



ЕВРОПЕЙСКИ СЪЮЗ



ОПАК. Експерти в действие



Европейски социален фонд
Инвестиции в хората

МЕТОДИКА ЗА ПОСЛЕДВАЩА ОЦЕНКА НА ВЪЗДЕЙСТВИЕТО

Е-УПРАВЛЕНИЕ

АДМИНИСТРАТИВНО
ОБСЛУЖВАНЕ



ЕДНА ПОСОКА
МНОГО ВЪЗМОЖНОСТИ



ИНСТИТУТ ПО ПУБЛИЧНА АДМИНИСТРАЦИЯ



ЕВРОПЕЙСКИ СЪЮЗ

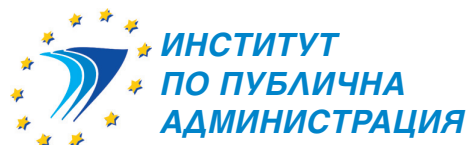


ОПАК. Експерти в действие



Европейски социален фонд
Инвестиции в хората

МЕТОДИКА ЗА ПОСЛЕДВАЩА ОЦЕНКА НА ПРОГРАМИ И НОРМАТИВНИ АКТОВЕ



Настоящият документ е създаден в рамките на проект „Изграждане на капацитет на ИПА за научни изследвания, обучение и прилагане на иновативни европейски практики в доброто управление“, Договор № С13-22-1/16. 04.2014, финансиран по Оперативна програма „Административен капацитет“, съфинансирана от Европейския съюз чрез Европейския социален фонд

Съдържание

1. Понятия и дефиниции	5
1.1. Ретроспективно оценяване спрямо наблюдение и одит	6
1.2. Цели на оценяването	7
1.3. Стъпки в процеса на оценяване	8
2. Логика на интервенцията	10
2.1. Определение и основни термини (входящи ресурси, резултат, последствия, въздействие)	10
2.2. Как се изработва/преработва логиката на интервенцията?	13
3. Критерии за последваща оценка	15
3.1. Постигане на целите	15
3.2. Ефективност	17
3.3. Ефикасност	18
3.4. Устойчивост	19
3.5. Полезност	20
4. Въпроси за оценка	21
4.1. От логика на интервенцията и критерии към въпроси за оценка	21
4.2. Как да се изработят добре насочени въпроси за оценка	21
5. Методи и техники за оценяване	23
5.1. Строги подходи	23
5.1.1. (Квази-) експериментални подходи	23
5.1.2. Прочуване на казуси	33
5.2. Прагматични подходи	46
5.2.1. Наблюдение на изпълнението	46
5.2.2. Теория на промяната като рамка за данни	46
6. Източници и събиране на данни	48
6.1. Първични източници	48
6.2. Вторични източници	50
6.3. Триангулация	51
7. Докладване	52
7.1. Насоки 52	
7.2. Разпространяване и използване на резултатите	53
8. Организация, планиране и контрол на качеството	54
8.1. Вътрешно спрямо външно оценяване	54
8.2. Необходимо време и ресурси	55
8.3. Качество на (процеса) оценяване	56

Списък с таблици и графики

Фигура 1: Видове оценяване в рамките на цикъла на политиката	6
Фигура 2: Общ логически модел	11
Фигура 3: Сложен логически модел	12
Фигура 4: Ситуация с и без политика	17
Фигура 5: Повишване на ефикасността	19
Фигура 6: Ефекти и устойчивост	20
Фигура 7: Причинно-следствени връзки в количествените изследвания	24
Фигура 8: Наблюдаван напредък	25
Фигура 9: Съпоставката позволява оценяване на въздействието	25
Фигура 10: Най-близко налично съвпадение на оценките за предразположение	28
Фигура 11: Разлика в разликата	30
Фигура 12: Прекъсване на регресията	31
Фигура 13: Непрекъснатата верига от действия и противодействия във Фаза 1 на процесите на личностното развитие (ПЛР)	35
Фигура 14: Разлика между анализ на съгласуваност и проследяване на процеса	37
Фигура 15: Механизъм за подпомагане от бюджета	38
Фигура 16: ССТV КМР конфигурации	39
Фигура 17: Таблица за анализ на съгласуваност	41
Фигура 18: Теория на промяната за анализ на приноса	42
Фигура 19: Други разяснителни фактори в анализа на приноса	43
Фигура 20: ПЛР като теория на промяна	44
Фигура 21: Таблица на истинност при качествен сравнителен анализ	45
Фигура 22: Рамка на ОИСП за оценяване на бюджетна помощ	47

1. ПОНЯТИЯ И ДЕФИНИЦИИ

Теорията за оценяването и практиката на оценяване се характеризират с многообразие. Същото важи и за определенията за оценяване. Предлагаме две определения, които илюстрират това многообразие.

Скривен (1980), влиятелен американски учен в областта, посочва същността на оценяването:

„Оценяването е систематично и обективно определяне на стойността или достойнствата на даден обект.“

С други думи, това означава, че оценяването е насочено към определяне на добавената стойност на „евалуандума“ (оценявания обект).

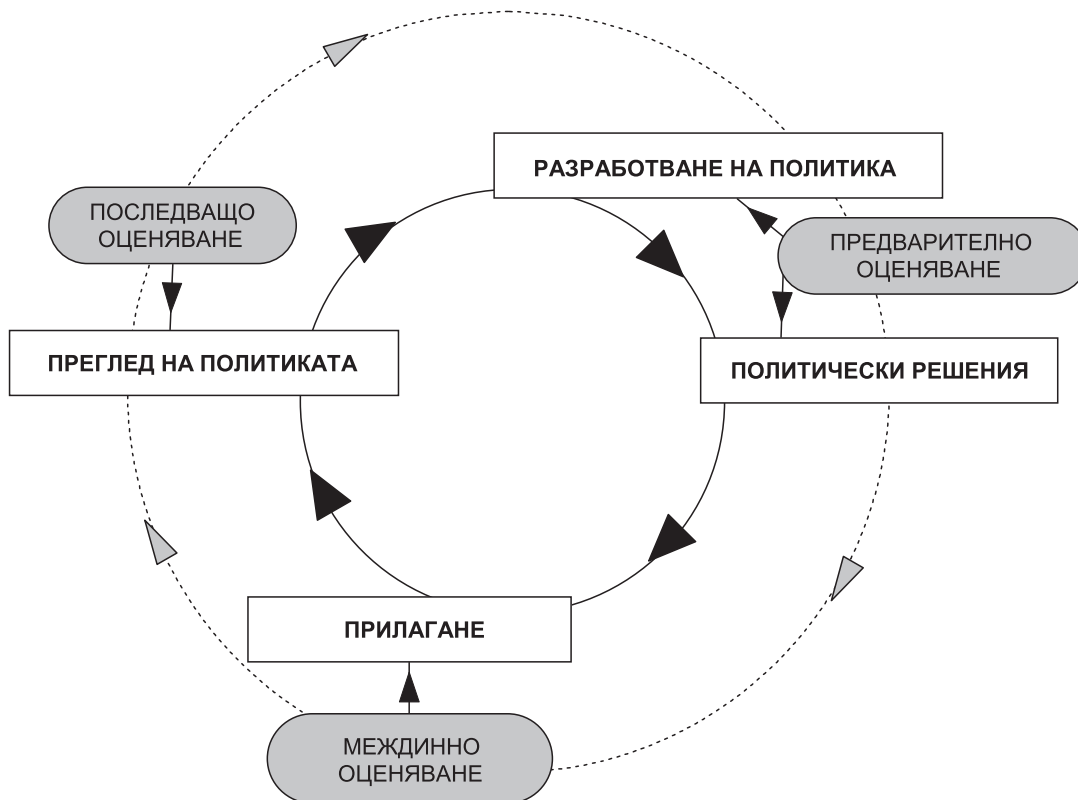
Макдейвид, Хюс и Хоторн (2012) наблягат повече на характеристиките и функциите на оценяването като го определят, както следва:

„Оценяването може да се разглежда като структуриран процес, който създава и синтезира информация, предназначена за намаляване на нивото на несигурност за заинтересованите страни за дадена програма или политика. Той е предназначен да отговори на въпроси или да тества хипотези, като след това резултатите се включват в информационните бази, използвани от тези, които имат интерес от програмата или политиката.“

Тяхното определение ни дава по-добра представа за целите, процеса, общия метод и използването на оценяването. Ще се върнем към тези въпроси в следващите параграфи с повече подробности за въпросите и критериите за оценяване.

Често оценяването се намира в края на така наречения цикъл на политиката, като последна стъпка. Такова описание само по себе си не е погрешно, но то не представя напълно реалната практика, тъй като оценяването може да допринесе и за другите етапи на цикъла на политиката като подходящ източник на информация за нея. В етапа на планиране, оценяването може да подпомогне решенията за избор между няколко възможности или алтернативи. По време на прилагането на политиката, оценяването може да осигури поглед върху напредъка и качеството на процеса или да разкрие моменти, изискващи допълнително внимание и подобряване. Затова на фигура 1 е дадено по-точно позициониране на оценяването в цикъла на политиката. Тя разкрива също и класическата типология на оценяването, като се разграничават предварително оценяване (*ex ante*), междинно оценяване (*ad interim*) и последващо оценяване (*ex post*). Предварителното оценяване дава поглед в перспектива за бъдещи резултати от политиката, които биха могли да се реализират, но все още не са осъществени. Междинното оценяване протича успоредно с етапа на прилагане на политиката. В техния смисъл на думата, последващото оценяване трябва да дава ретроспективен поглед и предполага оценяване на изцяло изпълнена политика.

На практика, обаче, между тези категории оценяване съществуват тънки (времеви) линии. Тъй като често дадена политическа мярка не е изрично или напълно приключила, разликата между междинното и последващото оценяване е свързана със зрелостта на политиката към момента на оценка и следователно се отнася до различни цели на оценяването и въпроси (вж. по-долу). По същия начин, оце-



Фигура 1: Видове оценяване в рамките на цикъла на политиката

няването може да се извърши чрез оценка на ограничено изпробване на дадена мярка, по възможност преди да се вземе решение за нейното окончателно и широкомащабно прилагане.

Оттук нататък, настоящото ръководство се съсредоточава върху оценяване от ретроспективна гледна точка, т.е. оценка, която може да бъде определена като последваща, но и като междинна.

И последната концептуална забележка се отнася до термина „политика“. Отново многообразието е на ход, тъй като оценяването на политиката може да е насочено върху еднаединствена мярка, инструмент или действие, но и върху комплексна програма, проект или план за политика. Освен това, оценяването може да има за предмет законодателен акт. Оценяването на нормативен акт може да бъде от специфична правна гледна точка, за търсене на правното качество на дадена разпоредба или закон, но може също да се съсредоточи върху съдържанието на политиката. Във втория случай, нормативът следва по-скоро да се разглежда като опаковка, носител или инструмент на оценяваната политика.

Оттук нататък, настоящото ръководство се съсредоточава върху оценяване на нормативен акт като инструмент или средство на политическата интервенция.

1.1. Ретроспективно оценяване спрямо наблюдение и одит

Важна разлика, която трябва да бъде подчертана, е разликата между (ретроспективно) оценяване и наблюдение. И двете са инструменти за информиране и са в основата на вземане на политически и управленски решения. Вън от това

важно сходство, те трябва да се разглеждат като различни, но все пак допълващи процеси с различаващи се стремежи и основни въпроси.

Наблюдението, т.е. периодично и системно измерване на базата на показатели, по същество осигурява описателна информация за ситуации или развития и се отнася до въпроси: „какво, каква е ситуацията, какво се е променило“?

Оценяването започва от конкретни въпроси, свързани със „защо?“ и „как?“; за ефекта от политическата мярка, за обстоятелствата и условията за успех или провал на политиката, за причините за странични ефекти или неочаквани резултати и т.н.

Следователно, наблюдението може да се разглежда като ценен (но не единствен) източник на информация в контекста на оценяването. Най-важното ограничение е във факта, че само описва, но не тълкува данните. Така, наблюдението може да подава сигнал, а оценяването се опитва да разчете и да обясни сигнала.

Например, статистиката на пътнотранспортни произшествия в даден район може да показва внезапно увеличение с течение на времето. Това е сигнал, предоставен от наблюдението на произшествията във времето. Оценяването ще трябва да разкрие причината за подобно развитие и в каква степен то може да бъде свързано с политическата намеса (или промени в политическата намеса, например, намаляване на контрола на движението в съответния район).

Вниманието към измерването в контекста на обществената политика и управление е придружено от различни метафори, като табло за управление или термометър. Светлините на таблото показват нещо, но често причината за проблема остава да бъде изследвана и анализирана.

Въпреки че представляват различни дейности, на практика, наблюдението и оценката понякога погрешно се считат за взаимозаменяеми термини.

Оценяването следва да се разграничи също и от одита. И двете понятия имат общи аспекти, например, акцентът върху процесите или резултатите. Възможно е разликите между оценяване и одит да са станали по-размити през годините с развитието на така наречените одити на изпълнението. И все пак, докато последният се съсредоточава върху *предпоставките* за ефективна политика или управление, оценяването се опитва да провери дали политиките са ефективни. Отново атрибутивният въпрос – „До каква степен (наблюдаваните) резултати са следствие от дадена мярка или политическа намеса?“ – е централен за оценяването и оставащ в периферията за дейността на одита.

1.2. Цели на оценяването

Практиката на оценяване е довела до широк спектър от стремежи, а именно подобряване на управлението, укрепване на демократичните принципи и, от плуралистична гледна точка, стремеж за подпомане постигането на съгласие между различни заинтересовани страни. Многобройните цели на оценяването могат да се сведат по четири главни цели, свързани с оценяване:

- Укрепване на основата на разработването на политиката;
- Подобряване на прилагането на политиката (организация и процеси);
- Отчетност;
- Поуки от политиките.

Докато първата цел е свързана предимно с предварителното оценяване, то другите три мотива могат да бъдат основни цели на ретроспективното (междинно или последващо) оценяване.

Първо, оценяването може да осигури подкрепа за разработването на политиката и вземането на решения. Например, в центъра може да бъде уместността на евентуална политическа мярка в светлината на проблема, който трябва да бъде решен. Често даден проблем може да бъде решен чрез различни мерки, подходи или инструменти. А съответният въпрос, на който трябва да се отговори, е коя алтернатива представлява най-добрият избор спрямо набор от критерии за сравнение на възможностите. Тези въпроси обикновено се разглеждат при предварителното оценяване при оценките на въздействието.

Второ, (ретроспективното) оценяване може да бъде мотивирано от целта да се управлява и подобри изпълнението. В този случай вниманието е насочено повече върху наличния капацитет, ефективното използване на ресурсите, гладкото функциониране на процесите и първи поглед върху (междинни) резултатите, измерени спрямо времевата линия на очакваното изпълнение. Определянето на факторите за успех/неуспех може да даде тласък на коригиращи действия за справяне с променящите се обстоятелства или пък ограничаващи фактори да бъдат разкрити в резултат от дейността по оценяване.

Трета основна цел на оценяването засяга отчетността в контекста на управление или в демократичен контекст. Качеството на обслужване, постигането на целите, както и ефективното и ефикасно използване на ресурсите, са легитимни очаквания от политическите интервенции и могат да бъдат оценени чрез ретроспективно оценяване.

Накрая, но в никакъв случай не по-малко важна цел на оценяването, са поуки от политиката. Какви поуки могат да се извлекат от опита с дадена мярка или подход? Кога и за кого дадена политика е ефективна? Оценяването, насочено към поуки от политиката, се съсредоточава върху връзките между проблема, логиката на интервенцията и резултатите от политиката. Защо и как определена политика успява/не успява в осъществяването на промяна?

Важно е да има яснота за целта (целите) от самото начало на оценяване, защото тя ще обуслови ползата, която може да се очаква от оценката (Челимски, 1997). Трябва да се отбележи, че оценяването може да произтича от повече от една от посочените по-горе цели.

1.3. Стъпки в процеса на оценяване

Независимо дали се провежда предварително, междинно или последващо оценяване и независимо от неговата насоченост, оценяването може да се разглежда като процес, който се състои от набор от стандартни стъпки.

Първата стъпка се отнася до определяне на обхвата или структуриране на оценяването. Определянето на обхвата ще определи много от следващите стъпки. В тази стъпка насоката на оценяването се определя чрез избор на въпроси и критерии за оценяване, както и чрез решенията за периода от време, който ще бъде разглеждан. Освен същественото структуриране на оценяването, се подготвят и следващите стъпки, а именно събиране и анализ на данни. Това предполага избор на методи, техники за провеждане на оценяването и източници на инфор-

мация, които да бъдат използвани. Добре подготвеното събиране и анализ на данни е важно, за да се излезе накрая с точни и съответстващи заключения.

Следващата важна стъпка включва събирането на данни. В тази стъпка могат да бъдат комбинирани няколко техники и източници, които могат да бъдат подходящи, в зависимост от поставяните въпроси. Един източник на данни може да бъде полезен в контекста на повече от един въпрос за оценка. Например, проучване сред целевата група, която е предмет на политическата мярка, може да осигури на оценителя информацията относно възприемането на политиката от целевата група както и относно мотивите за реакция на целевата група към съответната политика.

Даден въпрос за оценка може да изисква информацията от няколко източника, например, исторически статистически данни и интервюта. Важно е да се подготви етапът на събирането на данни чрез вглеждане в това как би могло да се отговори на всеки въпрос, чрез кои техники и с използване на какви източници. Следователно, оценителят трябва да знае кои източници ще бъдат на разположение и достъпни по време на оценяването. Най-често, при оценяването ще се наложи частично събиране на допълнителни данни, извън тези, които са на разположение. В глава 6 ще се върнем на различните източници на информацията.

Третата стъпка се състои от анализ на събраните данни. Тази стъпка не е характерна само за оценяването; тя е също важна стъпка в класическите изследвания, като превръща данните в информация, свързана е с отговора на поставените въпроси, чрез тълкуване, обобщаване, изчисляване и/или синтез.

Четвърта стъпка е преценяването. В този етап се извличат заключения, до което се свежда същността на оценяването: приписване стойност на предмета на оценяването чрез сравняване на резултатите от анализа с референтната рамка, т.е. нормите и ценностите, осигурени от целите на политиката. Този етап може да се развие в повече от една стъпка, например, преценка по всеки отделен критерий, която да бъде последвана от обща преценка.

При всяка стъпка от процеса на оценяване могат да се прилагат една или повече техники. Най-често, оценяването по този начин се основава на повече от една техника. Някои техники съчетават стъпките на анализ и преценка, например, многокритериен анализ. Това означава, че тяхното приложение не само осигурява систематичен начин за анализиране на данни, но също и заключение, например, класация на алтернативи въз основа на сравнителна преценка спрямо определен набор от критерии за оценяване.

Описаните дотук стъпки представляват основата на процеса на оценяване. Въпреки това, остават и по-нататъшни стъпки като изготвяне на доклада за оценка и организиране на разпространението на резултатите. Добре направеният доклад за оценка и целенасоченото разпространение са от решаващо значение за въздействието на оценяването. В глава 7 тези въпроси се разглеждат по-задълбочено.

2. ЛОГИКА НА ИНТЕРВЕНЦИЯТА

2.1. Определение и основни термини (входящи ресурси, резултат, последствия, въздействие)

За да се проведе последваща оценка е важно да се разбере логиката зад мярката или политиката, която е предмет на оценяване. Тази „логика на интервенцията“ трябва да бъде ясно и изцяло обяснена във вид на верига от предвидени причинно-следствени връзки. Само с ясна логика на интервенцията може да се вземат добри решения по отношение на фокуса и изпълнението на оценяването чрез подбор на подходящи въпроси за оценка, критерии, показатели и техники.

Сред широката гама от подходи за оценяване, вниманието към логиката на интервенцията е станало широко разпространена практика. Тя може да се разглежда като отправна точка за структуриране на процеса на оценяване и за определяне на неговата насока. Първо ще опишем основните елементи на логиката на интервенцията, както и нейното значение за оценяването. Второ, ще установим как логиката на интервенция може да бъде изработена/преработена.

Логиката на интервенцията може да се определи като мотивите на създателите на политики и/или други заинтересовани страни, които стоят зад дадена политическа намеса. Тя показва как се предполага интервенцията да постигне определени цели или да генерира промяна. Следователно, изразът „теория на промяната“ често се използва в контекста на логиката на интервенция.

Логиката на интервенцията описва очакваните причинни механизми, които стоят зад даден обществен проблем и причинните процеси, които се очаква да възникнат, когато се изпълнява дадена политика. С други думи, логиката на интервенция дава представа за факторите, които се смятат за податливи на намеса на политиката по начин, който води до желаните последствия или резултати от политиката.

Важно е да се отбележи, че логиката на интервенцията често е имплицитна; мотивите зад избора на политика рядко са разписани подробно в официални документи. С други думи, логиката на интервенцията често съществува в съзнанието на създателите на политиките и обикновено не е формализирана в официални документи или не е пряко достъпна за обществеността.

