

Univerzita Karlova v Praze
Fakulta sociálních věd

DISERTAČNÍ PRÁCE

Březen 2010

Mgr. Jiří Remr

Univerzita Karlova v Praze
Fakulta sociálních věd
Institut sociologických studií

TEORIE A PRAXE EVALUAČNÍCH ŠETŘENÍ V ČESKÉ REPUBLICE

Vypracoval:

Mgr. Jiří Remr
Březen 2010

Školitel:

Prof. PhDr. Hynek Jeřábek, CSc.

Potvrzuji, že jsem disertační práci vypracoval samostatně, a to pouze s využitím uvedených pramenů.

V Praze, 5.3. 2010

Mgr. Jiří Remr

Autor děkuje školiteli Prof. PhDr. Hynku Jeřábkovi, CSc., za vedení práce a cenné náměty a připomínky k prvním verzím textu. Autor současně děkuje všem recenzentům jednotlivých textů, které předcházely vypracování této práce.

Kontaktní adresa:

Fakulta sociálních věd UK
Institut sociologických studií
Mgr. Jiří Remr
U Kříže 8
155 00 Praha 5

e-mail: remr@centrum.cz

Abstrakt

Předkládaná práce je zaměřena na analýzu způsobů, jakými jsou v ČR uplatňovány vybrané prvky evaluační metodologie. Výchozím bodem řešení vytyčeného výzkumného úkolu je identifikace konkrétních oblastí, jimiž se postup provádění evaluačních šetření odlišuje od ostatních činností a postupů. Na základě srovnávací analýzy s takovými oblastmi jako je základní výzkum, monitoring či audit jsou tak definovány některá specifika, jež jsou dále doplněna o další charakteristické rysy procesní povahy.

Navazující série empirických výzkumů soustředila pozornost na zjištění míry, v jaké jsou evaluační teorie definované metody, techniky a postupy reálně uplatňovány v každodenní praxi. V tomto ohledu byly zvoleny tři rozdílné, vzájemně spolu nesouvisející oblasti, ve nichž se evaluace běžně aplikuje. S ohledem k dalším cílům předkládané práce navíc každá z oblastí byla zvolena tak, aby reprezentovala jinou formu provádění evaluačních šetření. Výsledky provedeného desk research tak poukazují na konkrétní podobu evaluačních šetření prováděných externími realizátory evaluačních šetření v oblasti operačních fondů EU, dále na vlastní hodnocení, jež odrážejí současnou praxi interního způsobu realizace evaluačních šetření a v neposlední řadě jen také dokumentován způsob hybridního provádění evaluačních šetření, který vychází z expertního posuzování (v daném případě jde o posuzování odborné kvality a úrovně bakalářských prací). Další okolnosti provádění evaluačních šetření byly zjišťovány pomocí hloubkových rozhovorů s realizátory a zadavateli evaluačních šetření. S ohledem ke zřejmým okolnostem se dané šetření omezilo výlučně na oblast externě prováděných evaluačních šetření.

Průběžná konfrontace poznatků z prováděných výzkumů s relevantními jevy popisovanými v teorii umožnila identifikovat a detailně popsat několik okolností, kterými se uplatňování evaluace v ČR vyznačuje. Jsou jimi (1) mimořádný důraz kladený na doporučení vycházející z provedené evaluace, (2) nedostatek zpětné vazby k provedeným evaluačním šetřením, (3) exaktní základ prováděných evaluačních šetření, (4) klamný charakter vynucených evaluačních šetření a (5) vyhýbavý přístup ke specifickým evaluačním technikám a nástrojům.

Získané poznatky lze považovat za významné zejména proto, že poskytují zpětnou vazbu samotným evaluátorům v ČR. Dále přinášejí nové poznatky o způsobu, jakým jsou zde evaluační šetření prováděna a v neposlední řadě mohou pomoci při rozvoji nových evaluačních přístupů.

Résumé

The main focus of presented thesis is the in-depth analysis of major patterns of how the evaluation methodology is applied within the Czech society environment. The starting point is identification of key processes and activities that distinguish evaluation from other activities and disciplines. Based on comparative analysis among evaluation, primary research, monitoring and auditing the crucial features are defined and further supplemented by other characteristics of the procedural nature.

The whole series of researches focus on the extent in which the evaluation methods, techniques and procedures are applied in everyday practice. In this respect, there were identified three different and uncorrelated branches that serve as application environment for routine deployment of evaluation inquiries. Following secondary goals of our thesis, each of the branches represent different example of how the evaluation could be applied/implemented. Desk-research informs about actual patterns of how the evaluation is performed by external evaluators who work for Managing Authorities of EU Operational Programs, it also shows how internally based self-evaluation is performed and last but not least, how the hybrid evaluation based on expert judgment principles could be utilized for evaluation of bachelor thesis. Specific in-depth interview with evaluation providers and clients were also conducted in order to get other aspects of evaluation performance. This part piece of research is for obvious reason focus only on external evaluations.

On-going comparison of research results with theoretical description of the phenomenon led to identification and specification of five topics that are typical for the current situation in the Czech Republic. It is above all (1) extraordinary emphasis put on recommendations resulting from the evaluation inquiry, (2) lack of feedback to conducted evaluations, (3) rigor of conducted evaluations, (4) of evaluations that are required and (5) low utilization of specific evaluation tools and techniques.

Research results could be found important mainly because it provides the necessary feedback to evaluators in the Czech Republic. It also brings new information about how the evaluations are performed and it also helps to introduce new approaches in the field of evaluation.

Stručný obsah

<i>Abstrakt</i>	5
<i>Résumé</i>	6
<i>Stručný obsah</i>	7
<i>Podrobný obsah</i>	8
<i>1. Úvod</i>	12
<i>2. Vymezení sledované oblasti</i>	14
<i>3. Identifikace specifík evaluačních šetření</i>	30
<i>4. Analýza přístupů k realizaci evaluačních šetření</i>	54
<i>5. Popis metodiky provedených výzkumů</i>	104
<i>6. Uplatnění evaluace v praxi</i>	111
<i>7. Realizace evaluačních šetření v praxi</i>	152
<i>8. Závěry</i>	166
<i>Přílohy</i>	168
<i>Seznam použité literatury</i>	184

Podrobný obsah

<i>Abstrakt</i>	5
<i>Résumé</i>	6
<i>Stručný obsah</i>	7
<i>Podrobný obsah</i>	8
1. Úvod	12
1.1 Motivace k provedení výzkumu	12
1.2 Cíle práce	12
1.3 Význam předkládané práce	13
2. Vymezení sledované oblasti	14
2.1 Definice a rozbor základních pojmů	15
2.2 Předmět evaluace dle zkoumané oblasti	21
2.2.1 Evaluace intervenčních programů	21
2.2.2 Evaluace výrobků (product evaluation)	27
2.2.3 Evaluace pracovníků (personnel evaluation)	28
2.2.4 Evaluace výkonu (performance evaluation)	28
2.2.5 Evaluace nabídek (proposal evaluation)	28
2.2.6 Evaluace politik (policy evaluation)	29
2.2.7 Souhrn	29
3. Identifikace specifík evaluačních šetření	30
3.1 Věcná specifika evaluačních šetření	30
3.1.1 Základní výzkum ve společenských vědách	31
3.1.2 Monitoring	35
3.1.3 Audit	37
3.1.4 Souhrn	39
3.2 Procesní specifika realizace evaluačních šetření	40
3.2.1 Zapojení zainteresovaných stran	40
3.2.2 Konceptualizace a tvorba konceptuálních modelů	47
3.2.3 Metaevaluace	51
3.2.4 Souhrn	53
4. Analýza přístupů k realizaci evaluačních šetření	54
4.1 Teoretické přístupy k realizaci evaluačních šetření	54
4.1.1 Cílené evaluace (goal-based evaluation; Ralph Tyler)	55
4.1.2 Explorační přístup k provádění evaluačních šetření (goal-free evaluation; Michael Scriven)	59
4.1.3 Responsivní evaluace (responsive evaluation; Robert Stake)	62
4.1.4 Užitizační přístup k realizaci evaluačních šetření (utilization-focused evaluation; Michael Quinn Patton)	66
4.1.5 Konstruktivistický přístup k realizaci evaluačních šetření (constructivist evaluation; Egon Guba, Yvonna Lincolnová)	70
4.1.6 Přístup k provádění evaluačních šetření vedený teorií (theory-driven evaluation; Huey-Tsyh Chen, Peter Rossi)	73
4.1.7 Realistický přístup k provádění evaluačních šetření (realistic evaluation; Ray Pawson, Nick Tilley)	77
4.1.8 Souhrn	81

