. Приклад даних в таблиці, яка має забагато ліній. Демографічна інформація про учасників

№ учасника	Зріст	Вага	Вік	Район
1	44	30	7.2	Північ
2	46	35	7.1	Схід
3	40	20	7.6	Північ
4	32	22	7.2	Південь
5	29	23	7.0	Південь
6	50	38	7.8	Північ
7	44	30	7.3	Захід
8	44	28	7.3	Захід
9	42	30	7.5	Схід
10	48	45	7.9	Південь
Середнє	38.09	27.36	6.72	

Джерело: Автори.

Таблиця 13.6. Приклад даних в таблиці, яка не перевантажена лініями. Демографічна інформація про учасників

№ учасника	Зріст	Вага	Вік	Район
1	44	30	7.2	Північ
2	46	35	7.1	Схід
3	40	20	7.6	Північ
4	32	22	7.2	Південь
5	29	23	7.0	Південь
6	50	38	7.8	Північ
7	44	30	7.3	Захід
8	44	28	7.3	Захід
9	42	30	7.5	Схід
10	48	45	7.9	Південь
Середнє	38	27	7.0	

Джерело: Автори.

Таблиця 13.7. Приклад класифікаційної таблиці. Стратегії подолання бідності: країни, де проводилося дослідження

Країна	Дата початку	Роки впровадження	Чи була закінчена перевірка
Ефіопія	17-09-2002	4,7	Так
Гвінея	25-07-2002	4,9	Так
Мавританія	06-02-2001	6,3	Так
Мозамбік	25-09-2001	5,7	Так
Танзанія	30-11-2000	6,3	Так

Джерело: Adapted from World Bank 2004.
Примітка: дані станом на травень 2007 року.

Підвищення ефективності візуальної інформації

За Едвардом Тафтом (Edward Tufte), експертом з візуалізації інформації, візуальні образи мають досягати наступних цілей:

- Показувати дані.
- Заохочувати думати скоріше про зміст графіки, ніж про методологію, графічний дизайн, технологію створення графіки чи щось інше.
- Не спотворювати того змісту, який мають передавати дані.
- Представляти багато цифр на малій площі.
- Змістовно представляти великі набори даних.
- Заохочувати порівнювати різні масиви даних.
- Показувати дані на кількох рівнях деталізації, від загального огляду до складної структури.
- Бути тісно інтегрованими зі статистичним та словесним описом даних (Tufte 1983).

Тафт називає заповненість простору фарбою, що використовується для презентації графічних даних, «коефіцієнтом корисності чорнила». Більшість чорнила, що використовується у графічному елементі, має бути корисним; координатна сітка, підписи та інші відволікаючі елементи повинні бути якомога менш помітними (блок 13.2). Він називає декоративні елементи, що не несуть додаткової інформації, «графічним сміттям».

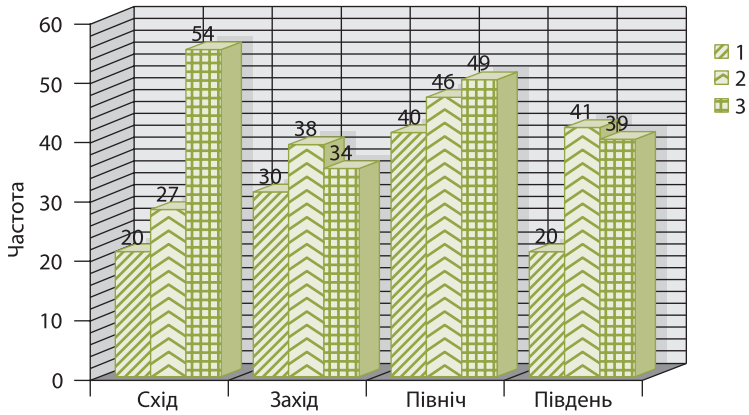
Блок 13.2 «Поради щодо створення ефективних рисунків»

Наступні поради допоможуть збільшити кількість значущих елементів, «коефіцієнта корисності чорнила», та створити ефективні малюнки:

- Представляйте достатньо даних, щоб передати повідомлення, але обмежуйте кількість даних, що представлені у кожному окремому малюнку.
- Графіка повинна обов'язково передавати певне повідомлення; повідомлення має бути зрозумілим.
- Використовуйте читабельну гарнітуру (шрифт), розміром не менш як 10 пунктів. Не використовуйте занадто багато різних шрифтів і типів написання. Вони повинні бути читабельними для аудиторії (розташовуйте надписи за можливості горизонтально та уникайте перетинання з іншими елементами графіка).
- Концентруйтеся на даних, а не на тому, що їх передає.
- Уникайте жирних координатних сіток; перевантажених координатних сіток; надлишку поміток та галочок; надлишкового представлення простих даних, блоків, відтінків та вказівників, а також легенд.
- Замінюйте блоки навколо графіків підписами, орієнтованими по осі *x* чи *y*.
- Приховуйте координатну сітку в лінійчастих діаграмах.
- Використовуйте галочки (без ліній), щоб показати фактичне розташування тих чи інших даних.
- Оптимізуйте графіку, замінюючи смужки одинарними лініями, вилучаючи те, що не є «корисним чорнилом», видаляючи лінії з осей та починаючи відлік по осі *x* чи *y* з фактичного значення даних [рамки діапазону].
- Дотримуйтеся точності шкал та чітко підписуйте їх. Надавайте якомога більше інформації стосовно шкал, але включайте її у графік таким чином, щоб не загромаджувати її.
- Уникайте занадто деталізованих та непотрібних послідовностей елементів.
- Враховуйте, що у надрукованому документі відтінки можуть відрізнятися від того, що ви бачили на моніторі; підбирайте їх так, щоб вони не зливалися.
- Не шкодуйте простору, що дозволяє очам відпочити.
- Вказуйте джерело інформації.

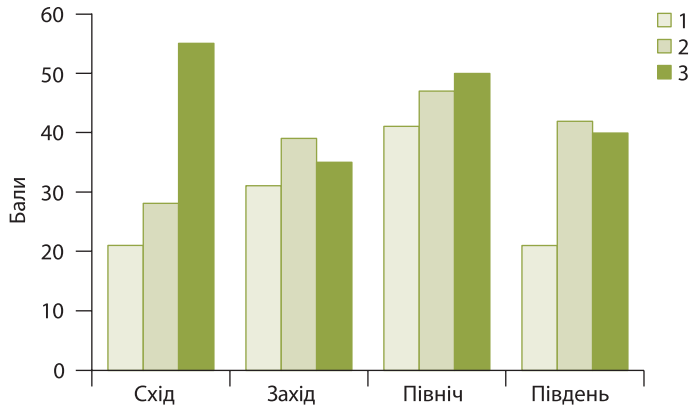
На рис. 13.10 показано приклад графіка, який загромаджено «графічним сміттям». Відтінки та координатна сітка ніяк не посилюють повідомлення, яке передає графік. Підписані дані вгорі кожної колонки відволікають читача від того, щоб побачити закономірність зміни. На рис. 13.11 наведено той самий графік без «графічного сміття».

Рис. 13.10. Приклад рисунка з надмірною кількістю елементів



Джерело: Автори.

Рис. 13.11. Приклад оптимізованого графіка, на якому більшість площі займає «корисне чорнило».



Джерело: Автори.

Як провести усну презентацію

Багато людей страшенно бояться говорити на публіці. Страх перед публічним виступом можна полегшити, якщо добре підготуватися та заздалегідь потренуватися.

Плануючи презентацію, поміркуйте над наступними питаннями:

- Хто є аудиторією? Чого вони чекають? Наскільки деталізовану презентацію вони хочуть бачити?
- Який ключовий зміст моєї презентації? Які три найважливіші повідомлення, що їх має запам'ятати моя аудиторія? Що я хочу, щоб вона зробила з інформацією, яку я представлю?
- Чи існують мовні або технічні труднощі спілкування з цією аудиторією?
- Яким чином я можу заздалегідь дізнатися, як аудиторія відреагує на мою презентацію?
- Скільки часу повинна займати моя презентація?
- Які аудіовізуальні ресурси я зможу використати (слайди, проектор)?

Готуючи презентацію, пам'ятайте про аудиторію, концентруйтеся на основних повідомленнях та дотримуйтеся перевіреного часом правила: «Скажіть їм, що ви збираєтеся сказати, скажіть їм це, а потім повторіть, що ви їм щойно сказали».

Практика є одним із найкращих способів вдосконалення якості презентації. Проведіть репетицію презентації спочатку самостійно, потім перед іншими особами. Після репетиції попросіть надати відгуки та виправте презентацію відповідним чином. Перевірте, щоб презентація заповнювала виділений час та не затягувалася довше, ніж це необхідно.

Під час презентації говоріть з аудиторією, а не своїми записами. Якщо ви використовуєте екран і проектор, роздрукуйте всі слайди і покладіть перед собою, щоб не повертатися спиною до аудиторії для читання з екрана.

Використання візуальних елементів

Візуальні елементи можуть «оживити» презентацію. Вони мають бути у вигляді слайдів, проекторів, постерів та роздаткових матеріалів. (Якщо ви збираєтеся використовувати слайди чи проектор, придумайте запасний план в разі вимкнення електропостачання чи відмови обладнання.)

Письмова інформація посилює усну. Через це, особливо якщо презентація робиться для людей, для яких мова презентації не є рідною, корисно перед початком роздати трохи роздаткових матеріалів (блок 13.3). Це також дозволить слухачам робити нотатки на їхніх паперових примірниках. Без роздатків вони намагатимуться у шаленому темпі конспектувати презентацію! Недоліком при роздаванні матеріалів є те, що люди переглядають слайди наперед, а не концентруються на тому, що говорить доповідач.

Якщо з цієї причини роздаткові матеріали видати в кінці презентації, важливо повідомити аудиторію, що вони все ж таки будуть. Якщо презентація включає складні дані і таблиці, роздавайте їх в ході презентації теми.

Блок 13.3. «Поради щодо підготовки презентацій у PowerPoint та роздаткових матеріалів»

Можна вдосконалити роздатки, скориставшись наступними порадами:

- Мінімізувати кількість слів. Обмежити текст не більш як вісьмома рядками на кожен слайд чи кадр проектора.
- Додавайте більше порожнього місця.
- Використовуйте чіткі графічні елементи.
- Для забезпечення читабельності друкуйте не більш як два слайди на кожному аркуші роздатків.

Використання ПЗ для презентацій

Програми для презентацій, такі як PowerPoint, можуть надати слайдам більш професійного вигляду. Однак ними треба користуватися розсудливо, щоб не перетворити бесіду на звичайне слайд-шоу. Щоб підтримувати інтерес до презентації, доповідач повинен обмежувати кількість слайдів, відображаючи на них лише найбільш важливі моменти (блок 13.4).

Надлишкове використання слайдів може зробити презентацію надзвичайно важкою. Це також створює ризик того, що вона перетвориться на рекламування того, про що йдеться, або не дає аудиторії можливості висловитися (Tufte 2003). Як зазначає Е. Тафт:

> Аудиторія постійно має витримувати стомлюючий потік інформації, один клятий слайд за іншим. Якщо багато інформації зосереджено у короткому відрізку часу, їй важко сприймати та розуміти логічні взаємозв'язки. Візуальне представлення ефективніше, якщо інформація у слайдах подається послідовно. Часто ясність та зрозумілість тим більша, чим чіткіші деталі. Це найбільш релевантно для статистичних даних, головною метою яких є порівняння (Tufte 2003).

Блок 13.4. «Поради щодо використання слайдів»

Наступні поради можуть підвищити ефективність слайдів у презентації:

- Починайте зі слайду, який притягує увагу аудиторії. Дайте аудиторії приблизно 10 секунд, щоб прочитати слайд; потім приблизно п'ять хвилин треба витратити на пояснення того, чому важливо те, про що йдеться у слайді.
- Використовуйте слайди, які ілюструють важливі питання, твердження, які важко сприйняти на слух, та важливі висновки.
- Спочатку покажіть слайд, а потім розкажіть аудиторії, про що саме йдеться мова.

Джерело: Based on Taylor 2007.

Резюме

Повідомлення результатів оцінювання є виключно важливим для самого оцінювання: якщо результати не будуть оприлюднені, оцінка як така має небагато сенсу. Незалежно від того, представлена презентація в письмовому чи усному вигляді, вона має бути пристосованою до потреб аудиторії.

Як письмові, так і усні звіти може збагатити використання графіки. Дотримуючись принципів, викладених у цій главі, спеціалісти з оцінювання зможуть посилити вплив, який матиме їхня оцінка.