Въпреки това, добре изработената логика на интервенцията е от значение: разхвърляните разсъждения или недостатъчно подкрепената логика на интервенцията подкопава шансовете за постигане на добре насочени и търсени резултати от политиката.

Корените на много логики на интервенция произтичат от прозрения, убеждения и очаквания на създателите на политики, практикуващи и групи по интереси. Те имат свои собствени възприятия и предпочитания при разработването или предлагането на политически интервенции въз основа на собствения си опит, конвенционални знания, дискусии с колеги или партньори, съвети от експерти, запознатост с даден подход или дори просто интуиция. Това обяснява защо логиката на интервенцията често е имплицитна. Логиката на интервенция може също да се основава на по-широки научни теории, които са емпирично доказани. След това създателите на политиките използват научно потвърдени доказателства, когато решават кои фактори да бъдат цел на интервенцията. От една страна, използването на научни доказателства може да подобри точността и дори легитимността на интервенцията. От друга страна, логиката, основана на научни изследва-

ния, може да се сблъска с критики, че не се основава в достатъчна степен на практическите условия и ще бъде ограничителна на местно равнище. Като алтернатива, логиката на интервенция може да се основава на документираните резултати от други и предишни политически интервенции.

Степента, в която логиката на интервенция се основава на научни постижения или на опит и експертни познания на създателите на политиките ще зависи от наличието на научни доказателства и на отвореността на създателите на политиките към научна обосновка на техните политики.

Като отправна точка за оценяване е полезно да се разработи логиката на интервенция, която да се визуализира в така наречения „логически модел“ или „модел на промяната“.

Първо трябва да дефинираме набор от ключови термини, които са от значение за политическите интервенции: вход, резултат, последствия и въздействие.

ВХОДЯЩИ РЕСУРСИ се отнася до ресурсите, които могат да бъдат използвани за изпълнение на дадена мярка: финансови ресурси, персонал, логистика и т.н.

РЕЗУЛТАТ отразява сърцевината на интервенцията, т.е. действията, настъпили по време на интервенцията. Те могат да са свързани с безброй резултати, като изграждане на път, издаване на паспорт, осигуряване на обучение, събиране на битови отпадъци, и т.н.

ПОСЛЕДСТВИЯ се отнасят до непосредствения или краткосрочен ефект(и) на интервенцията.

ВЪЗДЕЙСТВИЕ се отнася до последствията в дългосрочен план след интервенцията. На практика, разликата между последствия (краткосрочни или преки ефекти) и въздействие (по-дългосрочни ефекти) не може да се определи с установен интервал от време. Поради различията в отстъпите във времето, по-лесно е да се докаже връзката между резултати и преки ефекти (последствия), в сравнение с доказването на връзките между резултатите и въздействието.

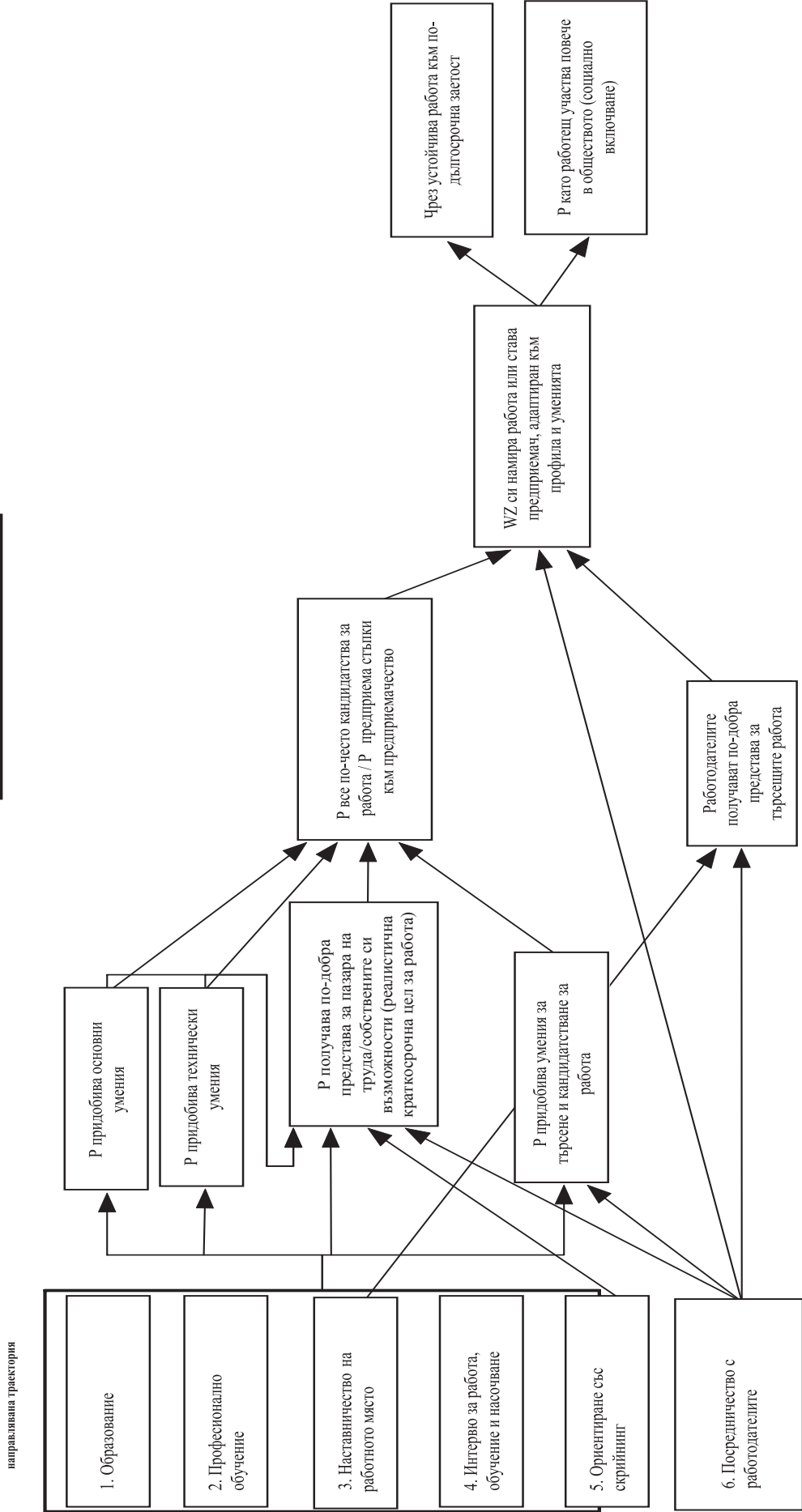
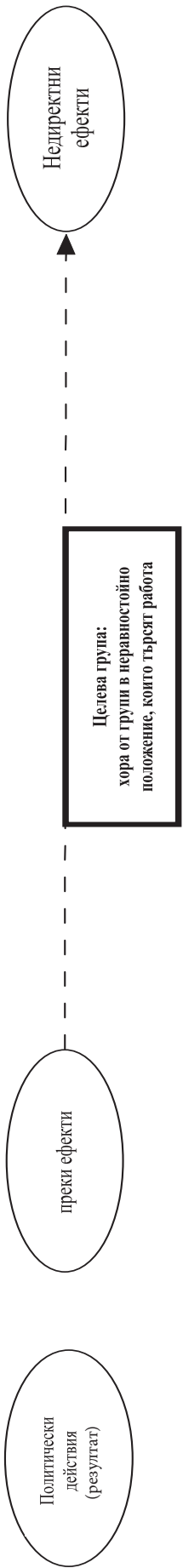
Логиката на интервенцията обикновено съдържа елементите на резултат, последствия и въздействие. Теорията на промяната предполага, че резултатите водят до последствия с трайни въздействия. Логиката на интервенцията се визуализира като се започне със сърцевината на интервенцията, т.е. генерираният резултат (вж. фигура 2).



Фигура 2: Общ логически модел

Фигура 2 изобразява обща и опростена версия на логиката на интервенция. На практика, логиката на интервенцията може да съдържа няколко резултата, няколко последствия (преки ефекти) и няколко въздействия (косвени ефекти). В зависимост от политическата интервенция и свързаните с нея мотиви и разсъждения, тези резултати, последствия и въздействия са взаимосвързани по различни начини, които отразяват очакваните причинни вериги на логиката на интервенция.

Моделът на промяната се чете от ляво на дясно: дадените един или повече резултати – според логиката на интервенция – се предполага, че те ще доведат



Фигура 3: Сложен логически модел

до един или повече посочени преки ефекти (последствия), което от своя страна следва да доведе до споменатите косвени ефекти (въздействия). По този начин моделът промяна се разглежда като причинна верига, която се очаква да се случи, когато е изпълнена дадена политика.

Алтернативни термини за компонентите на модела на промяната са съответно действия (резултати), детерминанти (преки ефекти) и резултати на политиката (непреки ефекти).

Фигура 3 илюстрира по-сложен модел промяна, който предполага множество набори от резултати, последствия (преки ефекти) и въздействия (косвени ефекти). Представена е логиката на интервенция на Европейския социален фонд във Фландрия за безработните, търсещи работа.

2.2. Как се изработва/преработва логиката на интервенцията?

Тъй като логиката на интервенцията обикновено е имплицитна или не е напълно установена, оценителят ще трябва често да разработва или да преработва логиката на интервенцията. Могат да бъдат от полза различни източници, като се започне от официални документи, като бележки по политиката, годишни отчети, пояснителни бележки, ръководни нормативи, до интервюта със създателите на политиката, участвали в разработването или осъществяването на мярката/мерките на дадена политика, както и с експерти в областта с ясно познаване на политиката.

Могат да се приложат разсъждения в права и обратна посока, за да се визуализира и преработи логиката с очакваните причинно-следствени връзки между резултат, последствия и въздействие. Обратното разсъждение „превърта назад“ очакваните причинни връзки. То започва с идентифицирането на целевите въздействия на политиката, след това разглежда последствията и най-накрая – резултатите.

Последователността на въпросите е следната:

- 1) Кои резултати (въздействия) бяха целеви за тази политика и за кого?
- 2) На кои детерминанти на преки ефекти (последствия) политиката искаше да съсредоточи усилията си, за да се получат тези резултати?
- 3) С какви резултати политиката се е опитвала да получи директните ефекти (последствия)?

Разсъждението в посока напред следва хронологичния поток на модела на промяната. То започва с определяне на резултатите от политиката, след това разглежда кои са очакваните преки ефекти и най-накрая – какво въздействие беше преценено, че ще се случи.

Последователността на въпросите е следната:

- 1) Какви са резултатите от оценяваната политика?
- 2) Кои са преките ефекти, които политиката се стремеше да предизвика с тези резултати?
- 3) Какви по-дългосрочни въздействия бяха очаквани в крайна сметка от предизвикването на преките ефекти (последствия)?

Когато се преработва логиката на интервенцията могат да се комбинират и двата подхода на разсъждения в посока напред и назад. На практика и двете групи въпроси обикновено се използват в дискусии със заинтересованите страни, за да се реконструира логиката на интервенция.

Като цяло, повтарящите се въпроси, за да се преработи логиката на интервенцията, са:

- Какво е това, което искахме да постигнем в крайна сметка? (въздействие)
- Какви са лостовете за постигане на въздействието? (преки ефекти)
- Какво можем да очакваме от конкретна намеса?

Две важни неща трябва да се имат предвид, когато се преработва логиката на интервенция. На първо място е важно да се гарантира, че логиката на интервенция, както е визуализирана на схемата (логически модел или модел на промяната), в действителност отразява логиката на създателите на политиката. Това изглежда просто, но на практика процесът на преработване може лесно да предизвика дискусия сред заинтересованите страни и да се поставят под съмнение някои очаквани връзки между резултат, преки ефекти и косвени ефекти. Някои могат да посочат примери, при които предполагаемите връзки не са се осъществили и да се постави под въпрос валидността на причинните връзки. Други могат да се позовават на данни, които показват, че връзката може да се появява само при определени благоприятни условия и т.н. Въпреки че това може да бъде много ценна информация за оценителя, в по-нататъшните стъпки за структуриране и провеждане на оценяването, оценителят трябва да се съсредоточи върху преработване на логиката на интервенция, както е била първоначално изградена и използвана при осъществяването на оценяваната политика. Самото оценяване по същество ще се опита да провери до каква степен очакваните връзки се или не се случват на практика. Това обаче не е лесна задача за оценителя, тъй като много от заинтересованите страни ще имат свой собствен прочит и интерпретация на разглежданата политика. Така оценителят трябва да постигне общоприетата версия на логиката на интервенция, в смисъл, че правилно отразява това, което политиката и нейните очаквани преки и косвени ефекти са били първоначално предназначени да постигнат. Въпреки това, възможните забележки във връзка с предпоставки, обстоятелства и вероятността на ефектите са важна информация, която трябва да се вземе под внимание при подготовката на оценяването по принцип и по време на етапа на анализ в частност. Предположенията и съмненията могат да бъдат проверени в процеса на оценяване.

Вторият важен въпрос се отнася до степента на подробност, с която логиката на интервенция е преработена. Степента на подробност, определена в логиката на интервенцията ще повлияе значително на степента на подробност при оценяването на причинни връзки, резултати, последствия и въздействия. Нека разгледаме като пример политика в областта на безопасността на движението по пътищата с целенасочено въздействие за намаляване на броя на жертвите на пътя. Казано по този начин, оценката, основана на тази логика на интервенцията, ще провери дали броят на жертвите е увеличен или намален. Обаче, анализът на въздействието на политиката (и заключенията от оценяване на интервенцията) биха могли да бъдат по-разнообразни и нюансирани, ако се разглеждат видове жертви (например, смъртни случаи, сериозни наранявания, леки наранявания и т.н.) и/или видове участници в пътното движение (например, пешеходци, колоездачи, коли, камиони, обществен транспорт и т.н.). По този начин трябва да се провери дали логиката на интервенция прави такива разграничения и дали конкретни желани въздействия са определени за различните видове участници в пътното движение. Ако е така, в схемата на оценяване трябва да се вземат под внимание съответните разграничения. Ако не е така, оценяването може, но без да бъде задължително, да направи по-подробен анализ на въздействието върху възможните произшествия и видовете участници в движението по пътищата.

3. КРИТЕРИИ ЗА ПОСЛЕДВАЩА ОЦЕНКА

При последващо оценяване могат да се прилагат различни критерии, за да се преценят (междинните) резултати от политиката. В следващите параграфи разглеждаме най-често използваните критерии за ретроспективна оценка. Общото правило е, че всеки критерий ще предизвика конкретни въпроси за оценка, които определят посоката на оценяването.

3.1. Постигане на целите

Оценяването може да се съсредоточи на първо място върху степента, в която целите на политиката са постигнати. На този въпрос може да се отговори без да се поглежда към дълбоките причини, които обясняват успеха или провала. Въпросът за степента, до която наблюдаваните промени наистина могат да бъдат приписани на извършените действия (или други фактори са повлияли положително или отрицателно), се оставя извън обхвата на оценката. Оценяването се съсредоточава само върху постигането на заявените цели и няма стремеж да покаже причинни връзки между факторите, нито пък се стреми да обясни тези причинни връзки.

Прилагайки този критерий към модела на промяната, ние се концентрираме предимно върху отделните компоненти на модела, а не върху взаимоотношението между тях. Въпреки това, всеки от компонентите може да бъде разгледан. Всъщност, постигането на целта може да се проверява на различни нива: на ниво резултат, ниво последствия (преките ефекти) и на ниво въздействие (косвени ефекти). Например, може да искаме да проверим постигането на целите по отношение на интервенция, състояща се в по-голям контрол на скоростта на движението, за да се наложат ограничения на скоростта (пряк ефект) и в крайна сметка да се подобри безопасността на движение по пътищата (въздействие). Постигането на целта на ниво резултат би означавало, че предвижданите (часове на) засилени проверки на скоростта по пътищата са реализирани. Постигането на целта на ниво последствия би означавало, че полицията е регистрирала по-малко нарушения на ограничението на скоростта (очаквано последствие или пряк ефект) в сравнение с предходните периоди. Постигането на целта на ниво въздействие или косвен ефект би означавало, че пътната безопасност се е повишила, например изразено в по-малко пътнотранспортни произшествия.

С терминологията на планирането, това означава, че можем да разглеждаме както постигането на оперативни цели (резултат), така и стратегически цели (преки и косвени ефекти).

Тази разлика между постигане на оперативните цели и постигане на стратегическите цели е от значение, тъй като оперативните цели (резултат) могат да бъдат по-пряко свързани с функционирането и управлението на обществените органи. Правителствата могат само косвено да направляват ефектите, тъй като факторите на контекста също могат да влияят на обществения ефект.

Оценяването на постигането на целите зависи от резултата и /или измерването на ефекта (наблюдение). За да се извърши оценяване по критерия за постигане на целите, са важни три елемента:

- 1) ясни и измерими цели (SMART цели);
- 2) събиране на данни за резултати и/или ефекти;
- 3) сравняване с избрана мярка за постигане на цели.

Наличието на SMART цели е важна, но на практика често проблематична предпоставка за оценяване. SMART е съкращение за конкретни, измерими, съгласувани, реалистични и срочни цели. По-специално, често проблематични са характеристиките на целите като измерими и срочни. Целите обикновено са формулирани доста неясно и не винаги им се приписват ясни времеви хоризонти. Например: „Искаме по-малко жертви на пътнотранспортни произшествия“, казва нещо за посоката на резултатите, които политиката иска да постигне, но нищо конкретно за нивото на амбициите (цел), нито за сроковете, в които следва да се постигне желаният резултат. Например, по-SMART формулирана цел, би била: „Искаме намаление от 5% на пътнотранспортни произшествия до края на следващата година“ (като за произшествие се счита всяка жертва на пътнотранспортни произшествия, която умира в рамките на 6 седмици след инцидента).

Събирането на данни може да се организира чрез различни източници в зависимост от вида на разглежданите резултати и ефекти. Отново степента на подробност е от значение: трябва ли да разглеждаме броя на жертвите като цяло или трябва да разглеждаме конкретни видове жертви?

Третата стъпка се отнася до тълкуването на данните спрямо база за сравнение. Най-общо са възможни три ситуации:

- *първоначалната ситуация* (I) преди началото на изпълнението на политиката;
- *целевата или очаквана ситуация* (E) след определен период на прилагане на политиката;
- *идеалната, т.е. най-оптималната ситуация* (O).

В зависимост от избраната база за сравняване (I, E или O) и отчитайки действително постигнатата и измерена ситуация (A), понякога след интервенцията са възможни заключения за:

- Постигнатата промяната $(A - I)$;
- Разликата между очакваната и постигнатата ситуация $(E - A)$;
- Полученият дял от очакваната промяна $(A - I) * 100 / (E - I)$;
- Разликата между оптималната и постигнатата ситуация $(O - A)$;
- Разликата между полученото и очакваното развитие $(E - A) * 100 / (E - I)$.

Например, интервенция, целяща да повиши качеството на повърхностните води, може да се опита да вдигне нивото от начална степен (30) до очаквано ниво (60), докато оптималното ниво е 100 (при индекс от 0–100). Когато интервенцията е била осъществена, ново измерване може да покаже, че качеството на водата се е увеличило до 50. Въпреки че това без съмнение е подобрение, все още не е постигнато нито очакваното ниво, нито оптималното ниво. В зависимост от мярката за постигане на целите, като се вземе първоначалното, очакваното или оптималното ниво, заключението ще бъде или по-скоро положително или отрицателно.

