

УДК 303.44

СЕРГІЙ ДЕМБІЦЬКИЙ,

*аспірант відділу історії, теорії і методології
соціології Інституту соціології НАН України*

Історія розвитку методів оцінювання в соціальних і поведінкових науках за кордоном

Анотація

У статті розглянуто історію розвитку оцінних методів за кордоном, що має три основні етапи. Перший етап характеризується домінуванням експериментальної традиції, другий — появою нових альтернатив (якісні дослідження, метааналіз і програмна теорія), третій — інтеграцією якісних і кількісних методів.

Особливу увагу приділено метааналізу та програмній теорії. Метааналіз являє собою сукупність статистичних методів для об'єднання результатів різних кількісних досліджень. Його головною метою є отримання вичерпного висновку стосовно емпіричних даних, наявних з певної тематики.

Програмна теорія полягає в моделюванні (зазвичай у графічній формі) взаємозв'язків між різноманітними елементами програми, що підлягає оцінюванню. Складена модель визначає дослідницький дизайн і компоненти програми, на яких слід сфокусуватися найретельніше.

Наприкінці статті розглянуто дослідницькі дизайни в дослідженнях змішаного типу, призначених для проведення оцінювання. Також розглянуто ситуацію, що склалася у сфері оцінних досліджень в Україні.

Ключові слова: *методи оцінювання, метааналіз, програмна теорія, теоретично спрямоване оцінювання, дослідження змішаного типу*

У своїх відомих рекомендаціях, присвячених засадам дослідницького дизайну і соціального вимірювання, Д. Мілер виокремив три типи емпіричних досліджень — базові, прикладні та оцінні. Метою базових досліджень є вироблення нових знань, що висвітлюють природу взаємовідносин між

різноманітними явищами. Прикладні дослідження спрямовані на перевірку придатності знань, здобутих у перебігу базових досліджень, стосовно різноманітних предметних сфер, здатності їх розв'язувати соціальні проблеми. Своєю чергою, оцінні дослідження дають змогу проаналізувати ефективність тієї чи іншої програми, призначеної для розв'язання соціальної проблеми [Miller, 2002: р. 3].

Якщо перші два типи доволі відомі у вітчизняній соціології¹, то останній не став об'єктом пильної уваги. Так, на I Конгресі Соціологічної асоціації України (Харків, 2009) з-поміж багатьох виступів лише один безпосередньо стосувався оцінних досліджень [Кропива, 2009: с. 260]. Причину такої ситуації може прояснити історія розвитку оцінних методів за кордоном.

Історія розвитку методів оцінювання у своїх головних рисах відповідає еволюції методологічних поглядів у західній соціології. Ідеться про початкове домінування кількісного підходу до вивчення соціальної дійсності та подальшу легітимацію опозиційного йому якісного підходу. Але якщо у загальнометодологічному контексті розвиток обох підходів відбувався багато в чому на ґрунті зіткнення двох видів дослідницького пуризму, то у сфері оцінювання, що має прикладний характер, джерелом розвитку були практичні ускладнення, що потребують прагматичного розв'язання [Rallis, Rossman, 2003: р. 491–492].

Початок: експериментальна парадигма

Стрибок інтересу до оцінних методів стався у США в середині 1960-х років. Його безпосередньою причиною була реалізація програми “Велике суспільство” президента Ліндона Джонсона. Конгрес США ухвалив низку нових програм, а в найважливіших із них було затверджено також систематичне оцінювання [Rossi, Wright, 1984: р. 332]. Основні напрями політичної програми Л.Джонсона передбачали “війну з бідністю”, розвиток соціального законодавства, програму “Медікер”, федеральну допомогу освіті й культурі, створення Міністерства житлового будівництва і міського розвитку [Американа, 1996: с. 386].

Відповідно були створені нові державні установи, частково вкомплектовані вченими соціального профілю, зацікавленими у проведенні прикладних досліджень. Серед них були економісти, соціологи, психологи, а також спеціалісти у сфері освіти.