Глава 13 Вправи

Практична вправа 13.1: Вдосконалення звіту з оцінювання

Розробіть список критеріїв того, наскільки доступно у звіті з оцінювання викладена його методологія, виявлені факти, висновки та рекомендації для цільової аудиторії. Спираючись на ці фактори, оцініть звіт з тих питань, над якими ви працюєте. Поставте йому бали залежно від оцінки за кожним критерієм (А – чудовий, В – дуже добре, С – достатньо та ПВ – потребує вдосконалення). Визначте, які вдосконалення треба зробити, щоб звіт краще передавав вкладені у нього повідомлення.

Практична вправа 13.2: Адаптація звіту до аудиторії

Для звіту, який ви перевіряли у попередній вправі, визначте різні типи цільової аудиторії, яка може зацікавитися виявленими фактами чи методологією. Поміркуйте, які аспекти можуть становити найбільший інтерес для певної аудиторії, який рівень деталізації потрібен і яким чином можна якнайкраще передати основні повідомлення звіту для кожної групи. На основі вашого аналізу створіть контрольну відомість, що показуватиме, яка комунікація, коли та від кого потрібна для кожного типу аудиторії.

Посилання та рекомендована література

Busuladzic, Dino, and James Trevelyan. 1999. *Demining Research: An Ergonomic Aspect of Humanitarian Demining.* University of Western Australia, Perth. http://www.mech.uwa.edu.au/jpt/demining/tech/dino/ergonomics.html.

Cummings, Harry. 2003. "Using Graphics in Development Evaluations." International Program for Development Evaluation Training (IPDET) presentation, Carleton University, Ottawa, June 16–July 11.

Druker, Phil 2006. *Advanced Technical Writing.* University of Idaho, Moscow, Idaho. http://www.class.uidaho.edu/adv_tech_wrt/week14/conclusion_recommendation_final_report.htm.

Ehrenberg, A. S. C. 1977. "Rudiments of Numeracy." *Journal of the Royal Statistical Society* 140: 277–97.

Independent Evaluation Group of the International Finance Corporation. 2008. *Enhancing Monitoring and Evaluation for Better Results: Biennial Report on Operations Evaluation in IFC 2008.* Washington D.C.: World Bank.

http: www.ifc.org/.../BROE2008_FullReport/$FILE/BROE+for+disclosure-final+V8-final.pdf.

Kaplan, Craig S. 2003. "Presentation Tips." Craig Web Experience. http://www.cgl.uwaterloo.ca/~csk/presentations.html.

Lawrenz, Frances, Arlen Gullickson, and Stacie Toal. 2007. "Dissemination: Handmaiden to Evaluation Use." *American Journal of Evaluation* 28 (3): 275–89.

Lentz, Michelle 2007. "Tufte and Visual Presentation." http://www.writetech.net/2007/02/presentation_st.html.

Lester, P. M. 2000. *Visual Communication: Images with Messages.* 2nd ed. Kentucky: Wadsworth Cengage Learning.

Levin, J. R., G. J. Anglin, and R. N. Carney. 1987. "On Empirically Validating Functions of Pictures in Prose." In *the Psychology of Illustration,* vol. 1, eds. D. A. Willows and H. A. Houghton, 51–85. London: Springer-Verlag.

MIGA. 2007. *Annual Report 2007–08.*
http://www.miga.org/documents/IEG_MIGA_2007_Annual_Report.pdf.

Scriven, Michael. 2007. "Key Evaluation Checklist." http://www.wmich.edu/evalctr/checklists/kec_feb07.pdf.

Taylor, Dave. 2007. "Use PowerPoint to Enhance Your Presentation, Not Cripple It." http://www.intuitive.com/blog/use_powerpoint_to_enhance_your_ presentation_not_cripple_it.html.

Torres, R., H. S. Preskill, and M. E. Piontek. 1997. *Evaluation Strategies for Communicating and Reporting.* Thousand Oaks, CA: Sage Publications.

Tufte, Edward R. 1983. The *Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press.
———. 1990. *Envisioning Information.* Cheshire, CT: Graphics Press.
———. 1997. *Visual Explanations: Images and Quantities, Evidence and Narrative.*
Cheshire, CT: Graphics Press.
———. 2003. "PowerPoint Is Evil: PowerPoint Corrupts. PowerPoint Corrupts Abso- lutely." *Wired* 11.09, September. http://www.wired.com/wired/archive/11.09/ppt2.html.
———. 2006. *Beautiful Evidence.* Cheshire, CT: Graphics Press.
———. 2008. The *Work of Edward Tufte.* http://www.edwardtufte.com/tufte/. Wallgren, Anders, Britt Wallgren, Rolf Persson, Ulf Jorner, and Jan-Aage Haaland.
1996. *Graphing Statistics and Data.* Thousand Oaks, CA: Sage Publications.
World Bank Operations Evaluation Department. 2004. "The Poverty Reduction Strategy Initiative: An Independent Evaluation of the World Bank's Support through 2003." World Bank, Washington, DC.

Веб-сайти

Statistics Canada. *Using Graphs.* http://www.statcan.ca/english/edu/power/ch9/ using/using.htm.
Torok, George. *Presentation Skills Success.* http://www.presentationskills.ca/.

ГЛАВА 14

Керівництво для фахівця з оцінювання: етика, політика, стандарти та керівні принципи оцінювання

Перед фахівцями з оцінювання стоїть багато завдань: планування, організація і вибір плану оцінювання, а також збір, аналіз та представлення даних. Їм також доводиться мати справу з внутрішніми та зовнішніми тисками. Крім того, їх можуть просити внести зміни до плану, організації або звітності оцінки відповідно до потреб інших сторін. Іноді запропоновані зміни не викликають проблем, а іноді можуть викликати заперечення етичного чи політичного характеру.

Етика й політика – проблеми для всіх фахівців з оцінювання, особливо для тих, хто працює в країнах з неналежним врядуванням і високим рівнем корумпованості. Внутрішні тиски всередині організації розвитку також можуть спричинити етичні проблеми. В цій главі обговорюються етичні проблеми й політичні питання в оцінці.

Ця глава складається з трьох основних частин:

- Етична поведінка
- Політика й оцінювання
- Стандарти й керівні принципи оцінювання

Етична поведінка

Етика – це набір цінностей та переконань, що диктують вибір в різних ситуаціях. Етика розглядає глобальні питання, наприклад: «Як ми маємо жити?» Замисліться над цим питанням. Хто визначає «як» (метод), «ми» (група, що шукає згоди щодо цього питання), «маємо» (прагнення) та «жити» (існувати як істоти з тілами) (World IQ 2008)? Етика залежить від людей, які приймають рішення. В її основі – їхні погляди щодо того, що правильно, що неправильно, що добре, а що погано.

Закони пишуться для того, щоб спробувати виявити й контролювати етичну поведінку, але закони не здатні передбачити кожну можливу ситуацію. Більше того, поведінка може бути законною, але, тим не менш, неетичною (наприклад, прийняти невеличкий подарунок від людини, що має незабаром проходити оцінку, або змінити тон звіту, зробивши його більш позитивним чи негативним, незважаючи на аргументи, що свідчать проти цього).

Фахівці з оцінювання часто стикаються з ситуаціями, коли правильний вчинок не є очевидним; існує багато «сірих» зон. Але завжди очікується, що вони мають проводити оцінки етично. Через те, що судження щодо етики виносяться на основі цінностей і переконань, різні культури встановлюють різні закони. Тому важливо знати, що законно і що етично в країні, де проводиться оцінювання.

Фахівці з оцінювання також повинні знати про всі закони, що вимагають повідомляти про злочин чи підозру на злочин. Наприклад, у Сполучених Штатах Америки сексуальне домагання вважається злочином. Воно загалом визначається як небажані знаки уваги, схиляння до сексу, а також інші словесні чи фізичні дії сексуального характеру, якщо згода чи відмова явно або неявно впливає на працевлаштування особи чи її роботу та створює залякуючу, ворожу або образливу атмосферу на роботі (U.S. Equal Employment Opportunity Commission 2002.) Неповідомлення про підозру щодо сексуальних домагань карається законом.

Фахівцям з оцінювання слід знати закони країни, в якій вони працюють, а також керівні вказівки й стандарти своєї власної організації. Чимало організацій розробили етичні стандарти та керівні вказівки.

Корумпованість і омани під час оцінки

Фахівці з оцінювання мають бути здатними говорити правду можновладцям. Обов'язковою рисою фахівця з оцінювання є незалежний, допитливий, аналітичний розум. Фітцпатрік, Сандерс і Уортен (Fitzpatrick, Sanders, and Worthen 2004) представляють п'ять форм корумпованого оцінювання:

- готовність викривити правду й пред'явити позитивні висновки в результаті конфлікту інтересів або можливості отримання вигід чи понесення втрат (така готовність може бути свідомою або несвідомою);
- нав'язування необґрунтованих поглядів в результаті застосування недбалих, непостійних та непрофесійних практик оцінки;
- «згущення фарб» у звіті про оцінювання через упередженість або необ'єктивність фахівця з оцінювання;
- спонукання клієнтів або учасників до співпраці шляхом обіцянок, які не можуть бути виконані;
- невиконання зобов'язань, які можна було виконати.

Хаус (House 1995) розглядає корумпованість дещо з іншої точки зору. Він зазначає, що фахівці з оцінювання можуть просто неправильно розуміти свої обов'язки. Він визначає п'ять «оман оцінки»:

- *Клієнтизм:* омана стосовно того, що будь-які запити клієнта та все, що піде на користь клієнту – це етично;
- *Контрактуалізм:* оманлива думка про те, що фахівець з оцінювання повинен виконувати письмовий договір без запитань, навіть якщо це йде врозріз із суспільним благом;
- *Методологізм:* переконання, що якщо фахівець з оцінювання дотримується припустимих методів дослідження, його поведінка є етичною, навіть якщо застосування деяких методик ставить перед фахівцем з оцінювання етичні дилеми;
- *Релятивізм:* оманлива думка, що всі дані про погляди респондентів, які фахівець з оцінювання отримує від різних учасників, повинні мати однакову вагу, наче немає можливості призначити думкам членів периферійних груп менший пріоритет, ніж думкам членів основних груп;
- *Плюралізм/елітизм:* оманливе рішення надавати голосу впливових груп вищий пріоритет через те, що фахівець з оцінювання вважає, начебто вони більш престижні й активні, ніж безправні або безвладні групи.

Визначення етичних проблем

Морріс і Кон (Morris and Cohn 1993) опитали членів Американської асоціації оцінювання про їхні погляди на етичні проблеми. Наступний перелік етичних проблем сформований за результатами цього опитування:

- Клієнт вирішує, якими мають бути висновки, ще до початку оцінювання або планує використати висновки в сумнівний з точки зору етики спосіб.
- Клієнт забороняє розглядати якісь питання в ході оцінювання, незважаючи на їх актуальність.

- Клієнт навмисно змінює висновки перед їх публікацією.
- Клієнт тисне на фахівця з оцінювання, вимагаючи змінити представлення висновків.
- Клієнт приховує або ігнорує висновки.
- Клієнт тисне на фахівця з оцінювання, вимагаючи порушити конфіденційність.
- Клієнт допускає нецільове використання результатів.
- Справжні зацікавлені особи не підключені до процесу планування.
- Фахівець з оцінювання виявляє незаконну, неетичну або небезпечну поведінку.
- Фахівець з оцінювання з якихось причин не бажає представляти висновки в повному обсязі.
- Фахівець з оцінювання не певен щодо своєї здатності об'єктивно чи справедливо представити висновки.
- Хоча ні клієнт, ні зацікавлені сторони не вимагають порушення конфіденційності, фахівець з оцінювання розуміє, що розкриття деяких висновків може її порушити.
- Результати оцінювання використовуються як докази проти когось.

На додачу до цих етичних проблем існують інші, пов'язані з планом оцінювання. Етичні проблеми можуть виникати в рандомізованих планах, коли вирішується, хто має бути охоплений дією інтервенції, а хто має опинитися в контрольній групі. Якщо в ході реалізації проекту інтервенція видається перспективною, члени контрольної групи можуть вимагати доступу до інтервенції. Надання такого доступу призведе до зміни плану оцінювання та його результатів.

Для того, щоб оцінювання було корисним для керівників, організацій розвитку, учасників і громадян, воно має бути чесним, об'єктивним та неупередженим. Обов'язок фахівця з оцінювання полягає в тому, щоб збір, аналіз і представлення даних були виконані чесно й неупереджено.