Когато планираните цели се вземат за отправна точка за преценката за постигане на целите, трябва да се има предвид, че съществуват няколко вида планирани цели:

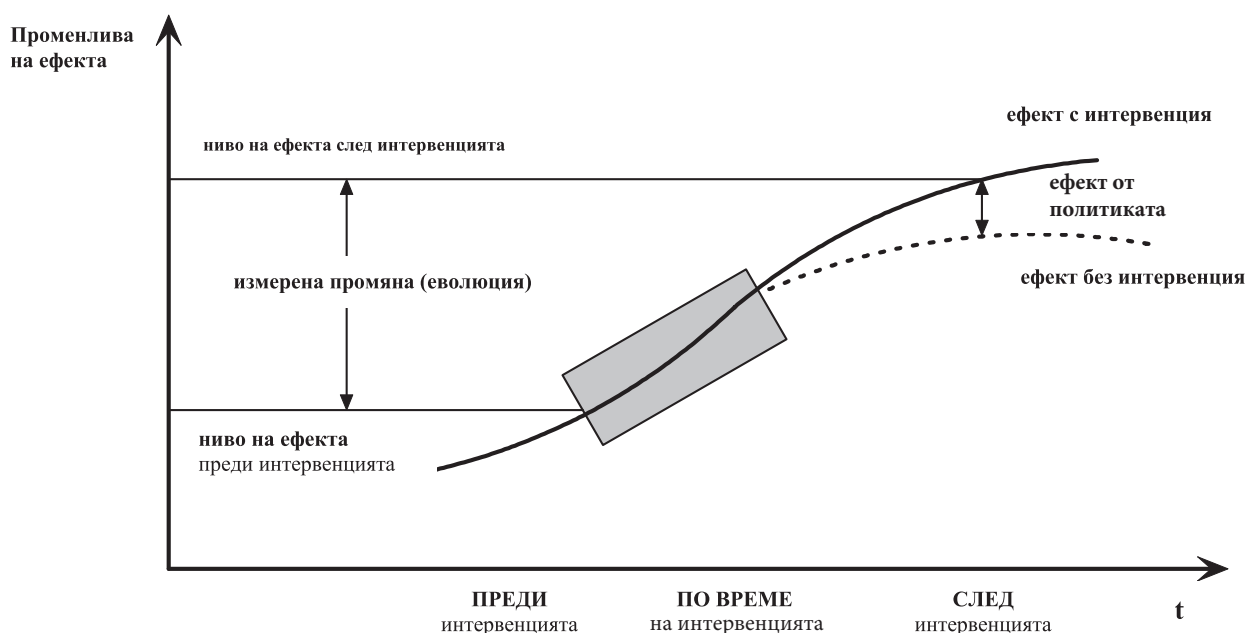
- Политически избраното ниво на амбиция за мотивиране на различните заинтересовани страни в усилията им;
- Технически стандарт (например, определяне на достъпа до пазара);
- Изискване за минимална прагова стойност;
- Приета максимална стойност;
- Интервал от стойности.

За да извърши оценяване на постигането на целите, оценителят се нуждае от по-голяма яснота по отношение на целите на политиката и на изисквания фокус сред целите (оперативно и / или стратегическо ниво), докато само конкретни планирани цели ще осигурят ясен ориентир за формулиране на заключенията.

3.2. Ефективност

Критерият за ефективност е в основата на последващата оценка и отразява съществения въпрос, който стои пред създаващите политиките: до каква степен дадена политическа интервенция е допринесла за постигане на желаната промяна? От гледна точка на логическия модел, фокусът тогава е върху очакваната причинна динамика между резултатите и преките и косвени ефекти. Причинно-следствената връзка между причина А и следствие В предполага три условия. Първо, А следва хронологично да предхожда В. Второ, между А и В трябва да има взаимна връзка: това обикновено означава, че ако А се увеличава, В трябва да се увеличи; и когато А намалява, В следва също да намалее. И накрая, възможни алтернативни обяснения за връзката между причина и следствие не са потвърдени.

Така, въпросът дали дадена политика е ефективна се отнася до въпроса дали интервенцията на политиката е постигнала нещо различно и дали е причина за наблюдаваното въздействие. Това сравнение между ситуация „с политика“ (наличие на интервенция) и ситуация „без политика“ е представено схематично на фигура 4.



Фигура 4: Ситуация с и без политика

Графиката показва нивото на променливата на ефекта (ос Y) в рамките на определен период от време (ос X). Променливите са измерими фактори, които могат да имат различни и развиващи се стойности. Например, качеството на водите, натовареността на трафика, удовлетвореността на клиента или броя на учениците, които пушат в дадено училище, могат да варират с течение на времето. Политиката се стреми към определени ефекти, които могат да бъдат преведени в променливи и измерени чрез показатели. Графиката показва еволюцията на стойността на променливата на ефекта преди, по време и след приключване на интервенцията на политиката. Плътната линия отразява нивото на ефекта въз основа на периодични измервания. Това е крива, показваща, че променливата на ефекта се променя и това може да се дължи на няколко обстоятелства.

Нивото на ефекта не предоставя директно информация за ефективността на политиката. Плътната линия показва само еволюция и обща посока (увеличение).

Пунктираната линия отразява еволюцията на променливата на ефекта в сценария без интервенция на политиката. От сравнението на двете линии, можем да заключим, че нивото на ефекта също би се увеличило и без интервенция на политиката, обаче в по-малка степен. Разликата между двете нива на ефекта отразява добавената стойност в резултат от интервенцията на политиката. Това е разликата, която с основание може да бъде посочена като „ефект на политиката“, т.е. може да бъде приписана на интервенцията на политиката след сравняване на сценариите с и без политика.

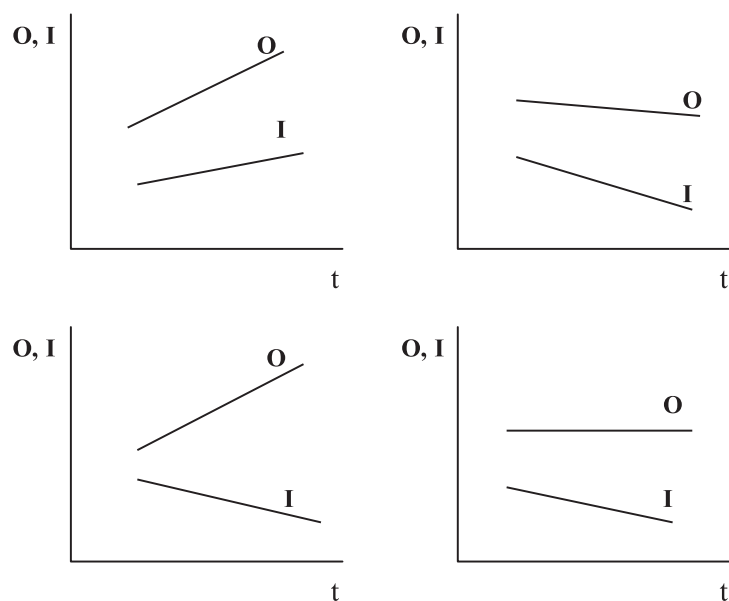
Например, субсидиите за фирми, инвестиращи в управлението на качеството на водата, могат да допринесат за подобряване качеството на повърхностните води. Въпреки това, дори и без такава интервенция качеството на вода може покаже положителна тенденция. Въпросът е до каква степен интервенцията е добавена стойност, т.е. дава ли допълнителен тласък на подобряването на качеството на водата?

Определянето на ефекта от политиката в този смисъл е най-трудният въпрос за оценка. Интуитивно ще стане ясно, че на практика не е лесно да се направи сравнение между ситуация с и без политика. В глава 5 се обсъждат в по-голяма дълбочина възможни стратегии за оценяване, където ефективността е основен критерий за оценка.

3.3. Ефикасност

Понятието за ефикасност се отнася до връзката между вложените ресурси (входящи) и доставените стоки или услуги (резултат). Например, времето, необходимо за вземане на решение по заявление за субсидии, паричните средства, необходими за изграждане на тунел и т.н. Оценяването на критерия за ефикасност обикновено се осъществява от целите на подобрението или отчетността. Тъй като подобрението може да обхваща различни аспекти, повишаването на ефикасността се приема като подобрение. Може ли използването на ресурсите да бъде по-добре организирано, за да се създаде резултат, или изразходвани ли са ресурсите по икономичен начин при създаването на резултата? Аналогът на ефикасност е производителност. Ефикасност се отнася до количеството ресурс на входа, което е необходимо за получаване на единица продукция, а производителността е мярка, която показва колко продукция може да бъде получена на

единица входящ ресурс. Ефикасността е променлива и следователно може да се увеличава или намалява. Повишаването на ефикасността може да бъде постигнато чрез няколко сценария, които са посочени на фигура 5. Първият квадрант показва повишаване на ефикасността, когато продукцията/резултатът нараства сравнително повече от ресурса на входа. По същия начин, продукцията/резултатът може да намалява относително по-малко от ресурса на входа, което също води до по-голяма ефикасност. Третата графика показва увеличаване на продукцията/резултата, съчетано с намаление на входящия ресурс. Накрая, както е показано на четвъртата графика в дясната част долу, продукцията/резултатът може остане на същото ниво, като намалява ресурса на входа, което също води и до по-голяма ефикасност.



Фигура 5: Повишване на ефикасността

Оценяването на критерия за ефикасност изисква информация за ресурсите на входа и резултатите на изхода, които са свързани с интервенцията на оценяваната политиката.

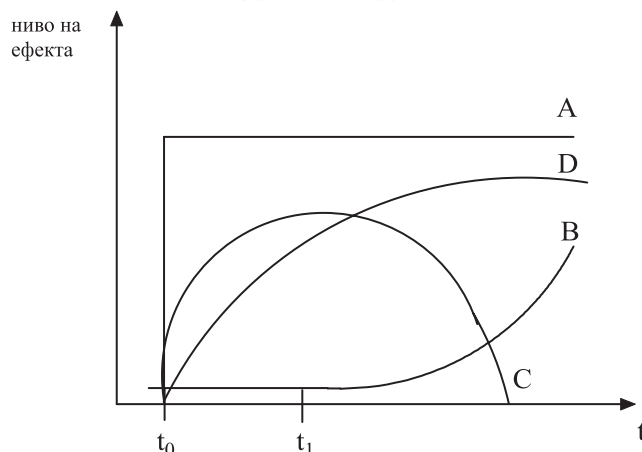
3.4. Устойчивост

Освен въпросите за постигането на целите, ефективността и ефикасността, устойчивостта на резултатите от политиката също е от значение. Често финансовите средства са ограничени по обем и време, а ние търсим трайно решение на проблемите. Една политика може да бъде поставена под въпрос за това дали дава устойчиви резултати. Във връзка с това е въпросът дали усилието на политика може да се ограничи, без да има въздействие върху наблюдаваната и постигната промяна. Например, субсидии могат да бъдат използвани от правителството за промяна на поведението на някаква целева група, например, при потреблението на енергия. Докато в краткосрочен план субсидиите може да убедят хората, въпросът е дали тази промяна в поведението ще продължи да расте, само ако се поддържа от субсидиите.

Въздействията се развиват като следват различни пътища и потоци. Следващата фигура 6 представя възможните еволюции на ефекта. Ефект А представлява

ва ефект, който се задейства веднага след началото на интервенцията и който остава на едно и също ниво. На практика това може да бъде броят на водачите на превозни средства, които спазват ограничението на скоростта след монтирането на камера по пътя.

Ефект В представлява ефект, който се изгражда известно време след началото на интервенцията; необходимо е време преди интервенцията да предизвика реакция сред целевата група, например, чрез подаване на заявления за субсидии. Ефект С се разгръща незабавно и затихва малко след края на интервенцията. Това е много обичайно за кампании за повишаване на информираността. Ефект D нараства и намалява постепенно след по-дълъг период от време. Такива, например, могат да бъдат последствията от политика, която има за цел рециклиране на отпадъци или пречисване на отпадъчни води.



Фигура 6: Ефекти и устойчивост

От тази графика става ясно, че подходящото време за оценка на устойчивостта на ефектите ще зависи от вида на очакваните ефекти. Някои ефекти могат да бъдат видими само след като прилаганата политика придобие известна зрялост.

3.5. Полезност

Друг критерий за оценка, който е от значение за ретроспективното оценяване е полезност. До каква степен предоставяната услуга отговаря на очакванията на гражданите и клиентите? Политиката може да се промени по време на изпълнението по начин, че да не отразява напълно политиката във вида, в който е била разработена на масата. За да разбереш какъв е пудингът, трябва да ядеш от него, затова оценяването може да провери дали приложената политика отговаря на планираната интервенция.

4. ВЪПРОСИ ЗА ОЦЕНКА

4.1. От логика на интервенцията и критерии към въпроси за оценка

Изборът и формулирането на критериите за оценка е от ключово значение за определяне на обхвата на оценяването. Вземането на решения относно въпросите за оценка е от решаващо значение, но не е лесна задача. Изборът трябва да бъде направен сред потенциално широк спектър от въпроси. Освен това, конкретната формулировка на въпроса е от решаващо значение. Една малко по-различна формулировка може да доведе до различен фокус и по този начин да наложи друг подход или различни техники в следващите стъпки на процеса на оценяване.

Преработената логика на интервенцията и избора на критерии за оценка ни води до въпросите за оценка. В съответствие с критериите за оценка, посочени в глава 3, следните общи въпроси са от значение в рамките на последващото оценяване:

- Постигнати ли са целите на политиката?
- Какви ефекти и промени могат да се наблюдават?
- Ефективна ли е политиката, какви са последствията от политиката (краткосрочни) и въздействието (в по-дългосрочен план)?
- Задействала ли е интервенцията странични ефекти?
- Устойчиви ли са ефектите от политиката?
- Има ли обществена подкрепа за приложената политика?

Тези въпроси, разбира се, трябва да бъдат „преведени“ или приложени към конкретната политика и логика на интервенцията, която е обект на оценяване.

4.2. Как да се изработят добре насочени въпроси за оценка

Изборът на въпросите за оценка може да следва подход на два етапа. В първи кръг на консултации със заинтересованите страни може да бъде съставен дълъг списък от въпроси и точки, заслужаващи внимание. Обикновено се разработва един доста широк спектър от потенциални въпроси. Във втория кръг трябва да се направи добър подбор, с консенсус между заинтересованите страни, които са поискали оценяването и / или които ще бъдат включени в последващите дейности в процеса на оценяване. Подборът на въпроси за оценка с участие ще допринесе за подпомагане на обхвата на оценяването. Разбира се, заинтересованите страни ще имат различни приоритети и власт. Важно е да се гарантира принос от всяка от избраните заинтересовани страни, които участват в структурирането на оценяването.

Освен процесът на подбор на въпроси за оценка, от съществено значение е и въпросът за качеството. Потенциалните въпроси за оценка трябва да бъдат подложени на проверка на качеството: какво представлява добре изработен въпрос за оценка? Следващият контролен списък предвижда някои ключови точки, на които трябва да се обърне внимание:

- 1) Отговорът ще даде информация на тези, които ще използват резултатите от оценяването.

- 2) Отговорът ще генерира нова информация и прозрения.
- 3) Отговорът ще предостави важна информация за вземане на решения по политиката.
- 4) Въпросът се вписва последователно и разбираемо в обхвата на оценяването и е в съответствие с избраните критерии за оценка.
- 5) Отговорът ще има въздействие върху политиката.
- 6) Осъществимо ли е от оперативна гледна точка (бюджет, персонал, време, опит) да се отговори на въпроса?
- 7) Възможно ли е да се отговори на въпроса в настоящия контекст на политиката (зрелост на политиката, наличието или достъпността на данни, чувствителност и т.н.)?

Всеки потенциален въпрос за оценка може да бъде преценен спрямо тези въпроси, като се вземат под внимание възможните предложения или забележки на заинтересованите страни, участващи в процеса на подбор. Последните два въпроса се отнасят до това, което се разбира под „оценимост“; оперативно и по същество възможно ли е да се отговори на предложения въпрос за оценка? Когато практически или контекстуални бариери могат да потискат отговора на въпрос, то точен и полезен отговор априори ще бъде невъзможен.

5. МЕТОДИ И ТЕХНИКИ ЗА ОЦЕНЯВАНЕ

5.1. Строги подходи

Строги подходи за повлияване на оценяването са подходи, които имат добре разбрана и приета основа за твърдения по отношение на причинно-следствените връзки. Прагматичните подходи се отклоняват от тези основи (те не могат да отстояват причинност), но са по-лесни за изпълнение и разбиране.

5.1.1. (Квази-) експериментални подходи

5.1.1.1. Основа на причина и следствие при тези подходи

Оценяването на „въздействие“ означава оценка на това дали интервенцията (като проект, регламент, съобщение и т.н.) може да се разглежда като (една от) причините за даден ефект, който представлява интерес (например, увеличаване броя на заетите, превенция на престъпност, увеличаване на данъчните приходи и т.н.). По същество сме заинтересовани от установяване на причинно-следствени връзки въз основа на доказателства.

Квази-експерименталните подходи, които разглеждаме по-долу принадлежат на областта на количествените изследователски методи. По традиция, причинно-следствените връзки са представени с променливи X и Y и посока на промяна (вж. фигура 7).

Стрелката показва връзката между X и Y . X се предполага да бъде причина за Y и оттам посоката на стрелката.

В количествените изследвания, X и Y са променливи, които може да имат различни стойности. Полът, например, е променлива, която може да има две стойности: мъжки или женски.

За интервенции, X обикновено означава наличие на интервенцията ($X = 1$) или отсъствие ($X = 0$). Възможните стойности на Y , променливата за резултата, зависят от това, което Y представлява, например Y се отнася до намиране на работа ($Y = 1$) в сравнение с ненамиране на работа ($Y = 0$), след участие в интервенцията.

Причинно-следствените връзки са изразени като тенденции в количествените изследвания: ако имаме $X = 1$ (наличие на интервенция) очакваме да има средно повече случаи на $Y = 1$ (хора на работа), отколкото ако $X = 0$ (отсъствие на интервенция).

Количествените изследвания изискват сравнително голям брой наблюдения на X и Y , тъй като е доста трудно да се посочат тенденции въз основа на само няколко наблюдения.

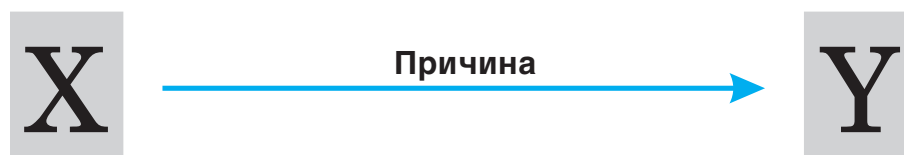
Представете си, че разполагаме с данни само за три безработни лица от 1000, които са участвали в проект ($X = 1$) за намиране на работа ($Y = 1$). Установяваме, че две лица са намерили работа и едно не е намерило. Но това не ни позволява за изразим тенденция за всички 1000 участници. Ако имаме информация за още един човек и установим, че този човек също не си е намерил работа, тенденцията става неутрална, а не в полза на намирането на работа. Броят на субектите в това, което е посочено като „извадка“, трябва да е достатъчно висок и необхо-

димият размер състав на извадката трябва да бъде предварително оценен от обучен статистик (като се използва процедура, посочена като „анализ на статистическата мощност“).

Въпреки това, ако голям брой наблюдения показват, че в действителност средно по-голям брой лица си намират работа ($Y = 1$), ако участват в X ($X = 1$) отколкото, ако не участват ($X = 0$), все още не ни позволява да бъдем сигурни, че X е причина за Y .

Защо не? Защото хората, които участват в или са изложени на интервенция (например, в проект за трудова заетост или кампания за обществено здраве) са различни от гледна точка на техните характеристики живеят при различни обстоятелства, което означава, че те също са засегнати и от други влияния, вън от интервенцията. Тези различни влияния, несвързани с интервенцията също биха могли да допринесат към наблюдението на стойността на Y (независимо дали е заетост, ниво на доходи или някаква друга „цел“).

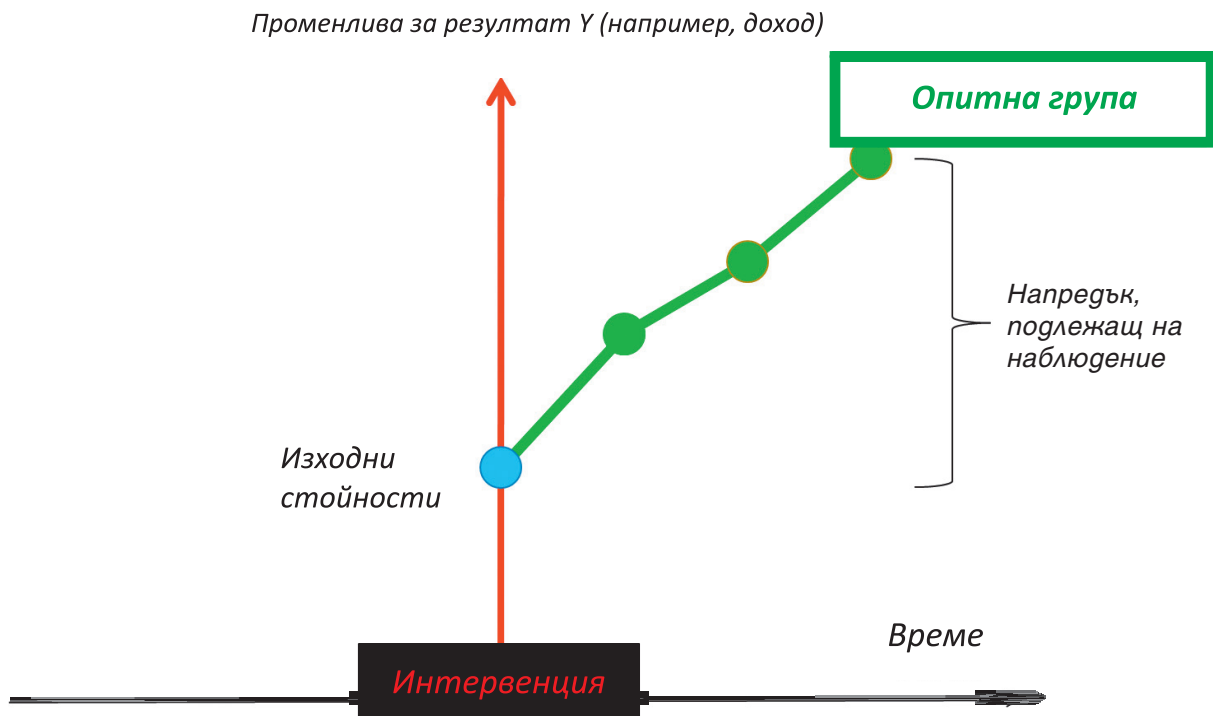
Това се нарича „отклонение в избора“. Представете си, че се случи хората, които участват в проект за заетост да бъдат като средно по-интелигентни, с по-добри кариери в предишните си работи и всички живеят в райони с повече възможности за работа, в сравнение с тези, които не са участвали в проекта. По същество това не надценява ли реалното влияние на проекта върху състоянието на заетостта на хората? Всъщност, можем да очакваме, че групите на участниците, поради техните благоприятни характеристики, ще имат по-добри резултати по отношение на трудовата заетост след определен период от време, в сравнение с неучастващите, дори и ако изобщо не е имало интервенция.