Міждисциплінарний характер цієї дослідницької активності заслуговує на особливу увагу. По-перше, спеціалісти з різноманітних сфер мусили конкурувати один з одним, а це, своєю чергою, сприяло обміну знаннями, професійними навичками та взаємному визнанню. По-друге, дослідницькі фірми вузького профілю почали звертати увагу на спеціалістів з інших предметних областей з метою підвищення своєї конкурентоспроможності. По-третє, сформувалися міждисциплінарні наукові спільноти — Спільнота оцінних досліджень та Оцінна мережа [Rossi, Wright, 1984: р. 332–333]. У 1986 році ці організації об'єдналися в Американську оцінну асоціацію, яка є,

¹ Під “вітчизняною” я розумію суто українську соціологію.

власне кажучи, міжнародною організацією. Зараз асоціація об'єднує близько п'яти з половиною тисяч членів, які представляють усі 50 штатів США і 60 країн світу [Encyclopedia of Evaluation, 2005: p. 17].

Найпасивніше сприйняли нові можливості, пов'язані з оцінними дослідженнями, науковці університетів. Приватний сектор, навпаки, оперативно відреагував на їхню появу. Дослідницькі фірми створили або суттєво розширили наявні підрозділи, відповідальні за проведення соціальних досліджень. Також на цьому етапі з'явилося чимало нових фірм, що продемонстрували разючі успіхи у сфері оцінних досліджень. Ще одним серйозним учасником, який почав надавати відповідні послуги, виявилися великі науково-дослідні інститути (наприклад, у Чикаго і Мічигані) [Rossi, Wright, 1984: p. 333–334].

Посилення інтересу до оцінних досліджень позначилося й на відповідних методологічних розробках. Оpubлікована 1966 року праця Дональда Кемпбела і Джуліан Стенлі [Campbell, Stanley, 1966], присвячена дослідницькому дизайну для оцінювання ефективності освітніх програм, забезпечила зовсім нову термінологічну базу для класифікації дослідницьких дизайнів і обговорення проблем валідності. А експериментальна парадигма, ґрунтована передусім на рандомізованому контрольованому експерименті, визначила тон у причинно-наслідковому аналізі. Значна частина оцінних досліджень на цьому етапі орієнтувалися саме на стандарти експериментальних досліджень.

Від початку мета оцінювання полягала у вимірюванні результативних ефектів програми. Відповідно, головна методологічна проблема зводилася до визначення тих ключових чинників, котрі за решти рівних умов уможливили би валідне оцінювання цих результативних ефектів. У рамках цього підходу рандомізований контрольований експеримент набув домінуючого статусу. Його концептуальні засади були закладені куди раніше у психології та біології, де використання експериментальних досліджень було поширене повсюдно. Головною відмінністю між експериментами в оцінних дослідженнях та експериментами у психології й біології було те, що перші були польовими дослідженнями і вимагали інтеграції опитувальних технік, тоді як другі — лабораторними. Для багатьох учених соціальної сфери, які дотримувалися технократичних поглядів, рандомізований польовий експеримент обіцяв заступити місце неефективного методу проб і помилок і підвищити адекватність рішень, ухвалюваних задля поліпшення соціально-політичного становища суспільства.

На початку 1970-х були профінансовані й стартували великомасштабні польові експерименти. Парадоксально, але більшість із них були сплановані й реалізовані економістами — тими, хто був далеким від експериментальної традиції.

Проте невдовзі вочевидилося, що рандомізований контрольований експеримент можна коректно здійснити лише в обмежених умовах, а потреба в оцінюванні існує у багатьох програмах, які попросту не можуть бути оцінені експериментальними методами. Бар'єри полягали не тільки в етичних і правових обмеженнях рандомізації, а й у тому, що численні тодішні програми, які цілком охоплювали свою цільову аудиторію, просто не можна було

оцінити контрольованим експериментом — через неможливість створення належної контрольної групи. Крім того, польові експерименти тривали надто довго — від 3 до 5 років і більше (від складання дизайну до підсумкового звіту). Останнє обмеження було просто неприйнятним для більшості політиків і державних чиновників, які орієнтуються на значно коротшу часову перспективу.

