

История развития методов оценивания в социальных и поведенческих науках за рубежом

Аннотация

В статье рассматривается история развития оценочных методов за рубежом, состоящая из трех основных этапов. Первый этап характеризуется доминированием экспериментальной традиции, второй — появлением новых альтернатив (качественные исследования, метаанализ и программная теория), третий — интеграцией качественных и количественных методов.

Особое внимание уделено метаанализу и программной теории. Метаанализ представляет собой совокупность статистических методов для объединения результатов различных количественных исследований. Его главной целью является получение исчерпывающего заключения относительно эмпирических данных, имеющих по определенной тематике.

Программная теория заключается в моделировании (обычно в графической форме) взаимосвязей между различными элементами программы, подлежащей оцениванию. Составленная модель определяет исследовательский дизайн и компоненты программы, на которых следует сфокусироваться наиболее тщательно.

В заключение в статье рассматриваются исследовательские дизайны в исследованиях смешанного типа, предназначенных для проведения оценки, а также ситуация, сложившаяся в сфере оценочных исследований в Украине.

Ключевые слова: *методы оценивания, метаанализ, программная теория, теоретически направляемое оценивание, исследования смешанного типа*

В своем известном руководстве, посвященном основам исследовательского дизайна и социального измерения, Д.Миллер выделил три типа эмпирических исследований — базовые, прикладные и оценочные. Целью базовых исследований является производство новых знаний, раскрывающих

природу взаимоотношений между различными явлениями. Прикладные исследования направлены на проверку пригодности знаний, полученных с помощью базовых исследований, в отношении различных предметных сфер, их способности решать социальные проблемы. В свою очередь, оценочные исследования позволяют проанализировать эффективность той или иной программы, предназначенной для решения социальной проблемы [Miller, 2002: p. 3].

Если первые два типа хорошо известны в отечественной социологии¹, то последний не стал объектом пристального внимания. Так, на I Конгрессе Социологической ассоциации Украины (Харьков, 2009) из множества сделанных выступлений лишь одно непосредственно касалось оценочных исследований [Крапива, 2009: с. 260]. Причину сложившейся ситуации может прояснить история развития оценочных методов за рубежом.

История развития методов оценивания, в своих главных чертах отвечает эволюции методологических взглядов в западной социологии. Речь идет о первоначальном доминировании количественного подхода к изучению социальной действительности и последующей легитимации оппозиционного ему качественного подхода. Но если в общеметодологическом контексте развитие обоих подходов происходило во многом на основе столкновения двух видов исследовательского пуризма, то в сфере оценивания, имеющей прикладной характер, источником развития были практические трудности, требующие прагматического разрешения [Rallis, Rossman, 2003: p. 491–492].

Начало: экспериментальная парадигма

Скачок интереса к оценочным методам произошел в США в середине 1960-х годов. Его непосредственной причиной стала реализация программы “Великое общество” президента Линдона Джонсона. Конгресс США одобрил много новых программ, а в наиболее важных из них было утверждено и систематическое оценивание [Rossi, Wright, 1984: p. 332]. Основные направления политической программы Л.Джонсона включали “войну с бедностью”, развитие социального законодательства, программу “Медикэр”, федеральную помощь образованию и культуре, создание Министерства жилищного строительства и городского развития [Американа, 1996: с. 386].

Соответственно были созданы новые государственные учреждения, частично укомплектованные учеными социального профиля, заинтересованными в проведении прикладных исследований. Среди них были экономисты, социологи, психологи, а также специалисты в сфере образования.

Междисциплинарный характер этой исследовательской активности заслуживает особого внимания. Во-первых, специалисты из различных сфер были вынуждены конкурировать друг с другом, а это, в свою очередь, способствовало обмену знаниями, профессиональными навыками и взаимному признанию. Во-вторых, исследовательские фирмы узкого профиля стали обращать внимание на специалистов из других предметных областей с целью повышения своей конкурентоспособности. В-третьих, сформировались междисциплинарные научные сообщества — Сообщество оценочных исследований и Оценочная сеть [Rossi, Wright, 1984: p. 332–333]. В 1986

¹ Под “отечественной” понимается исключительно украинская социология.