Вплив на початку оцінювання може бути ледве помітний. Іноді на фахівця з оцінювання здійснюють тиск з метою не допустити розгляду певних питань оцінки або підштовхнути до розгляду менш делікатних тем. Можуть не підніматися певні питання, що виставляють в невигідному світлі організацію або програму. Клієнти можуть говорити: «І нам, і вам відомо, що в нас є проблема, і ми вже її вирішуємо, але не варто про неї розповідати й ставити під загрозу підтримку програми». Дослідники, учасники програми та громадяни можуть стикатися зі спротивом через те, що на поверхню випливають делікатні (негативні) теми. Бувають також випадки, коли певних людей не допускають на зібрання чи інтерв'ю, або ж обмежують виїзну роботу через «брак часу». Фахівець з оцінювання має прагнути піднімати питання, яких всі уникають, не підтримувати жодну із сторін, вислуховувати й брати до уваги всі точки зору.

Також часто трапляються спроби вплинути на оцінювання під час його проведення. І хоча більшість фахівців з оцінювання швидко розпізнають спробу дати хабар, вони не завжди можуть розгледіти тонші форми впливу. Пропозиції дружби, вечеря або спільний відпочинок можуть бути приємним жестом для людини, що опинилася далеко від дому. Це також може бути спробою вплинути на точку зору фахівця з оцінювання і врешті на звіт.

Іноді хтось надає фахівцю з оцінювання відомості про корупцію або шахрайство. Він не повинен намагатись дізнатися, чи має він справу зі спробою відвернути його увагу від інших питань, чи інформатор намагається таким чином з кимось поквитатися, чи справді мова йде про корупцію або шахрайство. Фахівці з оцінювання – не слідчі. Тому вони не повинні проводити розслідування, щоб встановити істину. Вони мають повідомляти про звинувачення у відповідні слідчі органи.

У більшості організацій розвитку існують процедури звітування про підозри у посадовому злочині або шахрайстві. Приміром, Відділ внутрішніх розслідувань Світового банку розглядає інформацію про посадові злочини персоналу, а також звинувачення щодо шахрайства й корупції в операціях банку. В подібних департаментах часто є «гарячі лінії» для прийому відомостей про посадові злочини, шахрайства та зловживань.

Члени групи оцінювання повинні бути обізнані з політиками та стандартами своїх організацій стосовно дій у випадку отримання відомостей про можливе шахрайство, зловживання чи посадовий злочин. Ці процедури можуть бути викладені в інструкції для персоналу, спеціальній брошурі або трудовому договорі. Фахівцям з оцінювання варто пам'ятати, що вони зобов'язані дотримуватися політик і стандартів своєї організації.

Девіз «не нашкодь», безумовно, може застосовуватися до оцінювання. За винятком тих випадків, коли існує підозра щодо вчинення шахрайства або зловживання, про що поінформовано слідчі органи, оцінка не повинна завдавати шкоди учасникам. Це означає недопустимість ідентифікування осіб, які беруть участь в оцінці, або створення загрозливих для них ситуацій.

Якщо ви гарантували учасникам конфіденційність, а потім розкрили їх імена, то завдали їм шкоди. Захист конфіденційності має першочергове значення, але трапляються випадки, коли забезпечити його складно. Уявімо, що фахівці з оцінювання освітньої програми під час кількох інтерв'ю почули від співрозмовників, що директор витрачає гроші програми для власного зиску. Що мають робити фахівці з оцінювання? Розкривши ці відомості, вони можуть «підставити» тих, хто повідомив їм конфіденційну інформацію. Відмова від звернення до слідчих органів потенційно означатиме співучасть у злочині. Ця ситуація ілюструє плюси анонімних гарячих ліній. Також може бути корисно поспілкуватися з супервайзером або керівником (якого не обвинувачують) стосовно ситуації й можливого виходу з неї.

Фахівцям з оцінювання варто бути особливо обережними, відповідаючи на питання репортерів, керівників програми або навіть членів керівної ради «не для протоколу». Потрібно щонайменше з'ясувати, що саме означає для цієї особи вираз «не для протоколу». Фахівцю з оцінювання варто щось говорити лише тоді, коли його влаштовує тлумачення цього терміну.

Політика й оцінювання

Оцінювання завжди здійснюється в політичному контексті; вже сам той факт, що оцінка проводиться, може бути використаний для просування персональних або інституційних інтересів. Часто кажуть, що знання – сила. Оцінювання надає знання, що можуть використовуватися як кнут чи пряник[9] для впливу на поведінку людей. Це діяльність, що нерідко приносить успіх або розчарування.

Політика може порушити об'єктивність оцінювання. Вона здатна визначати, як і якою мірою використовується оцінювання. Позитивні оцінки дозволяють отримати більше коштів, розгорнути пілотні проекти в повноцінні програми, покращити репутацію й збільшити вплив, а також сприяти кар'єрі осіб, які виконують інтервенцію. Оцінювання, що викривають серйозні проблеми, можуть покращити інтервенцію й майбутні результати, але вони створюють ризик скорочення фінансування або скасування програм та втрати іміджу, влади і впливу людьми, які здійснюють втручання.

Причини втручання політики в оцінювання

Оцінювання – форма організаційного знання. Через це процесу оцінювання властива складність визначення реального стану справ.

Маррей (Murray 2002) наполягає, що політика неодмінно відіграватиме свою роль в оцінюванні через те, що сама оцінка залишає багато можливостей для проявів суб'єктивності. Суб'єктивність призводить до розбіжностей між фахівцями з оцінювання. Фахівці з оцінювання збирають враження про реальний стан речей у зацікавлених осіб та осіб, які є об'єктами оцінювання. Враження можуть бути різними, що нерідко викликає незгоду на різних етапах оцінювання й спонукає людей до політичної поведінки.

Маррей (Murray (2002, p. 2)) пояснює причину протиріч «невід'ємними проблемами технічних елементів методів оцінювання та поширеними людськими слабкостями». Він також описує технічні й людські слабкості в оцінюванні, що можуть мати політичний ефект. Кожна з наведених дослідником слабкостей описана нижче.

---

[9] *В оригіналі «carrot or stick approach».*

Технічні недоліки

Більшість оцінювань демонструють найвищу ефективність, коли мова йде про порівняння із заданими цілями, завданнями й стандартами. Але фахівцям з оцінювання, клієнтам та іншим зацікавленим особам може бути важко дійти згоди щодо того, що потрібно вимірювати. Крім того, може бути важко визначити фокус оцінювання. Належне оцінювання передбачає визначення теорії змін та основних припущень програми. Якщо теорія змін розроблялася спільно із зацікавленими особами програми, вона здатна забезпечити спільне розуміння цілей, завдань і заходів програми, а також результатів, яких програма має досягти. Таким чином, теорія змін здатна запобігти потенційним конфліктам стосовно розуміння програми до того, як вони перетворяться на політичні проблеми.

Маррей (Murray 2002) називає і другу найпоширенішу технічну проблему, що призводить до проблем політичних: оцінювання одного рівня організацій та перенесення висновків на інший рівень. Це спричиняє проблеми, коли зв'язки між ефективністю роботи окремих осіб, програм або функціональних підрозділів і організації в цілому визначені нечітко. Знову-таки, теорія змін – спосіб, за допомогою якого можна виявити основні припущення.

Людські недоліки

Люди часто роблять те, що відповідає їхнім власним інтересам. Вони можуть мати несвідомі упередження, що впливають на те, як вони проводять оцінку та аналізують її результати.

Кат і Маррей (Cutt and Murray 2000) визначають три людських фактори, які можуть позначитися на оцінці:

- ментальність «зберігай обличчя – уникай вини» (ЗОУВ);
- явище «суб'єктивної інтерпретації реальності» (СІР);
- фактори довіри.

Ментальність ЗОУВ визначає спільну людську рису: люди прагнуть успіху та не хочуть бути пов'язаними з невдачею. Більшість оцінювань спрямовані на виявлення як успіхів, так і невдач. Однак фахівці з оцінювання можуть не повідомляти про повний масштаб невдач, не бажаючи бути пов'язаними з невдачами. В ситуації ЗОУВ люди винахідливо доводять, що не мали впливу на негативні результати. Другий варіант: вони можуть оспорювати обсяг і підхід оцінки (це одна з причин, чому попередні угоди щодо плану оцінювання настільки важливі).

Явище СІР виникає під час інтерпретації даних, що використовуються для оцінювання. Дві особи, які спостерігають одну й ту саму подію, можуть описати її по-різному: «вчитель втрачає контроль за класом» та «вчитель,

який повністю заволодів увагою класу». У фахівців з оцінювання є власні переконання й думки стосовно ефективності того, що працює. Результати оцінки можуть інтерпретуватися суб'єктивно, на основі цих переконань і думок. Наприклад, фахівець з оцінювання, який вважає, що хороший учитель – це той, у котрого всі діти сидять на своїх місцях, розмовляючи тільки якщо їх просять, більш схильний бачити у подіях хаос та втрату контролю. Фахівець з оцінювання, власна дитина якого відвідує факультативні заняття, більше схильний розглядати ту саму подію як приклад вчителя, який вміє заволодіти увагою учнів.

Явище СІР – одна з причин того, чому ми в нашій книзі говоримо, що фахівці з оцінювання мають аналізувати літературу, щоб більше дізнатися про рішення, які працюють, й питання, виявлені під час оцінки аналогічних програм. Ще один спосіб зменшити упередженість суб'єктивної реальності – використовувати декілька методів збору даних: анкети, інтерв'ю і структуровані спостереження, що збільшує кількість джерел та є різновидом стратегії тріангуляції.

Ще один фактор, на який звернули увагу Кат і Маррей – фактор довіри. Він може запустити в дію фактори ЗОУВ і СІР. Довіра – це переконання у чесності або здібностях особи. Якщо люди вважають, що іншій особі бракує чесності або здібностей, вони можуть їй не довіряти. Вони можуть побоюватися, що ця особа завдасть їм шкоди. Довіра вимірюється за рівнем: від часткової довіри (лише в деяких контекстах або в певних питаннях) до повної (в усіх ситуаціях). Якщо виникає недовіра, ЗОУВ або явище СІР можуть викликати втручання політики у стосунки. В попередніх главах пропонувалися різноманітні стратегії побудови довіри – від залучення зацікавлених осіб до планування оцінювання і конкретних технік побудови довіри.

Виявлення політичних ігор

Неможливо повністю відокремити оцінювання від політики. Проблема полягає в управлінні політичною ситуацією, в яку потрапляють фахівці з оцінювання. Для допомоги фахівцям з оцінювання Маррей (Murray 2002) класифікує ігри відповідно до ролей людей, які беруть в них участь. Розуміння цих ігор допомагає фахівцям з оцінювання справлятися з ситуацією.

Політичні ігри, в які грають оцінювані люди

Люди, яких оцінюють, часто прагнуть уникнути небажаного офіційного спостереження за їх діями. Вони можуть реагувати:

- заперечуючи потребу в оцінюванні;
- заявляючи, що оцінка забирає забагато часу, а вони зараз працюють (можливо, з підвищеним навантаженням), і тому оцінювання поки що варто відкласти;

- заявляючи, що оцінювання – це добре, але відтягуючи його здійснення;
- прагнучи встановити тісні особисті стосунки з фахівцями з оцінювання і завоювати їхню довіру.

Після початку оцінювання й збору даних починаються інші політичні ігри. Постачальники інформації можуть:

- замовчувати або перекручувати інформацію, про яку питають фахівці з оцінювання, щоб не втратити обличчя;
- надавати фахівцям з оцінювання величезну кількість інформації, через що їм складно відсортувати релевантну інформацію від нерелевантної;
- видавати нові дані ближче до завершення оцінювання.

Після збору даних, коли фахівець з оцінювання готовий інтерпретувати їх, оцінювані люди можуть:

- заперечувати існування проблеми;
- принижувати важливість проблеми, заявляючи, що про неї вже було відомо і відповідні зміни вже вносяться, або ж перекладаючи відповідальність на інших осіб та непідвладні обставини;
- доводити, що інформація неактуальна, тому що обставини змінилися.

Політичні ігри інших зацікавлених осіб

Інші зацікавлені особи також можуть здійснювати вплив на політику оцінювання. В різних зацікавлених осіб є різні інтереси й прагнення.

Якщо зацікавлені особи не брали участі у виявленні найважливіших питань оцінювання, вони можуть вирішити, що розглядалося щось не те. Окрім того, вони можуть намагатися залучити інші сторони (наприклад, ЗМІ) до критики організації та диктування способу проведення оцінювання.

Політичні ігри, що в них грають фахівці з оцінювання

Фахівці з оцінювання також можуть грати в ігри. Під час збору даних деякі фахівці можуть відхилятися від процесу, збираючи власну інформацію «не для протоколу». Ця неофіційна інформація потім інтерпретується в рамках оцінювання.