Д.Кемпбел і Дж.Стенлі запропонували одне із можливих рішень цих проблем, що полягає у проведенні квазіекспериментальних досліджень, які не потребують рандомізації. Вони розглянули також відповідні методологічні питання.

Але квазіекспериментальний дизайн також мав суттєві обмеження. По-перше, хоч і не такою мірою, як справжні, квазіексперименти потребували значних фінансових і часових витрат, а також залучення персоналу зі спеціальними навичками. По-друге, чимало правильно проведених квазіекспериментів (особливо модель часових рядів) передбачали наявність даних, отримуваних упродовж тривалого часу — в ідеалі були потрібні дані, зібрані до початку програми і по її завершенні упродовж тривалого часу. У переважній більшості випадків даних, зібраних до початку проведення програми, не було. По-третє, в таких дослідженнях завжди існувала серйозна загроза отримати артефакти замість необхідної інформації з приводу ефекту програми [Rossi, Wright, 1984: р. 334–335]. Останнє пов'язано із загрозами валідності. Для внутрішньої валідності головною є історична загроза¹, а для зовнішньої — загроза взаємодії процесу тестування і незалежного чинника² (докладніше див.: [Малхотра, 2007: с. 346–347]).

Коригування досвідом: нові альтернативи

Унаслідок розглянутих вище проблем дуже швидко виникла потреба в таких методах, які б дали змогу здійснювати своєчасне, порівняно недороге і “чутливе” до потреб чиновників оцінювання. Останнє значно підсилило інтерес до якісних методів, використання яких не потребувало спеціальних технічних навичок. Ці методи завжди мали своїх прихильників у соціальних науках і особливо в соціології, позаяк забезпечували розуміння соціальних процесів завдяки наближеності до польових умов.

Важливою характеристикою у плані оцінювання була також їхня гнучкість — одного разу включені в оцінювання певної програми, вони вможлилювали зміну фокусування їх від одного аспекту до іншого замість дотримання лише заздалегідь заданих дослідницьким дизайном цілей. Крім того, були й такі програми, де цілі не визначалися достатньо чітко. Тому кожне агентство, що брало участь у реалізації таких програм, самостійно форму-

¹ Специфічні події, які є побічними щодо експерименту, але відбуваються одночасно із ним. Такі події можуть впливати на значення залежних змінних, тобто на противагу незалежній змінній бути причинним чинником.

² Ефект, який виявляється в тому, що попереднє оцінювання впливає на реакцію одиниць спостереження на незалежну змінну. За наявності такого ефекту не завжди коректно поширювати здобуті результати на генеральну сукупність.

лювало завдання, а також мало можливість змінювати їх у процесі роботи. Якісні методи давали змогу долати ускладнення такого штибу.

Але якісні методи були не єдиною альтернативою експериментальному підходу. Поряд із ними почали широко застосовувати метааналіз і програмну теорію [House, 2005: р. 1070]. Оскільки ці методи практично не висвітлено у вітчизняній літературі, пропоную спинитися на них докладніше.

Метааналіз — це сукупність статистичних методів для об'єднання результатів різних кількісних досліджень із метою формулювання вичерпного висновку щодо емпіричних даних із певної тематики. Такий підхід використовують для аналізу центральних тенденцій і варіації сукупності незалежних досліджень, а також для коригування наявних помилок і зсувів. При цьому результати всіх відібраних для метааналізу досліджень належить перетворювати відповідно до спільної метрики (так званий ефективний розмір), після чого об'єднувати у загальний масив [Corcoran, Littell, 2010: р. 299].

Поетапне застосування метааналізу розглянуто у праці Чіех-Чена Бовена [Bowen, 2008: р. 705–720]. Автор виокремлює вісім етапів метааналізу: 1) концептуалізація взаємозв'язку між двома змінними, що становлять інтерес для дослідника; 2) добір сукупності досліджень, в яких перевіряють взаємозв'язок, виокремлений на попередньому етапі; 3) дизайн програмного бланка (coding sheet) для реєстрації умов, у яких проводилися відібрані дослідження; 4) розгляд кожного дослідження із використанням програмного бланка для фіксації відповідних умов; 5) обчислення міжкодувальної надійності стосовно результатів попереднього етапу; 6) розрахунок “ефективного розміру” для кожного дослідження; 7) статистичний аналіз характеристик та ефективних розмірів; 8) дослідницький звіт.