4.2. Praktické přístupy k realizaci evaluačních šetření	83
4.2.1 Dle účelu: formativní a sumativní	83
4.2.2 Dle předmětu: evaluace výsledku, procesu a výzkumu potřeb	91
4.2.3 Dle vztahu zadavatele a realizátora: externí, interní a hybridní	97
4.2.4 Souhrn	103
5. Popis metodiky provedených výzkumů	104
5.1 Obecný výzkumný rámec	104
5.2 Stanovení cílových skupin	105
5.3 Technika výběru	106
5.4 Výzkumné instrumenty	107
5.5 Sběr dat	108
5.6 Etické souvislosti sběru dat	109
6. Uplatnění evaluace v praxi	111
6.1 Externí forma provádění evaluačních šetření	111
6.1.1 Kontext případové studie	111
6.1.2 Popis případové studie	113
6.1.3 Výsledky	116
6.1.4 Diskuse výsledků	123
6.1.5 Dílčí závěry	123
6.2 Interní forma provádění evaluačních šetření	125
6.2.1 Kontext případové studie	125
6.2.2 Popis případové studie	127
6.2.3 Výsledky	129
6.2.4 Diskuze výsledků	137
6.2.5 Dílčí závěry	137
6.3 Hybridní forma provádění evaluačních šetření	139
6.3.1 Popis případové studie	139
6.3.2 Výsledky	141
6.3.3 Diskuse výsledků	148
6.3.4 Dílčí závěry	149
6.4 Souhrn poznatků případových studií	151
7. Realizace evaluačních šetření v praxi	152
7.1 Strana poptávky	152
7.2 Strana nabídky	155
7.3 Způsob provádění evaluačních šetření	158
7.3.1 Dilema metody a výsledku	158
7.3.2 Dilema administrátora programu a zadavatele evaluačních šetření	161
7.3.3 Dilema evaluace a pseudoevaluace	164
8. Závěry	166
Přílohy	168
Seznam použité literatury	184

Seznam obrázků

Obrázek č. 1	Věcný základ a přínos	16
Obrázek č. 2	Schéma horizontálních interakcí	17
Obrázek č. 3	Schéma vertikálních interakcí	17
Obrázek č. 4	Schéma temporálních interakcí	18
Obrázek č. 5	Hlavní oblasti dle předmětu evaluace	27
Obrázek č. 6	Vztah evaluačních šetření k ostatním disciplínám	39
Obrázek č. 7	Ukázka explanačního modelu	48
Obrázek č. 8	Ukázka implikačního modelu	49
Obrázek č. 9	Mapa jednotlivých evaluačních přístupů	54
Obrázek č. 10	Obecný princip cílených (goal-based) evaluací	56
Obrázek č. 11	Návrh konceptuálního schématu	77
Obrázek č. 12	Princip realistických evaluací	79
Obrázek č. 13	Význam formativních a sumativních evaluací během působení programu	87
Obrázek č. 14	Srovnání vzájemného vztahu mezi potřebami a osobními preferencemi	94
Obrázek č. 15	Formy provádění evaluačních šetření v aplikační praxi	101

Seznam grafů

Graf č. 1	Rozsah závěrečných zpráv (počet stran)	115
Graf č. 2	Specifikace zaměření evaluačního šetření	116
Graf č. 3	Provedení analýzy sekundárních dat a primárního výzkumu	118
Graf č. 4	Využití metod kvalitativního a kvantitativního výzkumu	119
Graf č. 5	Konkrétní použité metody kvalitativního výzkumu	120
Graf č. 6	Konkrétní použité metody kvantitativního výzkumu	120
Graf č. 7	Identifikace jednotlivých zainteresovaných stran	121
Graf č. 8	Funkce autora autoevaluační zprávy	127
Graf č. 9	Typ subjektu	128
Graf č. 10	Rozsah autoevaluačních zpráv (počet stran)	128
Graf č. 11	Specifikace úspěchů a neúspěchů školy v daném období	129
Graf č. 12	Obsah jednotlivých částí autoevaluační zprávy	131
Graf č. 13	Míra podrobnosti jednotlivých částí autoevaluační zprávy	132
Graf č. 14	Specifikace kritérií vlastního hodnocení školy	133
Graf č. 15	Vytyčení standardů vlastního hodnocení školy	134
Graf č. 16	Využívání sekundárních dat a výsledků primárních výzkumů při autoevaluaci	136
Graf č. 17	Rozsah posudků (počet znaků včetně mezer)	140
Graf č. 18	Srovnání rozsahu posudků konzultantů a oponentů	141
Graf č. 19	Komplex zpětné vazby	142
Graf č. 20	Komplement zpětné vazby	143
Graf č. 21	Silné a slabé stránky posuzovaných bakalářských prací	144
Graf č. 22	Klasifikace bakalářské práce	145
Graf č. 23	Návrh na zlepšení způsobu analýzy poznávacího problému	146
Graf č. 24	Návrh na zlepšení závěrů, ke kterým autor/ka dospěl/a	147
Graf č. 25	Návrh na zlepšení použité metodiky a jejího popisu	147
Graf č. 26	Náměty na obhajobu	148

Seznam tabulek

Tabulka č. 1	Srovnání vybraných parametrů základního výzkumu a evaluačních šetření	34
Tabulka č. 2	Srovnání vybraných parametrů monitoringu a evaluačních šetření	36
Tabulka č. 3	Srovnání vybraných parametrů auditu a evaluačních šetření	38
Tabulka č. 4	Zájmy vybraných zainteresovaných stran	43
Tabulka č. 5	Maticе zainteresovaných stran a jejich zájmů	45
Tabulka č. 6	Porovnání některých charakteristik jednotlivých teoretických přístupů k realizaci evaluačních šetření	82
Tabulka č. 7	Rozdíly mezi formativní a sumativní evaluací	88
Tabulka č. 8	Vztah mezi formou provedení evaluace a způsobem využití jejích výsledků	89
Tabulka č. 9	Počty posuzovaných závěrečných zpráv	115
Tabulka č. 10	Tematické Operační programy 2007 – 2013	153
Tabulka č. 11	Regionální Operační programy 2007 – 2013	153
Tabulka č. 12	Seznam realizátorů evaluačních šetření	155

1. Úvod

1.1 Motivace k provedení výzkumu

Evaluace je dnes především pomocníkem při komplexních rozhodovacích procesech zahrnujících mnoho alternativ, mnoho účastníků častokrát navíc s rozličnými zájmy, a mnoho kritérií, podle nichž je možné jednotlivá rozhodnutí provádět. S evaluací se tak setkáváme ve vědě, školství, veřejné správě a v mnoha dalších oblastech, kde evaluace naplňuje některé ze svých typických funkcí.

Kromě již uvedené podpory rozhodování, evaluace především (1) pomáhá stanovit odpovědnost (*accountability*); (2) pomáhá identifikovat dobré praktiky, a rozšiřuje tak rozvojový potenciál jednotlivých společností a v neposlední řadě je (3) nástrojem soustavné reflexe prováděných činností. V těchto případech jsou používány postupy objektivní, systematické a kontrolované *evaluace*, jež se zásadně odlišuje od nesystematického, intuitivně prováděného a subjektivního *hodnocení*.

I mnozí realizátoři evaluačních šetření jsou laiky! Požadavky na provedení evaluací zastihly současné evaluátory v situaci, kdy již byli etablováni v jiných oborech a jen málokdo měl možnost systematicky studovat evaluační metody a techniky. V důsledku toho je aktuální úroveň povědomí o základech evaluačního myšlení nedostatečně rozvinutá, základní principy evaluační činnosti nejsou vždy uplatňovány, evaluace se mnohdy zaměřuje s jinou činností, a to se promítá do způsobu, jak jsou evaluační šetření v současné době v ČR prováděna.