Більшість ігор, в які грають фахівці з оцінювання, відбувається протягом етапу інтерпретації. Це, зокрема:

- зміна або невстановлення стандартів вимірювань;
- застосування незаявлених критеріїв прийняття рішень;
- застосування незаявлених цінностей та ідеологічних фільтрів для інтерпретації даних, наприклад рішення не довіряти одному з джерел інформації;
- ігнорування даних, що суперечать висновку фахівця з оцінювання.

Робота з політичною складовою під час оцінювання

Оскільки присутність політики неминуча, важливо навчитися з нею працювати. Протягом усього процесу фахівець з оцінювання має будувати довірчі стосунки. Найкраще буде на кожному етапі оцінювання проводити відкриті обговорення, що дадуть всім гравцям можливість обговорити свої справи та принаймні погодитись або відмовитись погоджуватися щодо наявності розбіжностей. Маррей (Murray 2002) говорить про важливість того, щоб усі пов'язані з оцінкою сторони повністю розуміли логіку подій. Він вважає, що теорія змін (або логічна модель) – один зі способів сформулювати логіку таким чином, щоб не залишилося місця для непорозумінь.

Як іще можна будувати довірчі стосунки? Зазвичай для цього потрібний час і багато взаємодій. Маррей рекомендує залучати до процесу оцінювання усі зацікавлені сторони, зокрема тих осіб, що мають стати її об'єктом.

Врівноваження інтересів зацікавлених сторін за допомогою переговорів

Одна з найсерйозніших проблем для фахівців з оцінювання – робота з кількома зацікавленими сторонами. Фахівці з оцінювання повинні добре вміти вести переговори, щоб враховувати інтереси зацікавлених сторін та їхні цілі, що нерідко є суперечливими. Маркевіц (2005) наводить перелік принципів і практик для ведення переговорів в інтересах оцінювання.

Принципи ведення переговорів в інтересах оцінювання

Запропоновані нижче принципи взято зі списку принципів ведення переговорів в інтересах оцінювання Маркевіц (Markiewicz 2005):

- Визнайте невід'ємну політичну природу оцінки.
- Цінуйте внесок зацікавлених осіб.
- Оцініть позиції зацікавлених осіб і виконайте планування оцінювання.
- Зробіть так, щоб фахівець з оцінювання був активним гравцем у спільноті зацікавлених осіб.
- Розвивайте у фахівця з оцінювання навички посередника, що вміє реагувати на конфлікт.
- Вчіться розв'язувати конфлікти зацікавлених осіб.

Одна з ключових стратегій полягає у тому, щоб об'єднати зацікавлених осіб у референтні групи, керівні або дорадчі комітети для нагляду за процесом оцінювання. Цей підхід відповідає моделі «фахівець з оцінювання як координатор» (вона зазвичай не може бути застосована для незалежного оці-

нювання, основною метою якого є оцінка відповідальності). Важливо чітко визначити ролі й функції цих груп. Групи або комітети повинні встановити базові правила, що визначатимуть ступінь активності їхніх членів у процесі оцінювання.

За словами Маркевіц (Markiewicz 2005), після того, як фахівець з оцінювання завоює довіру зацікавлених осіб, йому доведеться розв'язувати їхні конфлікти та суперечки. Фахівець з оцінювання має діяти як каталізатор, що допомагає зацікавленим особам знаходити власні рішення. Для цього йому знадобиться вміння спілкуватися, тобто активно вислуховувати співрозмовників, ставити доречні запитання, переконуватись у правильному їх розумінні. Фахівець з оцінювання також має підтримувати цілеспрямованість переговорного процесу, а також сприяти взаємодії зацікавлених осіб.

Фахівцям з оцінювання потрібно розвивати навички ведення переговорів. У деяких випадках фахівцям з оцінювання варто проходити додаткову теоретичну й практичну підготовку щодо ведення переговорів (для моделі «фахівець з оцінювання як координатор»). Інший спосіб отримати необхідні навички – працювати з колегами й обмінюватися досвідом (як успішним, так і негативним) розв'язання конфліктів.

Петон (1997) рекомендує проводити щонайменше чотири зустрічі (або більше для довгострокових проектів):

- перші збори: фокус оцінювання;
- другі збори: методи й інструменти вимірювання;
- треті збори: розробка інструментів перед збором даних;
- четверті збори: аналіз даних з метою встановлення консенсусу щодо інтерпретацій, які приведуть до висновків.

Ці кроки аналогічні запропонованим у нашій книзі.

Маркевіц говорить про активну роль, що її фахівець з оцінювання має відігравати для зацікавлених осіб. Вона називає дві важливі риси: чуйність та гнучкість, необхідні для залучення зацікавлених осіб до процесу. Вона також згадує про складнощі, що виникають, якщо фахівець з оцінювання занадто зближується із зацікавленими особами та занадто тісно з ними спілкується.

Петон пропонує бути завжди зосередженим на емпіричному процесі та допомагати зацікавленим сторонам дотримуватись такого ж підходу. Завдяки цьому стосунки залишаться об'єктивними, а упередженості висновків та маніпулюванню ними вдасться запобігти.

Практика ведення переговорів в інтересах оцінювання

Маркевіц (Markiewicz 2005) визначає три етапи переговорів щодо оцінювання:

- Початковий етап: сторони викладають свої позиції.
- Середній етап: ведуться активні переговори.
- Останній етап: сторони роблять кроки для досягнення консенсусу.

Для використання цієї моделі фахівцю з оцінювання потрібно вміти бути співчуваючим і наполегливим. Співчуття створює атмосферу, що позитивно впливає на переговорний процес. Наполегливість додає цьому процесу структури.

Співчуття – «процес демонстрації точного, неосуджуючого розуміння потреб, інтересів і позицій іншої сторони» (Mnookin, Peppet, and Tulumello 2000, p. 46). Воно складається з двох компонентів: здатності бачити світ очима іншої людини та вміння висловити цю точку зору словами. Співчуття означає перетворення розуміння позиції іншої сторони в спільну реакцію.

Маркевіц (Markiewicz 2005) вважає співчуття важливим для отримання інформації про цілі, цінності та пріоритети інших людей. Ця якість фахівця стає каталізатором відкритості для інших та інструментом переконання на переговорах.

Добре зрозумівши погляди кожної зацікавленої особи, фахівець з оцінювання має трансформувати це розуміння у рішення, прийнятне для всіх сторін (Hale 1998). Потім він повинен поцікавитись, чи правильно він їх розуміє, та попросити з'ясувати розбіжності, якщо вони є. Здатність активно слухати співрозмовника допомагає фахівцю з оцінювання зрозуміти, про що йде мова, ставити доречні питання й перевіряти своє розуміння слів зацікавлених осіб.

Впевненість – це здатність висловлювати й захищати власні потреби, інтереси та позиції (Mnookin, Peppet, and Tulumello 1996). В контексті ведення переговорів під час проведення оцінювання це можна оцінювати як впливовість координатора. Підтримувати рівновагу між співчуттям та впевненою поведінкою може бути нелегко. Мнукін, Пепет та Тулумелло (Mnookin, Peppet, and Tulumello 1996) розглядають співчуття та впевненість як два взаємозалежні виміри поведінки на переговорах. Використані разом, ці дві якості здатні забезпечити краще розуміння потреб зацікавлених осіб.

Стандарти й керівні принципи оцінювання

Фахові об’єднання розробляють стандарти й керівні вказівки, покликані допомагати їхнім членам приймати етично правильні рішення. Фахові об’єднання Північної Америки, Європи та Австралії запровадили етичні кодекси для фахівців з оцінювання.

Американська асоціація оцінювання (AAO) розробила набір стандартів і принципів та виклала їх у двох документах: «Стандарти оцінювання про-

грам» та «Керівні принципи для фахівців з оцінювання». Ці стандарти й принципи нерідко слугували базою для інших груп, що адаптували їх до місцевих умов.

ААО розробила «Керівні принципи для фахівців з оцінювання» для того, щоб надати фахівцям з оцінювання орієнтир у повсякденній роботі. Основна різниця між стандартами й принципами ААО полягає у їхній меті. Стандарти стосуються професійної діяльності, в той час як керівні принципи зосереджуються на професійних цінностях. В центрі уваги стандартів – продукт оцінювання, в центрі уваги керівних принципів – поведінка фахівця з оцінювання. Обидва документи визначають правильні та етичні шляхи проведення оцінювання.

Стандарти оцінювання програм

Стандарти оцінювання програм ААО охоплюють чотири категорії:

- практичність;
- доцільність;
- правильна поведінка;
- точність.

До категорії правильної поведінки входять вісім стандартів:

- *Орієнтація на надання послуг:* Мова йде про те, що фахівці з оцінювання в своїй роботі повинні орієнтуватися не лише на інтереси спонсора оцінки, а й на потреби в інформації учасників програми, громади й суспільства.
- *Формальні угоди:* Охоплено такі питання, як виконання протоколу, отримання доступу до даних, чітке попередження клієнтів про обмежені можливості оцінювання, утримання від обіцянок, які не можна виконати.
- *Права людей-учасників оцінки:* Отримання інформованої згоди, забезпечення права на приватність, збереження конфіденційності.
- *Взаємодія з людьми:* Продовження теми прав людей-учасників оцінювання. Стандарт визначає: взаємодіючи з ними, фахівці з оцінювання мають поважати людську гідність, а також жоден учасник оцінювання не повинен зазнавати шкоди чи приниження.
- *Повна й чесна оцінка:* Стандарт вимагає точного відображення сильних і слабких сторін програми. Фахівець з оцінювання повинен переконатися, що він не спотворює висновки дослідження, щоб догодити спонсору або іншим групам.
- *Оприлюднення результатів:* Мова йде про обов’язок фахівця з оцінювання слугувати інтересам широкого загалу – людей, які отримують вигоду як від програми, так і від її точної оцінки – а не тільки клієнтів або спонсорів. Результати оцінювання мають бути оприлюднені.

- *Конфлікт інтересів:* Фахівці з оцінювання мають чесно й відкрито повідомляти про свої переконання і цінності, щоб клієнти могли робити поправку на потенційну упередженість, яка може проникнути в роботу навіть найчесніших фахівців з оцінювання.
- *Фінансова відповідальність:* Всі витрати (в тому числі витрати часу й зусиль для організації, координації і проведення збору інформації, необхідної фахівцям з оцінювання та потрібної для роз'яснення оцінки різним сторонам) повинні бути адекватними, розважливими та добре задокументованими.

ААО підкреслює, що її стандарти оцінювання програм пристосовані виключно до американських умов та можуть не підходити для застосування в інших країнах без доопрацювання. В 2000 році Фонд Келлога профінансував зустріч регіональних і національних фахівців з оцінювання на Барбадосі. Там було представлено кілька міжнародних організацій з оцінювання. Одним з питань порядку денного було: чи можуть стандарти оцінювання програм ААО застосовуватися в інших країнах та яким чином.

В результаті було опубліковано кілька документів, присвячених цьому питанню (Russon 2000). У першій публікації Таут (Taut 2000) говорить про те, що Стандарти оцінювання програм, розроблені ААО, базуються на певних цінностях, а такі цінності в різних культурах різні. Вона розглядає, яким чином різниця в цінностях впливає на придатність стандартів ААО до використання поза межами США.

Таут порівнює цінності, виражені в стандартах, з культурними цінностями, інформацію про які можна знайти в крос-культурній літературі. Найважливішими вимірами культурних цінностей вона вважає:

- протиставлення індивідуалізму та колективізму;
- протиставлення ієрархічності та егалітаризму (або дистанція влади);
- протиставлення консерватизму та автономії;
- протиставлення панування й гармонійного співіснування;
- уникнення невизначеності.

До інших відмінностей культур, які відзначає дослідниця, відносяться протиставлення прямого та непрямого спілкування, високого й низького контексту, а також важливість старшинства.

Таут робить висновок, що «уявлення про корисність та етику відрізняються в різних культурах сильніше, ніж уявлення про доцільність та точність… Стає очевидно, що аспекти, висвітлені в Стандарті правильної поведінки, значною мірою залежать від політичних та культурних впливів» (2000, p. 24). Вона рекомендує фахівцям з оцінювання описувати свої суспільства в контексті культурних вимірів, розглянутих в її статті. Також

вона рекомендує їм консультуватися з колегами щодо своїх вражень, а також з експертами з культур і проводити аналіз з урахуванням результатів цих консультацій.