На першому етапі дослідник має визначити незалежну (X) і залежну (Y) змінні як у теоретичних, так і в операціональних термінах. Це визначення окреслює межі для літератури, яку варто розглянути. Змінні-посередники (moderator variables, W) або характеристики дослідження також важливі. Вони являють собою змінні, що можуть вплинути на напрям і/або силу взаємозв'язку між X і Y . Ці змінні слід визначити якомога чіткіше.

На наступному етапі необхідно отримати доступ до всіх або принаймні до максимально можливої кількості релевантних дослідницьких звітів. За можливості треба включити до метааналізу неопубліковані дослідження (тези, дисертації, технічні звіти і робочі документи). Вибірка в метааналізі належить до числа ефективних розмірів, наявних у попередніх дослідженнях. Її рекомендований розмір становить тридцять і більше одиниць.

Незважаючи на те, що в кожному відібраному дослідженні перевіряють взаємозв'язок між X і Y , умови, в яких вимірювався цей взаємозв'язок (W), можуть варіювати від одного дослідження до іншого. На третьому етапі здійснюють підготовку до фіксації таких умов. Для цього створюють програмний бланк. Надалі змінні-посередники можна використовувати для пояснення суперечливих результатів різних досліджень. З метою належного дизайну програмного бланка дослідник має: 1) добре знатися на всіх уключених до метааналізу дослідженнях; 2) використовувати придатну теорію як керівництво з визначення змінних-посередників, здатних вплинути на

силу і/або напрям досліджуваного взаємозв'язку; 3) бути впевненим, що програмний бланк покриває всі релевантні характеристики розглядуваних досліджень.

На четвертому етапі фіксуються всі важливі характеристики (*W*) кожного дослідження. Ці характеристики можуть вимірюватися як за безперервними, так і за дискретними шкалами. Дискретні шкали відображають якісні відмінності між дослідженнями (наприклад, тип організації, в якій проводили дослідження), а безперервні — відмінності, котрі можна виміряти в одиницях однакової розмірності (наприклад, відсоток службовців чоловічої статі).

На наступному етапі необхідно обчислити міжкодувальну надійність, яка є аналогом надійності кодувальників у контент-аналізі. Міжкодувальна надійність являє собою кількісне оцінювання узгодженості результатів, отриманих у перебігу визначення змінних-посередників двома незалежними дослідниками. У разі дискретних змінних для її обчислення можна скористатися коефіцієнтом Капа¹, а в разі безперервних — коефіцієнтом кореляції Пірсона.

На шостому етапі здійснюють розрахунок ефективного розміру, що передбачає стандартизацію сили взаємозв'язку між залежною та незалежною змінними. Мета обчислення ефективного розміру полягає в конвертації підсумкових статистик індивідуальних досліджень в одиниці стандартного відхилення, котрі можна інтегрувати в єдиний масив. Останнє відкриває можливість вичерпного опису і статистичних висновків стосовно досліджуваного взаємозв'язку.

Коли незалежна змінна є дихотомічною, а залежна — безперервною, для розрахунку ефективного розміру використовують стандартизовану розбіжність середніх. Якщо обидві змінні належать до безперервної шкали, необхідно використовувати коефіцієнти кореляції. У разі належності незалежної і залежної змінної до дихотомічних шкал використовують підхід на основі ймовірності успішного виходу (odds ratio).

На сьомому етапі аналізується рівень узгодженості ефективних розмірів, а також визначається загальне середнє значення їх. Погіршення узгодженості може зумовлюватися чи то змінними-посередниками, чи то розміром вибірок. У разі суттєвої розбіжності в розмірах вибірок, перш ніж продовжувати подальший аналіз, слід здійснити зважування ефективних розмірів, зворотно пропорційне мінливості незалежної змінної.