Фигура 7: Причинно-следствени връзки в количествените изследвания

За да се преодолее горепосоченото отклонение, ние се стремим да създадем „съпоставителна“ обстановка: трябва да бъдем в състояние да поставим една и съща група хора както в групата, участваща в интервенцията, така и в групата, която не е участвала в интервенцията, за да имаме надеждно сравнение. Очевидно е, че това не е възможно и може да се направи само приблизително. Хората не могат да бъдат I да не бъдат участници примерно в проект за заетост. Въпреки това, методите, които ще бъдат описани по-долу се опитват да създават група в интервенцията, обикновено наричана „опитна група“, и група извън интервенцията, наричана обикновено „контролна група“.

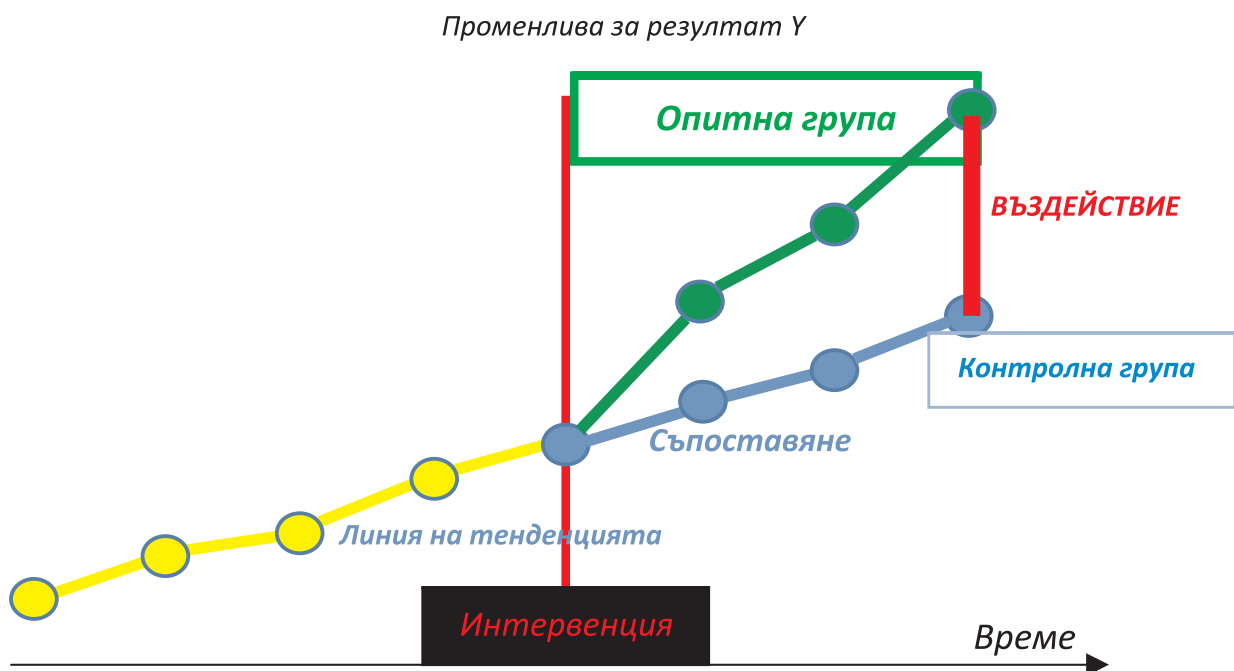
Фигура 8 показва, че измерването, например за участници в интервенцията, на средното равнище на доходите в началото на интервенцията и сравняването му със средното ниво след определен период от време, не дава добра представа за въздействието. То осигурява само „наблюдаван“ напредък. Но ние не знаем каква част от този напредък се дължи на това, че хората имат определени благоприятни характеристики и /или са част от благоприятна среда. Този забелязан напредък е възможно да е постигнат независимо и в отсъствие на интервенцията, просто защото тези хора ще са предприели определени стъпки за увеличаване



Фигура 8: Наблюдаван напредък

на доходите си, или защото икономиката е процъфтявала по време на наблюдения период.

Фигура 9 показва как можем да получим добра представа за въздействието чрез съпоставяне на две групи, които са идентични за всичко, с изключение, че едната група получава интервенцията (опитна група), а другата – не (контролна група). Тъй като изначално няма разлика между групите (те са „еквивалентни“), причината за забелязаната разлика в резултатите може да бъде приписана на интервенцията.



Фигура 9: Съпоставката позволява оценяване на въздействието

Фигура 9 показва, че средно, стойностите на променливата на резултата Y (например, ниво на доходи) са идентични и за двете групи. Едва след интервенцията стойностите започват да се отклоняват. Контролната група продължава да покачва, повече или по-малко продължава предишният тренд, докато опитната група изживява прекъсване на тенденция и покачва по-бързо. Въздействието е разликата в нивата на Y между двете еквивалентни групи, едната от които е била обект на интервенцията, а другата – не, след определен период от време.

Единствената разлика между различните методи, които ще бъдат описани по-долу, се отнася до КАК точно групите са направени еквивалентни и сравнявани. Идеалният тип е наречен рандомизирано контролирано проучване (RCT) или „истински“ експеримент.

5.1.1.2. Рандомизирани контролирани проучвания (RCT)

Рандомизираните контролирани проучвания се считат за „златен стандарт“ при проучванията за оценка на въздействието, тъй като предлагат най-добрия начин за изграждане на еквивалентни групи. По същество се избира популация от субекти (участници, целева аудитория, и т.н.), обект на интервенцията. От тях се прави достатъчно голяма извадка (въз основа на предварителен анализ на статистическа сила). Избраната извадка се разделя на случаен принцип на две отделни групи. Тъй като хората към всяка група се разпределят на случаен принцип, можем да предположим, че средно двете групи ще са еквивалентни (ако групите са достатъчно големи).

Запазване на заетостта в Обединеното кралство и демонстрация на напредъка

Демонстрационният проект на Обединеното кралство за запазване на заетостта и напредък (ERA) включва изследване последиците за дългосрочно безработните и икономически неактивните лица от продължаването на помощта и подкрепата, както и финансови стимули за тези, които са напуснали социалното подпомагане и са започнали работа. По този начин проектът ERA удължава подкрепата, осигурена чрез активни политики на пазара на труда за групите с ниски доходи, които работят.

Тези, които отговарят на условията за две от основните активни програми на пазара на труда във Великобритания в момента – Новият курс за дългосрочно безработните и Новият курс за самотни родители – са били разпределени на случаен принцип в опитни и контролни групи. Контролната група влезе в програмите за Новия курс както обикновено. Опитната група получи подкрепа преди наемане на работа (по подобен начин с контролната група), но продължи да получава консултации и помощ при напускане на социалното подпомагане и започване на работа. По време на проучването във Великобритания социалните помощи приключваха при започване на работа. Участниците също имаха право на различни видове финансова подкрепа и стимули за насърчаване на обучението и задържането на работа. Целта беше да се насърчат участниците да останат извън системата на социални помощи и да напредват като повишават доходите си и с други възможности и условия за заетост. Като цяло, около 16 000 лица са били

разпределени на случаен принцип в опитни и контролни групи в около петдесетина офиси на обществената служба по заетостта. Процесът на случайно разпределение дава възможност за опитни и контролни групи, които в момента на разпределение са много сходни. В резултат на това всички разлики между двете групи по ключови резултати, като започване на работа, доходи, часове и качество на работните места, отчетени след влизането в интервенцията, може с увереност да бъдат приписвани на намесата на ERA. Резултатите от проучването показват, че интервенцията е била особено успешна сред трайно безработните, като е спомогнала за повишаване както на равнището на заетостта, така и на доходите.

Вж.: <http://statistics.dwp.gov.uk/asd/asd5/rports2011-2012/rrep765.pdf>

Преди да се избере този подход, трябва да бъде съобразено дали:

- достатъчно големи извадки са осъществими на практика;
- структурата на оценяването трябва да бъде създадена **преди** началото на интервенцията, защото рандомизацията е нужна за определяне на състава на опитната група;
- искате да знаете само дали има въздействие или искате да се установи размера на въздействието, или дали въздействието варира за различните подгрупи (например, по-възрастни в сравнение с по-млади). В двата последни случая, методът изисква базисно измерване, което трябва да се планира преди започване на интервенцията.

5.1.1.3. Сравняване на оценките за предразположение (Propensity Score Matching)

Сравняване на оценките за предразположение (PSM) е метод, използван за случаи, когато не е възможно да се разпределят на случаен принцип субекти за опитни и контролни групи. Той предлага статистическа процедура за приближаване до рандомизирано разпределение.

Представете си, че вече сме избрали безработни участници в интервенция, която има за цел връщането им на работа. Също така, имаме достъп до доста информация за тези участници. Например, знаем, тяхната възраст, място на пребиваване, пол, тяхната предишна работа и т.н. Освен това, имаме достъп до същата информация за субекти, които **не са били** избрани за интервенцията, но по принцип са отговаряли на условията.

След това можем да разберем, чрез специфичен регресионен анализ, до каква степен изборът за участие в тази интервенция е свързан с характеристиките на лицата (например, ако сте мъж, е по-вероятно да бъдете в интервенцията, отколкото ако сте жена). Формата на такъв регресионен анализ е уравнение:

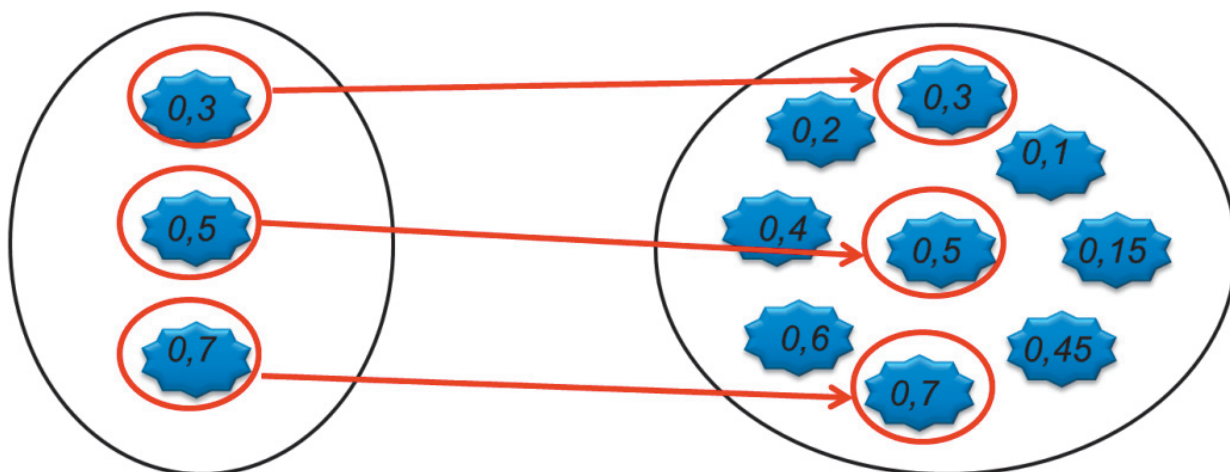
$$Y = v_0 + v_1 X_1 + v_2 X_2 + \dots + e$$

където X обхваща характеристики като пол, възраст и др., а Y се отнася до това дали лицето е в интервенцията (като $Y = 0 = \text{НЕ}$ и $Y = 1 = \text{ДА}$). Резултатът от такава регресия е например $Y = 0,1 + 0,3X_1 + 0,4X_2 \dots + e$ с прогнозни стойности за v_0 , v_1 , v_2 и т.н. Тези прогнозни стойности обозначават относителното влияние на свързаните променливи X_1 , X_2 и т.н. върху Y (лицето е в или извън интервенцията).

Регресионният анализ, който използвахме, за да разберем влиянието на характеристиките върху избирането на лицето, сега може да се използва за оценка на „предразположението“ за всеки човек, като въведем в уравнението точните характеристики за всяко лице. След това получаваме, например, за първото лице в нашия набор от данни, резултат $0,6 = 0,1 + 0,3 * 1 + 0,4 * 0,5...$ където 0,6 се отнася за 60% шанс за включване в множеството на участниците. За друго лице получаваме резултат $0,4 = 0,1 + 0,3 * 0,7 + 0,4 * 0,6 + ...$

С една дума, изчисляваме резултати за всички лица в нашата база данни както за тези, които вече са били избрани, така и за лицата, които не са били избрани за интервенцията.

Следващата стъпка е да се създаде група от неучастващи, която да бъде възможно най-близка до групата, която вече е избрана за участие. За това съществуват различни начини, но тук ще се спрем на най-простия подход, наричан „най-близко налично съвпадение“. Примерът по-долу показва как всеки избран за участник е съчетан с неизбран със същата оценка за предразположение.



Фигура 10: Най-близко налично съвпадение на оценките за предразположение

Ограниченията на този метод следва да бъдат ясни. Методът може да вземе под внимание само тези променливи, за които действително са налични данни. Ако например, се предполага, че възрастта е важна променлива при определяне на вероятността за избор за участие в интервенцията, но ние нямаме данни за нея, тогава възможното отклонение, произтичащо от разлики във възрастта между опитната и контролната група няма да бъде отчетено. При рандомизираните контролирани проучвания, рандомизацията гарантира, че всички променливи са взети под внимание и стойности на тези променливи ще бъдат разпределени средно по еквивалентен начин в опитната и контролната група. Това подчертава значението на доброто разбиране за това кои променливи са наистина важни за избор за участие в интервенция. Може да имаме данни за много от характеристиките на субекта, но да пренебрегнем това, което може да е абсолютно ключово (например, удобен обществен транспорт до мястото на интервенцията).

В заключение, условие за правилното използване на този метод е:

- наличие на колкото е възможно повече съответни данни относно характеристиките на субектите, свързани с избора за участие както за групата,

която вече е избрана, така и за групата, която не е избрана. В идеалния случай, лицата, които не са избрани, следва да отговарят на условията, но да не знаят за интервенцията или да са пропуснали възможността да кандидатстват;

- възможност да се направи достатъчно голяма извадка както за групата на вече избраните, така и за групата на неизбраните, с разбирането, че доста данни от групата на неизбраните няма да се окажат полезни, тъй като оценките им за предразположение изобщо не съвпадат с тези на участниците. Следователно трябва да има много повече налични данни за неучастниците отколкото за участниците;
- както при рандомизираните контролирани проучвания, ще бъде необходимо базисно измерване, ако искаме да се отчете размера на въздействието и/ или разликите по подгрупи.

Предимството на сравняване на оценките за предразположение (PSM) в сравнение с рандомизираните контролирани проучвания (RCT) е, че не се изисква произволно разпределяне на субекти преди започване на интервенцията. Следователно, този подход може да се използва дори ако интервенцията вече е започнала. Разбира се, все още се нуждаем от данни за изходното състояние, за да се оцени размерът на въздействието. Ако изходните данни не съществуват (например, данни за трудов статус обикновено са налични в съществуващите бази данни), пак ще трябва да се проведе базисно измерване както за групата на участниците, така и за контролната група преди започване на интервенцията, което означава, че методът губи предимството си като приложим дори след започване на интервенцията.

Ефект от ползването на вода от водопровод

Jalan и Ravallion провеждат оценка на въздействието, която измерва ефекта на достъп до вода от водопровод върху честотата и продължителността на диарията при децата под 5-годишна възраст в 16 щата в Индия. Изследването използва проучване на домакинствата, проведено през 1993–1994 г. и използва PSM да създаде съпоставими опитни и контролни домакинства от по-голямата извадка. За целите на изследването, променливите за предварителна експозиция (например, щата на пребиваване, състава на домакинството, активи, религия, достъп до обществени блага и характеристики на населените места) бяха включени в оценките за предразположение чрез модел на логит регресия. Опцията за петте най-близки съседни съвпадения е използвана за създаване на извадка за анализ на 33 000 наблюдения. Приблизително 650 домакинства са били изключени след процеса на сравняване, поради липса на достатъчно сходни домакинства. Авторите заключават, че достъпът до вода от водопровод намалява разпространението на болестта с 21% и продължителността на боледуване с 29%.

Източник: <http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-2664>

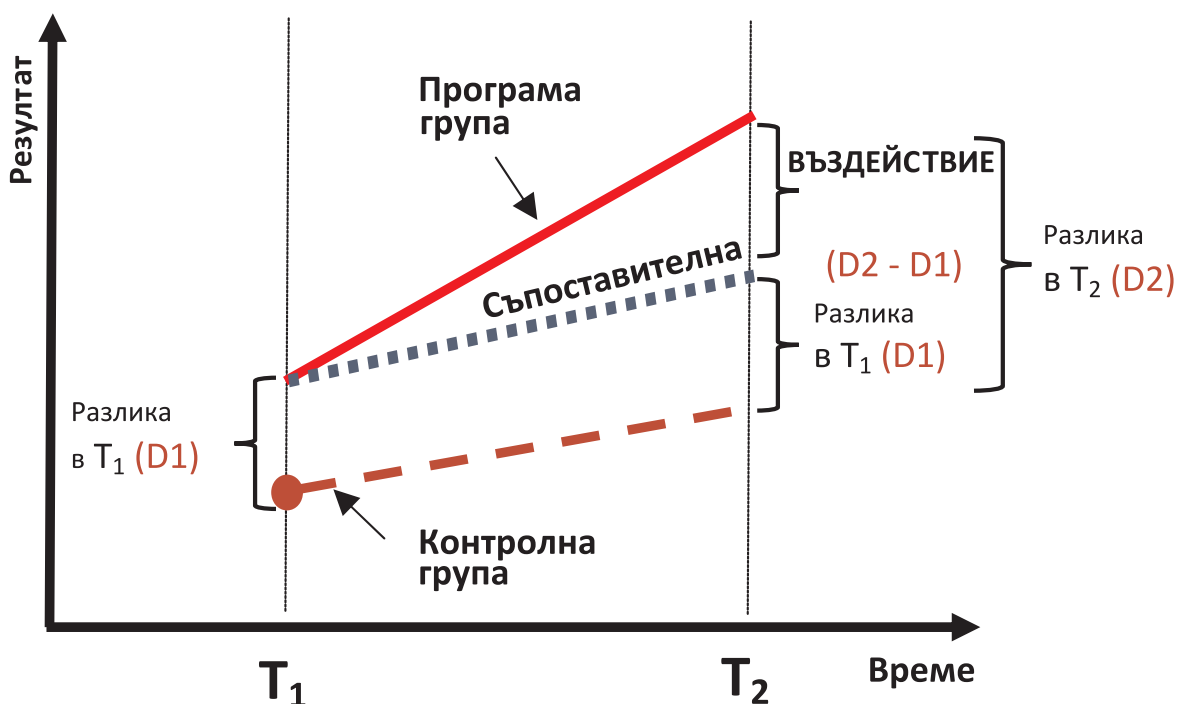
5.1.1.4. Разлика в разликата

Разлика в разликата е още един квази-експериментален метод. Този път, вместо създаване на две еквивалентни групи (една, която участва в интервенцията в сравнение с тази, която не участва), методът изрично признава разликите между двете групи преди интервенцията да се състои.

За да се отчетат разликите между двете групи, се провежда базово измерване на резултата, който представлява интерес. То обикновено показва, че участниците и неучастниците в действителност нямат един и същ начален резултат (на фигура 11 разликата е обозначена с D_1 в начален момент t_1). Затова се потвърждава, че двете групи не са еквивалентни, защото ако бяха, те би трябвало да имат статистически сходни резултати за изходното състояние.