году эти организации объединились в Американскую оценочную ассоциацию, являющуюся, по сути, международной организацией. На данный момент ассоциация включает около пяти с половиной тысяч членов, представляющих все 50 штатов США и 60 зарубежных стран [Encyclopedia of Evaluation, 2005: p. 17].

Наиболее пассивно восприняли новые возможности, связанные с оценочными исследованиями, ученые университетов. Частный сектор, напротив, оперативно отреагировал на их появление. Исследовательские фирмы создали либо значительно расширили имеющиеся подразделения, отвечающие за проведение социальных исследований. Также на этом этапе появилось множество новых фирм, продемонстрировавших впечатляющие успехи в сфере оценочных исследований. Еще одним крупным участником, начавшим предоставлять соответствующие услуги, стали крупные научно-исследовательские институты (например, в Чикаго и Мичигане) [Rossi, Wright, 1984: p. 333–334].

Рост интереса к оценочным исследованиям отразился и на соответствующих методологических разработках. Опубликованная в 1966 году работа Дональда Кэмпбелла и Джулиан Стенли [Campbell, Stanley, 1966], посвященная исследовательскому дизайну для оценивания эффективности образовательных программ, обеспечила совершенно новую терминологическую базу для классификации исследовательских дизайнов и обсуждения проблем валидности. А экспериментальная парадигма, основывающаяся в первую очередь на рандомизированном контролируемом эксперименте, задавала тон в причинно-следственном анализе. Значительная часть оценочных исследований на данном этапе ориентировалась именно на стандарты экспериментальных исследований.

Изначально цель оценивания заключалась в измерении результирующих эффектов программы. Соответственно, главная методологическая проблема сводилась к определению тех ключевых факторов, которые при прочих равных условиях позволили бы валидно оценить эти результирующие эффекты. В рамках этого подхода рандомизированный контролируемый эксперимент получил доминантный статус. Его концептуальные основы были заложены намного раньше в психологии и биологии, в которых использование экспериментальных исследований было распространено повсеместно. Главным отличием между экспериментами в оценочных исследованиях и экспериментами в психологии и биологии было то, что первые были полевыми исследованиями и требовали интеграции опросных техник, а вторые — лабораторными. Для многих ученых социальной сферы, придерживающихся технократических взглядов, рандомизированный полевой эксперимент обещал заменить неэффективный метод проб и ошибок и повысить адекватность решений, принимаемых для улучшения социально-политического состояния общества.

В начале 1970-х были профинансированы и начались впечатляющие по своему масштабу полевые эксперименты. Парадоксально, но большинство из них были спланированы и реализованы экономистами — теми, кто был далек от экспериментальной традиции.

Однако вскоре стало очевидно, что рандомизированный контролируемый эксперимент может быть корректно осуществлен только в ограниченных условиях, а потребность в оценке существует во многих программах,

которые попросту не могут быть оценены экспериментальными методами. Барьеры заключались не только в этических и правовых ограничениях рандомизации, но и в том, что многие существующие программы, полностью охватывающие свою целевую аудиторию, просто нельзя было оценить с помощью контролируемого эксперимента — по причине невозможности создания подходящей контрольной группы. Кроме того, полевые эксперименты занимали очень много времени — от 3 до 5 лет и больше (от составления дизайна до итогового отчета). Последнее ограничение было просто неприемлемым для большинства политиков и государственных чиновников, ориентирующихся на куда более короткую временную перспективу.

Д.Кэмпбелл и Дж.Стенли предложили одно из возможных решений этих проблем, заключающееся в проведении квазиэкспериментальных исследований, не требующих рандомизации. Они рассмотрели и соответствующие методологические вопросы.