Далі обчислюють середнє значення ефективних розмірів і будують його 95%-й довірчий інтервал, а також оцінюють силу, напрям і гомогенність² використовуваних ефективних розмірів. Якщо всі ефективні розміри є гомогенними, то відкривається можливість підсумкового висновку стосовно досліджуваного взаємозв'язку. Коли 95%-й довірчий інтервал включає нуль, середнє значення ефективних розмірів не відрізняється від нуля. У цій

1 Для ознайомлення з коефіцієнтом Капа див.: [Cohen's kappa, s.a.].

2 Докладніше про гомогенність/гетерогенність ефективних розмірів див.: [Corcoran, Littell, 2010: p. 303–305].

ситуації дослідник має зробити висновок про відсутність взаємозв'язку між X та Y . Якщо ж 95%-й довірчий інтервал не включає нуля, можна зробити висновок про існування статистично значимого взаємозв'язку між X та Y .

На останньому етапі здійснюється написання дослідницького звіту, що відображає процес метааналізу і його результати. Такий звіт зазвичай містить анотацію, вступ, опис методів, результатів і проблемних питань (discussion).

Програмна теорія полягає в моделюванні (здебільшого в графічній формі) взаємозв'язків між різноманітними елементами програми, що підлягає оцінюванню. До таких елементів належать вхідні дані програми (program inputs), дії та процеси (реалізація програми), а також результати [Patton, 2002: p. 162].

Дизайн і реалізація програми зазвичай базуються на сукупності дескриптивних і прескриптивних допущень стосовно дій, необхідних для розв'язання тієї чи іншої соціальної проблеми. Дескриптивні допущення (модель зміни) описують каузальні процеси, що ведуть до досягнення поставлених цілей. Прескриптивні допущення (модель дії) визначають програмні компоненти і дії, які дають змогу здійснити програму. Програмна теорія реалізується у вигляді систематичної конфігурації прескриптивних і дескриптивних допущень, що належать зацікавленим у розв'язанні проблеми учасникам. Отже, відповідне оцінювання програми базується на концептуальній схемі, що забезпечує інформацію не тільки про ефективність чи неефективність програми, а й про підстави її ефективності.

Модель зміни програмної теорії включає три елементи: 1) цілі та результати; 2) детермінанти (для досягнення цілей програми необхідно визначити механізм, що є ключовим у розв'язанні проблеми); 3) інтервенція (дії, що їх здійснюють у рамках програми для досягнення змін).

Своєю чергою, модель дії має шість елементів: 1) організації, відповідальні за виконання програми; 2) протоколи інтервенції і надання послуг (у протоколі інтервенції визначено сутність, зміст та основні види діяльності, а в протоколі послуг — конкретні кроки, завдяки яким інтервенція буде здійснюватися на практиці); 3) реалізатори програми: добір, навчання і супровід (люди, відповідальні за надання послуг клієнтам: консультанти, керівники справ, соціальні працівники, спеціалісти у сфері соціального забезпечення, освіти й охорони здоров'я); 4) установа співпраці з партнерськими організаціями і партнерами у спільноті; 5) екологічний контекст (деякі програми особливо потребують контекстуальної підтримки, що полягає в отриманні доступу до супутніх можливостей навколишнього середовища¹); 6) цільова група: визначення, включення до програми, надання послуг і моніторинг.

Спираючись на ці моделі, оцінювач формулює ідеї стосовно дослідницького дизайну і компонентів програми, на яких слід сфокусуватися найре-

¹ Наприклад, програма реабілітації молодих людей групи ризику функціонуватиме ефективніше за участі в ній членів їхніх сімей і знайомих.

тельніше. Цей різновид досліджень називають теоретично спрямованим оцінюванням (theory-driven evaluation) [Chen, 2005, 1].

Мета теоретично спрямованого оцінювання полягає у забезпеченні інформації не лише про ефективність програми, а й про те, як і чому програма досягає певних результатів. Існує чотири основні види такого оцінювання:

- теоретично спрямоване процесуальне оцінювання¹ (цілісне оцінювання узгодженості між основними компонентами програмної теорії, особливо в частині моделі дії);
- оцінювання проміжного механізму² (оцінювання моделі змін у рамках програмної теорії);
- оцінювання опосередковувального механізму³ (оцінювання одного чи більше чинників, що впливають на реалізацію програми, зумовлюють ефект інтервенції й ведуть до відповідних змін результатів);
- інтегративне оцінювання процесу/результатів⁴ (вичерпне систематичне оцінювання ключових допущень, що лежать в основі реалізації програми і відповідних каузальних процесів).