1.2 Cíle práce

Předkládaná práce se v první řadě snaží analyzovat způsoby uplatňování vybraných prvků evaluační metodologie v praxi. Proto, aby bylo možné vytyčeného cíle dosáhnout, je třeba nejprve přiblížit podstatu evaluace a objasnit teoreticky zakotvené přístupy k evaluačním šetřením. Druhým cílem předkládané práce je tedy definice evaluace a rozbor jejích principů; tento cíl naplňují kapitoly druhá, třetí a čtvrtá.

Na základě obsahového vymezení a předmětného zúžení naší práce je možné prezentovat empirická data dokumentující situaci v České republice. Třetím cílem této práce je popsat

současné praktiky při provádění evaluačních šetření a přispět k objasnění důvodů využívání konkrétních postupů a metodik; výsledky daného cíle jsou prezentovány v šesté a sedmé kapitole, přičemž osmá kapitola je pokusem o souhrnnou charakterizaci aktuálně používaných evaluačních praktik.

Práce si neklade za cíl dokazovat, že správné provádění evaluačních šetření posiluje jejich utilizaci; nepoukazujeme ani na výhody plynoucí z adekvátního používání závěrů evaluačních šetření a nesnažíme se dospět v tomto ohledu k ultimativním závěrům. Usilujeme však o rozkrytí podstaty praktického uplatnění principu evaluace a o podporu rozvoje evaluačních schopností.

1.3 Význam předkládané práce

Předkládaná práce přináší údaje, které dokumentují schopnosti relevantních subjektů v ČR provádět evaluační šetření v souladu s teoreticky zakotvenými postupy a používat jejich výsledky. Jde o údaje, které lze považovat za unikátní; existují sice i jiná šetření zaměřená na oblast evaluací, ta se však omezují buď jen na vybraný segment (např. na evaluaci Strukturálních fondů EU), nebo se zaměřují na zkoumání obecnější roviny, kdy sledují evaluační kapacitu celé země. Doposud však chybí informace o způsobu provádění evaluačních šetření či obecněji: o formách uplatnění evaluace v praxi. Získané údaje a závěry, ke kterým jsme dospěli tak poskytují reflexi jednotlivým aktérům, kteří se spolupodílejí na faktickém naplňování evaluačního poslání. Na základě empirického šetření tak bylo identifikováno celkem pět aktuálně významných témat. Jsou jimi (1) mimořádný důraz kladený na doporučení vycházející z provedené evaluace, (2) nedostatek zpětné vazby k provedeným evaluačním šetřením, (3) exaktní základ prováděných evaluačních šetření, (4) klamný charakter vynucených evaluačních šetření a (5) vyhýbavý přístup ke specifickým evaluačním technikám a nástrojům.

2. Vymezení sledované oblasti

V této podkapitole je proveden nástin typických oblastí, ve kterých se evaluace založené na posuzování věcného základu a přínosu využívá. V předmětném vymezení evaluace se snažíme poukázat na portfolio základních oborů a disciplín, v jejichž rámci vykazuje přístup k evaluaci velmi podobné rysy, a to i přes značnou pestrost a různorodost jednotlivých oblastí. Není cílem naší práce odhadovat, jaký význam mají jednotlivé oblasti; z hlediska aplikačního potenciálu evaluací neuvádíme na tomto místě ani analýzu diferencí či podobností v metodologických přístupech uplatňovaných v rámci samostatných evaluačních oborů, a dokonce ani neprovádíme rozbor evaluační kapacity, kterou jednotlivé oblasti představují. Uvedení širokého spektra oblastí využívajících evaluaci vytváří předpoklad pro následné podrobné zpřesnění a specifikaci konkrétního předmětu zájmu, který v této práci sledujeme.

2.1 Definice a rozbor základních pojmů

S výstižnou charakterizací evaluace přišel na samém počátku rozvoje evaluačního oboru Scriven (1967), který navrhl rozlišovat dvě klíčové komponenty evaluace. Jeho pojetí se tak opírá o posouzení věcného základu (*merit*¹) a přínosu (*worth*). **Věcný základ** vystihuje a charakterizuje samotnou podstatu objektu evaluace. Jeho stanovení je ve velké většině případů relativně nekomplikované, neboť věcný základ přímo odkazuje k funkčnosti posuzovaného objektu.

Věcným základem např. rekvalifikačního kursu je umožnit jeho účastníkům osvojení nové kvalifikace; věcným základem systému recyklace je opětovné využití odpadu.

Prvním kokem evaluace je tak podle M. Scrivena ověření věcného základu, tedy stanovení, zdali je předmět evaluace skutečně tím, za co je považován, resp. zdali a v jaké míře předmět evaluace naplňuje svůj věcný základ.

Je skutečně daný kvalifikační kurs nástrojem, který umožňuje dosáhnout jeho účastníkům nové kvalifikace? Nedochozí tedy k tomu, že si účastníci zdokonalují a prohlubují znalosti a dovednosti ve své stávající kvalifikaci?

Z uvedených příkladů je zřejmé, že hledání odpovědí na otázky po věcném základu směřuje vždy k charakterizaci posuzovaného předmětu a k jeho odlišení od objektů jiného druhu.

Vedle toho **přínos** odkazuje k tomu, jak je věcný základ vnímán konkrétními cílovými skupinami.

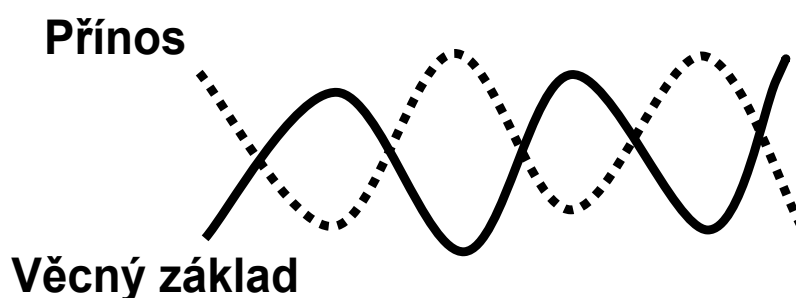
Přínosem rekvalifikačního kursu pro jeho účastníky je osvojení nové kvalifikace a zvýšení šancí na uplatnění na trhu práce; přínos téhož rekvalifikačního kursu pro jeho lektory může spočívat jejich vlastním udržením na trhu práce; přínosem rekvalifikačního kursu pro zaměstnavatele v dané lokalitě je snazší dostupnost pracovních sil s potřebnou kvalifikací.

Otázkami po přínosu se tedy hodnotitel obrací k hodnotě, kterou může předmět posuzování poskytnout konkrétním skupinám dotčených. Do centra pozornosti je tedy umístěn nikoliv

¹ Některé jiné prameny používají jako synonymum také „quality“.

samotný předmět evaluace a jeho charakteristické vlastnosti, ale soulad těchto vlastností s očekáváním či potřebami těch, kteří jsou (či alespoň teoreticky mohou být) daným objektem evaluace dotčeni. Posouzení přínosu může být poměrně složitým procesem, který vyžaduje ze strany evaluátora detailní obeznámenost s potřebami relevantních subjektů. Podobně jako při posuzování věcného základu, i posouzení přínosu je v tomto pojetí oproštěno od proliferace vlastních potřeb či individuálních postojů evaluátora.

Obrázek č. 1 Věcný základ a přínos



Zdroj: autor

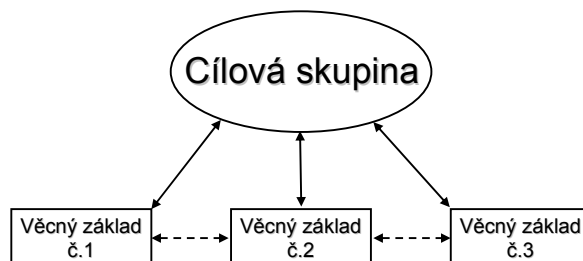
Z tohoto přiblížení je patrné, že **věcný základ** je objektivní daností předmětu evaluace a je nezávislý na vnějším kontextu, v němž se vyskytuje (popř. v němž je hodnocen); je nezávislý na způsobech použití a nezávisí ani na dalších charakteristikách daného objektu. Věcný základ rekvalifikačního kursu zůstává stejný bez ohledu ke skutečnosti, že jeho absolventi nakonec najdou zaměstnání, ve kterém využijí svou původní, nikoliv nově nabytou, kvalifikaci; daný rekvalifikační kurs zůstane rekvalifikačním kursem bez ohledu na počet jeho frekventantů, i bez ohledu na kvalitu a důkladnost odborné přípravy. V takových případech lze identifikovat případné zdroje či příčiny nízké efektivity rekvalifikačního kursu, lze hledat faktory, které přispívají k jeho dysfunkčnosti, nicméně přisouzení jiného věcného základu je v těchto případech jen velmi těžko představitelné.