При подхода разлика в разликата след това се приема, че при липса на интервенция, двете групи биха се развивали по идентична траектория. Това е отразено на фигура 11 от „съпоставителната линия“, която върви съвършено успоредно на линията, която показва напредъка на неучастниците. Това показва, че ако не е имало интервенция, бихме могли да очакваме, че в по-късен момент t_2 , разликата между двете групи пак ще бъде D_1 , тъй както тенденцията на променливата на резултата се предполага, че ще бъде еднаква.

Обаче, ако има последваща интервенция в момент t_2 , въздействието на тази интервенция се изчислява като се извади първоначалната разлика D_1 от наблюдаваната разлика D_2 .



Фигура 11: Разлика в разликата

Този подход има някои очевидни ограничения. Най-важното от тях е, че се приема, че и двете групи биха следвали обща тенденция при отсъствие на интервенция. Това предполага, че разликите между групите са „инвариантни по отношение на времето“, което означава, че се признава, че има първоначални разлики по отношение на индивидуалните и контекстуални характеристики на субектите и се признава също, че стойностите на тези характеристики могат да се проме-

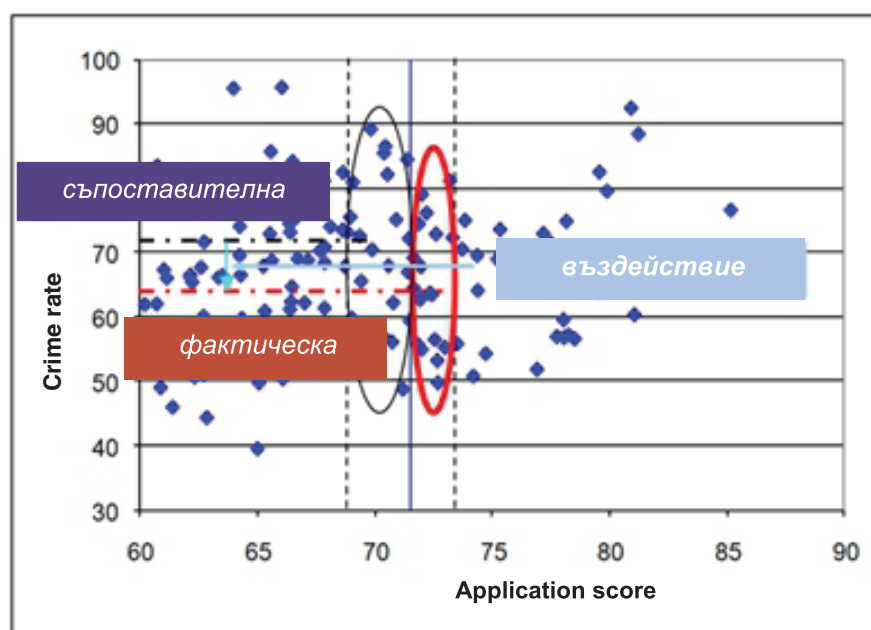
нят с течение на времето. Обаче промените в стойностите на тези характеристики биха възникнали по подобен начин и за двете групи, като по този начин се неутрализира всякакво влияние, което тези промени биха могли да имат. Например: при започване на интервенцията двете групи могат да имат различна средна възраст. Разбира се, с течение на времето хората и в двете групи ще остаряват, но разликата в средната възраст остава една и съща. Въпреки че това със сигурност важи за променливата за възраст, съществуват много други променливи, за които това предположение не е очевидно.

В сравнение с рандомизираните контролирани проучвания (RCT) и сравняването на оценките за предразположение, подходът на разликавразликата не се опитва да създаде еквивалентни групи, а по-скоро признава разликата с течение на времето при допускане на инвариантност по време.

В заключение, за да изберем метода разлика в разликата са ни необходими само изходни данни за достатъчно голяма извадка от участници и от неучастници. Трябва обаче да сме наясно, че в този метод съществува силно и понякога неоправдано допускане за инвариантност по време. Следователно, подходът на разликавразликата обикновено не се използва като самостоятелен метод за оценка на въздействието, а в комбинация със сравняване на оценките за предразположение.

5.1.1.5. Прекъсване на регресията

Прекъсване на регресия може да се използва, когато участието в интервенцията се определя от някаква прагова стойност в рамките на постоянна мярка. Например, достъпът до обучение може да се определи от резултата от изпит за пригодност, като тези, които са изкарали оценка над определения праг (или прекъсване) преминават през обучение, като в същото време тези, които са изкарали оценка под прага, не се обучават. Друг пример е използването на определен праг на оценка за селектиране на заявления от затворници за включване в програмата за изучаване на занаят.



Фигура 12: Прекъсване на регресията

Да вземем примера по-горе, когато заявление с резултат от 0–100 се използва като граничен праг. Ако имаме достатъчно голяма извадка, можем да считаме, че тези, попадащи в границите от 69 до 74 не са съществено различни.

Фигура 12 онагледява този хипотетичен пример, където всяка точка представлява човек. Граничен резултат от 72 разделя субектите на две групи: едната група, която не е включена в интервенцията и следователно представлява съпоставителната група (състояща се от субекти с резултат 69–72), спрямо групата, която е включена в интервенцията и следователно представлява фактическата група (състояща се от субекти с резултат 72–74).

Като вземем средната престъпност в някакъв определен срок след интервенцията за всички лица, за които имаме данни, можем да видим, че тази средна стойност за лицата с резултат от 72–74 (които поради това са включени в интервенцията) е по-ниска от средната за тези, които са получили 69–72 (които не са включени в интервенцията). Тази разлика отразява въздействието.

Съществуват по-сложни подходи, които използват регресионен анализ вместо сравняване на средни стойности, откъдето произхожда и терминът „прекъсване на регресия“.

Ограничението на този подход е, че само тези, които са точно над прага и тези, които са точно под прага, са достатъчно сходни и могат да бъдат сравнени. Няма начин да разберем въздействието на интервенцията за субекти, които са по-далеч от граничния резултат (прекъсване) от 69 (в ляво) и 72 (в дясно).

В заключение, изискванията за правилното използване на този метод са следните:

- непрекъснат индекс на допустимост (например, резултатите от тестовете, възраст и т.н.), за да се класира популацията;
- ясна прагова стойност, която определя допустимостта, която е част от структурата на програмата и по този начин е предварително зададена (и която не съвпада с причина, различна от интервенцията);
- достатъчно голяма извадка, за да се вземе предвид, че само данните около граничната стойност могат да се използват за оценка на въздействието;
- интерес към обобщаване само около граничната стойност.

Измерване на въздействието на Европейската регионална политика върху икономическия растеж

Оценителите са използвали подход на регресия с прекъсване, за да се оцени въздействието на регионалните фондове на ЕС върху икономическия растеж. Използвани са данни за периода 1995–2005 г., както и факта, че регионите на ЕС-15 са получили средства, ако техният БВП на глава от населението е по-малко от 75 на сто от средния за ЕС. Така класацията, използвана за определяне на опитната група е БВП на глава от населението, а точката на прекъсване или прагът е 75 на сто от средния за регионите на ЕС като цяло. Стратегията за идентификация се позовава на факта, че регионите в близост до точката на прекъсване от двете страни на прага, са сходни с изключение на това, че тези, които са били под прага са полу-

чили средства, докато тези, които са били над прага – не са я получили. Констатациите са, че регионалните фондове на ЕС имат малко положително въздействие върху икономическия растеж.
Източник: http://www.dps.tesoro.it/documentazione/uval/materiali_uval/european_regional_policy_Muval20.pdf

5.1.2. Проучване на казуси

5.1.2.1. Основа за причина и следствие в тези подходи

Проучването на казуси е различно от изброените (квази) експериментални подходи. Докато последните са насочени към наблюдения, свързани със „субекти“ (участници или „планирани цели“ на интервенцията, независимо дали това са лица, групи или организации), първите могат да бъдат насочени към „явление“, което представлява интерес, например, събитие, проект, процес, определен вид промяна, зависимост и т.н. Това „явление“ се нарича „единица за анализ“. Надеждните проучвания на казуси започва с идентифициране на единицата за анализ. Не е сигурно, че при извършването на оценяване „инициативата“ винаги е най-подходящата единица за анализ, както е видно от примера, представен по-долу.

Единица за анализ при оценяване на програма на Европейския социален фонд

В контекста на оценяването на програмата на Европейския социален фонд във Фландрия е проведена оценка на въздействието на една конкретна интервенция в рамките на програмата, наречена „процеси на личностно развитие“ (ПЛР). Проучването определя този ПЛР като „поддържащ процес за подобряване на ориентирано към пазара на труда личностно развитие на индивида“. ПЛР е отворен както за безработни, така и за работещи фламандски граждани, но ние тук се концентрираме само върху заетите лица. За тях ПЛР по същество означава, че по един ментор работи с всеки участник, за да им помогне да определят по какъв начин биха искали кариерата им да се развива в бъдеще и да предприемат необходимите действия, за да стане това реалност. Причината за подкрепа на тази интервенция с публични финанси е, че се приема, че участниците ще бъдат по-активни в изграждането на кариерата им, което ще ги направи по-самостоятелни и разчитащи на себе си пред лицето на нещастия като внезапни масови съкращения.

Единицата за анализ е процесът на личностно развитие, както е приложен конкретно от три различни проекти. Тъй като проектите са независими и отразяват различен контекст и целеви групи, се очаква, че проведените ПЛР ще бъдат различни, но все пак ще съответстват на цялостния модел на ПЛР. В рамките на всеки един от тези три проекта са проучени четири отделни ПЛР. Тъй като теорията за това как се предполага, че ПЛР би трябвало да работи се намира на нивото на отделните ПЛР, нивото на проекта не е правилното ниво за проучване на

тези ПЛР и затова, тези четири отделни процеси се разглеждат като вградени казуси (с поединици за анализ ПЛР, в рамките на обща единица за анализ на по-високо ниво, т.е на ниво организация). Въз основа на тези четири вградени ПЛР казуси могат да бъдат направени заключения за цялата организация (която съвпада с проекта), в която се случват. Следователно „казусът“ е „процес на личностно развитие в рамките на конкретна организационна среда“. Един проект може да изпълнява ПЛР в повече от една организация. В този случай, проектът вече не съвпада с казуса.

Източник: въз основа на проучване, поръчано от HIVA (Католически университет в Льовен) за агенцията на ЕСФ Фландрия (www.esf-agentschap.be)

Единицата за анализ в примера по-горе е свързана с процес на личностното развитие (ПЛР) в рамките на по-широк организационен контекст. Следователно, специфичният за организацията процес е единица за анализ и образува „казуса“.

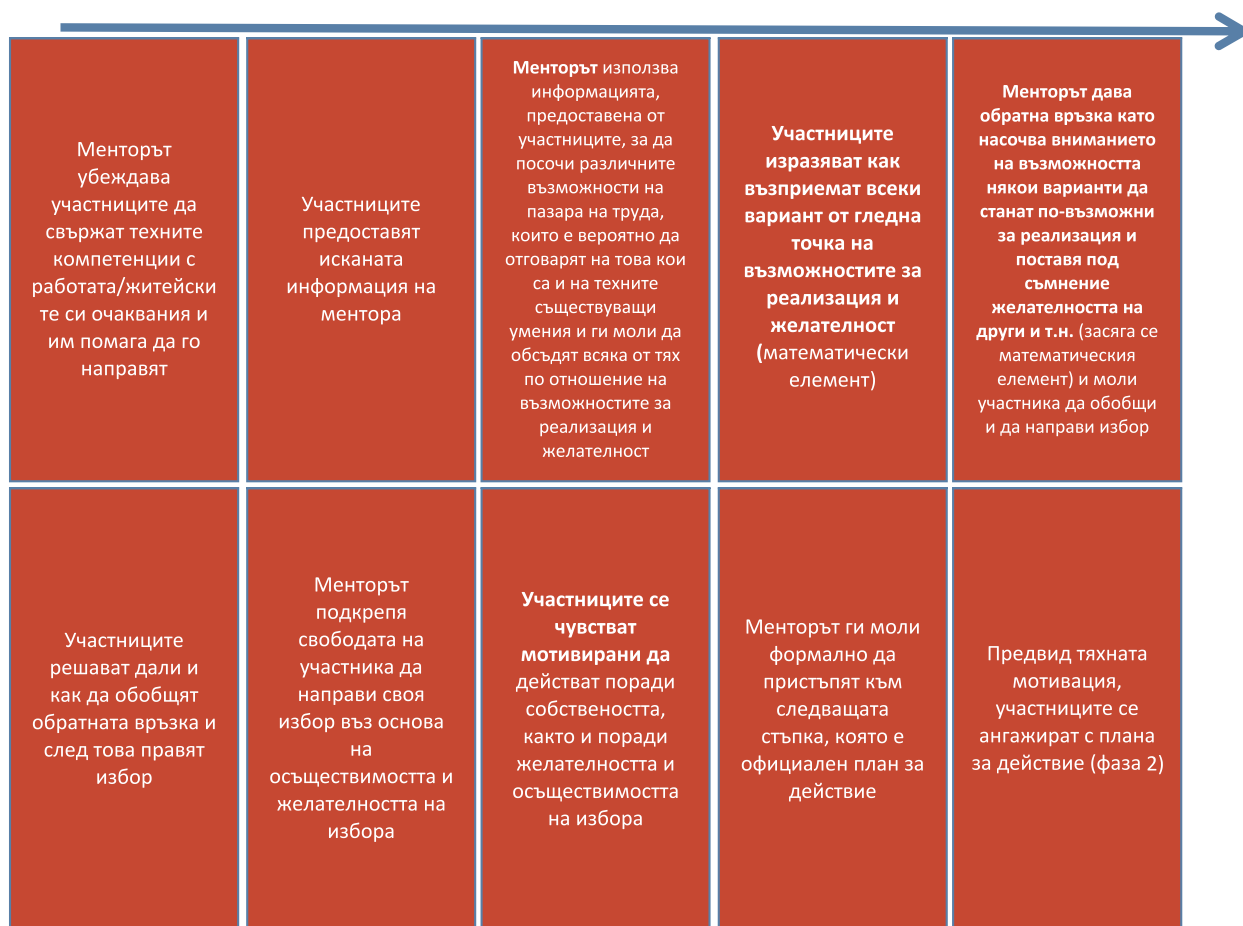
При проучване на единицата за анализ, проучването на казуси се основава на разнообразни методи за събиране на данни, вариращи от интервюта от отворен тип, фокус групи и наблюдения, до анкетни карти. За разлика от тях, (квази) експерименталните структури обикновено разчитат само на един количествен инструмент за измерване.

Съществуват няколко подхода за обосноваване на причинно-следствени връзки в проучването на казуси. Общото на всички подходи е, че причините са определени по отношение на достатъчност и необходимост или комбинация от двете. Достатъчност означава, че всеки път, когато има причина също присъства и следствие. Необходимост означава, че всеки път, когато има следствие, трябва да присъства причина. Нека да изясним това с един пример, свързан с тютюнопушенето. Пушенето не е нито достатъчно, нито необходимо условие за рак на белия дроб. Много пушачи никога не се разболяват от рак на белия дроб и не всички хора, които имат рак на белия дроб са пушачи. Въпреки това, пушенето в комбинация с редица други фактори (например, генетични) е достатъчно за рак на белия дроб, но не е необходимо, тъй като хората се разболяват от рак на белите дробове и по други причини, различни от пушенето (например, от дишане на замърсен въздух). Това прави тютюнопушенето недостатъчна, но необходима част от ненужна, но достатъчна причина за рак на белия дроб (състояние INUS). Тези понятия за причинно-следствената връзка са ключът към разбирането на проучването на казуси.

В допълнение е полезно да се въведе понятието за проследяване на процес, което използва идеята за непрекъсната верига на действие и реакция или „механизъм“, който свързва причината със следствието. Всъщност проследяването на процеса предполага, че всяка стъпка от процеса сама по себе си е недостатъчна, но необходима част от голяма, всеобхватна, но може би (не винаги) необходима или достатъчна причина. За изясняване на понятието, на Фигура 13 е използван същият пример за ПЛР.

Веригата на действие и реакция се отнася само за първата фаза на ПЛР, но тя може да се проучи като частична интервенция сама по себе си, с резултат ангажимент за Фаза 2.

Стъпка 2 „Участниците предоставят исканата информация на ментора“ е предизвикана от стъпка 1 и провокира следващата стъпка. Ако пропуснем тази стъпка и прескочим от стъпка 1 на стъпка 3, това ще представлява загадка: не би било



Фигура 13: Непрекъсната верига от действия и противодействия във Фаза 1 на процесите на личностното развитие (ПЛР)

ясно защо се случва този „скок“ от стъпка 1 на 3. В този смисъл, стъпка 2 е „необходима“. Това сочи към основната логика на проследяване на процеса: ако не можем да покажем, че например, стъпка 2 наистина присъства, тогава ние фалшифицираме целия механизъм, както е предположен от оценителите. В действителност, всяка стъпка е необходима, но недостатъчна част от целия механизъм. Ако не можем да намерим доказателства за всяка стъпка, тогава веригата на действие и реакция спира да бъде непрекъсната и следователно X (интервенцията) в началото на веригата не може да е довела до Y (резултата) в края на веригата.

В такъв случай механизмът ще трябва да се преосмисли (с изграждане на някаква нова теория) и изследването трябва да се проведе отново, за да обоснове всяка стъпка на новия механизъм.

Като алтернатива е възможно да намерим доказателства за всички тези стъпки. Все пак, това не означава, че целият механизъм е достатъчно обяснение за наблюдавания ефект. Може да са необходими допълнителни механизми. Също така, не можем да установим дали механизмът е необходим. Същият резултат може да се получи и без да е налице избраният механизъм.

За да сме в състояние да правим изявления относно необходимостта и достатъчността на механизма като цяло, можем да използваме Качествен сравнителен анализ, който изисква наличието на повече от един казус (вж. точка 5.1.2).

Анализът на съгласуваността е алтернативен метод за проучването на казуси, който обработва проследяването като използва идеята за „модел за съвпадение“.

Това налага намиране на доказателства за набор от хипотези, които се отнасят за повече от една конкуриращи се или алтернативно, допълващи се теории, които се опитват да обяснят как интервенцията може да произведе ефекта.

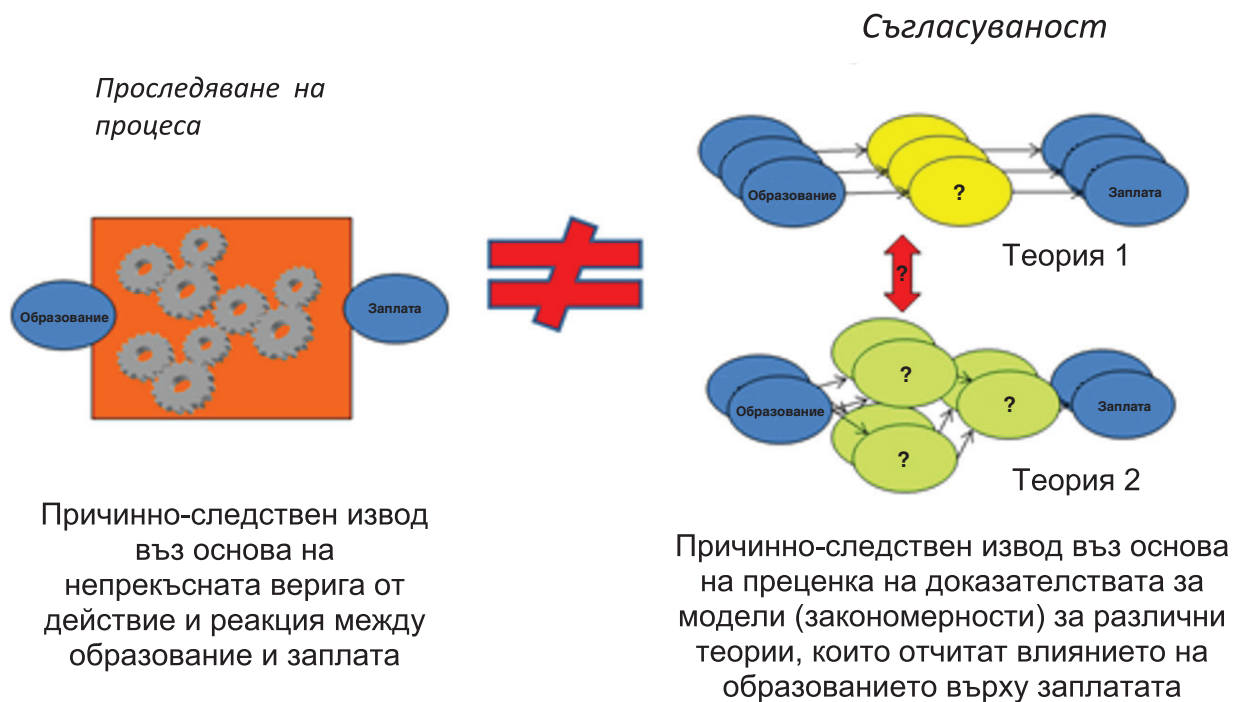
Например, както фигура 13 с непрекъснатата верига от действие и реакция, показва типичен подход за рационален избор (информацията се обработва и това води до рационално решение), може също така да бъде, че това, което е действително най-важно е, че ПЛР за първи път дава възможност на лица в организацията да получат действителен достъп до възможности за развитие, които те винаги са търсили и желали. В тази връзка, хората в организацията действително не вземат никакви решения за кариерата си, а просто преминават с ментора през процеса, който е до голяма степен безполезен, но е неизбежно препятствие, за да стигне до фаза 2 на ПЛД, където те получават достъп до истински и вече търсени възможности за развитие.