Но и квазиэкспериментальный дизайн имел существенные ограничения. Во-первых, хотя и не в такой степени как истинные, квазиэксперименты требовали значительных финансовых и временных затрат, а также привлечения персонала со специальными навыками. Во-вторых, многие правильно проведенные квазиэксперименты (особенно модель временных рядов) предусматривали наличие данных, получаемых в течение длительного времени — в идеале требовались данные, собранные до начала программы и после ее завершения на протяжении длительного времени. В подавляющем большинстве случаев данных, собранных до начала проведения программы, не было. В-третьих, в таких исследованиях всегда существовала серьезная угроза получить артефакты вместо необходимой информации относительно эффекта программы [Rossi, Wright, 1984: p. 334–335]. Последнее связано с угрозами валидности. Для внутренней валидности основной является историческая угроза¹, а для внешней — угроза взаимодействия процесса тестирования и независимого фактора² (более подробно см.: [Малхотра, 2007: с. 346–347]).

Коррекция опытом: новые альтернативы

Вследствие рассмотренных выше проблем очень быстро возникла потребность в таких методах, которые бы позволили осуществлять своевременную, сравнительно недорогую и “чувствительную” к потребностям чиновников оценку. Последнее значительно усилило интерес к качественным методам, использование которых не требовало специальных технических навыков. Эти методы всегда имели своих приверженцев в социальных науках и особенно в социологии, поскольку обеспечивали понимание социальных процессов благодаря близости к полевым условиям.

¹ Специфические события, посторонние по отношению к эксперименту, но происходящие одновременно с ним. Такие события могут воздействовать на значение зависимых переменных, то есть в противовес независимой переменной являться причинным фактором.

² Эффект, проявляющийся в том, что предварительная оценка влияет на реакцию единиц наблюдения на независимую переменную. При наличии такого эффекта не всегда корректно распространять полученные результаты на генеральную совокупность.

Важным качеством, с точки зрения оценивания, была и их гибкость — однажды включенные в оценку какой-либо программы, они позволяли менять их фокусировку от одного аспекта к другому, а не придерживаться лишь заранее заданных исследовательским дизайном целей. Кроме того, были и такие программы, где цели не определялись достаточно четко. Поэтому каждое агентство, участвующее в реализации таких программ, само формулировало задачи, а также имело возможность менять их в процессе работы. Качественные методы позволяли разрешать затруднения подобного рода.

Но качественные методы были не единственной альтернативой экспериментальному подходу. Наряду с ними стали широко применять метаанализ и программную теорию [House, 2005: р. 1070]. Поскольку эти методы практически не освещены в отечественной литературе, предлагаю остановиться на них более подробно.

Метаанализ представляет собой совокупность статистических методов для объединения результатов различных количественных исследований с целью получения исчерпывающего заключения относительно эмпирических данных по определенной тематике. Данный подход используется для анализа центральных тенденций и вариации совокупности независимых исследований, а также для коррекции имеющихся ошибок и смещений. При этом результаты всех отобранных для метаанализа исследований преобразуются в соответствии с общей метрикой (так называемый эффективный размер), после чего объединяются в общий массив [Cocoran, Littell, 2010: р. 299].

Поэтапное применение метаанализа представлено в работе Чиех-Чена Бовена [Bowen, 2008: р. 705–720]. Автор выделяет восемь этапов метаанализа: 1) концептуализация взаимосвязи между двумя переменными, интересующими исследователя; 2) отбор совокупности исследований, в которых проверяется взаимосвязь, выделенная на предыдущем этапе; 3) дизайн программного бланка (coding sheet) для регистрации условий, в которых проводились отобранные исследования; 4) рассмотрение каждого исследования с использованием программного бланка для фиксации соответствующих условий; 5) вычисление межкодировочной надежности в отношении результатов предыдущего этапа; 6) вычисление «эффективного размера» для каждого исследования; 7) статистический анализ характеристик и эффективных размеров; 8) исследовательский отчет.