Керівні принципи для фахівців з оцінювання

ААО активно просуває етичний підхід до оцінювання програм, персоналу й політики. Для цього Асоціацією було розроблено керівні принципи (1995), покликані допомагати фахівцям з оцінювання у їхній професійній діяльності. Керівні принципи охоплюють такі питання:

- систематичне дослідження;
- компетентність;
- неупередженість/чесність;
- повага до людей;
- відповідальність за загальний добробут.

Комітет з етики ААО в 2004 році організував масштабний перегляд і оновлення Принципів. (Відповідну публікацію й додаткову інформацію про Керівні принципи та Стандарти оцінки програм можна знайти на веб-сайті ААО. Посилання на обидва ці документи наведено в кінці даної глави в розділі «Організації оцінки».)

Ось інші організації з оцінювання, які розробили (або розробляють) керівні вказівки чи стандарти, що їх можна знайти в мережі Інтернет:

- Африканська асоціація з оцінювання (проект стандартів і керівних вказівок з оцінки);
- Австрало-азійська спілка з оцінювання (керівні вказівки з етики);
- Канадська спілка з оцінювання (Керівні вказівки з етичної поведінки);
- Німецька спілка з оцінювання (стандарти);
- Італійська асоціація з оцінювання (керівні вказівки);
- Швейцарська спілка з оцінювання (стандарти);
- Британська спілка з оцінювання (керівні вказівки з добросовісної практики).

На додачу до керівних вказівок для спеціалістів-практиків британські керівні вказівки визначили правила для замовників і учасників проведення оцінювання.

Етика оцінювання в системі ООН

Група оцінювання ООН сформулювала свою позицію з питань етики оцінювання в «Керівних вказівках з етики» (UNEG 2008). Ці керівні вказівки базуються на «розповсюджених та міжнародно визнаних професійних ідеалах» (UNEG 2008, p. 3). Вони стосуються проведення оцінювання членами персоналу, зовнішніми консультантами й фахівцями з оцінювання з партнерських організацій в усіх агенціях ООН.

Далі наводиться перелік етичних принципів з цих керівних вказівок:

- цільовий характер оцінювання, що передбачає наступне:
  - практичність
  - необхідність
- зобов'язання фахівців з оцінювання:
  - незалежність
  - неупередженість
  - достовірність
  - уникнення конфліктів інтересів
  - чесність та непідкупність
  - відповідальність
- зобов'язання перед учасниками:
  - повага до гідності й етнічної приналежності
  - права
  - конфіденційність
  - недопущення завдання шкоди
- процес і продукт оцінювання:
  - точність, повнота й надійність
  - прозорість
  - звітування
  - упущення й помилки (UNEG 2008).

Група оцінювання ООН також встановила Стандарти оцінювання в системі ООН.

Її стандарти з етики передбачають наступне:

- Фахівці з оцінювання повинні поважати вірування, традиції та звичаї, бути чесними й непідкупними у своїх стосунках з усіма зацікавленими особами.
- Фахівці з оцінювання повинні виявляти до людей таку ж повагу, яку хотіли б отримувати на свою адресу.
- Фахівці з оцінювання мають захищати анонімність і конфіденційність своїх інформаторів.
- Фахівці з оцінювання відповідають за якість і результати своєї роботи.

Стандарти КСР

У главі 1 цієї книги розповідалося про Мережу Комітету сприяння розвитку з оцінювання розвитку ОЕСР – головну групу з формування норм і стандартів оцінювання розвитку. Оновлений виклад основних норм і стандартів КСР можна знайти за адресою http://www.oecd.org/dac/evaluationnetwork.

Особливо актуальним для даної глави є проект стандартів якості оцінювання КСР, що містить вказівки стосовно процесу й продукту оцінювання. Ці стандарти будуть остаточно оформлені в 2009–2010 рр. після завершення трирічного етапу тестування. Розділ 6 керівних вказівок зосереджується на незалежності фахівців з оцінювання та на тому, наскільки вільно й без втручання може працювати команда фахівців з оцінювання. Розділ 7 присвячений етиці оцінювання. Цей нормативний проект вимагає, щоб процес оцінки, а отже, і сам фахівець з оцінювання «делікатно ставився до гендеру, переконань, традицій і звичаїв усіх зацікавлених осіб, та щоб оцінка проводилася чесно й неупереджено. Права й добробут учасників оцінювання мають бути захищені. Анонімність та конфіденційність інформаторів має захищатися тоді, коли про це просять та/або коли цього вимагає закон» (р. 22).

Конфлікт інтересів

Явний чи неявний конфлікт інтересів – серйозна проблема, яка може завдати шкоди довірі до фахівців з оцінювання та якості оцінювання. Під час кожного оцінювання фахівці повинні вказувати, чи немає в них конфлікту інтересів та чи не здається, що такий конфлікт інтересів може мати місце.

Деякі організації розробили керівні вказівки, що регулюють ситуації з конфліктом інтересів. Група оцінювання операцій Азійського банку розвитку, наприклад, забороняє своїм співробітникам або консультантам оцінювати роботи, з якими вони були пов'язані. Це хороша практика для підрозділів оцінювання.

Підсумки

Етика – це набір цінностей та переконань, що диктують вибір в різних ситуаціях. Фахівці з оцінювання можуть порушувати етичні стандарти, проводячи оцінювання незважаючи на конфлікт інтересів, виконуючи свою роботу недбало або непрофесійно, діючи упереджено, даючи обіцянки, які не можуть виконати, або ж не виконуючи свої зобов'язання. Непорозуміння, що виявляється в таких формах, як клієнтизм, контрактуалізм, методологізм, релятивізм або плюралізм/елітизм, також може спричинити неетичну поведінку.

Мережа Комітету сприяння розвитку з оцінювання розвитку ОЕСР, Група оцінювання ООН, ААО та інші агенції й об'єднання розробили стандарти, керівні вказівки й норми для визначення та вимірювання якості, етики й нормативної поведінки. Фахівці з оцінювання розвитку повинні бути обізнані з відповідними стандартами, керівними вказівками й нормами та застосовувати їх на практиці.

Завдання до Глави 14

Практична вправа 14.1: Етика: Роза й сільськогосподарське оцінювання

Роза Маркес розповідає вам наступну історію та просить поради. В чому полягають основні етичні проблеми, та як би ви порадили Розі вирішити їх?

Роза зустрілася з місцевими посадовцями, офіційними особами програми та землевласниками, щоб розповісти про заплановану оцінку місцевої сільськогосподарської програми. Громада роками отримувала значні кошти для будівництва систем зрошення, закупівлі добрив, прокладання доріг та придбання обладнання.

Роза вперше приїхала в цю місцевість. Член місцевої команди, Едуардо, вже бував тут кілька разів і був знайомий з багатьма землевласниками. Він запропонував їм усім зібратися на вечерю, щоб налагодити спілкування.

Під час вечері Роза слухала розмови між Едуардо й землевласниками. Виявилося, що землевласники приятелюють з Едуардо. Вони подарували йому коробку сигар. Вони обговорили потреби району. Землевласники вважали, що їм потрібно більше ресурсів, щоб ефективно використовувати землю. Вони прагнули закупити додаткове обладнання, щоб замінити в господарстві робітників. Також вони хотіли застосувати більше добрив, ніж було дозволено екологічним законодавством. Едуардо погодився з землевласниками й сказав їм, що майбутня оцінка може стати їм у нагоді, тому що може рекомендувати зробити для них виключення з правил.

Наприкінці вечері Розі запропонували показати довколишні місця й пообідати з сім'єю одного із землевласників. Роза подумала, що відмовлятися буде неввічливо, і домовилась зустрітися наступного дня. Вона коротко обговорила вечерю з Едуардо і спитала, чому він погодився із землевласниками. Едуардо сказав, що землевласники будуть активніше співпрацювати, якщо розраховуватимуть отримати від оцінки щось корисне для себе.

Протягом поїздки наступного дня землевласник розказував, як тяжко вони працюють та яких успіхів досягли, незважаючи на труднощі. Він сказав Розі, що розраховує на те, аби вона підтримала їхні зусилля. Якщо висновки оцінки будуть негативні, він і його сім'я не виживуть. На знак вдячності він подарував їй намисто, яке, за його словами, передавалося в його родині з покоління в покоління.

Після поїздки й обіду в сім'ї землевласника Роза зустрілася з керівником програми. Він уже спланував для неї розклад зустрічей на три наступні дні. Також він запланував два зібрання громади. На ці збори були запрошені землевласники, співробітники служби з поширення сільськогосподарських знань, представники ділових кіл, що продають сільськогосподарське обладнання й добрива, а також експортери сільськогосподарської продукції. Коли вона спитала, чому не було запрошено робітників, які працюють на фермах, та членів їхніх родин, керівник сказав, що вони не можуть повідомити нічого корисного для оцінки дієвості проекту. Тоді Роза спитала, чи варто їй поспілкуватися ще з кимось із представників громади. Керівник сказав, що доклав максимум зусиль, аби організувати їй зустрічі з усіма необхідними людьми, щоб їй було легко провести оцінку програми.

Посилання та рекомендована література

Cutt, James, and Vic Murray. 2000. *Accountability and Effectiveness Evaluation in Nonprofit Organizations.* London: Routledge.

Fitzpatrick, Jody L., James R. Sanders, and Blaine R. Worthen. 2004. *Program Evaluation: Alternative Approaches and Practical Guidelines.* New York: Pearson Education Inc.

Hale, K. 1998. "The Language of Co-Operation: Negotiation Frames." *Mediation Quarterly* 16 (2):147–62.

House, E. R. 1995. "Principles Evaluation: A Critique of the AEA Guiding Principles." In *Guiding Principles for Evaluators,* eds. R. Shadish, D. L. Newman, M.

A. Scheirer, and C. Wye, 27–34. New Directions for Program Evaluation No. 66. San Francisco: Jossey-Bass.

JCSEE (Joint Committee on Standards for Educational Evaluation). 1994. *The Program Evaluation Standards: How to Assess Evaluations of Educational.* Programs Thousand Oaks, CA: Sage Publications.

Markiewicz, Anne. 2005. "'A Balancing Act': Resolving Multiple Stakeholder Interests in Program Evaluation." *Evaluation Journal of Australasia* 4 (1–2): 13–21.

Molund, Stefan, and Göran Schill. 2004. *Looking Back, Moving Forward: SIDA Evaluation Manual.* Stockholm: Swedish International Development Cooperation Agency.

Mnookin, Robert H., Scott R. Peppet, and Andrew S. Tulumello. 1996. "The Tension between Empathy and Assertiveness." *Negotiation Journal* 12 (3): 20–35.

———. 2000. *Beyond Winning: Negotiating to Create Values in Deals and Disputes.*

Cambridge, MA: Harvard University Press.

Morris, M., and R. Cohn. 1993. "Program Evaluators and Ethical Challenges: A National Survey." *Evaluation Review* 17: 621–42.

Murray, Vic V. 2002. *Evaluation Games: The Political Dimension in Evaluation and Accountability Relationships.*

http://www.vserp.ca/pub/CarletonEVALUATIONGAMES.pdf.

Patton, M. Q. 1997. U*tilization-Focused Evaluation: The New Century Text.* 3rd ed.

Thousand Oaks, CA: Sage Publications.

Russon, Craig. 2000. The *Program Evaluation Standards in International Settings.*

http://www.wmich.edu/evalctr/pubs/ops/ops17.pdf.

Stufflebeam, D. L. 1986. "Standards of Practice for Evaluators." Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April.

Tassie, A. W., V. V. Murray, J. Cutt, and D. Gragg. 1996. "Rationality or Politics: What Really Goes on When Funders Evaluate the Performance of Fundees?" *Nonprofit and Voluntary Sector Quarterly* 25 (3): 347–63.

Taut, Sandy. 2000. "Cross-Cultural Transferability of the Program Evaluation Standards." In *The Program Evaluation Standards in International Settings,* ed. Craig Russon. Occasional Papers Series, Western Michigan University, Evaluation Center, Kalamazoo.
http://www.wmich.edu/evalctr/pubs/ops/ops17.pdf.

UNEG (United Nations Evaluation Group). 2005. *Norms for Evaluation in the UN System.* http://www.uneval.org/index.cfm?module=UNEG&Page=UNEGDocuments&LibraryID=96.

———. 2008. *Ethical Guidelines.*
http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=102.

U.S. Equal Employment Opportunity Commission. 2002. http://www.eeoc.gov/facts/fs-sex.html.

World IQ. 2008. "Simple View of Ethics and Morals." http://www.wordiq.com/definition/Simple_view_of_ethics_and_morals.

Веб-сайти

Department for International Development. http://www.keysheets.org/red_7_ swaps_rev.pdf.