Начинът, по който се доказва причинно-следствената връзка при анализ на съгласуваността е по-лесен в сравнение с проследяване на процеса. Наборът от хипотези, които съставят една теория за това как интервенцията предизвиква ефект, не се нуждае от непрекъснатата верига от действие и реакция. По-скоро, съгласуваността работи с основните и най-вече различни очаквания (хипотези) по отношение на реалността, която очакваме да наблюдаваме във връзка с различни теории за това кое е причината за наблюдавания резултат. Например, съгласно теорията на рационалния избор, можем да очакваме, че процесът на разсъждение и вземане на решение е доста труден и изисква известно време. Но по силата на нашата алтернативна теория, където хората просто чакат възможност, бихме могли да очакваме, че процесът на цялата фаза 1 ще протече много бързо и с по-голяма лекота.

На практика, анализът на съгласуваност търси доказателства за всяка хипотеза във всяка теория и впоследствие оценява коя теория има най-голяма „тежест“ (конкуриращи се теории) или може би най-добре допълва друга теория (допълващи се теории). Ненамирането на доказателства само за една хипотеза сред многото не обезсилва теорията като цяло. Въпросът е по-скоро коя от теориите е сравнително по-обоснована от другите. Това означава, че не се прави преценка по отношение на отделна теория сама по себе си. Преценката е винаги по отношение на друга теория: дали една теория обяснява по-добре (конкурентни) или пък обяснява нещо друго (допълващи) в сравнение с друга теория. Докато по принцип потвърдената хипотеза е необходима за една теория, за да издържи, това което има значение в анализа на съгласуваността е дали сравнително повече (конкурентни или допълващи се) хипотези са подкрепени с доказателства, в едната или по-скоро в другата (конкурираща/ допълваща) теория.

За разлика от проследяването на процеса, анализът на съгласуваност изисква хипотезата да бъде подкрепена с няколко емпирични наблюдения, за да се различи „модела“, т.е. през процеса на ПЛР няколко от участниците следва да преминат по много бърз начин, което не може да се обясни с теорията на рационалния избор, но се вписва добре в алтернативното обяснение, че хората вече знаят какво искат, но просто се нуждаят от достъп до възможности за развитие. Един участник, който преминава бързо през процеса, не може да се разглежда като „модел“. Необходими са повече наблюдения, които да са в съответствие с тези теоретични очаквания.

В тази връзка, анализът на съгласуваност споделя изискването за закономерност на количествените подходи. Обратно, предимството на проследяването на процеса е, че може да се докаже, че дадена теория издържа (или не издържа), дори и ако се проследи само един човек. Разбира се, не можем да кажем за колко други участници тази теория е приложима (в други случаи друг механизъм може да обяснява резултата), но със сигурност знаем, че е приложима за един участник, за който сме събрали данни.



Фигура 14: Разлика между анализ на съгласуваност и проследяване на процеса

Анализът на съгласуваност, изобразен на фигура 14, не се опитва да разбере, например, как точно дадена образователна програма засяга заплата. Той само се опитва да намери повече от един случай (модел), потвърждаващ някои хипотези (със стрелките между мехурчета, изобразяващи логически връзки „ако-тогава“ между ключови хипотези), които или са свързани с теория 1, или алтернативно – с теория 2. Обратно, проследяване на процеса се съсредоточава върху това как точно възниква връзката между стъпките и следователно е достатъчно да наблюдава това само веднъж, за да докаже, че връзката, както е предположена на теория, действително съществува.

Следващите глави разработват в дълбочина различни подходи и се отнасят до вариации, които често се срещат в различни проучвания за оценяване.

5.1.2.2. Проследяване на процеса и реалистично оценяване

Както бе посочено по-горе, проследяването на процеса изисква, като първа стъпка, теория изобразена като механизъм, който е непрекъснатата верига от действие и реакция на действащи лица.

По-долу е даден друг пример, свързан с оценяване на бюджетното подпомагане на развиващите се страни.

X: Нефинансови приноси на бюджетна помощ, свързани с надзорните органи са изпълнени		
↓	↓	↓
Донорите и правителството установяват <i>политически диалог</i> , основан на доверие	Донорите прилагат условия за по-добро интегриране на надзорните органи в процеса на бюджетиране	Донорите предоставят техническа помощ на персонала на надзорните органи с цел подобряване на капацитета
↓	↓	↓
Правителството предоставя информация за бюджета (приходи, разходи) за диалог по въпроса как да подобри изпълнението на собствената си стратегия за намаляване на бедността	Правителството спазва изискванията и подобрява официалния статут на надзорните органи	Персоналът на надзорните органи възприема прилаганите мерки като полезни за укрепване на капацитета им
↓		
Информацията е споделена с различни заинтересовани страни (парламент, организации на гражданското общество, медии)		
↓		
→ Бюджетният процес е по-прозрачен (политически диалог), официалната роля на надзорните органи се подобрява (условия) и тяхната способност да контролират бюджетните процеси се увеличава (техническа помощ) № 1		
→ По-добри, официално интегрирани и обучени надзорни органи все повече искат информация по бюджетни въпроси от правителството № 2		
→ Правителството предоставя повече и по-висококачествена информация по бюджетните въпроси № 3		
→ Надзорните органи извършват ценен и надежден анализ на въпросите, свързани с бюджета и дават препоръки за политиката на правителството № 4		
→ Правителството реагира положително на тези препоръки и променя своята политика		
→ Връзките между държавни и контролни органи по отношение на бюджетния контрол са укрепени: Y		

Фигура 15: Механизъм за подпомагане от бюджета

Горният пример показва как различните участници действат и реагират един на друг. Сега за всяка стъпка в този механизъм трябва да се търсят доказателства, които да установят, че тези стъпки наистина се случват в реални условия. Също така е много важно да се помисли върху контекста, в който е възможно механизмът да функционира както е предвидено на теория. Всъщност, не всеки контекст е еднакво благоприятен за функциониране на механизма. Изискваните характеристики на контекста в подходите за проследяване на процеса се наричат „условия за обхват“. Например, за да функционира механизмът за бюджетна подкрепа (вж. фигура 15), ангажиментът на правителството за ефективно изпълнение на съответните политики и неговият политически и административен капацитет са важни предварителни условия. В случая с ПЛР се предполага, че участниците имат минимални умения за разсъждение и езикови умения, ниво на мотивация, доверие в менторите и др. Контекстът има значение, защото неблагоприятните контекстуални условия, могат да попречат на функционирането на очаквания причинно-следствен механизъм. Въпреки това, липсата на причинно-следствен механизъм не означава, че механизмът изобщо не функционира. Като се има

предвид наличието на подходящи контекстуални условия механизмът все пак може да се материализира.

При проучванията за оценяване откриваме подход, наречен „реалистично оценяване“, който споделя много от характеристиките на проследяване на процеса и може да се определи като негова ‚light‘ версия.

Реалистичното оценяване се съсредоточава върху намирането на доказателства за различните комбинации от Контекст–Механизъм–Резултат или КМР. По-долу ще открием някои комбинации на КМР, които се опитват да дадат обяснение на това, което монтирането на видеокамери за наблюдение (CCTV) провокира в различен контекст.

(1) КОНТЕКСТ Персоналът се страхува от крадци и няма увереност, че може да се справи с тях	+	МЕХАНИЗЪМ CCTV може да даде на персонала по-голяма увереност да следи за заподозрени	=	РЕЗУЛТАТ Приближавайки се до магазина крадците се възпират от кражба и това намалява кражбите като цяло
+ (2) КОНТЕКСТ Персоналът възприема CCTV сам по себе си да бъде ефективен срещу кражби от магазина	+	МЕХАНИЗЪМ CCTV може да намали бдителността на персонала, когато започне да разчита на него	=	РЕЗУЛТАТ Нивото на кражби се е увеличило тъй като наблюдението от страна на персонала е намаляло
(3) КОНТЕКСТ Удовлетворението на клиентите може да се увеличи	+	МЕХАНИЗЪМ CCTV се използва като инструмент за управление, за да се повиши удовлетвореността на клиентите	=	РЕЗУЛТАТ В резултат, повече клиенти посещават магазина и осигуряват естествено наблюдение
(4) КОНТЕКСТ Персоналът не може да забележи подозрително поведение	+	МЕХАНИЗЪМ Чрез наблюдение на мониторите на камерите за видеонаблюдение, служителите се разполагат ефективно в зони, където се наблюдава подозрително поведение	=	РЕЗУЛТАТ Персоналът действа възпиращо и може да задържи нарушителя, ако е необходимо

Фигура 16: CCTV КМР конфигурации

В примера по-горе са посочени различни обяснения на различни хипотетични резултати. Фокусът на проучването не се ограничава само до намаляване на кражбите в магазина, а отчита и други последици (например, привличане на клиенти, бдителност на служителите). Впоследствие се събират доказателства, за да се провери кои от тези КМР ще възникнат в действителност.

Контекстът представлява фактор, който влияе. Контекстуалните условия могат да позволяват или възпрепятстват възникването на различните механизми.

При реалистичното оценяване има четири вида контекстуални условия:

- индивидуални възможности на ключови фактори като интереси, нагласи, способности, достоверност и др .;
- междуличностни отношения в подкрепа интервенцията: правила за комуникация, управление и административна подкрепа, съюзни споразумения, професионални договори и др.;

- институционална среда: култура, лидерство и др. на прилагащата организация;
- широкообхватна (инфра-)структурна система и система за социално подпомагане като политическа подкрепа, наличие на средства и др.

В заключение за основните разлики между проследяването на процес и реалистично оценяване: проследяването на процес разглежда причинно-следствения механизъм и различните му стъпки по задълбочен и сложен начин. Разглеждат се малко на брой причинно-следствени механизми, но отделните стъпки в причинно-следствената верига са задълбочено проучени. Реалистичното оценяване поставя по-слаб акцент на непрекъснатата верига от очаквани действия и реакции. Причинно-следствената връзка се счита за даденост, но не е системно мотивирана. Въпреки тези различия, двата метода са сходни в подхода и предположенията.

5.1.2.3. Анализ на съгласуваност и анализ на продължението

В анализа на съгласуваност различни, конкуриращи се теории могат да се превърнат в категорични или да се операционализират чрез набор от хипотези. В примера с ПЛР, обсъден по-горе, дадена теория може да се превърне в следния набор от хипотези:

- силно мотивирани служители (СМС) в организацията, които искат да напреднат в кариерата и вече имат добра представа за това как искат да развият кариерата си, но не могат да продължат поради липса на структурирани възможности;
- ПЛР задейства организацията да създаде процеси на вътрешна мобилност;
- СМС бързо ще заявят доброволно участие в ПЛР, за да се възползват от тази възможност;
- СМС ще се придвижат по-бързо (в ПЛР) през фазата на размисъл относно това, което искат и техните компетенции, без да се налага голяма намеса от ментора;
- СМС ще изпълнят своите планове за действие по-систематично и по-бързо от другите служители;
- СМС ще реагират и ще кандидатстват по-бързо за нови или свободни позиции.

Алтернативен набор от хипотези, основани на конкурентна теория, би могъл да се състои от следното:

- Участниците в ПЛР получават по-добро разбиране за собствените си компетенции;
- Получават по-добро разбиране за собствените си интереси и за това какво ценят в работата;
- Подобряват разбирането си за възможно бъдещо кариерно развитие;
- Въз основа на направените изводи, участниците ще направят информиран избор относно проблемите на развитието, с които трябва да се справят;
- След това участниците ще изготвят планове за действие, които са насочени към идентифицираните проблеми на развитието;

- Участниците изпълняват тези планове за действие и придобиват или усъвършенстват необходимите компетенции;
- Участниците кандидатстват за други работни места или изпълняват по-добре настоящата си работа.

Вторият набор от хипотези отразява теорията на рационалния избор, при която съзнателно оценената информация определя действията. И двете теории и техните съответни хипотези отразяват различна логика, но не е задължително те да бъдат напълно противоположни. Някои хипотези могат да покриват частично или напълно различни елементи. Разбира се, основният въпрос е да се търсят доказателства и след това да се прецени коя от двете теории и техните съответни хипотези са потвърдени или отхвърлени от емпирични наблюдения.

Таблицата по-долу е полезен инструмент за тълкуване смисъла на наблюденията.

	Наблюдение(я) в съответствие с очакванията, изведени от теория В	Наблюдение(я) в противоречие с очакванията, изведени от теория В	Наблюдение(я) отвъд очакванията, изведени от теория В
Наблюдение(я) в съответствие с очакванията, изведени от теория А	<p>Заклучение А:</p> <p>Връзки към други наблюдения, необходими за дискриминационни доказателства</p>	<p>Заклучение В:</p> <p>Силно доказателство за предпочитане на А пред В</p>	<p>Заклучение С:</p> <p>Доказателството подчертава обяснителната сила на А</p>
Наблюдение(я) в противоречие с очакванията, изведени от теория А	<p>Заклучение D:</p> <p>Силно доказателство за предпочитане на В пред А</p>	<p>Заклучение Е:</p> <p>Силно доказателство за необходимост от други теории</p>	<p>Заклучение F:</p> <p>Доказателството подкопава обяснителната сила на А</p>
Наблюдение(я) отвъд очакванията, изведени от теория А	<p>Заклучение G:</p> <p>Доказателството подчертава обяснителната сила на В</p>	<p>Заклучение H:</p> <p>Доказателството подкопава обяснителната сила на В</p>	<p>Заклучение I:</p> <p>Доказателство за необходимост от разширени или други теории</p>

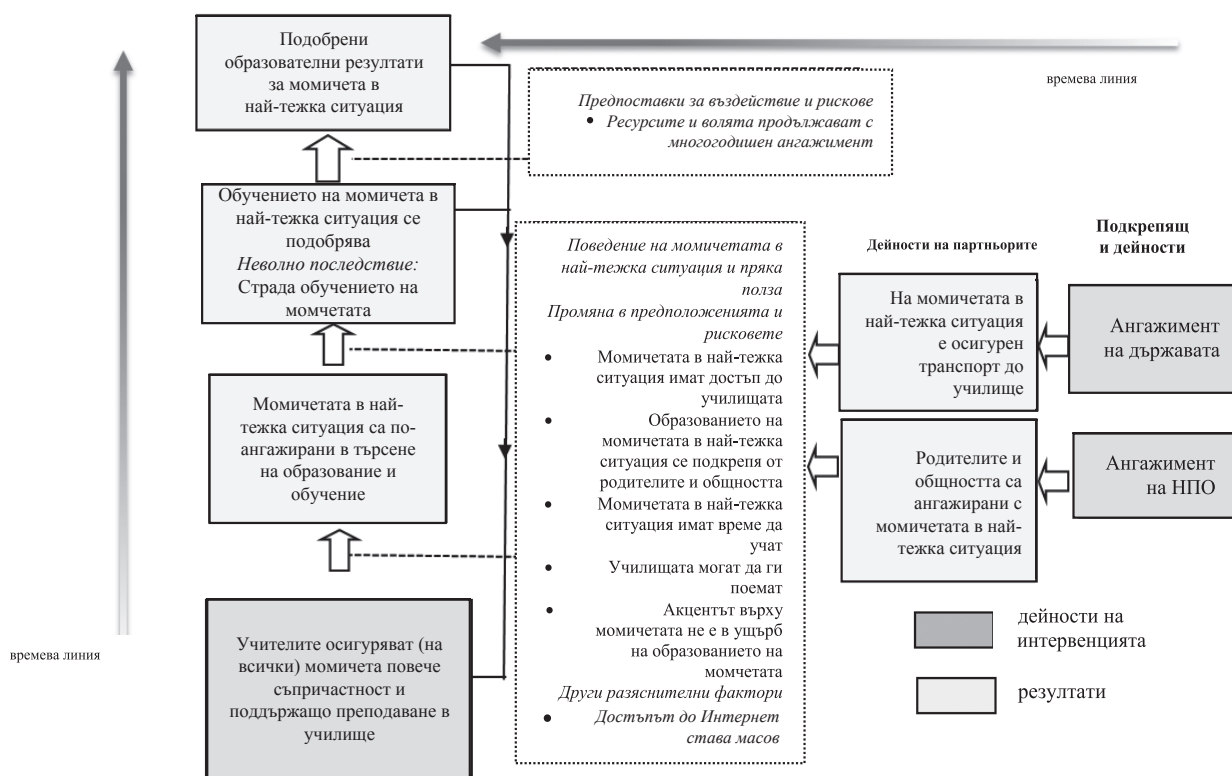
Фигура 17: Таблица за анализ на съгласуваност

Представете си, че открием потвърждение на факта, че много от участниците в проекта ПЛР преминават много бързо през процеса на размисъл. Този бърз процес на размисъл противоречи на това, което бихме очаквали по теорията на рационалния избор (теория В), но е в съответствие с обяснението за съперничеството (теория А). Следователно, имаме сериозни доказателства да предпочитаме втората пред първата теория (заклучение В в таблицата по-горе). Независимо от това, трябва да се подчертае, че преценката за относителната съгласуваност на теориите с наблюдения може да се направи единствено въз основа на доказателства, свързани с всички включени хипотези, а не само въз основа на една–единствена хипотеза.

Както и в случая с подхода при проследяването на процес, при анализа на съгласуваност контекстът също има значение. Определен контекст може да осигури по-стимулиращи условия „на простор“ в сравнение с други.

Един алтернативен подход, който често се цитира в проучвания за оценяване, е така нареченият „анализ на приноса“. Анализът на приноса има много общи характеристики с анализа на съгласуваността. Идеята е първо да се формулира „теория на промяната“ и да ѝ се припишат съответните рискове, свързани с нея.

Пример за това е даден по-долу от създателя на „анализ на приноса“, Джон Майн.

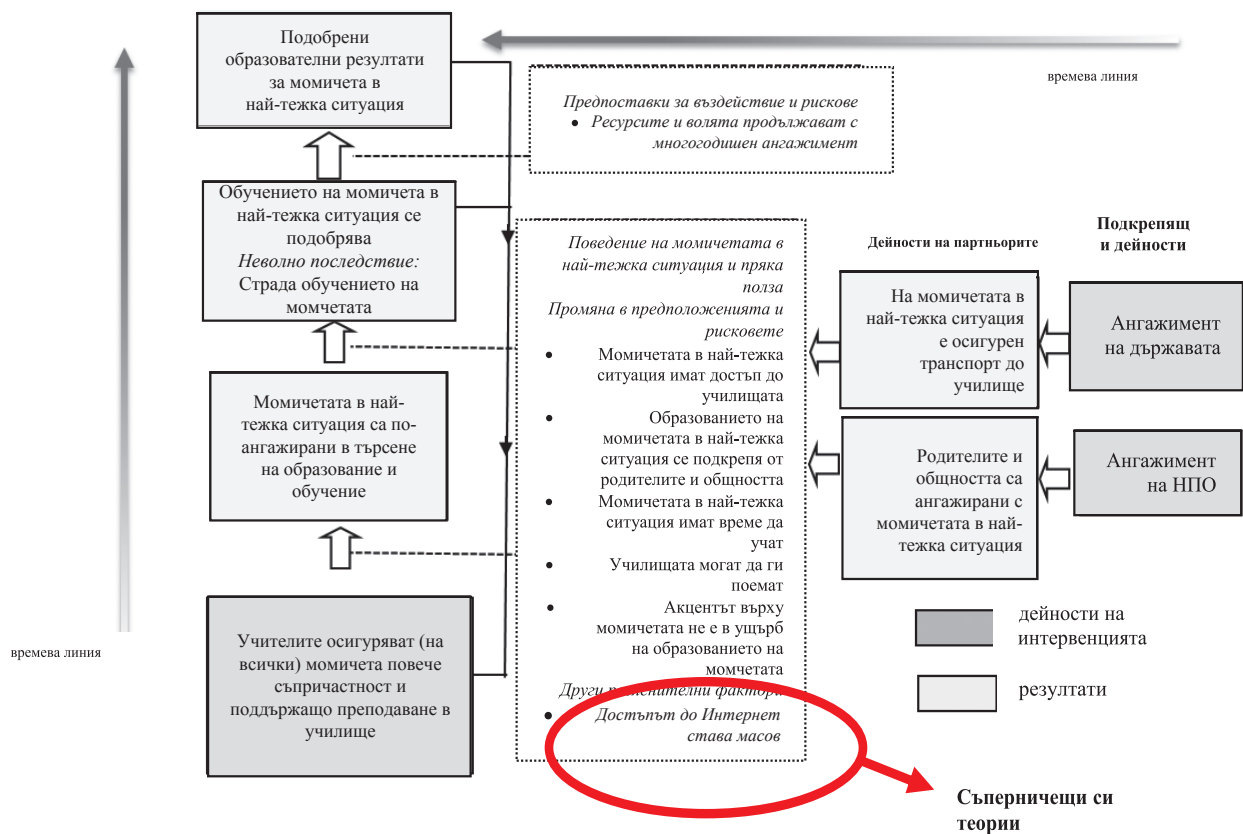


Фигура 18: Теория на промяната за анализ на приноса

Лявата страна на Фигура 18 изобразява теорията на промяната с поредица от изявления „Ако... Тогава“. Например: ако учителите осигуряват на всички момичета повече съпричастност и поддържащо преподаване в училище, тогава момичетата в най-тежка ситуация ще бъдат повече ангажирани в търсене на образование.