Naproti tomu **přínos** je parametrem relativním; jeho míra vždy záleží na interakci věcného základu s daným okolím. V tomto ohledu lze evaluaci opřít o následující analytické přístupy vymezující vztah věcného základu ke svému kontextu:

a) horizontální interakce

Věcný základ se nenachází ve vakuu, ale vždy interaguje s věcnými základy jiných opatření či zkrátka s jinými stimuly, které mohou ovlivňovat chování jednotlivých cílových skupin.

Obrázek č. 2 Schéma horizontálních interakcí



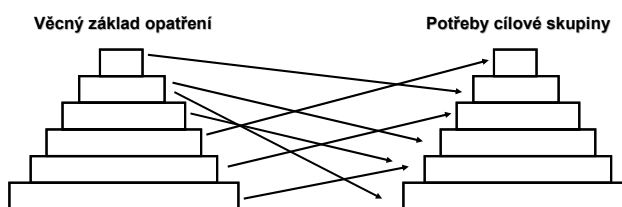
Zdroj: autor

Příklad: Účast v rekvalifikačním kursu se může u konkrétních zástupců cílové skupiny (tj. aktuálně nezaměstnaných) střetávat s možností výkonu pracovní činnosti v rámci šedé ekonomiky, s nabídkou jiného rekvalifikačního kursu, s možností okamžitého nástupu do zaměstnání s nižší kvalifikační náročností atd.

b) vertikální interakce

Každý jednotlivý zástupce cílové skupiny má svůj individuální hodnotový systém, na jehož základě vnímá přínos, který mu poskytuje předmět evaluace. Daný jednatel tak identifikuje u rekvalifikačního kursu klíčové hodnoty, na nichž je daný kurs vystaven (a na které apeluje) a poměřuje je se svou individuální hierarchií hodnot. Na základě toho volí své další jednání, tedy zapojení do daného kursu, svou aktivní participaci v něm apod.²

Obrázek č. 3 Schéma vertikálních interakcí



Zdroj: autor

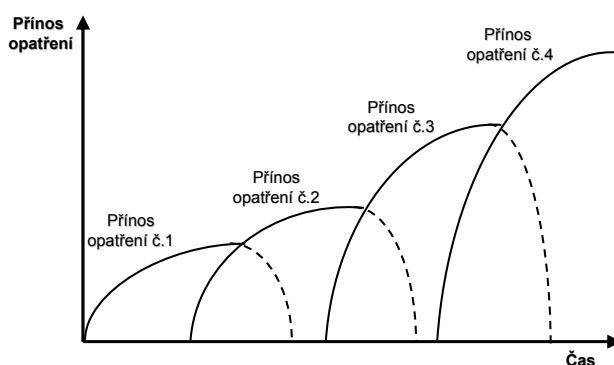
Příklad: Potřeba okamžitého výtěžku na straně zástupců cílové skupiny nemusí vždy korespondovat s charakterem (dlouhodobě koncipovaného) kursu. Skutečnou potřebou některých zástupců cílové skupiny navíc nemusí být faktické osvojení nových dovedností, ale jen formálních splnění podmínek nároku na příslušné dávky.

² Předpokládáme na tomto místě, že se daný jedinec může o svých krocích samostatně rozhodovat.

c) *temporální interakce*

Oblasti, v nichž jsou jednotlivá opatření zaváděna, mohou již nějakým způsobem být regulovány a může se tak k nim vztahovat nějaká jiná forma intervence. Zaváděný kurs tedy může stát na základech (či na troskách) svého předchůdce. V podmínkách funkční evaluační kultury by nově zaváděný rekvalifikační kurs měl být rozvíjen a dále modifikován tak, aby dodržoval, ba zvyšoval efektivitu intervence v dané oblasti a aby překonával své vlastní limity.

Obrázek č. 4 Schéma temporálních interakcí



Zdroj: autor

Příklad: Soustavná evaluace může vést jednak k zavádění nových forem výuky, a dále také k lepší provázanosti nabídky rekvalifikačních kursů s potřebami podnikatelského sektoru v dané oblasti.

Scrivenovu definici evaluace přejala rovněž Joint Committee on Standards for Educational Evaluation jako svou „oficiální“ definici; v jejich Program Evaluation Standards³ (cit. dle Sanders 1994) je totiž evaluace definována jako „systematické šetření věcného základu a přínosu daných objektů“. Přestože tento způsob vymezení evaluace je dále akceptován mnoha dalšími autory (mj. Patton 2008; Clarke 1999; Alkin 2004; Gupta 2007; Mohr 1995), nezůstává nehybným pomníkem pionýrských dob evaluačního oboru a je dále rozvíjen a precizován. Příkladem je Fitzpatricková et al. (2004), která v rámci své definice evaluace poukazuje na vysoký význam takových procesů, jakými jsou identifikace, vyjasnění a použití obhajitelných kritérií⁴, s jejichž pomocí je stanoveno, zdali je dosaženo vytyčených cílů či nikoliv. Podle J. Fitzpatrickové přitom tato kritéria zásadně omezují subjektivismus samotného posuzování a umožňují kontrolu adekvátnosti jejich definice.

³ Jde o soubor pravidel a doporučení, který upravuje způsob realizace evaluací. Dále v textu uvádíme jen „Standardy“.

⁴ Obhajitelnost kritérií je v daném kontextu třeba chápat jako imperativ omezující subjektivitu hodnotitele při jejich aplikaci.

Patrně nejkompexnější definice evaluační činnosti pochází od Fournierové (2005), jež definuje evaluaci jako aplikovaný proces zjišťování, sběru a syntézy důkazů, které vedou k závěrům o věcném základu a o přínosu; o významu či kvalitě posuzovaného programu, produktu, osoby, politiky, návrhu nebo plánu. Její definice tedy zahrnuje jak empirický aspekt (proces zjišťování sběru a syntézy důkazů), tak také aspekt normativní (závěr či úsudek o věcném základu a o přínosu), přičemž je to právě normativní prvek, jenž odlišuje evaluaci od jiných typů šetření, jako např. od základního výzkumu, od klinické epidemiologie, od investigativní žurnalistiky či od výzkumu veřejného mínění.

Scriven (1991) svou základní definici evaluace později rozvinul o obecný nástin metodologického rámce specifikujícího způsob dosažení závěru o věcném základu a o přínosu. Navrhuje systematický a kontrolovaný přístup k jejich stanovení, rozvíjí konkrétní postup a specifikuje jeho jednotlivé etapy. Na počátku stojí (1) identifikace kritérií, tj. stanovení vlastností, jež mají být posuzovány a (2) definování standardů, tedy určení, jaké intenzity (jakého stupně) mají vlastnosti hodnoceného objektu v daných kritériích dosáhnout. Pokračuje (3) získáváním relevantních informací, tedy aplikací příslušných metod sběru dat, které umožní ověřit skutečný stav sledovaných kritérií, a končí (4) integrací či syntézou získaných výsledků, jež porovnává získané údaje s předem stanovenými standardy.

V případě evaluace rekvalifikačního kursu provedení prvního kroku znamená vytyčení konkrétních vlastností, které se vztahují k efektivitě daného kursu. Může být např. stanoveno, že kurs bude posuzován z hlediska: počtu absolventů, průměrné doby hledání zaměstnání jeho absolventy, počtu vyučovaných hodin atd. Druhý krok prováděné evaluace předpokládá stanovení konkrétních standardů pro každé sledované kritérium. V našem případě může být např. stanoveno, že kurs by mělo absolvovat 150 účastníků v průběhu jednoho kalendářního roku; jeho absolventi by neměli hledat nové zaměstnání v průměru déle než 2 měsíce; každý z účastníků kursu by měl absolvovat nejméně 400 hodin výuky. Třetí krok je sběr potřebných údajů, které dokumentují, zdali bylo dosaženo stanovených standardů. V rámci čtvrtého a posledního kroku by byly empirické poznatky porovnány s předem stanovenými standardy a podle dosažených výsledků by evaluátor mohl konstatovat, zdali je kurs skutečně efektivní, či nikoliv, popř. ve kterých ohledech dosahuje očekávání a ve kterých konkrétních ohledech zůstává za očekáváním.