Теоретично спрямоване оцінювання передбачає спільне використання якісних і кількісних методів. Проте в термінах дослідницького дизайну, збирання й аналізу даних певні методологічні патерни втілюються, виходячи зі специфіки теоретично спрямованого оцінювання. Кількісні методи найбільшою мірою придатні при оцінюванні проміжного механізму й оцінюванні опосередковувального механізму, а теоретично спрямоване процесуальне оцінювання та інтегративне оцінювання процесу/результатів передбачають збирання й аналіз і якісних, і кількісних даних [Chen, 2005, 2].

Завершення “золотого століття” оцінних досліджень і їхній сучасний стан

Підвищений інтерес до оцінних досліджень припинився так само раптово, як і розпочався. Це було пов'язане з тим, що адміністрація Рональда Рейгана, яка прийшла до влади 1981 року, почала згортати соціальні програми, реалізовані впродовж останніх 20 років. Разом із тим потреба в таких дослідженнях не зникла, і вони стали одним із інструментів уряду не лише на федеральному рівні, а й на рівні штатів і міст.

Парадоксально, але саме на цьому етапі дістала завершення інституціоналізація оцінних досліджень у США. Багато університетів почали здійснювати підготовку спеціалістів прикладного профілю і забезпечили можливість захисту докторських дисертацій з відповідної тематики. Президен-

1 Theory-driven process evaluation.

2 Intervening mechanism evaluation.

3 Moderating mechanism evaluation.

4 Integrative process/outcome evaluation.

ти Американської соціологічної асоціації 1981–1982 років також сфокусували свою діяльність на розвитку прикладної соціології [Rossi, Wright, 1984: р. 336–337].

Утім, розвиток методів оцінних досліджень і далі йде у фарватері еволюції емпіричної методології соціальних і поведінкових наук. На сучасному етапі він пов'язаний зі зміцненням статусу досліджень змішаного типу (mixed methods research), спрямованих на інтеграцію якісної і кількісної стратегій дослідження.

Залишаючись, власне, прикладним типом досліджень, сучасне оцінювання спрямоване, першою чергою, на задоволення інформаційних потреб замовників. Це відбивається в необхідності отримання високоякісної і водночас придатної для всіх зацікавлених сторін інформації. Така специфіка розмежовує оцінні та академічні дослідження і дістає виявлення в прагматичній спрямованості дослідників відповідного профілю.

Останнє пов'язане також із тим, що в більшості досліджень, здійснюваних у сфері оцінювання, необхідно реалізовувати різноманітні за своєю природою завдання, що мають міжпарадигмальні¹ межі. Оцінювач при концептуалізації та проведенні дослідження має обирати із сукупності наявних дослідницьких дизайнів і методів ті, які дадуть змогу отримати релевантні й точні результати. Ця вимога згодом найкраще виконується за використання досліджень змішаного типу.

Відповідно до типології дизайнів досліджень змішаного типу, застосовуваної в оцінюванні², можна виокремити два основні типи — компонентний та інтегративний. У компонентному дизайні якісні та кількісні методи хоча й використовують у рамках одного дослідження, але окремо одне від одного. Проте в інтегративному дизайні методи, що належать до різних парадигм, навпаки, використовують укупі.

Компонентний тип включає три види дизайну: триангуляційний, комплементарний та експансійний. За триангуляційного дизайну результати, отримані за одним методом, використовують для підтвердження результатів, отриманих за допомогою інших методів. У разі комплементарного дизайну результати, отримані за основним методом, конкретизують і доопрацьовують на підставі результатів, отриманих методами, що мають другорядне значення. При використанні експансійного дизайну застосовують різноманітні методи для отримання інформації з приводу різних аспектів оцінювання, тобто кожен метод відповідає за специфічну частину інформації.