На первом этапе исследователь должен определить независимую (X) и зависимую (Y) переменные как в теоретических, так и в операциональных терминах. Это определение очерчивает границы для литературы, подлежащей рассмотрению. Переменные-посредники (moderator variables, W) или характеристики исследования также важны. Они представляют переменные, которые могут повлиять на направление и/или силу взаимосвязи между X и Y . Эти переменные следует определить настолько четко, насколько это возможно.

На следующем этапе необходимо получить доступ ко всем или, по крайней мере, к максимально возможному количеству релевантных исследовательских отчетов. По возможности следует включить в метаанализ неопубликованные исследования (тезисы, диссертации, технические отчеты и рабочие документы). Выборка в метаанализе относится к числу эффективных

размеров, имеющих в предыдущих исследованиях. Ее рекомендуемый размер составляет тридцать и более единиц.

Несмотря на то, что в каждом отобранном исследовании проверяется взаимосвязь между X и Y , условия, в которых измерялась эта взаимосвязь (W), могут варьировать от одного исследования к другому. На третьем этапе осуществляется подготовка к фиксации таких условий. Для этого создается программный бланк. В дальнейшем переменные-посредники могут быть использованы для объяснения противоречивых результатов различных исследований. С целью надлежащего дизайна программного бланка исследователь должен: 1) быть хорошо знаком со всеми включенными в метаанализ исследованиями; 2) использовать пригодную теорию в качестве руководства по определению переменных-посредников, которые бы могли повлиять на силу и/или направление исследуемой взаимосвязи; 3) быть уверенным, что программный бланк покрывает все релевантные характеристики рассматриваемых исследований.

На четвертом этапе фиксируются все важные характеристики (W) каждого исследования. Эти характеристики могут измеряться как с помощью непрерывных, так и дискретных шкал. Дискретные шкалы отражают качественные различия между исследованиями (например, тип организации, в которой проводилось исследование), а непрерывные — различия, которые можно измерить с помощью единиц одинаковой размерности (например, процент служащих мужского пола).

На следующем этапе необходимо вычислить межкодировочную надежность, являющуюся аналогом надежности кодировщиков в контент-анализе. Межкодировочная надежность представляет собой количественную оценку согласованности результатов, полученных в ходе определения переменных-посредников двумя независимыми исследователями. В случае дискретных переменных для ее вычисления можно использовать коэффициент Каппа¹, а в случае непрерывных — коэффициент корреляции Пирсона.

На шестом этапе осуществляется расчет эффективного размера, предполагающий стандартизацию силы взаимосвязи между зависимой и независимой переменными. Цель вычисления эффективного размера заключается в конвертации итоговых статистик индивидуальных исследований в единицы стандартного отклонения, которые могут быть интегрированы в единый массив. Последнее открывает возможность исчерпывающего описания и статистических выводов относительно исследуемой взаимосвязи.

Когда независимая переменная является дихотомической, а зависимая — непрерывной, для расчета эффективного размера используется стандартизированное различие средних. Если обе переменные относятся к непрерывной шкале, необходимо использовать коэффициенты корреляции. В случае принадлежности независимой и зависимой переменной к дихотомическим шкалам используется подход на основе вероятности успешного исхода (odds ratio).

На седьмом этапе анализируется степень согласованности эффективных размеров, а также определяется их общее среднее значение. Ухудшение согласованности может быть связано либо с переменными-посредниками,

¹ Для ознакомления с коэффициентом Каппа см.: [Cohen's kappa, s.a.].

либо с размером выборок. В случае существенного различия в размерах выборок, прежде чем проводить дальнейший анализ, следует осуществить взвешивание эффективных размеров, обратное пропорциональное изменчивости независимой переменной.

Далее рассчитывается среднее значение эффективных размеров и строится его 95%-й доверительный интервал, а также оценивается сила, направление и гомогенность¹ используемых эффективных размеров. Если все эффективные размеры являются гомогенными, то открывается возможность итогового вывода относительно исследуемой взаимосвязи. Когда 95%-й доверительный интервал включает ноль, среднее значение эффективных размеров не отличается от нуля. В этой ситуации исследователь должен сделать вывод об отсутствии взаимосвязи между X и Y . Если же 95%-й доверительный интервал не включает ноль, можно сделать вывод о существовании статистически значимой взаимосвязи между X и Y .