Evaluation Center, Western Michigan University. http://www.wmich.edu/evalctr/

Human Rights Education. http://www.hrea.org/pubs/EvaluationGuide/.

Institute of Development Studies. 1998. *Participatory Monitoring and Evaluation: Learning from Change,* IDS Policy Briefing 12 (November). http://www.ids.ac.uk/ids/bookshop/briefs/PB12.pdf.

Inter-American Development Bank. 2004. *Proposal for Sectorwide Approaches (SWAps).* http://enet.iadb.org/idbdocswebservices/idbdocsInternet/IADBPublicDoc.aspx?docnum=509733.

MEASURE Evaluation Project. *Monitoring and Evaluation of Population and Health Programs.* University of North Carolina, Chapel Hill. http://www.cpc.unc.edu/ measure.

National Aeronautics and Space Act of 1958. http://www.hq.nasa.gov/office/pao/ History/spaceact.html.

Treasury Board of Canada. *Linkages between Audit and Evaluation in Canadian Federal Developments.* http://www.tbs-sct.gc.ca/pubs_pol/dcgpubs/TB_h4/ evaluation03_e.asp.
United Nations Population Fund. *UNFPA List of Evaluation Reports and Findings.*
http://www.unfpa.org/publications/index.cfm.
World Bank. *The World Bank Participation Sourcebook.* Washington, DC: World Bank. http://www.worldbank.org/wbi/sourcebook/sbhome.htm.

Організації з оцінювання

African Evaluation Association. http://www.geocities.com/afreval/.
American Evaluation Association. http://www.eval.org.
Australasian Evaluation Society. http://www.aes.asn.au/.
Canadian Evaluation Society. http://www.evaluationcanada.ca.
European Evaluation Society. http://www.europeanevaluation.org.
German Society for Evaluation. http://www.degeval.de/.
Government and Nongovernmental Organizations. http://www.eval.org/Resources/govt_orgs_&_ngos.htm.
Institute of Internal Auditors. http://www.theiia.org.
International Organization of Supreme Audit Institutions. http://www.gao.gov/ cghome/parwi/img4.html.
Italian Evaluation Association. http://www.valutazioneitaliana.it/aiv/news.php. Swiss Evaluation Society. http://www.seval.ch/en/.
United Kingdom Evaluation Society. http://www.evaluation.org.uk/.

Стандарти та керівні принципи

American Evaluation Association. 2004. *Guiding Principles.* http://www.eval.org/ Publications/GuidingPrinciples.asp. (Abbreviated version in brochure form: http://www.eval.org.)
———. 2005. *Program Standards.* http://www.eval.org/EvaluationDocuments/ progeval.html.
Asian Development Bank. *Guidelines.* Operations Evaluation Group. http://www.adb.org/documents/guidelines/evaluation/independent-evaluation.pdf.

Australasian Evaluation Society. *Ethical Guidelines for Evaluators.* http://www.aes.asn.au/about/Documents%20-%20ongoing/guidelines_for_ the_ethical_ conduct_of_evaluations.pdf.

Canadian Evaluation Society. *Guidelines for Ethical Conduct.* http://www.evaluationcanada.ca.

German Society for Evaluation. *Standards.* http://www.degeval.de/standards/ standards.htm.

Italian Evaluation Association. *Guidelines.* http://www.valutazioneitaliana.it/ statuto.htm#Linee).

Organisation for Economic Cooperation and Development. *Principles for Evaluation of Development Assistance.* Development Assistance Committee. http:// www.oecd.org/dataoecd/31/12/2755284.pdf.

———. *DAC Criteria for Evaluating Development Assistance.* Development Assistance Committee. http://www.oecd.org/document/22/0,2340,en_2649_34435_ 2086550_1_1_1_1,00.html.

———. DAC Network on Development Evaluation. *Evaluating Development Cooperation: Summary of Key Norms and Standards.* http://www.oecd.org/dac/ evaluationnetwork.

United Kingdom Evaluation Standards. 2008. http://www.mfcr.cz/cps/rde/xbcr/ mfcr/EvalStandards_UK.pdf and http://www.evaluation.org.uk/resources/ guidelines.aspx.

ГЛАВА 15

Погляд у майбутнє

Попередні 14 глав описували «Шлях до результатів». Розглядалися основи оцінювання розвитку: як готувати, планувати й проводити оцінювання, а також що таке професійна поведінка. Ці глави допомогли нам пройти певний шлях. У главі 15 ми спочатку розберемося, де ми були і де опинилися, а потім поговоримо про майбутнє оцінювання розвитку.

Ця глава складається з двох основних частин:

- Від минулого до сьогодення
- Майбутнє

Від минулого до сьогодення

Зміни в розвитку швидко знаходять своє відображення в його оцінюванні. Затвердження світовими лідерами Цілей розвитку тисячоліття (2000) та визначення ними розвитку як глобального партнерства для скорочення бідності означало зростання уваги до вимірювання результатів. Оцінювання розвитку відреагувало на виклик, але перед нами все ще довгий шлях.

Ця книга покликана сприяти активнішому використанню знань, що їх дає оцінювання, для прийняття рішень. Сім кроків здатні підвищити корисність оцінювання в контексті визначення ефективності розвитку:

1. *Перехід до оцінювання з орієнтацією на результат.* Все більше уваги приділяється досягненню довгострокових результатів, але цей підхід дотепер не став стандартним для розвитку та його оцінювання. Не знаючи, як має виглядати успіх інтервенції в коротко-, середньо- та довгостроковій перспективі, організації продовжуватимуть винагороджувати надання товарів або послуг, не маючи впевненості, що вони забезпечать зміни в поведінці та бажане стале інституційне покращення.
2. *Формулювання теорій змін, на яких ґрунтуються інтервенції розвитку.* Перше ключове питання під час оцінювання впливу – не «Яким був вплив інтервенції?», а, скоріше, «Чи була потрібна ця інтервенція та чому ми розраховували на те, що вона створить такий вплив?» Такий аналіз має стати стандартною практикою. Теорії змін мають наочно представляти логіку зв'язку між інтервенцією та її цільовими результатами, основні припущення стосовно того, як і чому інтервенція має працювати, а також інші помітні конкуруючі події в середовищі, що можуть зумовити зміни.
3. *Стимулювати застосування моделювання.* Для встановлення зв'язку між інтервенцією та довгостроковими результатами і впливами часто необхідні контрольні вибірки. Та хоча експериментальні (та квазі-експериментальні) плани оцінки потужні, це не «золотий стандарт». Фахівці з оцінювання мають зосереджуватися на питаннях, на які потрібно дати відповідь, та обирати план оцінювання відповідно до цих питань. Наприклад, корисніше дізнатися про додатковий вплив подовження тривалості програми або зміну інтервенції в реалізації внаслідок факторів контексту чи характеристик учасників, ніж розглядати різницю впливу в експериментальній та контрольній групах. Експериментальні та квазі-експериментальні плани оцінювання повинні бути частиною навчання будь-якого фахівця з оцінювання, але таке навчання має також охоплювати неекспериментальні плани (наприклад, вивчення практичних прикладів).

4. *Використання картування, оцінки вкладу та додатковості на рівні країни замість спроб віднести вплив на рахунок певних заходів.* Визначити, які результати розвитку дала робота однієї організації, важко вже на рівні проекту; в більш складних оцінках (на рівні сектору або країни) це просто неможливо. Після того, як розвиток визнали «глобальним партнерством», прийшла потреба зміни парадигм. Перший крок для оцінювання розвитку – картування діяльності основних донорів, кредиторів та неурядових організацій (НУО) в даній країні. Краще, щоб цей аналіз виконувався самою країною, а потім використовувався як база для оцінювання координації допомоги та визначення внеску окремої організації розвитку за допомогою методу додатковості.
5. *Зниження витрат та збільшення вигід оцінювання на рівні проекту.* Незважаючи на пишну риторику, більшість ресурсів оцінювання все ще пов'язані з оцінюванням проекту. Слід активніше використовувати кластерні оцінювання супутніх проектів протягом одного–трьох років після останнього раунду фінансування. Ці оцінки могли б вивчати закономірності формування довгострокових результатів та перевіряти наявність характеристик учасників і контекстних змінних, критичних для успіху.
6. *Нарощування потенціалу оцінювання.* Зростання уваги до вимірювання результатів само по собі призвело до збільшення попиту на спеціалістів з оцінювання розвитку. Паризька декларація та Аккрська угода допомогли зосередити увагу на потребах країн, що розвиваються, у нарощуванні потенціалу для проведення оцінок. Донори реагують збільшенням фінансування, але нарощування потенціалу – тривалий процес. Донорам також треба працювати на задоволення попиту на фахівців з оцінювання у міністерствах та інших організаціях країн, що розвиваються. Цей аспект зростаючого попиту ігнорувати не варто.

Міжнародні й двосторонні банки розвитку, а також інші організації, що надають допомогу, повинні власними силами сприяти нарощуванню кадрового потенціалу оцінки. Їх власним спеціалістам з оцінювання, яких зазвичай набирають з числа співробітників проекту розвитку, часто бракує відповідної підготовки. Іноді це економісти або дослідники іншого профілю, які пройшли підготовку щодо застосування загальних дослідницьких методів суспільствознавчих наук та економетрії, але не знайомі з якісними методами.

Навчання й інші види робіт з нарощування потенціалу також потребують оцінки. Якщо попит перевищує пропозицію, можуть виникати проблеми з якістю програм. Заяви про високу кваліфікацію в навчанні оцінюванню потребують перевірки.

7. *Використання знань, наданих оцінюванням.* Усі наведені вище кроки не дадуть жодних результатів, якщо висновки досконалого дослідження не будуть використовуватися під час розробки наступної інтервенції. Оцінювання потрібне не лише для забезпечення відповідальності, а й для навчання. Фахівці з оцінювання говорять про це багато років. Вже достатньо відомо про те, як підвищити ймовірність практичного застосування результатів оцінки в контексті розвитку. Давно пора перейти до втілення цих знань у життя (Rist and Stame 2006).

Майбутнє

Ми не можемо зазирнути в майбутнє і точно сказати, як воно складеться. Проте ми можемо робити обґрунтовані прогнози щодо тенденцій і викликів, з якими стикається оцінювання розвитку. Ми завершуємо свою книгу цими думками.

Перше і головне: ми передбачаємо постійне поширення систем моніторингу й оцінювання (МіО) буквально в кожній країні, що розвивається. Цікава тенденція – навчальні поїздки до країн, що розвиваються, з метою вивчення їх систем МіО. Ці поїздки розраховані на спеціалістів як з розвинених країн, так і з країн, що розвиваються. Ми вважаємо, що аналогічні явища спостерігатимуться не лише в міністерствах і відомствах, а й у парламентах, фондах, церквах, об'єднаннях та інших добровільних організаціях, а також структурах приватного сектору.

Попит на відповідальність за результати нікуди не дінеться. Все, що йому загрожує – це перевантаження системами моніторингу, які отримують занадто багато даних.

Внутрішнє і зовнішнє оцінювання

Оцінювання розвитку майже не зачепили дебати прибічників кількісних та якісних методів, що вирували в Американській асоціації оцінювання (ААО) в 1990-х роках. Справді, фахівці з оцінювання розвитку нерідко використовують змішані методи.

Більшою проблемою для оцінювання розвитку стало питання щодо того, яке оцінювання варто проводити – зовнішнє чи внутрішнє – та де організаційно мають бути розміщені підрозділи з оцінювання. Хоча системи моніторингу й персонал МіО явно розглядаються як штатні працівники, з оцінюванням не все так однозначно. Це особливо важливе питання для тих, хто починає створювати систему з нуля, а не використовує вже готову структуру.

Чи буде оцінювання розвитку внутрішньою функцією, поєднаною з орієнтованими на результат системами моніторингу? Чи стане вона більше відповідати стратегії Європейської Комісії, яка надає перевагу залученню до проведення оцінювання зовнішніх експертів і фірм?

Ми вважаємо, що подальший серйозний попит на результати означатиме, що оцінювання розвитку пошириться не лише в управлінських підрозділах МіО, а й у вигляді досліджень, які проводяться зовнішніми консультантами та фірмами. Ми також прогнозуємо зростання кількості закликів до проведення незалежного оцінювання.

Зовнішнє оцінювання не обов'язково є незалежним. Якщо воно проводиться сторонньою організацією, але фінансується й контролюється тими, хто керує програмою, це, насправді, внутрішнє оцінювання, яке не слід вважати незалежним. Незалежні оцінювання, незалежні підрозділи з оцінювання, а також незалежні фахівці з оцінювання повинні дотримуватися принципів незалежності, встановлених Групою співпраці з оцінювання міжнародних банків розвитку (див. главу 1), незважаючи на репутацію чи етику організації, що проводить оцінювання.