Podobný postup by byl využit i při evaluaci přínosu: evaluátor by nejprve stanovil jednotlivé cílové skupiny či zainteresované strany, pro které může daný rekvalifikační kurs představovat potenciální přínos⁵. Další činnosti by pokračovaly čtveřicí uvedených kroků, tedy (1) stanovením

⁵ Pozn.: „přínos“ je v této souvislosti třeba chápat jako významově neutrální pojem; přínosy totiž mohou být jak pozitivní, tak také negativní.

konkrétních přínosů a kritérií pro jejich posouzení z hlediska specifikovaných cílových skupin či zainteresovaných stran; (2) určení standardů pro tato kritéria; (3) provedení sběru potřebných dat a případné další evidence dokládající skutečné přínosy, které daný rekvalifikační kurs jednotlivým cílovým skupinám či zainteresovaným stranám skutečně poskytl a (4) vyhodnocení empirických dat založené na jejich porovnání s vytyčenými standardy.

Je tedy patrné, že hodnotitel neposuzuje soulad vlastností předmětu evaluace se svým vlastním očekáváním a svými vlastními potřebami, nýbrž posuzuje soulad těchto vlastností a charakteristik s očekáváním a potřebami relevantních cílových skupin (Khakee 2003). Důsledná aplikace uvedeného postupu je tak konceptuální obranou proti nežádoucímu subjektivismu při provádění evaluace.

Souhrn

Účelem této podkapitoly bylo oddělit *evaluaci* založenou na posouzení věcného základu a přínosu od neformalizovaného, laického a subjektivně založeného *hodnocení*. Tento obecný úvod je klíčem k chápání cílů naší práce, v níž chceme na příkladu několika různých oblastí ukázat, jak je evaluace prováděna v konkrétní aplikační praxi.

Evaluace coby obecný princip předznamenává univerzalitu základních stavebních prvků procesu posuzování bez ohledu na různost kontextů a charakterů předmětu evaluace, kterými mohou být jak předměty konkrétní (jako např. knihy, telefony či stavby), tak také velmi abstraktní (kupř. projekty, programy či politiky). Tato univerzalita evaluace přivedla Scrivena (1980) k myšlence tzv. trans-disciplinarity, jíž se snaží vystihnout skutečnost, že evaluace není svázána s žádným konkrétním oborem a že hranice mnohých vědních oborů a disciplin překračuje⁶. M. Scriven sám v tomto ohledu přirovnává evaluaci ke statistice, která rovněž nachází uplatnění v mnoha odlišných oborech (ibid.).

V souladu s výše deklarovanými cíli této práce ponecháváme tedy stranou neformalizované hodnocení, provázející jednotlivce mnohými oblastmi každodenního života, a v dalších kapitolách zaměříme pozornost výlučně na uplatnění evaluace založené na posuzování věcného základu a přínosu.

⁶ O tom nejlépe svědčí složení obce evaluátorů, kde se setkávají pedagogičtí výzkumníci, sociologové, psychologové, ekonomové, regionalisté a mnozí další.

2.2 Předmět evaluace dle zkoumané oblasti

Snaha vtisknout obecně pojatým a abstraktně formulovaným postupům evaluace konkrétní podobu je v relevantních zdrojích uskutečňována v zásadě dvojím způsobem. První přístup rozpracovává příslušnou evaluační problematiku na půdě konkrétní aplikační oblasti, přičemž čtenář není ponechán na pochybách o kontextu a rámci, v němž je dané téma ukotveno. Tento přístup dokumentují takové tituly jako např. „Evaluation Health Promotion Programs“ (Valente 2000), „Evaluating in Social Work (Unrau 2007), „Evaluating Social Work Services and Programs“ (Weinbach 2005) či „R&D Programme Evaluation (Hong, Boden 2003). Druhý přístup klade do centra své pozornosti samotnou evaluaci, přičemž konkrétní aplikační oblasti jsou uváděny jakoby na pozadí principiálních témat týkajících se zejména metody evaluace a způsobu provádění evaluačních šetření. To je patrné zejména v pramenech typu: „Evaluation“ (Rossi, Lipsey, Freeman 2004) či „Evaluation Theory. Models & Applications“ (Stufflebeam a Shinkfield 2007). Tyto prameny obvykle poukazují na Scrivenem uváděných šest oblastí používajících evaluaci; každá z těchto oblastí má přitom svou vlastní dynamiku, své specifické instituce, školy a časopisy, častokrát dokonce také svou vlastní terminologii a specifické metodologické akcenty.

2.2.1 Evaluace intervenčních programů

S tzv. intervenčními programy jsou evaluace spojovány pravděpodobně nejčastěji, a proto jim věnujeme na tomto místě největší pozornost. Podle mnohých je tato oblast využívání evaluací nejrozsáhlejší a nejvýznamnější ze všech; tuto hypotézu podporuje mj. množství pramenů věnovaných evaluaci intervenčních programů oproti počtu pramenů zaměřených na jiné oblasti, ve kterých se evaluace uplatňuje a využívá).

V relevantních pramenech se lze setkat s řadou různých pojmů, jimiž jsou intervenční programy popisovány a s jejichž pomocí jsou definovány, např. Weissová (1972, 1998) či Rossi, Lipsey a Freeman (2004) píšou o „sociálních programech“ (*social programmes*); Unrauová (2007), Chen (1996) a někteří další se vyjadřují o „programech“ (*programmes*), lze se však setkat i s pojmem „intervence“ (*interventions*). Joint Committee on Standards (cit. dle Sanders 1994) definuje (intervenční) programy jako „průběžné aktivity, které jsou zaměřeny

na poskytování služeb“. Důraz na dlouhodobý charakter programů je v tomto ohledu zásadní, neboť odlišuje intervenční programy od projektů, jež mají vždy jednorázový či krátkodobý charakter, a to i navzdory skutečnosti, že mnohé z projektů mohou vyvolávat dopady, které se projevují až v dlouhodobém časovém horizontu. Tato oficiální definice je sice užitečným pomocníkem pro získání základní představy o tom, co mají evaluátoři obecně na mysli, když hovoří o intervenčních programech, nicméně její využitelnost pro bližší operacionalizaci intervenčního programu či pro další analýzu je z důvodu její vysoké míry abstraktnosti velmi omezená.

Pro používání pojmu intervenční program hovoří především jeho univerzalita, neboť navržený pojem neodkazuje k žádné dílčí oblasti evaluací, a přitom je srozumitelný i v ostatních oblastech využívajících evaluaci. Rovněž platí, že zvolený pojem explicitně zdůrazňuje intervenční charakter prováděných činností a odkazuje tak k intencionálním opatřením vyvolávajícím změnu. Navržený termín v neposlední řadě poukazuje na programovost takových kroků, tedy na jejich plánování, rozvržení, přípravu a implementaci. Z takovéto definice je patrné, že nejde ani o regulační aktivitu *ad-hoc*, nejde ani o izolovaný dílčí krok postrádající návaznost na široké spektrum dalších činností či opatření a nejedná se ani o krátkodobě časově vymezené působení. Je však třeba čtenáře na tomto místě zdvořile požádat, aby pojem „intervence“ přijal jako neutrální pojmenování, které na poli evaluací nemá žádné pejorativní konotace.

V evropském kontextu považujeme za nezbytné označovat takto chápané aktivity jako intervenční programy, a umožnit tak jejich odlišení od programů operačních.⁷ Nepřipojujeme se tedy k evaluátorům, kteří používají pojem „program“, neboť pravděpodobnost nežádoucí konfúze, či přinejmenším znejistění čtenáře je v tomto případě velmi vysoká; je třeba vzít v potaz, že evaluace jsou v rámci implementace operačních programů používány skutečně velmi často.