На последнем этапе осуществляется написание исследовательского отчета, отражающего процесс метаанализа и его результаты. Такой отчет обычно включает аннотацию, вступление, описание методов, результатов и проблемных вопросов (discussion).

Программная теория заключается в моделировании (обычно в графической форме) взаимосвязей между различными элементами программы, подлежащей оцениванию. К таким элементам принадлежат входные данные программы (program inputs), действия и процессы (реализация программы), а также результаты [Patton, 2002: p. 162].

Дизайн и реализация программы обычно базируются на совокупности дескриптивных и прескриптивных допущений относительно действий, требующихся для решения той или иной социальной проблемы. Дескриптивные допущения (модель изменения) описывают каузальные процессы, которые ведут к достижению поставленных целей. Прескриптивные допущения (модель действия) предписывают программные компоненты и действия, которые бы позволили осуществить программу. Программная теория реализуется в виде систематической конфигурации прескриптивных и дескриптивных допущений, принадлежащих заинтересованным в решении проблемы участникам. Следовательно, соответствующее оценивание программы базируется на концептуальной схеме, обеспечивающей информацию не только об эффективности либо неэффективности программы, но и об основаниях ее эффективности.

Модель изменения программной теории включает три элемента: 1) цели и результаты; 2) детерминанты (для достижения целей программы необходимо установить механизм, являющийся ключевым в разрешении проблемы); 3) интервенция (действия, предпринимаемые в рамках программы для достижения изменений).

В свою очередь, модель действия включает шесть элементов: 1) организации, отвечающие за выполнение программы; 2) протоколы интервенции и предоставления услуг (в протоколе интервенции устанавливаются сущность, содержание и основные виды деятельности, а в протоколе услуг —

¹ Более подробно о гомогенности/гетерогенности эффективных размеров см.: [Cocoran, Littell, 2010: p. 303–305].

конкретные шаги, посредством которых интервенция будет осуществляться на практике); 3) реализаторы программы: подбор, обучение и сопровождение (люди ответственные за предоставление услуг клиентам: консультанты, управляющие делами, социальные работники, специалисты в области социального обеспечения, образования и здравоохранения); 4) установление сотрудничества с партнерскими организациями и партнерами в сообществе; 5) экологический контекст (некоторые программы особо нуждаются в контекстуальной поддержке, заключающейся в получении доступа к сопутствующим возможностям окружающей среды¹); 6) целевая группа: определение, включение в программу, оказание услуг и мониторинг.

Опираясь на эти модели, оценщик формулирует идеи относительно исследовательского дизайна и компонентов программы, на которых следует сфокусироваться наиболее тщательно. Этот вид исследований называется теоретически направляемым оцениванием (theory-driven evaluation) [Chen, 2005: 1].

Цель теоретически направляемого оценивания заключается в обеспечении информации не только об эффективности программы, но и о том, как и почему программа достигает определенных результатов. Существует четыре основных вида такого оценивания:

- теоретически направляемое процессуальное оценивание² (целостная оценка согласованности между основными компонентами программной теории, особенно в части модели действия);
- оценивание промежуточного механизма³ (оценка модели изменений в рамках программной теории);
- оценивание опосредующего механизма⁴ (оценка одного или более факторов, влияющих на реализацию программы, обуславливающих эффект интервенции и ведущих к соответствующим изменениям результатов);
- интегративное оценивание процесса/результатов⁵ (исчерпывающая систематическая оценка ключевых допущений, лежащих в основе реализации программы и соответствующих каузальных процессов).