Інтегративний тип має чотири види дизайнів: ітеративний, нестований³, голістський і трансформативний. За ітеративного дизайну результати, отримані певним методом, підказують або спрямовують до використання інших методів, релевантних у даній ситуації. Нестований дизайн пов'яза-

¹ Під парадигмами я маю на увазі якісний та кількісний типи дослідження.

² Для порівняння цієї типології із типологіями досліджень змішаного типу, застосовуваними в академічних дослідженнях, див., напр.: [Дембицкий, 2010].

³ Від англ. “*nested*” — вкладений.

ний із ситуаціями, коли один із методів інтегровано в інший¹. Голістський дизайн передбачає спільне інтегроване використання якісних і кількісних методів із метою вичерпного оцінювання тієї чи тієї програми. При цьому обидві групи методів мають еквівалентний статус. Трансформативний дизайн має місце в тому разі, коли різноманітні методи застосовують у комплексі для фіксації ціннісних поглядів, що їх у подальшому використовують для реконфігурації діалогу, учасники якого дотримуються різних ідеологічних позицій [Rallis, Rossman, 2003: р. 492–496].

Таким чином, у методологічному сенсі парадигма досліджень змішаного типу сфокусована передусім на розв'язанні проблем, пов'язаних із дослідницькими дизайнами, ґрунтованими на застосуванні і якісних, і кількісних методів.

Висновки

Розглянувши основні віхи розвитку оцінних методів за кордоном, можна дійти висновків, чому такого штибу дослідження не стали об'єктом уваги у вітчизняній соціології.

Головна причина полягає у відсутності потреби в проведенні відповідних досліджень на рівні держави. Але відсутність держзамовлення сама по собі є лише частиною проблеми. Другий тривожний момент — це незацікавленість вітчизняних соціологів у розвитку методології прикладних досліджень. За кордоном, ще до того як сформувалося держзамовлення, проводили інтенсивні методологічні дослідження у сфері оцінювання, що мають важливе значення для розвитку всієї емпіричної соціології. Так само і по завершенні “золотого століття” оцінних досліджень багато важливих методологічних ідей були сформульовані вченими, котрі спеціалізувалися на оцінюванні. Ці ідеї мали важливе значення для розвитку якісних досліджень і досліджень змішаного типу.

Не проаналізувавши світового досвіду, вітчизняна соціологія залишається не готовою до виконання завдань, що рано чи пізно постануть перед нею. Тут доречно пригадати слова вдумливого угорського соціолога Пала Тамаша: “Займатися не високою політикою, а держуправлінням — це клопітно і потребує не розмов, а здатності “сидіти і робити” ... І от з'ясується, що в новій ситуації потрібна брудна робота. Що ж, ця верства соціологів, політологів, економістів, просто інтелектуалів-демократів виявляється не готовою до виконання брудної роботи, до праці, держуправління, створення механізмів управління системою, регулювання її відхилень і відтворення” [Тамаш, 2009: с. 27]. Це, ясна річ, не означає, що ми взагалі не зможемо здійснювати оцінювання. Питання в тому, наскільки адекватним та ефективним воно буде. У цьому контексті показовою є ситуація, що склалася в Росії.

¹ Типовий випадок такого дизайну становлять дослідження на підставі методу кейс-стаді, під час проведення котрих як додатковий інструмент використовують кількісні опитувальні техніки.

Зараз наші сусіди вже усвідомили, що держава не в змозі “тягти” на собі геть усе соціальне забезпечення. Виходячи з цього, було вжито низку відповідних заходів, серед яких — ухвалення закону про монетизацію, поява автономних установ соціального забезпечення (це, наприклад, заклади соціального захисту, охорони здоров’я, освіти тощо), стандартизація послуг та інше. Оскільки один із пріоритетів у сфері соціального забезпечення — впровадження ринкових механізмів надання соціальних послуг, остільки російська держава потребує проведення оцінювання діяльності недержавних установ, що претендують на розподіл бюджетних коштів. І хоча в Росії емпірична соціологія загалом і оцінні дослідження зокрема на порядок вищі, ніж в Україні, пріоритет традиційно віддають оцінюванню опитувальними методами. Останнє продиктоване зовсім не методологічними міркуваннями, а високою формалізованістю результатів таких методів (читай — зручністю у звітності). Про те, що опитувальні методи відповідають розв’язанню лише вузького кола завдань, годі й казати.