Това не представлява механизъм, еквивалентен на механизмите, обсъждани при проследяването на процес, тъй като все още има някои въпроси без отговор, за това как точно действията на учителите ще доведат до по-голяма ангажираност. Механизмът за проследяване на процеса би изяснил напълно причините за промени в поведението в непрекъсната верига от действия и реакции. Всъщност, множеството от изявленията в квадратчетата в лявата страна на фигура 18 съответстват повече на свободно свързания набор от хипотези, дискутирани в раздела за анализ на съгласуваността.

Анализът на приноса изисква последващо събиране на доказателства, за да се провери дали предполагаемите твърдения действително съществуват в реалността. Той налага също да се отдаде дължимото внимание на „съперническите теории“. Това е подобно на идеята за конкуриращите се или допълващите се теории в анализа на съгласуваността. На фигура 19 виждаме една съперническа теория под заглавието „други разяснителни фактори“, а именно наличието на достъп до интернет. Това може евентуално да повлияе за подобрене на ученето.



Фигура 19: Други разяснителни фактори в анализа на приноса

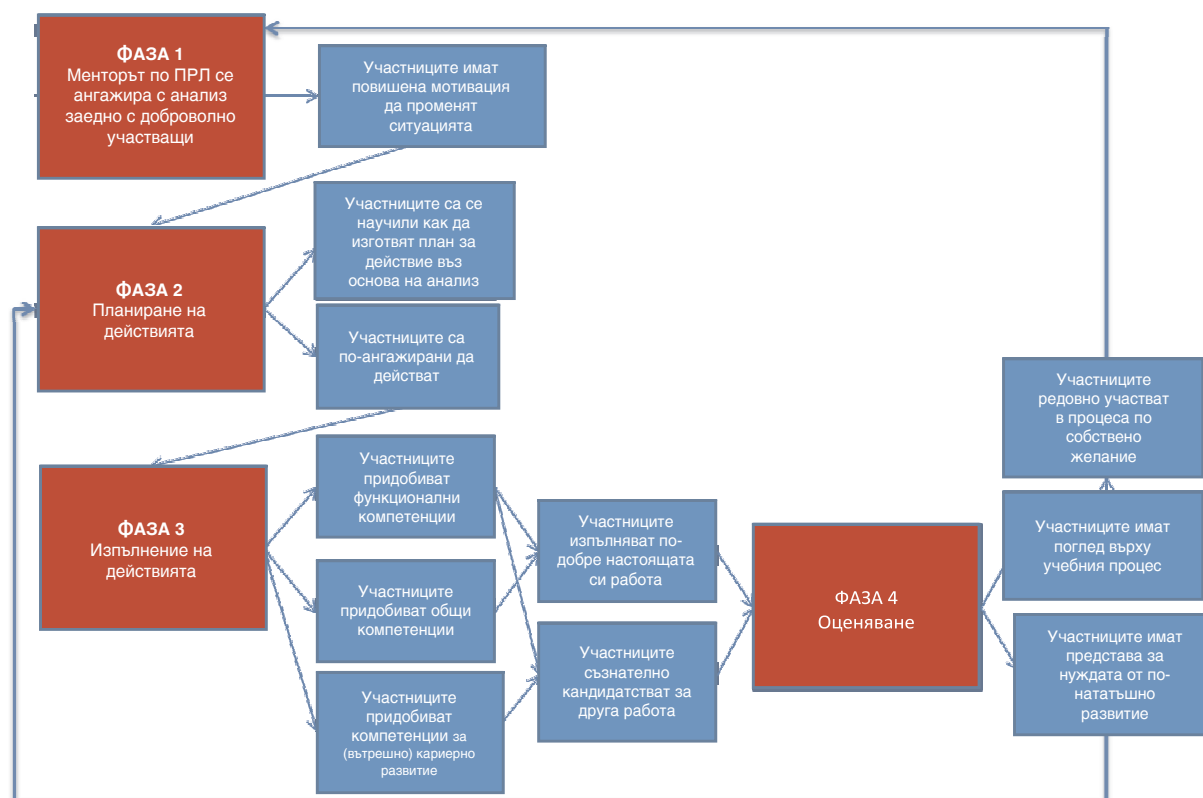
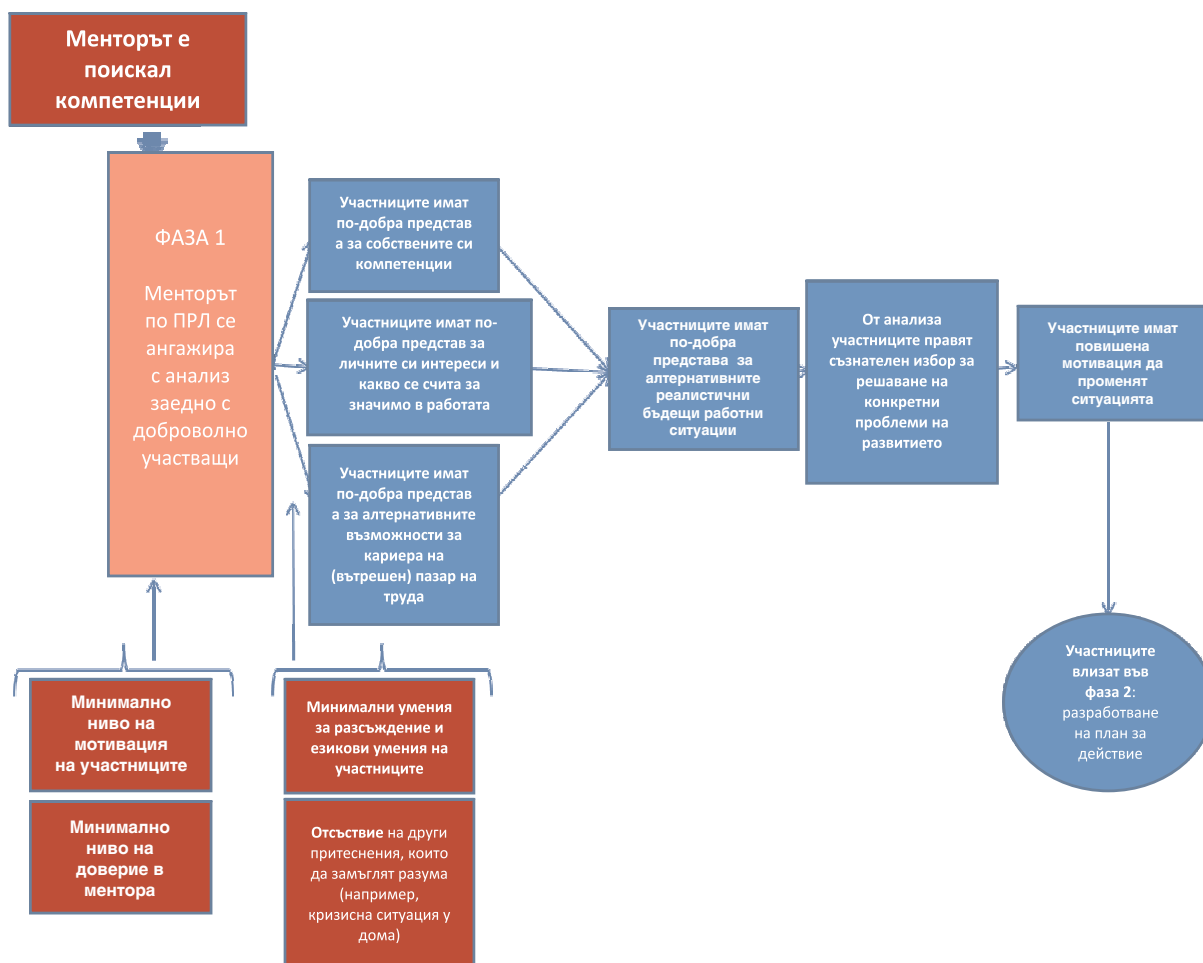
Основната разлика между анализа на приноса и анализа на съгласуваността е, че в последния са изследвани изцяло разработени съпернически си теории, а не само едно изявление като „наличие на широк достъп до интернет“ (вж. фигура 19).

Освен някои различия, двата подхода са доста сходни и тяхната комбинирана употреба може да укрепи качеството на анализа. Например, теорията на рационалния избор в ПЛР може да се изрази като „теория на промяната“, както е показано по-долу. Подобна теория на промяната може да се конструира като съперническа теория.

5.1.2.4. Качествен сравнителен анализ

Досега обсъдихме подходи, които са насочени върху конкретни казуси, например, процеса на личностно развитие ПЛР в контекста на определена организация и проект. Бихме могли да анализираме някои от тези казуси поотделно и да направим някои общи заключения от тях. Това води до въпроса как можем да направим заключение по отношение на голям брой казуси? Качественият сравнителен анализ (QCA) може да предложи полезно средство за агрегиране на изводите от няколко различни казуси.

QCA използва също разбирането за причинно-следствената връзка, чрез която причините може да се разглеждат като необходими или достатъчни, или и двете. Основен инструмент за установяване на причинно-следствени връзки е така наречената „таблица на истинност“. Пример за оценяване с използването на качествен сравнителен анализ и таблица на истинност, е представен по-долу.



Фигура 20: ПРЛ като теория на промяна

Идентификация	Институционален капацитет (ИК)	Мулти Донър Баскет Фонд (МДФФ)	Национално припознаване (собственост) (НС)	Успех (У)	Брой случаи
A	1	1	1	1	3
B	1	1	0	1	1
C	1	0	1	1	2
D	1	0	0	0	2
E	0	1	1	–	0
F	0	1	0	–	0
G	0	0	1	–	0
H	0	0	0	–	0

Фигура 21: Таблица на истинност при качествен сравнителен анализ

Източник: Befani, Between complexity and generalization (Между сложността и обобщението) Оценяване, 2013, 19, 3.

От лявата страна (Conf ID) са (видовете) казуси (маркирани от А до Н). Другите колони показват присъствието (обозначени с 1) или отсъствието (с 0) на три различни условия и резултати (успех). Виждаме, че възможните комбинации от условия Е до Н не се наблюдават или не са установени в действителност.

Първият въпрос би могъл да бъде: има ли условия, необходими или достатъчни сами по себе си за постигане успешен резултат? Институционалният капацитет (ИК) може да бъде необходим, тъй като всеки път, когато има успех, той също присъства. Но това не е достатъчно, тъй като знаем за (редица) случай (и), където има ИК, но не и успех (казус D). Фонд Multi-Донор Basket (MDBF) не е необходим (други причини могат също да доведат до успех), но може да бъде достатъчен, тъй като всеки път, когато присъства има и успешен резултат. Националната собственост (NO) може да бъде достатъчно условие, защото възниква за всеки случай с успех, но не е необходима причина.

Въпреки това, като разглеждаме отделните фактори, също има ограничения доколкото комбинация от фактори произвежда резултата, а не един-единствен фактор. QCA прилага Булева логика (има специфичен софтуер, който поддържа QCA), за да разберете кои комбинации произвеждат успеха или липсата му. В примера по-горе, отговорът е, че успехът се случва, когато са налице IC и MDBF или алтернативно, когато присъстват IC и NO. QCA показва, че различни комбинации от същите или различни условия, могат да доведат до същия резултат (това се нарича equifinality). Всъщност, ако отстраним IC от успешните случаи, може да се окаже, че нито един от случаите не води до успешен резултат. Няма начин да се провери значението на IC в нашия пример, тъй като нямаме случаи, когато в нашия масив от данни IC липсва.

Условията в QCA могат да бъдат както „контекстуални условия“ (това, което нарекохме „обхват“), така и въведени „причинни“ фактори като елементи на намета (например, интервенцията може да включва обучение, менторство, елемент на инфраструктурата и т.н.), които искаме да разглеждаме отделно и/или в комбинация с други елементи.

Въпреки че е мощен аналитичен инструмент, защото може да се справи с 5–50 казуси и широка гама от условия, качественият сравнителен анализ (QCA) не

трябва да се използва като самостоятелен метод за потвърждаване на причинно-следствената връзка. В идеалния случай, проследяване на процес трябва да се използва след QCA, за да се разбере какво точно се случва в казусите с комбинация от фактори, които водят до успешен резултат. Може също така да се приложи анализ на съгласуваност, тъй като конкурентни комбинации от условия, които обясняват резултата, биха могли да присъстват в един случай (както и в трите случая, етикетирани като „Conf ID A“ в таблицата по-горе).

От гледна точка на ограниченията, трябва да е ясно, че QCA е много чувствителен към пълнотата на съответните условия, както и към начина, по който се определя отсъствието или присъствието на условие.

5.2. Прагматични подходи

Следните методи са подходи, които не могат да се считат за действително оценяване на „въздействие“, но все пак често се използват под този знак.

5.2.1. Наблюдение на изпълнението Performance monitoring

Може би най-често използваният подход е измерването на изпълнението на базата на показатели за резултатите. Например, ако искаме да върнем безработните участници на работа, след известен предварително определен период от време (например, шест месеца след напускане на програмата), се прави преглед на статуса им на заетост (планираният резултат е, че те са намерили работа).

Въпреки това, само наблюдение на промяната чрез показател не предлага средство, за да се докаже, че тази промяна не се дължи на характеристиките на субектите или техния контекст, а на интервенцията. Това беше обяснено в предишния раздел за основата на причината и следствието при (квази) експериментални подходи.

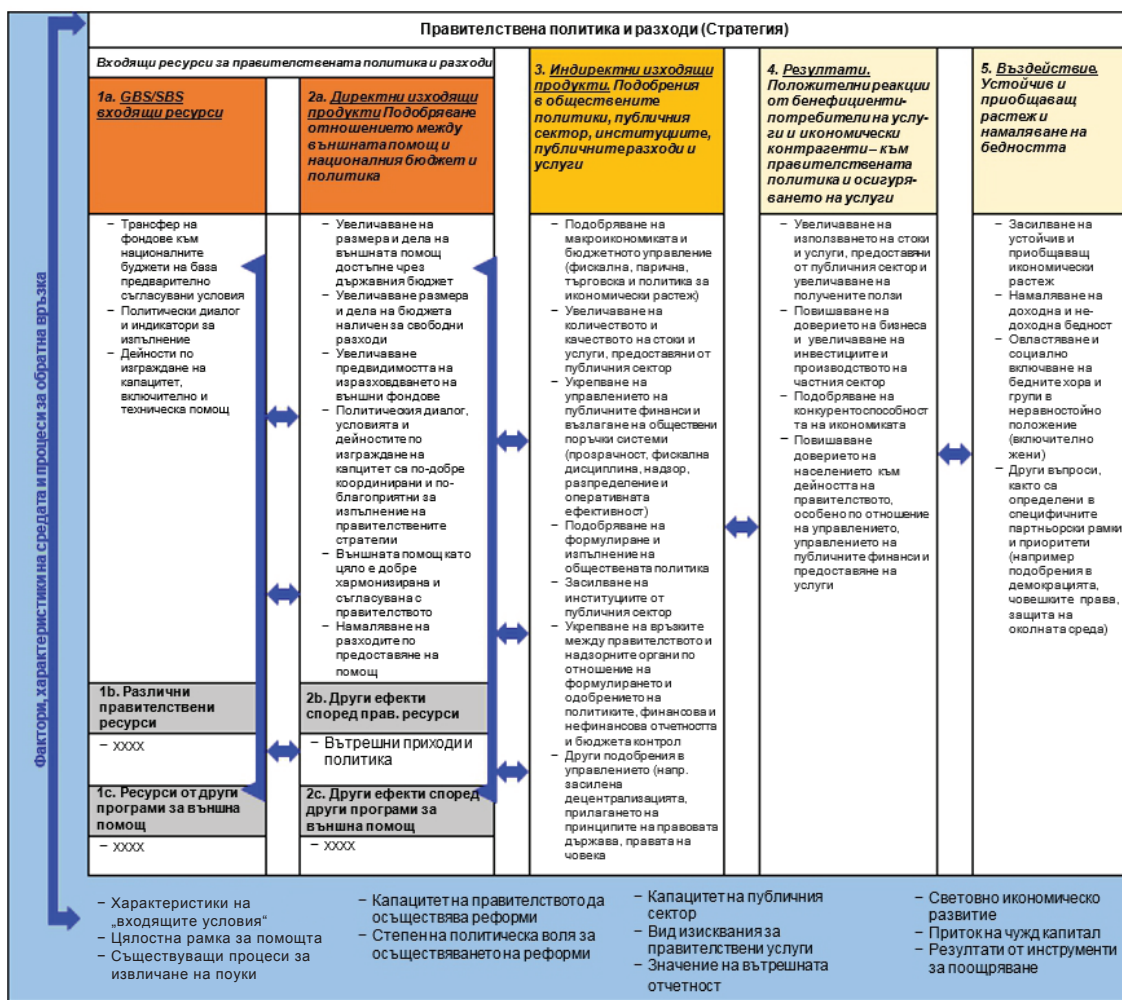
Такъв подход не се окачествява като оценяване на въздействието, а по-скоро като наблюдение. Той е широко разпространена практика, защото е лесен за изпълнение, но не предоставя достатъчно доказателства, за да се посочи, че дадена промяна се дължи на конкретна интервенция.

5.2.2. Теория на промяната като рамка за данни

По-усъвършенстван подход от простото наблюдение на индикатор за резултат е да се използва теорията на промяната като рамка за организиране на данните. Фигура 22 по-долу показва по-широката рамка, изложена от ОИСП за това как да се оценява „бюджетна помощ“.

Фигура 22 показва как „теорията“ (за това как се приема, че трябва да работи бюджетната помощ), се състои от цели, които са свързани с входни ресурси, резултати, последствия и въздействия.

Очевидно е, че за всички изброени цели може да се събере различна, обикновено вече съществуваща или лесна за събиране информация. Най-често са представени обобщени данни, което не позволява разграничаване между разликите в основните контексти и свързаните казуси. Това означава, че се прави присвояване на това, което в предишните глави нарекохме „условия за обхват“.



Фигура 22: Рамка на ОИСП за оценяване на бюджетна помощ

Източник: ОИСП /КПР, рамка за комплексна оценка за бюджетна помощ

Също така, обикновено не се разглеждат „съперничещи си теории“. Това може да е важно, не забравяйте, например, гореспоменатата програма за образование на момичета, където подобренията в обучението и ученето бяха може би (също или изключително) поради наличието на електронно обучение (вж. фигура 19). Объркващо, много източници наричат този вид подход оценяване на въздействие, въпреки че в действителност не трябва да бъде описван като нещо повече от „засилено наблюдение“, защото терминът „въздействие“ се използва също за обозначаване на цели по-високи на ниво, както на фигура 22, за които интервенцията обикновено може само да допринесе. Въпреки това, използването на термина „въздействие“ е много по-различно от понятието, което обсъждахме до този момент, където „въздействие“ се отнася за обосноваване на причинно-следствена връзка.

По-лесно е да се прилагат подобни прагматични подходи в сравнение с по-строгите такива, разгледани в предишните глави. Методът за проследяване на процес ще се съсредоточи само върху два от резултатите на фигура 22 (например, тези, маркирани с кръг) и взаимната им връзка. Освен това, връзките трябва да бъдат установени като ясен механизъм стъпка по стъпка, показващ как се очаква да се материализират резултатите (вж., например, фигура 15 за механизъм за бюджетна помощ). Установяването на подробно описан механизъм се изисква за всеки оценяван резултат.

6. ИЗТОЧНИЦИ И СЪБИРАНЕ НА ДАННИ

Следващата важна стъпка в процеса на оценяване се отнася до събирането на данни. Ръководните въпроси са:

- Кои данни и информация са необходими?
- Кои източници са на разположение?
- Кои методи и техники са възможни?
- Какво измерение във времето и честота на събиране на данни са от значение?

В случай на добре структуриран план за оценяване, отговорът на първия въпрос може да бъде получен от обхвата на оценката: избраните въпроси и критерии за оценка. Впоследствие може да се отговори на останалите въпроси. Оценяването може да бъде захранвано от два вида данни: първични и вторични. Вторичните данни съществуват още от началото на оценяването. Първичните данни са нови данни, които се събират в контекста на оценяването.