Незалежний підрозділ з оцінювання може існувати всередині організації розвитку (як, наприклад, Незалежна група оцінювання Групи Світового банку). Він може входити до складу парламенту або іншого законодавчого органу. Однак за будь-якого варіанту незалежний підрозділ з оцінювання має формально оцінювати свою незалежність за допомогою шкали оцінки (такої, наприклад, що наведена в главі 1).

Розвиток потенціалу оцінювання

Підвищений попит на МіО повинен супроводжуватися нарощуванням потенціалу оцінювання. Підрозділ з оцінювання, що покладається на кількох зовнішніх експертів, навряд чи задовольнить довгострокові потреби. Необхідно на всіх рівнях нарощувати потенціал оцінювання всередині країни. Ми прогнозуємо подальшу активізацію відповідних зусиль в країнах, що розвиваються.

Розвиток потенціалу оцінювання охоплює численні заходи з побудови та зміцнення систем МіО в країнах, що розвиваються. Він зосереджується на національному та секторальному рівнях. Мова йде про численні супутні концепції й інструменти, а саме: засоби для відстеження дієвості розвитку, специфікації цілей проектів і програм, ланцюжки формування результатів, інформацію про показники роботи (включно зі збором базових даних), МіО програм і проектів, опитування для оцінки бенефіціарів, секторальні аналізи та аудит ефективності (Kusek and Rist 2004; Mackay 2007).

Розвиток потенціалу оцінювання допомагає країнам визначати, з чого починати нарощування цього потенціалу (блок 15.1). Було створено різноманітні діагностичні інструменти (див., наприклад, World Bank 1999).

Блок 15.1 «Ключові питання діагностики урядових систем МіО»

1. Формування існуючої системи МіО: роль прибічників МіО, ключові події, що зумовили потребу в інформації МіО (наприклад, обрання уряду реформаторів, податково-бюджетна криза).
2. Міністерство або агенція, що відповідає за управління системою МіО й планування оцінювання. Функції та обов'язки основних учасників системи МіО (міністерства фінансів, міністерства планування, адміністрації президента, галузевих міністерств, законодавчого органу). Чинники, що спонукають зацікавлених осіб серйозно ставитися до МіО (інтенсивність попиту на інформацію МіО). Можливість функціонування кількох некоординованих систем МіО на національному й галузевому рівнях. Важливість загальнонаціональних, регіональних, місцевих проблем для системи МіО.
3. Стан державного сектору: як він впливає на роботу керівників? Легше чи складніше ефективно працювати й відповідати за свої показники? Існування реформ державного сектору – таких як стратегія скорочення бідності, бюджетування за результатами, покращення навичок аналізу політики, створення культури ефективності державної служби, покращення надання послуг (наприклад, впровадження стандартів обслуговування клієнтів), децентралізація уряду, активніше залучення громадянського суспільства, антикорупційна стратегія – для яких може бути корисним посилення оцінки ефективності роботи уряду.
4. Основні аспекти державного управління, що зазнають сильного впливу системи МіО: прийняття бюджетних рішень, управління програмою національного або секторального планування, а також стосунки звітування (перед міністерством фінансів, адміністрацією президента, парламентом, галузевими міністерствами, громадянським суспільством).
5. Роль інформації МіО на різних етапах бюджетного процесу (політичні консультації й планування, прийняття бюджетних рішень, аналіз ефективності й звітування про показники), а також можлива відсутність зв'язку між МіО галузевих міністерств та використанням такої інформації в бюджетному процесі. Порушення зв'язку між бюджетним процесом та національним плануванням. Можливості для зміцнення ролі МіО в бюджетному процесі.
6. Наскільки активно інформація МіО, замовлена ключовими зацікавленими особами (наприклад, міністерством фінансів), використовується іншими сторонами (наприклад, галузевими міністерствами).

Визначення перешкод використанню цієї інформації (якщо такі існують). Відомості про ступінь використання інформації різними зацікавленими особами (наприклад, діагностичний аналіз або опитування). Приклади великих проектів оцінювання, що мали значний вплив на рішення уряду.

7. Види інструментів, найактивніше застосовуваних у системі МіО (стандартні показники ефективності, швидкий аналіз чи оцінка, аудити ефективності, ретельні поглиблені оцінки впливу). Масштаб і вартість кожного з цих видів МіО. Спосіб визначення пріоритетів оцінювання: чи стосуються вони «проблемних програм», пілотних програм, дуже дорогих або дуже публічних програм? Чи спираються вони на дані систематичного дослідження, відповідаючи на питання про ефективність програми?
8. Відповідальність за збір інформації про ефективність роботи та проведення оцінювання (самі міністерства, науковці або консалтингові фірми). Проблеми з якістю чи надійністю даних, або ж з проведеними оцінками. Сильні й слабкі сторони місцевої системи МіО. Основні обмеження можливостей та пріоритети уряду в сфері нарощування потенціалу.
9. Масштаб донорської підтримки МіО в останні роки. Донорські проекти, спрямовані на підтримку МіО на загальноурядовому, секторальному або відомчому рівнях (надання технічної допомоги, інші заходи з нарощування потенціалу, а також фінансування масштабних оцінок – таких як ретельні оцінювання впливу).
10. Загальні висновки стосовно сильних і слабких сторін системи МіО, включаючи її стабільність в разі зміни складу уряду. Залежність системи від донорського фінансування та інших видів підтримки. Поточні плани щодо майбутнього зміцнення системи МіО.

Джерело: World Bank 2006.

Розвиток потенціалу оцінювання гарантує доступність результатів оцінки для допомоги країнам в чотирьох основних сферах:

- Виділення урядових ресурсів на планування, прийняття рішень та пріоритизацію, особливо в бюджетному процесі.
- Розуміння ефективності поточної діяльності на рівні сектору, програми та проекту з метою використання досвіду й забезпечення покращень.
- Забезпечення відповідальності керівників за ефективність діяльності, якою вони керують, що дає можливість контролювати роботу уряду. Ідея відповідальності охоплює визнання того, що економічне врядування й ефективний державний сектор – запорука економічної конкурентоспроможності країни. Країни, яким вдається управляти витратами й оптимізу-

вати їх, сповна винагороджуються ринками; оцінка надає інструмент, що дозволяє уряду ефективно управляти витратами.

- Демонстрація ступеня успішності заходів розвитку. Це стає все важливішим фактором для країн, які залучають зовнішні ресурси, зокрема з урахуванням того, що від агенцій міжнародного розвитку вимагають спрямовувати допомогу в країни, де попередні заходи розвитку пройшли успішно. Більше того, той факт, що агенції сприяння розвитку все частіше віддають перевагу загальноурядовому підходу, означає більшу вигоду для країни, де повсюдно доступні показники ефективності.

Для розвитку потенціалу оцінювання необхідно переконати осіб, які формують політику, в тому, що оцінка корисна; нав'язати цей процес уряду неможливо. Потрібна інформаційна інфраструктура, що дозволяє регулярно збирати дані та поширювати й використовувати результати оцінювання. Побудова практичних систем оцінювання зазвичай триває довго, адже вимагає політичних та інституційних змін (Schaumberg-Müller 1996). Приміром, у США 10 років пішло на те, щоб повною мірою втілити у життя Закон про роботу уряду та її результати, а також всі передбачені ним системи МіО (Kusek and Rist 2004). Посадовці мають вірити, що оцінювання – не система контролю, а інструмент для підвищення ефективності й оптимізації прийняття рішень.

Що можна зробити, щоб стимулювати розвиток потенціалу оцінювання? Які існують заохочення й перешкоди? Тулемонд (Toulemonde 1999) обговорює важливість використання як заохочень, так і примусу. Перший вид заохочення, який він виокремлює, пов'язаний з бюджетом. Оцінювання коштує грошей, доступність яких – необхідна умова розвитку потенціалу оцінювання. Другий стимулюючий фактор – привабливий кар'єрний шлях для тих, хто захоче стати фахівцем з оцінювання. Тулемонд визначає також шляхи формування примусового попиту на оцінювання. Основний з них – обов'язкове оцінювання. Якщо організації зобов'язані проводити оцінювання, вони повинні мати право ставити питання, а також мати повний доступ до польових даних. Має діяти вимога щодо використання результатів оцінювання.

Тулемонд визнає, що «закони й правила самі по собі нічого не варті» (1999, p. 159). Влада повинна забезпечувати їх виконання. Далі він обговорює доцільність поєднання заохочень і примусу.

Професіоналізм фахівців з оцінювання розвитку

Будь-хто може видати себе за фахівця з оцінювання розвитку. Поки що не впроваджено стандартного набору вмінь, які необхідно отримати через навчання або практику, немає атестації або сертифікації, яку треба пройти, перш ніж називати себе фахівцем з оцінювання розвитку. Чимало людей

вчаться вже в процесі роботи, без формальної фахової освіти. Управління підрозділами з оцінювання також не вимагає якихось підтверджень кваліфікації: знання й досвід в сфері оцінки не є обов'язковими.

Випробуваний часом метод навчання в ході роботи не слід відкидати; керівники без формальної підготовки можуть набрати необхідної кваліфікації. Однак через відсутність атестації замовникам послуг з оцінювання буває важко визначити кваліфікаційний рівень виконавців.

Попит на послуги з оцінювання розвитку зростає, і все частіше можна почути заклики до професіоналізму, які ми вважаємо реакцією на випадки погано спланованої або проведеної оцінки. Міжнародна асоціація оцінювання розвитку (IDEAS) відіграє все значнішу роль в зростанні професійності оцінювання розвитку в усьому світі.

Встановлення набору ключових професійних якостей

Важливим питанням оцінювання розвитку є те, чи варто встановлювати обов'язковий набір ключових професійних якостей фахівця з оцінювання (на кшталт наведеного в блоці 15.2). Американська асоціація оцінювання після початкового вивчення питання вирішила цього не робити. Канадська ж спілка розвитку, навпаки, впроваджує «професійні відзнаки» – різновид підтвердження кваліфікації. Одночасно вона прагне збільшити кількість підготовлених спеціалістів. Новозеландська асоціація оцінювання розглядає можливість визначення набору кваліфікаційних вимог.

Комітет казначейства Канади закликає до сертифікації очільників підрозділів з оцінювання, що працюють в канадських установах. Підготовлений секретаріатом цього Комітету звіт представляє результати дослідження, що «вивчало різні підходи до підвищення професіоналізму оцінювання у федеральному уряді» (Treasury of Board of Canada Secretariat 2005). Міжнародна рада зі стандартів фахової підготовки, ефективності роботи та навчання (IBSTPI) розробила набір професійних якостей для штатних працівників та зовнішніх консультантів, які виконують оцінку в організаціях (посилання на відповідний веб-сайт наведено в кінці даної глави).

Проект запровадження професійних відзнак Канадської спілки розвитку передбачає атестацію спеціалістів на основі поєднання освіти й досвіду. Причому вважається, що сертифікація вимагає проведення екзамену. Тому канадці обирають систему атестації без екзамену. Чимало механізмів, потрібних для запровадження атестаційної системи, вже розробляються (в тому числі визначення атестаційних вимог, атестація вже діючих спеціалістів-практиків, атестаційний комітет та механізм розв'язання суперечок). Початок демонстраційного проекту було заплановано на 2009 рік (AEA 2008). На сайті Канадської спілки розвитку, посилання на який наведено в кінці даної глави, містить додаткову інформацію про проект.

Поговоримо про стандарти для фахівців з оцінювання розвитку. IDEAS починає вивчати це питання. Великий обсяг підготовчої роботи й досліджень Канадської спілки оцінювання послужить базою для нового рішення. Також у нагоді може стати робота, виконана Робочою групою розвитку потенціалу оцінювання Групи оцінювання ООН (було розроблено набір кваліфікаційних вимог для персоналу з оцінювання на різних рівнях, затверджений в 2008 році) – посилання на веб-сайт Групи оцінювання ООН наведено в кінці цієї глави.