K samotné definici intervenčního programu přistupují jednotliví evaluátoři rozdílně a jejich pojetí je třeba vždy chápat na pozadí obecnějšího evaluačního přístupu, jež jednotliví autoři používají. Intervenční program tak lze chápat jako soubor plánovaných aktivit zamýšlených k tomu, aby způsobily změnu ve vymezené cílové skupině (Smith 1989 cit. dle Owena 2006). Toto pojetí zdůrazňuje nutnost existence jasného a podrobného plánu coby základního rámce

⁷ Operační programy jsou zpodobněním činnosti fondů (tj. Evropského fondu pro regionální rozvoj, Evropských sociálních fondů a Fondu soudržnosti) v rámci Národního strategického referenčního rámce.

projevujícího jednak záměr působit prostřednictvím intervenčního programu konkrétní změny, jednak přesvědčení, že se naplněním plánovaných kroků, činností a aktivit zamýšlené a požadované změny uskuteční. Smithova charakteristika intervenčního programu je zřetelně vázána na proces; v jeho pojetí ustupují jednotliví aktéři do pozadí a hlavní pozornost se upíná na jednotlivé podněty vyvolávající změnu a na schematické vyjádření postupu takovéto změny. Nastíněný přístup přitom není nijak ojedinělý, naopak – v evaluační literatuře zaujímá toto procesualistické pojetí poměrně významnou pozici. S jiným pojetím analýzy intervenčního programu vnímaného jako proces přišli např. McDavid a Hawthornová (2006), kteří chápou intervenční program jako „soubor aktivit, jež jsou vzájemně propojeny tak, aby směřovaly k dosažení vytyčených cílů“ (ibid.). Pakliže M. Smith věnuje aktérům intervenčního programu jen okrajovou pozornost a odkazuje na ně jen jako na obecně a vágně definovanou „cílovou skupinu“, v pojetí J. McDavida a L. Hawthornové se o aktérech v podstatě neuvažuje vůbec. Jejich zapojení do sítě procesních vazeb definujících intervenční program je přesto implicitně patrné v části, ve které J. McDavid a L. Hawthornová usuzují na smysl, účel a poslání intervenčních programů. Na pomyslnou otázku *Qui bono?* odpovídají poukazem na iniciátora daného intervenčního programu, který prostřednictvím jeho zavedení hodlá dosáhnout svých cílů, jež si vytyčil (popř. mu byly stanoveny). V tomto ohledu je patrná další z významných odlišností jejich pojetí; M. Smith totiž spatřoval smysl intervenčních programů v proměně podmínek cílové skupiny. Z obou uvedených přístupů je možné dovést dílčí záměr, že intervenční programy lze v obecné rovině popsat jejich (1) cíli (intervenční programy jsou vždy připraveny plnit definované cíle), (2) zvolenými prostředky k jejich dosažení (jaké konkrétní nástroje jsou použity k dosažení stanovených cílů) a také (3) průběhem implementace (tedy schopností zvolených prostředků vyvolávat očekávané změny a směřovat přímo k vytyčeným cílům).

Také Weissová (1989) zdůrazňuje na prvním místě jasný záměr, který intervenční programy sledují⁸, nicméně kromě toho C. Weissová rovněž poukazuje na reprodukovatelnost intervenčních programů předpokládající pečlivou a detailní dokumentaci všech jednotlivých prvků, popis klíčových procesů a rozbor jednotlivých funkcí. Charakteristickým rysem intervenčních programů je dále pevný mechanismus jejich fungování, realisté by řekli „pravidelnost“, která svými důsledky vede při jinak nezměněných podmínkách k identickým výstupům. Z analytického hlediska je v tomto ohledu významné, že působení intervenčního

⁸ Mimochodem je na místě poukázat na skutečnost, že tato intencionalita je ve větší či menší míře zřetelným prvkem téměř všech definic intervenčních programů.

programu může být kontrolovatelné a ověřitelné, a to jak z hlediska procesního (šetření totiž může být podrobena skutečné vykonávání jednotlivých činností), tak také z hlediska výstupů, výsledků a dopadů, jež daný intervenční program vyvolává (rovněž ty lze pomocí náležitě navrženého šetření ověřit).

Se zcela jiným vymezením intervenčních programů přicházejí Fitzpatricková, Sanders a Worthen (2004), kteří na rozdíl od M. Smithe, J. McDavida a L. Hawthornové či C. Weissové definují intervenční program jako komplex lidí, organizací a systému řízení, dostupných zdrojů; všechny tyto elementy přitom ve své spojitosti vyvíjejí podle autorů soustavný tlak a systematické úsilí k dosažení konkrétních výsledků (*outcomes*). Tato definice na rozdíl od všech předchozích s jednotlivými aktéry skutečně počítá, explicitně je pojmenovává a nezapomíná přitom ani na procesy zdůrazňované C. Weissovou. J. Fitzpatricková tak svou definicí vytváří představu intervenčního programu jako jakéhosi rámce, v němž jsou umístěni jednotliví aktéři spolu s jejich jednáním (úsilím vyvíjeným s cílem dosáhnout vytyčených výsledků). Její definice je vizí vycházející jednak z předpokládané informovanosti všech zapojených aktérů o dosahovaných cílech, jednak z předpokladu, že aktéři se shodují jak na samotných dosahovaných cílech, tak také na způsobech a formě jejich dosažení (Hyatt, Simons 1999). Takové pojetí však nachází svůj empirický korelát jen obtížně, neboť různost (a v praxi poměrně běžná antagoničnost) zájmů jednotlivých aktérů prakticky vylučuje obecnou shodu. Citovaná představa Fitzpatrickové dále vyvolává dojem, že intervenční program je společnou iniciativou zúčastněných aktérů. V této souvislosti je však třeba namítnout, že některé intervenční programy jsou připraveny a prosazeny jednou zainteresovanou stranou proti přání a preferencím jiných zainteresovaných stran. Dále není možné opomenout skutečnost, že individuální percepce cílů jednotlivých aktérů je v reálných podmínkách přirozeně odlišná (uživatelé sociálních služeb mohou celkem přirozeně sledovat jiné cíle než poskytovatelé těchto služeb), že v některých případech mohou být aktéři zapojeni do intervenčního programu, o jehož cílech nejsou informováni buď vůbec (např. pouliční kamerový systém), nebo mylně (např. Tuskegee experiment)⁹. Je nesporným přínosem, že J. Fitzpatricková ve svém pojetí intervenčních programů obrací pozornost k aktérům, nicméně její vize intervenčního programu může být efektivně naplněna jen v rámci některých

⁹ Na tomto místě ponecháváme stranou, zdali k těmto nedokonalostem v komunikaci mezi klíčovými aktéry dochází záměrně zkreslenými informacemi poskytovanými subjektem, který implementuje danou intervenci (např. kdy je zamlčován experimentální (pilotní) charakter testovaného intervenčního programu) či misinterpretací na straně příjemce daného sdělení. Rovněž zde ponecháváme stranou etický rozměr takového způsobu zavádění intervenčních programů.

evaluačních přístupů (blíže o nich viz samostatnou kapitulu), rozhodně však nelze její pojetí považovat za univerzálně použitelné.

Shadish, Cook a Levitonová (1991) poukazují při definici intervenčních programů na tři prvky, jež považují za klíčové. Prvním z nich je vnitřní struktura intervenčních programů, díky níž jsou dostupné zdroje využívány takovým způsobem, který vede ke konkrétním výsledkům.¹⁰ Nicméně způsob, jak intervenční programy pojímají W. Shadish, T. Cook a L. Levitonová, je mnohem komplexnější, neboť kromě těchto interních procesů vykonávaných v rámci daného intervenčního programu je rozlišován ještě další samostatný prvek, a to vnější vlivy, které napomáhají při (spolu)utváření intervenčního programu. Na jiném místě této práce je roli kontextu a vnějších podmínek věnována samostatná pozornost, a proto se zde omezíme jen na poznámku: zohlednění vnějších podmínek významným způsobem přispívá k realističtějším výsledkům evaluací intervenčních programů. Kontext totiž vždy zásadním způsobem ovlivňuje jak původní cíle, tak také funkce samotného intervenčního programu (Pawson, Tilley 1997). Třetím prvkem intervenčních programů je mechanismus, pomocí kterého dané intervenční programy dosahují svých cílů. Znalost tohoto mechanismu považují autoři za nezbytnou při optimalizaci intervenčního programu samotného, neboť bez ní by nebylo možné daný intervenční program adekvátně monitorovat, natož pak zlepšovat a dále rozvíjet.