6.1. Първични източници

Провеждането на оценяване често предполага събирането на нови или допълнителни данни, затова оценителите често ще трябва да използват първични източници.

Съществуват няколко източника, които са потенциално интересни за оценителя: официални документи, свързани с политиката или нормативни разпоредби в процес на оценяване, и заинтересованите страни в широкия смисъл на думата. Последната група може да бъде разделена на цяла гама от действащи лица: създатели на политики (политици), държавни служители, участващи в разработването на политиката, бюрократи на улично ниво, целевата група (граждани, фирми, други държавни органи и т.н.), трети страни, данъкоплатеца, експерти по политиката (академичната общност, частни мозъчни тръстове) и т.н.

Събирането на данни може се съсредоточи върху вербалните и невербалните актове. Вербалните актове съдържат мнения, идеи, нагласи, чувства и възприятия. Те не могат да се наблюдават пряко и информацията трябва да бъде събрана чрез запитване. Невербалните актове обхващат поведение, което може да се наблюдава.

Вербални актове могат да бъдат наблюдавани чрез устно или писмено запитване (уеб-базирано, лице в лице, по телефона, по пощата и т.н.). Подходящи техники са интервюта, фокус групи, анализ на документи и писмени въпросници.

За постигане на точни и прецизни констатации от оценките, запитванията сред избрани свидетели могат да бъдат много полезни, тъй като може да се очаква да бъдат добре информирани и да осигурят представителен преглед на събитията. В зависимост от контекста на анализа и нужната информация, необходимостта от добре насочени и точни данни, може да бъде разгледана по количествен или качествен начин. От количествена гледна точка, при създаването на извадка с единици за анализ (като лица, образувания, географски райони и т.н.) трябва да се вземат под внимание редица условия. На първо място, хетерогенността на „популацията“ (т.е. общият набор от потенциални единици за анализ) влияе на необходимия размер на извадката от данни, за да бъде тя представителна за съот-

ветната популация. Колкото по-разнородна е популацията, толкова по-голяма трябва да бъде извадката, за да се получи същата точност със същата степен на надеждност.

Надеждността се отнася до трудността да се гарантира, че заключенията на базата на данни от извадката важат също за цялата популация. Статистическата надеждност се посочва с помощта на така наречения доверителен интервал. Това е прогноза за стойността на параметър (променлива), определен от интервал с минимална и максимална стойност. Въз основа на извадката, се определя доверителен интервал, което означава, че се предполага, че реалната стойност на дадена променлива (в цялата популация), попада в стойностите на интервала (напр. преценка, че средната височина на популацията на българските възрастни попада в интервала 160–185 см). Въпреки това, изводът, че реалната стойност за популацията е в диапазона от стойностите на интервала, е придружен с вероятност. Общата вероятност в статистиката е 95%. Това означава, че ако направим няколко извадки (с еднакъв размер) и повторим тези извадки отново и отново, в 95% от случаите истинската стойност на параметъра (т.е. стойността му за популацията) попада в зададения интервал (напр. 160–185 см). Въпреки това 5% от извадките ще генерират интервали, които не съдържат истинската стойност за популацията (или действителната средна височина на възрастното население). Следователно съществува вероятност от 5% да направим погрешни заключения. Целевото равнище за надеждност може да се определи по-високо, например 99%, но това намалява точността, тъй като обикновено води до по-широк диапазон от стойности или по-широк доверителен интервал (например, интервал от 145–200 см би предложил по-висока ниво на доверие, но е по-неточен). Степента на точност се отнася до ширината на интервала, в който се намира резултатът за дадена променлива. Нивото на точност може да бъде определено от заинтересовани страни, участващи в оценяването. Въпреки това е важно да се помни, че по-голямата точност (например, по-малък интервал) често предполага по-голяма вероятност от грешки (например, интервал от 170–180 см, с 80% ниво на доверие), както и обратното. Един от възможните начини за увеличаване на точността е да се увеличи размерът на извадката. По-голямата извадка ще предложи по-добро отразяване на общата популация.

Накрая, размерът на извадката се определя и от броя на променливите, които ще бъдат анализирани, както и броя на категориите за всяка променлива. Например, ако вероятността за отношението (положително/ отрицателно) към задължителното ползване на автомобилни предпазни колани, се определя от степента на образование (ниска/ средна/ висока), възраст (–45 /+45), скорост (ниска/ висока) можем да предвидим, че са възможни 24 комбинации от променлива/ категории (напр. $2 \times 3 \times 2 \times 2 = 24$). После можем да организираме най-различни комбинации в таблица с 24 клетки. Като основно статистическо правило обичайната практика е да има най-малко 20 единици за наблюдение (напр. лица) на клетка. Надеждният размер на извадката в този пример ще изиска 480 лица (напр. $24 \times 20 = 480$).

Съществуват няколко варианта на създаване на случайни извадки. Например, може да се следва процедура на два етапа, като най-напред се раздели популацията на подкатегории въз основа на характеристика, свързана с анализа (в зависимост от пол, възраст, степен на образование и т.н.), последвано от извадка във всяка подкатегория.

Докато количествената извадка се ръководи от опасения за представителност, за да може да разшири заключенията въз основа на извадката върху цялата популация, качествените подходи за създаване на извадка акцентират върху придобиването на задълбочен поглед и разбиране на механизми, ситуации или развития.

Качествени казуси и извадки се избират въз основа на конкретните въпроси, които се контролират внимателно. Подходите могат да търсят да изберат най-сходни казуси, например, за да се разбере защо въпреки многото прилики казусите произвеждат различни резултати. Като алтернатива, могат да се избират най-различаващите се казуси, за да се намерят общи елементи или да се разбере защо много различни казуси все пак произвеждат сходни резултати.

За качествените извадки не съществуват ясни принципи за определяне на размера на извадката. Изискванията за извадката зависят от потребностите от информация, целта на събирането на данни, съответните въпроси за оценка, както и от това, което се счита за значимо и осъществимо в рамките на времевите и бюджетни ограничения на оценяването. Например, при провеждане на интервюта с определен вид заинтересовани страни по една и съща тема, постепенно диалозите ще разкриват все по-малко нови и актуални проблеми. Провеждането на интервюта може да стигне до момент, в който получаваната информация става повтаряща се и допълнителните интервюта имат малко смисъл. Вместо това, може да бъде полезно да се проведат допълнителни интервюта с други видове заинтересовани страни, за да се обхване многообразието от гледни точки и виждания по темата.

Ще стане ясно, че количествената и качествената извадка обслужват различни стремежи. Изборът между подходите трябва да бъде направен в светлината на избраните въпроси за оценка. Освен това и вън от избора между двата подхода към извадката, оценителят винаги ще трябва да се справя с компромиса между широтата и задълбочеността на анализа.

6.2. Вторични източници

В допълнение към данните, събирани от първични източници, оценителят може да намери съответната информация във вторични източници. Данните, получени от системи за наблюдение на организациите, участващи в политиката, която е предмет на оценяване, могат да предоставят полезна и систематизирана информация. Системите за мониторинг, с показатели както градивни елементи, могат да са съсредоточени върху входните ресурси, резултати, последствия, въздействие и контекстуални фактори, както и върху процеси (например, пропускателна способност). За да бъде от значение за целите на оценяването, фокусът на системата за наблюдение трябва да съвпада с фокуса на оценяването. Информацията от наблюдението често се представя в отчетните доклади за управление и политическите документи; достъпът до данни (база данни) очевидно е важна предпоставка за използването им в оценяването.

Друг вторичен източник на данни са публичните статистики и бази данни. Институции като Евростат, Националният статистически институт или изследователски отдели на отраслови организации обръщат внимание на надеждни данни, стандартизирано измерване и допълнителни контроли за качество. Отново оценителят трябва да прецени до каква степен статистическите данни са от значение за процеса на оценяване. Нива на детайлност, възможности за обобщаване и раз-

биване, както и времевият и географският обхват на измерванията са важни параметри в това отношение.

Съществуващите проучвания или доклади за оценка също могат да представляват подходящи информационни източници. Те могат да съдържат съответните модели, техники или източници на данни, или констатации, които са полезни за сравнение.

6.3. Триангулация

Качеството на анализа на събраните данни ще зависи също и от степента на така наречената триангулация. Триангулацията предполага, че към даден въпрос или проблем се подхожда от различни гледни точки, за да се получат възможно най-точни изводи и заключения. Съществуват три вида триангулация: триангулация на източниците на данни, триангулация на изследванията и триангулация на методите.

Триангулация на методите предполага използването на няколко изследователски техники. Структурата на оценяването обикновено съдържа подход на повече методи, при който се използват количествени и качествени методи и източници на данни, за да се отговори на въпросите за оценка. Докато количествените техники позволяват наблюдение и съпоставимост, качествените техники осигуряват вникване, обяснение и тълкуване.

Триангулация на източниците на данни съчетава видове информация, които може да произтичат от количествени и качествени техники. Например, първоначалните констатации въз основа на данните от мониторинга могат да бъдат проверени и обяснени чрез разглеждането им по време на интервюта с експерти и / или привилегировани свидетели. И обратното, изявления на респондентите при проучването може да бъдат сверени и подкрепени от количествени данни. Информацията, получена чрез различни качествени методи, например интервю, фокус групи и анализ на документи или преглед на литературата, могат да бъдат противопоставени помежду си, за да се получи последователно и надеждно становище по въпроса.

Триангулация на изследванията може да бъде постигната чрез мултидисциплинарен състав на екипа за оценяване, например, опит в публичната администрация и специфичен сектор на експертни познания както и комбиниране на академични знания със знания на практикуващите. Всяка гледна точка може да допринесе и да обогати анализа. Освен това, няколко членове на екипа за оценяване (с общ или различен експертен опит) могат индивидуално да направят отделен анализ въз основа на един и същ набор от информация и по-късно заключенията да бъдат сравнени и прецизирани.

По принцип, фазата на събиране и анализ на данни е един повтарящ се процес, включващ няколко кръга на анализ и допълнително събиране на данни, когато е нужно.

7. ДОКЛАДВАНЕ

7.1. Насоки

Стабилна структура, правилно събиране на данни и анализ, и достоверни констатации са основните съставки на добрата оценителска практика. Въпреки това, важна следваща стъпка е изготвянето на доклада за оценка. Докладът за оценка е краен продукт от процеса на оценяване и поради това е от съществено значение за съобщаването на констатациите от оценките по правилен, лесно четим и професионален начин. Посочваме някои основни принципи за правилно докладване на резултатите от оценяването.

- Докладът трябва да бъде написан по ясен и последователен начин. Важно е дължината на доклада да се ограничи до същността на всяка от стъпките в процеса на оценяване. Препоръчително е подробните описания на анализите да се дават в приложения към доклада.
- Резюмето на изпълнение е незаменима част. То се пише за създателите на политики и политици, които искат информация, насочена върху същността на подхода, резултатите, заключенията и препоръките.
- Ясна структура, която обхваща всички компоненти на процеса на оценяване.
- Фигурите и таблиците допринасят за яснотата на посланието.

Следващият прозорец съдържа основните съставки на доклада за оценка.

- ✓ Резюме
- ✓ Увод
 - Цел на оценката
 - Профил и обхват на предмета на оценяване (*Evaluandum*)
 - Самоличност на възложителя и оценителя
 - Структура на доклада
- ✓ Замисъл на оценяването
 - Описание на предмета на оценяване (*evaluandum*)
 - Обхват на оценяването: въпроси и критерии за оценка
 - Подход за събиране на данни
 - Подход за анализ на данни
 - Подход за преценка
 - Въвеждаща информация и бележки относно подхода
- ✓ Резултати от анализа
- ✓ Заключение (отговори на въпроси за оценка)
- ✓ Препоръки (ако е необходимо)
- ✓ Приложения

Разликата между резултатите, изводите и препоръките е важна.

Резултатите са свързани с анализа на данните. Те описват ситуацията и дават възможност за сравнение между установените и очакваните резултати и опреде-

лят разяснителни фактори за успех или провал. Следователно, резултатите представляват първата част от отговорите на въпросите за оценка.

Заклученията произтичат от същността на оценяването: определяне на добавената стойност на дадена политическа намеса. Те са изработени чрез проверка на резултатите от анализа спрямо избраните критерии и норми за оценка. Някои техники подпомагат и преценката, като я правят систематична и прозрачна. Заклученията са по-общи в сравнение с резултатите; те носят тълкуване и синтез на резултатите и посочват последиците от констатациите. Заклученията по същество дават отговорите на въпросите за оценка. Заклученията трябва да имат ясна връзка с резултатите и следователно всякакви методологични ограничения или проблеми трябва да бъдат споменати в доклада, особено когато те имат влияние върху обхвата и съдържанието на заклученията.

Препоръките са предложени действия, насочени към създателите на политики, за коригиране на политиката. Те са ориентирани към действие, а в идеалния случай са придружени с оценка на финансовото и организационно въздействие, както и с времеви хоризонт за изпълнение. Трябва да бъдат формулирани в съответствие с фокуса на оценката и да са свързани с резултатите и заклученията. Препоръките трябва да бъдат предложени за конкретни заинтересовани страни или ползватели, които са в състояние да доведат до предложената промяна. Въпреки че препоръките са обичайна практика в оценяването, те не винаги се изискват в заданието.

7.2. Разпространяване и използване на резултатите

За да може докладът за оценка да окаже значително въздействие, от решаващо значение е добре обмислената комуникационна стратегия. Комуникациите трябва да бъдат съобразени с конкретната аудитория, която може да бъде възложителят на оценяването, други пряко или косвено засегнати заинтересовани страни и обществеността като цяло. Политически чувствителни въпроси и личната неприкосновеност също трябва да бъдат взети под внимание при споделянето на информация. Оценителят има широк спектър от потенциални комуникационни канали:

- Лични дискусии
- Среци на ръководната група
- Междинни доклади
- Презентации
- Семинари
- Съобщения за медиите

На практика, използването на констатациите от оценяването може да доведе до нови управленски или политически решения или нова формулировка на въпроси на политиката и до промяна на идеите и гледните точки на създателите на политики.

8. ОРГАНИЗАЦИЯ, ПЛАНИРАНЕ И КОНТРОЛ НА КАЧЕСТВОТО

8.1. Вътрешно спрямо външно оценяване

Оценяването може да бъде вътрешно и да бъде извършено от лица, които са участвали в разработването и/ или изпълнението на политиката. Въпреки това, повечето от ретроспективните оценявания се извършват от външен оценител, на който това се възлага от името на организацията, управляваща политиката.

От една страна, ако организационното обучение е основната цел, вътрешното оценяване може да бъде предпочитано и по-специално, когато се разглеждат процесите на изпълнение. От друга страна, когато става въпрос за обща политика или оценяване, вдъхновено от опасения за отчетност или легитимност, външното оценяване може да бъде по-подходящо.

На практика съществуват хибридни видове, които могат да комбинират силните страни на вътрешното и външното оценяване и да се избегнат някои от недостатъците.

Някои от силните страни на вътрешното оценяване се отнасят до достъпа до данни и качеството на данните, на които вътрешните оценители, могат да разчитат. Вътрешният оценител също може да участва в разработването или изпълнението на политиката и по този начин да бъде по-чувствителен към контекста, да предложи по-добро разбиране за структури, отношения между субекти, техните гледни точки, зони на напрежение и т.н. Вътрешното оценяване често се възприема като по-малко заплашително и по-благоприятно за конструктивни подобрения. Вътрешните оценявания могат да играят важна роля в изграждането на култура на оценяване в рамките на организацията. Основният недостатък на вътрешното оценяване засяга потенциалната сложност на ролята на оценителя и потенциален конфликт на интереси. Собствеността на оценяваната политика би могла да възпрепятства критичната преценка. Отрицателните констатации могат да бъдат пренебрегнати или дори избегнати. Освен това, съществува риск да не се съобрази времето, необходимо за вътрешно оценяване, което води до недостиг на капацитет (персонал, умения, знания, и бюджет) за провеждане на оценяването.

Външното оценяване има предимството на (възприемана) по-висока степен на независимост, което обикновено се свързва с обективност и неутралност. Външният оценител също може да осигури нови перспективи, идеи и алтернативни формулировки на проблеми и решения. Външният оценител може да бъде и по-склонен да съобщава проблеми и негативни констатации. Други предимства на външното оценяване могат да бъдат в конкретния методически експертен опит и гъвкавост на персонала в екипа за оценяване, както и по-голяма прозрачност по отношение на пълната себестойност. Недостатъци на външното оценяване могат да бъдат ограничен достъп до данни или възможности за взаимодействие с лицата, отговорни за оценяваната политика. Липсата на взаимодействие може да доведе до проблеми със собствеността на обхвата и резултатите от оценяването.

Изборът между вътрешно и външно оценяване често включва компромис между предимствата и недостатъците на двата типа. Това може да доведе до хибридни форми, например вътрешно проектиран и изпълнен процес на оценяване,

съчетано с възлагане на външен изпълнител на събирането на данни сред целевата група.

За възложеното външно оценяване ще е необходимо ясно задание. Дори вътрешното оценяване може да се възползва от ясен набор от очаквания. Подолу е даден списък с ключови елементи на заданието:

1. Правна рамка и възложител на оценяването
2. Профил на предмета на оценяване (evaluandum) и съответния обхват
3. Цел на оценяването и съответни заинтересовани страни
4. Въпроси и критерии за оценка
5. Налични данни и информация
6. Изисквания или указания за събиране на данни, анализ и преценка
7. График и краен срок
8. Информация за бюджета (споменато максимален бюджет или част от критериите за възлагане)
9. Желана квалификация на оценителя
10. Указания за структурата на офертите
11. Указания за подаване на офертите и критерии за възлагане

8.2. Необходимо време и ресурси

Реалистичното планиране на процеса на оценяване ще допринесе за успеха и качеството на оценяването. В случай на външно или вътрешно оценяване, информацията за наличен бюджет и време е от съществено значение, за да се получи осъществимо и полезно оценяване. Определянето на бюджета може да се направи отгоре-надолу или отдолу-нагоре. В подхода отгоре-надолу бюджетът се разглежда като част от общия бюджет, свързан с интервенцията на политиката, макар че не е лесно да се посочи стандартен процент. На практика, разходите за оценяване варират от под 10% дори до 50% от стойността на интервенцията. Високият процент може да бъде оправдан, когато се отнася до оценяване на иновативни или експериментални интервенции. При подход отдолу-нагоре бюджетът може да се определи по различни начини: чрез обобщаване на очакваните разходи, свързани с всеки въпрос за оценка, чрез оценка на необходимото време за всяка дейност по оценяване (например, определяне на обхвата, събиране на данни и анализ, синтез и докладване), умножено по фиксирана ставка или разходи за персонала. В случай на тръжна процедура е препоръчително от съображения за сравнимост, да се поиска подробна информация относно стойността на офертата за оценяване, например, за всяка дейност или разбивка между цена на персонала, работа и режийни.

На практика, разходите за оценяване ще зависят от няколко елемента:

- Броят на въпросите за оценка;
- Изискваната дълбочина на анализа;
- Очакванията по отношение на надеждността на резултатите;
- Качество на съществуващите източници на информация и потребности или събиране на първични данни;
- Времето, изразходвано за проследяване на процеса на оценяване.

8.3. Качество на (процеса) оценяване

По време на процеса на оценяване е важен контролът на качеството. Както обективното качество на оценяването, така и възприятията на заинтересованите страни, че качеството има значително влияние върху крайната полезност и влияние на оценяването върху политическите решения. При оценката на качеството играят роля редица фактори. Първо, договорът за оценяването (външно оценяване) или планът за оценяване (вътрешно оценяване) съдържа насоки и изисквания по отношение на структурата на оценяване. На второ място, уменията на оценителя са от решаващо значение, както и качеството на наличната информация. Доброто взаимодействие между възложител и оценител и участие на съответните заинтересовани страни също ще допринесат за качеството на оценяване. Не на последно място, качеството на доклада за оценка като продукт на процеса, също определя цялостното качество на оценяването.

Това, което е от значение за процеса на оценяване, може да се намери в така наречените стандарти за оценяване, които са разработени от редица мрежи за оценяване в различните страни. Те са написани като насока както за възложители на оценяване, така и за оценители и заинтересовани страни. Много от проблемите могат да бъдат групирани в следните принципи:

- полезност (задоволяване на информационните потребности);
- осъществимост (реалистично и икономически ефективно изпълнение);
- собственост (като се вземат предвид исканията на заинтересованите страни);
- точност (строг метод и валидни заключения).



ИНСТИТУТ ПО ПУБЛИЧНА АДМИНИСТРАЦИЯ

Провежда обучение за служебно и професионално развитие на служителите от администрацията, конкурси за добри практики, конференции, годишни срещи и други форуми.