Блок 15.2 «Основні кваліфікаційні вимоги до фахівця з оцінки програм»

1.0 Професійна практика

1.1 Застосовує фахові стандарти оцінювання
1.2 Проводячи оцінювання, дотримується принципів етики, чесності й неупередженості
1.3 Застосовує особисті підходи й уміння в сфері оцінювання в роботі з клієнтами
1.4 З повагою ставиться до клієнтів, респондентів, учасників програм та інших зацікавлених осіб
1.5 Проводячи оцінювання, бере до уваги добробут суспільства
1.6 Розширює базу знань з оцінювання

2.0 Систематичні дослідження

2.1 Ознайомлений з базою знань з оцінювання (терміни, поняття, теорії, припущення)
2.2 Знайомий з кількісними методами
2.3 Знайомий з якісними методами
2.4 Знайомий зі змішаними методами
2.5 Проводить аналіз літератури
2.6 Формує теорію програми
2.7 Формує питання оцінювання
2.8 Розробляє плани оцінювання
2.9 Визначає джерела даних
2.10 Збирає дані
2.11 Оцінює валідність даних
2.12 Оцінює надійність даних
2.13 Аналізує дані
2.14 Інтерпретує дані
2.15 Виносить судження
2.16 Розробляє рекомендації
2.17 Обґрунтовує рішення, що приймаються під час оцінювання
2.18 Звітує про процедури й результати оцінювання
2.19 Відзначає сильні сторони й обмеження оцінювання
2.20 Проводить мета-оцінювання

3.0 Ситуаційний аналіз
3.1 Описує програму
3.2 Визначає оцінюваність програми
3.3 Визначає інтереси відповідних зацікавлених осіб
3.4 Задовольняє інформаційні потреби цільових користувачів
3.5 Бере до уваги конфлікти
3.6 Вивчає організаційний контекст оцінювання
3.7 Аналізує політичні аспекти, що мають відношення до оцінювання
3.8 Працює над питаннями використання оцінювання
3.9 Працює над питанням організаційних змін
3.10 Поважає унікальність місця оцінки й клієнта
3.11 Відкритий для зауважень і рекомендацій
3.12 Вносить необхідні корективи в дослідження

4.0 Керування проектами
4.1 Відповідає на запити пропозицій
4.2 Проводить переговори з клієнтами до початку оцінювання
4.3 Складає офіційні угоди
4.4 Підтримує зв'язок з клієнтами протягом усього процесу оцінювання
4.5 Готує бюджет оцінювання
4.6 Обґрунтовує витрати з урахуванням потреб в інформації
4.7 Визначає необхідні для проведення оцінювання ресурси (інформація, професійна кваліфікація, кадри, інструменти)
4.8 Застосовує відповідні технології
4.9 Супервізує інших учасників оцінювання
4.10 Навчає інших учасників оцінювання
4.11 Проводить оцінку конструктивно
4.12 Своєчасно представляє результати роботи

5.0 Професіоналізм
5.1 Знає свої якості фахівця з оцінювання (знання, вміння, схильності)
5.2 Аналізує власну практику оцінювання (сфери компетентності та сфери, що потребують вдосконалення)
5.3 Прагне професійного розвитку в сфері оцінювання
5.4 Прагне професійного розвитку у відповідних предметних областях
5.5 Будує професійні стосунки з метою покращення практики оцінювання

6.0 Міжособистісні навички
6.1 Використовує навички письмового спілкування
6.2 Використовує навички усного спілкування/активного слухання
6.3 Використовує навички ведення переговорів
6.4 Використовує навички розв'язання конфліктів
6.5 Сприяє конструктивній міжособистісній взаємодії (командна робота, координація групи й процесів)
6.6 Демонструє крос-культурну компетентність

Джерело: Stevahn and others 2005.

Дехто вважає, що галузь оцінювання в цілому ніколи не дійде згоди щодо обов'язкового набору професійних якостей фахівця з оцінювання, і кожна організація розроблятиме свій набір для фахівців з оцінювання з урахуванням власної ділової та культурної специфіки та традицій. Ми вважаємо, що для оцінювання розвитку буде сформовано загальний обов'язковий набір професійних якостей. Організації висуватимуть до фахівців з оцінювання додаткові вимоги з урахуванням своєї специфіки, але також очікуватимуть відповідності й загальним вимогам.

Також вважають, що володіти всіма якостями, необхідними для успішного проведення оцінювання в будь-якій ситуації, може тільки людина з надприродними здібностями. Звичайно, до тих, хто прагне спеціалізації в оцінюванні – наприклад соціального стану та стану довкілля – висуватимуться додаткові вимоги. Але основою буде базовий набір навичок. Це схоже на ситуацію в медицині: нейрохірург також проходить загальну медичну підготовку, але терапевт не вчиться нейрохірургії.

Сертифікація фахівців з оцінювання розвитку

Чи має бути запроваджений процес, за допомогою якого фахівці з оцінювання розвитку зможуть отримувати сертифікати, що визнаватимуться в усьому світі? Така сертифікація добре себе показала в сфері аудиту: підтверджений статус аудитора визнається й користується повагою в різних країнах. Проведене недавно IDEAS опитування своїх членів продемонструвало активну підтримку розвитку професіоналізму в різних сферах компетентності і, в остаточному підсумку, сертифікації (Morra Imas 2009).

Сертифікація фахівців з оцінювання уможливить захист споживачів від неякісної роботи та фахівців з оцінювання, що перебільшують свої досягнення. Це також механізм для відокремлення оцінки й фахівців з оцінювання від аудиту й аудиторів. Занепокоєння пов'язане з тим, що такий механізм створить перешкоди на вході в галузь, особливо в середовищі, де попит перевищує пропозицію, та в сфері, де традиційним методом підготовки спеціалістів є навчання в ході роботи.

Якщо важливими будуть якість оцінки й компетентність фахівця з оцінювання, особливо через високий зовнішній попит, високою є імовірність просування набору базових навичок. На наш погляд, якщо такий набір буде запроваджено, незабаром слідом за ним прийде й сертифікація.

Нарешті, головними викликами для оцінювання розвитку можуть бути оволодіння вмінням визначати реалістичні сподівання щодо інтервенцій, створення попиту на оцінку в суспільствах, що не мають традицій демократії, узгодження оцінювання, орієнтованого на результат, зі спільною оцінкою, а також заохочення оціночного мислення в усіх організаціях. Ми готові до цих викликів.

Завдання до Глави 15

Практична вправа 15.1: Нарощування потенціалу з оцінювання

Уряд попросив вас (або вашу групу) створити стратегічний план з нарощування потенціалу оцінювання у вашій країні. За допомогою наведених нижче питань підготуйте стратегічний план на одну–дві сторінки.

1. Які дві чи три найскладніші проблеми розвитку необхідно вирішити протягом кількох наступних років?
2. Чи відомо вам, яким потенціалом оцінювання країна володіє на даному етапі (наявність фахівців з оцінювання, навичок, ресурсів, інфраструктури)?
3. З урахуванням поточних і майбутніх потреб та проблем розвитку, а також вашої оцінки поточного потенціалу оцінювання складіть перелік з шести найважливіших змін, що допомогли б покращити потенціал оцінювання у вашій країні?
4. Що зумовлює потребу в побудові систем МіО у вашій країні?
5. Яка частина вашого уряду відповідає за результативну й ефективну реалізацію програм?
6. Чи існує офіційно оформлена (за допомогою статуту чи мандату) урядова стратегія чи організація для відстеження цілей оцінювання?
7. Хто у вашій країні володіє потенціалом і необхідними навичками для розробки й використання систем МіО? Як цей потенціал (чи його відсутність) впливає на використання МіО у вашій країні?
8. Визначте пріоритетність потреб вашої країни (критичні, дуже важливі, важливі).

Посилання та рекомендована література

AEA (American Evaluation Association). 2008. "Professional Designation for Evaluators in Canada: Where Are We Now?" Session at the annual conference of the American Evaluation Association, Denver, CO, November 5–8.

Beywl, W., and K. Harich. 2008. "University-Based Continuing Education in Evaluation: The Baseline in Europe." *Evaluation* 13 (1): 121–34.

Boyle, Richard. 1999. "Professionalizing the Evaluation Function: Human Resource Development and the Building of Evaluation Capacity." In *Building Effective Evaluation Capacity: Lessons from Practice,* eds. Richard Boyle and Donald Lemaire, 135-51. New Brunswick, NJ: Transaction Publishers.

Boyle, Richard, and Donald Lemaire, eds. 1999. *Building Effective Evaluation Capacity: Lessons from Practice.* New Brunswick, NJ: Transaction Publishers.

Dewey, J. D., B. E. Montrosse, D. C. Schroter, C. D. Sullins, and J. R. Mattox II. 2008. "Evaluator Competencies: What's Taught versus What's Sought." *American Journal of Evaluation* 29 (3): 268–87.

Ghere, G., J. A. King, L. Stevahn, and J. Minnema. 2006. "A Professional Development Unit for Reflecting on Program Evaluator Competencies." *American Journal of Evaluation,* 27 (1): 108–23.

Kusek, Jody Zall, and Ray C. Rist. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System.* World Bank, Washington, DC. http://www.oecd.org/dataoecd/23/27/35281194.pdf.

Mackay, Keith. 2007. *How to Build M&E Systems to Support Better Government.*

World Bank, Independent Evaluation Group, Washington, DC. http://www.worldbank.org/ieg/ecd/docs/How_to_build_ME_gov.pdf.

McDonald, K. E., and S. E. Myrick. 2008. "Principles, Promises, and a Personal Plea: What Is an Evaluator to Do?" *American Journal of Evaluation* 29 (3): 343–51.

Morra Imas, Linda. 2009. "The Competencies and Credentials Debate in Evaluation: What Is It all About?" Presentation to the International Development Evaluation Association's Global Assembly, Johannesburg, March.

Porteous, Nancy L., Barbara J. Sheldrick, and Paula J. Stewart. 1999. "Enhancing Managers' Evaluation Capacity: A Case Study for Ontario Public Heath." *Canadian Journal of Program Evaluation* (Special Issue): 137–54.

Rist, Ray C., and Nicoletta Stame, eds. 2006. From Studies to Streams: *Managing Evaluative Systems.* Piscataway, NJ: Transaction Publishers.
Schaumberg-Müller, Henrik. 1996. *Evaluation Capacity Building: Donor Support and Experiences.* Copenhagen: Organisation for Economic Cooperation and Development. http://www.oecd.org/dataoecd/20/52/16546669.pdf.
Sonnichsen, Richard C. 1999. "Building Evaluation Capacity within Organizations." In *Building Effective Evaluation Capacity: Lessons from Practice,* eds. Richard Boyle and Donald Lemaire, 53–73. New Brunswick, NJ: Transaction Publishers.
Stevahn, Laurie, Jean A. King, Gail Ghere, and Jane Minnema. 2005. "Establishing Essential Competencies for Program Evaluators." *American Journal of Evaluation* 26 (1): 43–59.
Toulemonde, Jacques. 1999. "Incentives, Constraints, and Culture-Building as Instruments for Development of Evaluation Demand." In *Building Effective Evaluation Capacity: Lessons from Practice,* eds. Richard Boyle and Donald Lemaire, 153–68. New Brunswick, NJ: Transaction Publishers.
Treasury of Board of Canada Secretariat. 2005. *Improving the Professionalism of Evaluation.* Final report.
http://www.tbs-sct.gc.ca/eval/dev/Professionalism/profession_e.asp.
World Bank. 1999. *Evaluation Capacity Development: A Diagnostic Guide and Action Framework.* ECD Working Paper 6, Operations Evaluation Department, Washington, DC.
———. 2006. *Diagnostic Guides.* Washington, DC. http://www.worldbank.org/ieg/ecd/diagnostic_guides.html.

Веб-сайти

Campbell, Patricia B., and Beatriz Chu Clewell. 2008. *Building Evaluation Capacity: Designing a Cross-Project Evaluation. Guide I.* Urban Institute, Washington, DC. http://www.uquebec.ca/observgo/fichiers/43799_GRA-1a.pdf.
———. 2008. *Building Evaluation Capacity: Collecting and Using Data in Cross-Project Evaluations. Guide II.* Urban Institute, Washington, DC. http://www.uquebec.ca/observgo/fichiers/42773_GRA-1b.pdf.
Canadian Evaluation Society. www.evaluationcanada.ca.

———. *Professional Designation Project.*
http://www.evaluationcanada.ca/site.cgi?s=5&ss=6&_lang=EN.

UNEG (United Nations Evaluation Group).2004a. *Evaluation Capacity Development: A Strategy Paper.*
http://cfapp1-docs-public.undp.org/eo/evaldocs1/ uneg_2006/eo_doc_350011048.doc.

———. 2004b. *Evaluation Capacity Development Task Force: Highlights.*
http:// cfapp1-docs-public.undp.org/eo/evaldocs1/uneg_2006/eo_doc_562044935.doc.

Weiss, Carol. *Evaluating Capacity Development: Experiences from Research and Development Organizations around the World.*
http://www.agricta.org/pubs/isnar2/ECDbood(H-ch7).pdf.