Теоретически направляемое оценивание предполагает совместное использование качественных и количественных методов. Однако в терминах исследовательского дизайна, сбора и анализа данных определенные методологические паттерны проявляются, исходя из специфики теоретически направляемого оценивания. Количественные методы наиболее пригодны при оценивании промежуточного механизма и оценивании опосредующего механизма, а теоретически направляемое процессуальное оценивание и интегративное оценивание процесса/результатов предполагают сбор и анализ как качественных, так и количественных данных [Chen, 2005: 2].

¹ Например, программа реабилитации молодых людей группы риска будет функционировать более эффективно при участии членов их семей и знакомых.

² Theory-driven process evaluation.

³ Intervening mechanism evaluation.

⁴ Moderating mechanism evaluation.

⁵ Integrative process/outcome evaluation.

Завершение “золотого века” оценочных исследований и их современное состояние

Повышенный интерес к оценочным исследованиям прекратился так же внезапно, как и начался. Это было связано с тем, что пришедшая в 1981 году администрация Рональда Рейгана начала сворачивать социальные программы, осуществляемые на протяжении последних 20 лет. Вместе с тем потребность в таких исследованиях не исчезла, и они стали одним из инструментов правительства не только на федеральном уровне, но и на уровне штатов и городов.

Парадоксально, но именно на этом этапе завершилась институционализация оценочных исследований в США. Многие университеты начали осуществлять подготовку специалистов прикладного профиля и обеспечили возможность защиты докторских диссертаций по соответствующей тематике. Президенты Американской социологической ассоциации 1981–1982 годов также сфокусировали свою деятельность на развитии прикладной социологии [Rossi, Wright, 1984: p. 336–337].

Развитие же методов оценочных исследований и дальше идет в фарватере эволюции эмпирической методологии социальных и поведенческих наук. На современном этапе оно связано с укреплением статуса исследований смешанного типа (*mixed methods research*), направленных на интеграцию качественной и количественной стратегий исследования.

Являясь, по сути, прикладным типом исследований, современное оценивание направлено, в первую очередь, на удовлетворение информационных потребностей заказчиков. Это отражается в необходимости получения высококачественной и в то же время пригодной для всех заинтересованных сторон информации. Такая специфика проводит главную границу между оценочными и академическими исследованиями и находит выражение в прагматичной направленности исследователей соответствующего профиля.

Последнее связано также с тем, что в большинстве исследований, осуществляемых в сфере оценивания, необходимо решать различные по своей природе задачи, имеющие межпарадигмальные¹ границы. Оценщик при концептуализации и проведении исследования должен выбирать из совокупности существующих исследовательских дизайнов и методов те, которые бы позволили получить релевантные и точные результаты. Это требование наилучшим образом выполняется при использовании исследований смешанного типа.

Согласно типологии дизайнов исследований смешанного типа, применяющейся в оценивании², можно выделить два основных типа — компонентный и интегративный. В компонентном дизайне качественные и количественные методы хотя и используются в рамках одного исследования, но по отдельности друг от друга. В интегративном же дизайне методы, принадлежащие к различным парадигмам, напротив, используются вместе.

¹ Под парадигмами подразумеваются качественный и количественный типы исследования.

² Для сравнения этой типологии с типологиями исследований смешанного типа, применяемыми в академических исследованиях, см., напр.: [Дембицкий, 2010].

Компонентный тип включает три вида дизайна: триангуляционный, комплементарный и экспансионный. При триангуляционном дизайне результаты, полученные с помощью одного метода, используются для подтверждения результатов, полученных с помощью других методов. В случае комплементарного дизайна результаты, полученные с помощью основного метода, конкретизируются и дорабатываются на основе результатов, полученных с помощью методов, имеющих второстепенное значение. При использовании экспансионного дизайна применяются различные методы для получения информации относительно различных аспектов оценивания, то есть каждый метод отвечает за специфическую часть информации.