Všem uvedeným pojetím je společný důraz na soustavnost působení (intervenční programy jsou dlouhodobé) a na zaměření na cíl, neboť intervenční programy mají svou konkrétní funkci (neinformovanost či neschopnost některých aktérů uvědomovat si tuto funkci, neschopnost ji popsat či kvantifikovat na tom nic nemění). Vyšší důraz na roli jednotlivých aktérů při provádění konkrétních procesů a při dosahování vytyčených cílů v definici J. Fitzpatrickové je z metodologického hlediska zásadní pro další analýzu, neboť upozorňuje na fakt, že jednotliví aktéři sehrávají roli samostatných entit, a později bude ukázáno, jak významná role v průběhu intervenčního programu a jeho analýzy jim přísluší. V případě některých intervenčních programů, jako je např. podpora regionálního rozvoje, vyžaduje identifikace a přesný a správný popis intervenčního programu častokrát větší úsilí, než samotné posouzení jeho výstupů a dopadů.

¹⁰ Tento prvek v zásadě odpovídá procesům, na něž upozornil již výše citovaný M. Smith.

Obecně lze tedy shrnout, že intervenční programy jsou souborem opatření, jejichž cílem je změna či formování chování jednotlivců. Přestože se lze s různými formami intervencí setkávat snad v každé oblasti společenského života, o intervenčních programech se obvykle hovoří v kontextu vzdělávání, sociální politiky, sociální práce, zdravotní politiky, kriminologie. Typickými příklady intervenčních programů jsou Medicare a Medicaid; za jistou formu intervenčního programu jsou považována i taková opatření, jako např. záporná daň z příjmu, rekvalifikační programy pro nezaměstnané či regulační poplatky ve zdravotnictví. V krajním případě by za intervenční programy mohly být považovány i takové nástroje a instituty, jako např. minimální mzda, státem podporované stavební spoření, novomanželské půjčky či bodový systém pro řidiče.

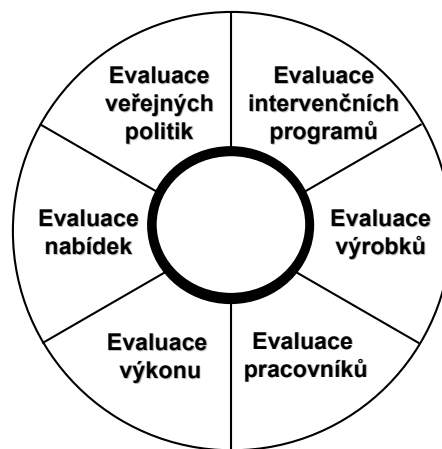
Evaluace v konkrétních souvislostech a ve spojitosti s aplikovanou oblastí tedy získává podstatně zřetelnější kontury, než jaké byly patrné v předchozí kapitole. Hallová a Hall (2004) považují evaluaci za svébytné šetření, které zjišťuje, do jaké míry byly naplněny záměry realizátorů intervenčních programů. Svým pojetím I. Hallová a D. Hall evaluaci notně zužují na pouhou evaluaci výsledků (*outcome evaluation*). Jak uvidíme později, takové pojetí však ani zdaleka nevyčerpává celou šíři úloh, jež jsou obvykle před evaluátory stavěny. Ucelenější pohled na podstatu evaluace lze zaznamenat v přístupu Guby a Lincolnové (1989), kteří definují evaluační činnosti jako aktivitu využívající vědecké postupy za účelem získávání a analyzování informací o obsahu, struktuře a výstupech zkoumaných programů, projektů a plánovaných intervencí. Extenzivita jejich pohledu na evaluaci je patrná hned ve dvou směrech: zaprvé ve funkčním vymezení, v jehož rámci se evaluace zaměřuje kromě výstupů také na obsah a strukturu intervenčního programu (čímž E. Guba s Y. Lincolnovou obrací pozornost k procesům, jež působí dokumentovatelné výsledky prováděných intervencí), zadruhé v obsahovém rozšíření povahy a typů subjektů, jež jsou předmětem evaluace. Zatímco I. Hallová a D. Hall vztahují aplikaci evaluací výlučně k intervenčním programům, E. Guba s Y. Lincolnovou považují za potenciální předmět evaluací rovněž projekty a další činnosti neintervenujícího charakteru (Remr 2008b).

Patton (2008) definuje evaluaci jako systematický sběr informací o působení, o charakteristikách a o výsledcích intervenčních programů s cílem provést jejich posouzení, zvýšit jejich kvalitu a efektivitu, poskytnout informace pro rozhodování o jejich další budoucnosti anebo zlepšit pochopení, jak tyto intervenční programy fungují. M. Patton sám vysvětluje, že svou definicí evaluace chtěl zdůraznit tři parametry, jež považuje za klíčové.

Jsou jimi (1) systematický sběr dat, (2) značná širší spektra problémů a témat, jež lze evaluaci podrobit a (3) pestrost jak v samotných evaluačních závěrech (tj. ve způsobech vyjádření věcného základu a přínosů), tak také v možnostech a formách jejich praktického použití. Tyto parametry by bylo možné vystihnout pojmy objektivita, univerzálnost a využitelnost.

V následujícím přehledu je stručně uvedena charakteristika jednotlivých oblastí spolu s klíčovými specifiky evaluace prováděné v jejich rámci.

Obrázek č. 5 Hlavní oblasti dle předmětu evaluace



Zdroj: autor

2.2.2 Evaluace výrobků (product evaluation)

Jedná se o oblast evaluací, která upíná svou pozornost k testování a srovnávání výrobků či služeb. Tato oblast evaluací poskytuje srovnávací zprávy o míře naplnění věcného základu definované skupiny výrobků či služeb. Obvykle se používají metodiky analýzy fyzikálních, chemických či senzorických vlastností a prostor pro uplatnění metod společenskovedního výzkumu je zde relativně malý. Příkladem evaluace výrobků jsou v USA velmi populární Consumer Reports, v České republice může jako příklad evaluace výrobků posloužit např. TEST DNES či některé testy výpočetní techniky prováděné ve speciálních testovacích laboratořích, kterými disponují vydavatelství některých odborných časopisů.

2.2.3 Evaluace pracovníků (personnel evaluation)

Tato oblast rozšířila evaluační portfolio poměrně nedávno, neboť starší evaluační literatura (z osmdesátých a z počátku devadesátých let) se touto oblastí nezabývá a nevěnuje jí žádnou pozornost (Fields 2002). Předmětem evaluace pracovníků jsou dovednosti a předpoklady zaměstnanců či uchazečů o zaměstnání vykonávat dané činnosti (Remr 2005). Tato oblast evaluací ve značné míře spoléhá na pozorování, metody měření pracovního výkonu (viz dále) a testy či interview, které osvědčují postojové orientace a znalosti respondentů o problematice, jež souvisí s vykonáváním dané pracovní činnosti.

2.2.4 Evaluace výkonu (performance evaluation)

Uvádí se, že tato oblast patří mezi nejdynamičtěji se rozvíjející oblasti evaluací (Knaap 2000). Zaměřuje se na poměrně specifickou a úzce vymezenou oblast metod měření a (dlouhodobého) sledování výkonu (Brinkerhoff, Dressler 1990). Některými autory bývá evaluace výkonu zaměřována s evaluací pracovníků, nicméně z dostupných pramenů je zřejmé, že oblast evaluace výkonu je širší a není s oblastí evaluace pracovníků totožná. Lze konstatovat, že evaluace výkonu se v obecné rovině zaměřuje především na ověřování, zdali bylo dosaženo vytyčených cílů, či nikoliv (Combs, Falletta 2000), přičemž své uplatnění nachází jak jako samostatná součást realizovaných intervenčních programů a projektů, které jsou vyhodnocovány pomocí tzv. KPI (*key performance indicators*), tak také v ostatních evaluačních oblastech, kterým napomáhá lépe stanovit míru dosažení definovaných výstupů a výsledků.