Таким чином, використовуючи досвід інших країн, вітчизняна соціологія могла б підготувати себе до майбутніх викликів. Та чи погодиться бодай хтось займатися цією “брудною роботою”?

Література

Американа : англо-русский лингвострановедческий словарь / под ред. Г.В. Чернова. — М. : Полиграмма, 1996.

Дембицкий С. Исследования смешанного типа в социологии / С. Дембицкий // Соціальні виміри суспільства. — 2010. — Вип. 13. — (Принято к печати).

Крапива И. Оценочные исследования: социологическое измерение / И. Крапива // Соціологія в ситуації соціальних невизначеностей : I Конгрес Соціологічної асоціації України : тези доп. учасників. — Х. : ХНУ ім. В.Н.Каразіна, 2009. — С. 260.

Малхотра Н. Маркетинговые исследования / Малхотра Н. — М. : ООО “И.Д. Вильямс”, 2007.

Тамаш П. Поражение “публичных социологов” в посткоммунистическом обществе — это поражение публичной социологии здесь и теперь? / П. Тамаш // Соціологія і суспільство сьогодні : II Міжнародні соціологічні читання пам’яті Н.В.Паніної : матеріали конф. / за наук. ред. Є.І.Головахи та О.Г.Стегнія. — К. : Інститут соціології НАНУ, 2009. — С. 20–29.

Bowen C.-C. Meta-Analysis / C.-C. Bowen // Handbook of Research Methods in Public Administration / ed. by G. Miller, K. Yang. — Boca Raton ; London ; New York : CRC Press, 2008. — P. 705–720.

Campbell D. Experimental and Quasi- Experimental Designs for Research / Campbell D., Stanley J. — Chicago : Rand McNally, 1966.

Chen H.-T. Program Theory [Електронний ресурс] / H.-T. Chen // Encyclopedia of Evaluation / ed. by S. Mathison. — Thousand Oaks ; London ; New Delhi : SAGE, 2005. — Режим доступу : http://www.sage-ereference.com/evaluation/Article_n448.html

Chen H.-T. Theory-Driven Evaluation [Електронний ресурс] / H.-T. Chen // Encyclopedia of Evaluation / ed. by S. Mathison. — Thousand Oaks ; London ; New Delhi : SAGE, 2005. — Режим доступу : http://www.sage-ereference.com/evaluation/Article_n542.html

Cohen’s kappa [Електронний ресурс]. — Режим доступу : http://en.wikipedia.org/wiki/Cohen’s_kappa

Corcoran J. Meta-Analyses / J. Corcoran, J. Littel // The Handbook of Social Work Research Methods / ed. by B. Thyer. — Los Angeles ; London ; New Delhi ; Singapore ; Washington DC : SAGE, 2010.

Encyclopedia of Evaluation / ed. by S. Mathison. — Thousand Oaks ; London ; New Delhi : SAGE, 2005.

Handbook of research design and social measurement / ed. by D.C. Miller, N.J. Salkind. — Thousand Oaks ; London ; New Delhi : SAGE, 2002.

House E. Qualitative Evaluation and Changing Social Policy / E. House // The Sage Handbook of qualitative research / ed. by N. Denzin, Y. Lincoln. — Thousand Oaks ; London ; New Delhi : SAGE, 2005.

Patton M. Qualitative evaluation and research methods / Patton M. — Thousand Oaks ; London ; New Delhi : SAGE, 2002.

Rallis S. Mixed Methods in Evaluation Context: A Pragmatic Framework / S. Rallis, G. Rossman // Handbook of Mixed Methods in Social and Behavioral Research / ed. by A. Tashakkori and C. Teddlie. — Thousand Oaks ; London ; New Delhi : SAGE, 2003. — P. 491–512.

Rossi P. Evaluation Research: An Assessment / P. Rossi, J. Wright // Annual Review of Sociology. — 1984. — № 10. — P. 332–352.