Интегративный тип включает четыре вида дизайнов: итеративный, нестированный¹, холистский и трансформативный. При итеративном дизайне результаты, полученные с помощью какого-либо метода, подсказывают или направляют к использованию других методов, релевантных в данной ситуации. Нестированный дизайн связан с ситуациями, когда один из методов интегрирован в другой². Холистский дизайн предусматривает совместное интегрированное использование качественных и количественных методов с целью исчерпывающей оценки той или иной программы. При этом обе группы методов имеют эквивалентный статус. Трансформативный дизайн имеет место в том случае, когда различные методы применяются вместе для фиксации ценностных взглядов, которые в последующем используются для реконфигурации диалога, участники которого придерживаются различных идеологических позиций [Rallis, Rossman, 2003: p. 492–496].

Таким образом, в методологическом смысле парадигма исследований смешанного типа сфокусирована в первую очередь на решении проблем, связанных с исследовательскими дизайнами, основывающимися на применении как качественных, так и количественных методов.

Выводы

Рассмотрев основные вехи развития оценочных методов за рубежом, можно сделать выводы о том, почему такого рода исследования не стали объектом внимания в отечественной социологии.

Главная причина заключается в отсутствии потребности в проведении соответствующих исследований на уровне государства. Но отсутствие госзаказа само по себе является лишь частью проблемы. Второй тревожный момент — это незаинтересованность отечественных социологов в развитии методологии прикладных исследований. За рубежом, еще до того как сформировался госзаказ, велись интенсивные методологические исследования в сфере оценивания, имеющие важное значение для развития всей эмпирической социологии. Так же и после завершения “золотого века” оценочных исследований, многие важные методологические идеи были сформулированы учеными, специализировавшимися на оценивании. Эти идеи имели

¹ От англ. “*nested*” — вложенный.

² Типичный случай такого дизайна представляют исследования на основе метода кейс-стади, при проведении которого в качестве дополнительного инструмента используются количественные опросные техники.

важное значение для развития качественных исследований и исследований смешанного типа.

Не проанализировав зарубежный опыт, отечественная социология будет не готова к решению тех задач, которые раньше или позже встанут перед ней. Здесь уместно вспомнить слова проницательного венгерского социолога Пала Тамаша: “Заниматься не высокой политикой, а госуправлением — это хлопотно и требует не разговоров, а способности “сидеть и делать” ... И вот выясняется, что в новой ситуации требуется грязная работа. Что ж, этот слой социологов, политологов, экономистов, просто интеллектуалов-демократов оказывается не готов к выполнению грязной работы, к труду, госуправлению, созданию механизмов управления системой, регулирования ее отклонений и воспроизводства” [Тамаш, 2009: с. 27]. Это, конечно, не означает, что мы вообще не сможем осуществлять оценивание. Вопрос в том, насколько адекватным и эффективным оно будет. В этом контексте показательна ситуация, сложившаяся в России.

В данный момент наши соседи уже осознали, что государство не в состоянии “тянуть” на себе все социальное обеспечение. Исходя из этого был осуществлен ряд соответствующих мер, среди которых — принятие закона о монетизации, появление автономных учреждений социального обеспечения (это, например, учреждения социальной защиты, здравоохранения, образования и другие), стандартизация услуг и прочее. Поскольку один из приоритетов в сфере социального обеспечения — внедрение рыночных механизмов предоставления социальных услуг, постольку российское государство нуждается в проведении оценивания деятельности негосударственных учреждений, претендующих на распределение бюджетных средств. И хотя в России эмпирическая социология в целом и оценочные исследования в частности на порядок выше, чем в Украине, приоритет традиционно отдается оценке с помощью опросных методов. Последнее продиктовано отнюдь не методологическими соображениями, а высокой формализуемостью результатов таких методов (читай — удобством в отчетности). Не стоит и говорить о том, что опросные методы удовлетворяют решению лишь узкого круга задач.

Таким образом, используя опыт других стран, отечественная социология могла бы подготовить себя к предстоящим вызовам. Но станет ли кто-то заниматься этой “грязной работой”?

Литература

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

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

Cohen's kappa [Электронный ресурс]. — Режим доступа : http://en.wikipedia.org/wiki/Cohen's_kappa

Corcoran J. Meta-Analyses / J. Corcoran, J. Littell // 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.