2.2.5 Evaluace nabídek (proposal evaluation)

Tato evaluační oblast rozvíjí postupy pro posuzování věcného základu a přínosu projektů či nabídek, které jsou předkládány v rámci poptávkových či výběrových řízení. Evaluace nabídek patří mezi oblasti evaluací, kde se primární výzkum (tak jak je chápán ve společenských vědách) téměř nevyskytuje. Pozornost se v této oblasti koncentruje na optimalizaci kritérií umožňujících adekvátní posouzení navrhovaných řešení, na metody stanovení nákladové efektivity a na způsoby dosažení porovnatelnosti předkládaných nabídek či projektů. Tento typ evaluace je přísně formalizovaný, dokonce i samotné posouzení věcného základu a přínosu je obvykle separováno do dvou samostatných částí/fází.

2.2.6 Evaluace politik (policy evaluation)

Oblast evaluace politik se zaměřuje na vybrané oblasti veřejných politik na základě předem definovaných kritérií (Remr, Kvapilová 1998). Svým zaměřením i mnohými metodami jde o oblast, která je velmi blízká evaluaci intervenčních programů, nicméně vzhledem k vyšší obecnosti a abstraktnosti politiky oproti intervenčnímu programu jsou zřejmé i mnohé diference.

2.2.7 Souhrn

Rozvoj evaluací je třeba chápat nejen jako produkt rozvoje v rámci jedné uvažované evaluační oblasti, ale především jako výsledný vektor, jehož velikost, směr a smysl je dán sumou dílčích pokroků v každé jednotlivé evaluační oblasti.

Přestože všech šest oblastí je mnohými považováno za jakýsi základ evaluačního oboru a i přes zjevnou podobnost všech uvedených oblastí, je třeba upozornit na skutečnost, že některé z nich k sobě mají blíže než jiné. V tomto ohledu lze poukázat na podobné evaluační mechanismy, které jsou patrné v oblastech evaluace projektů, evaluace intervenčních programů a evaluace politik. Zde jde vlastně jen o různou míru abstrakce, neboť projekt – program – politika leží na stejné vertikále. Podobnost mezi zbývajícími oblastmi již není tak patrná, neboť každá z dalších oblastí se vyznačuje zásadními specifiky.

Podobnost mezi evaluací projektů, evaluací intervenčních programů a evaluací politik má značné výhody v možnosti sdílení stejných metod a rozvíjení evaluace ve vzájemné součinnosti. Na druhé straně však existuje nebezpečí, že některé postupy, které jsou vhodné pro aplikaci na nižší úrovni obecnosti (např. u projektů), nebudou adekvátní při aplikaci na obecnější úrovni (politik).

3. Identifikace specifík evaluačních šetření

Tato část práce se zaměřuje na evaluační šetření, tedy na konkrétní formy aplikace obecného principu evaluace, který byl popsán v předchozí kapitole. Vzhledem k tomu, že výše uvedená abstraktní definice evaluace (tj. posuzování věcného základu a přínosu) neodkazuje ke konkrétním metodám a vzhledem ke značné diseminaci evaluace v celé řadě rozličných oblastí, dochází ke konfúzi její aplikované podoby (tj. evaluačních šetření) s jinými činnostmi podobného, leč nestejného charakteru. Cílem první podkapitoly v této části práce je proto poukázat na klíčové znaky evaluačních šetření, kterými se odlišují od některých jiných činností. Vysokou důležitost v tomto ohledu spatřujeme v odlišení evaluačních šetření od základního výzkumu, od monitoringu a od auditu. Druhá podkapitola této části pak přibližuje některá ze specifík, která provázejí realizaci evaluačních šetření. Pojednáno je o identifikaci zainteresovaných stran, o používání konceptuálních schémat a o provádění metaevaluací.

3.1 Věcná specifika evaluačních šetření

Jak poznamenává Stake (2004): „Každý člověk je sice rodilým evaluátorem, nikoliv však dobrým evaluátorem.“ Proto je při definici evaluace patrná snaha o to, aby se proces získávání podkladů (důkazů) a jejich vyhodnocení přiblížil formalizovaným postupům, které se vyznačují mj. systematičností, kontrolovatelností (přezkoumatelností) a vyloučením, či alespoň podstatnou redukcí subjektivismu. Tato snaha se v evaluačních šetřeních projevuje inkorporací některých principů vědeckého výzkumu, resp. metod a technik používaných ve společenskovědním výzkumu. Navzdory tomu však nelze evaluační šetření ztotožňovat s výzkumem. Druhou oblastí, vůči které se evaluační šetření vymezují, je monitoring. Ten s evaluačními šetřeními sdílí řadu metodických principů i některé aplikační akcenty. V některých případech je monitoring dokonce používán jako synonymum evaluací, což však považujeme za značnou simplifikaci, ospravedlnitelnou jen velmi specifickými podmínkami konkrétního kontextu. Třetí z oblastí je audit, který je pro některé evaluátory zdrojem inspirace např. v otázkách analýzy finančních toků a v otázkách udržitelnosti prováděných intervenčních programů. Empirismus a validitu důkazů používaných auditory netřeba s ohledem na výše uvedené dlouze rozvádět.

3.1.1 Základní výzkum ve společenských vědách

Evaluace již byly definovány v předchozích kapitolách, pokud jde o základní výzkum, ten může být pro účely této práce chápán jako snaha dospět k novým poznatkům zobecnitelným kladením jasně definovaných otázek při použití systematických a exaktních metod (Research Governance Framework 2005).

S vlastní definicí základního výzkumu ve společenských vědách přicházejí Singleton, Streits, Streits (1993), kteří poukazují v souvislosti s vědeckým poznáním na empirismus, objektivitu a ověřitelnost. V neposlední řadě také Kerlinger a Lee (2000) výstižně definují základní výzkum jako „systematické, kontrolované, empirické, hodnotově neutrální, veřejné a kritické zkoumání přirozených jevů [, které] je vedeno teorií a hypotézami o předpokládaných vztazích mezi těmito jevy“. Z uvedených definic je tedy patrné, že empirismus, objektivita a kontrolovatelnost jsou klíčovými charakteristikami, jež definují podstatu základního výzkumu. Zdá se tedy, že v těchto základních charakteristikách není mezi základním výzkumem a evaluačním šetřením patrný žádný podstatný rozdíl; obě oblasti se hlásí ke stejným hodnotám a usilují o jejich maximalizaci.

Diference se však objevují v oblastech souvisejících s účelem provádění základního výzkumu a evaluačních šetření. Stufflebeam a Shinkfield (2007) při porovnávání jejich cílů zdůrazňují, že smyslem evaluačních šetření není na rozdíl od výzkumu dokazovat (*prove*), ale zlepšovat (*improve*). Odkazují tím na aplikovaný charakter evaluačních šetření, která neslouží k budování teorie a k verifikaci formulovaných hypotéz, ale mají za cíl pomáhat v konkrétních situacích (např. při implementaci intervenčních programů či při zkvalitňování výuky). Ačkoliv tedy základní výzkum a evaluační šetření používají stejné metody (dotazníková šetření, interview, pozorování), spějí k rozdílným závěrům. Zatímco mezi hlavní cíle základního výzkumu patří získávání poznatků, kumulace vědění a rozvoj teorie, cílem evaluačních šetření je napomáhat v rozhodování. Alkin (1973) definuje evaluaci jako proces zjišťování, ve kterých oblastech je třeba učinit nějaké rozhodnutí, volby druhu vhodných informací, sběru a analýzy informací a přípravy souhrnných dat užitečných při rozhodování mezi dostupnými alternativami.

Cílem evaluací je „poskytovat zpětnou vazbu a umožnit dosažení praktických a konkrétních cílů“ (Isaac, Michael 1997). Evaluační šetření zahrnují takové činnosti, jako např. analýzu

efektivity či hodnocení adekvátnosti použitých prostředků vzhledem ke stanoveným cílům. Primárním cílem evaluačních šetření je poskytovat informace využitelné v dané rozhodovací situaci, v daném čase a přesně specifikovaným subjektům; bezprostřední využitelnost výsledků daného evaluačního šetření je v dané aplikační situaci tedy klíčová, zatímco v případě základního výzkumu není tento parametr nijak významný (Donaldson, Christie, Mark 2009). V tomto ohledu tedy není nijak překvapivé, že se evaluační šetření od základního výzkumu odlišují nejen svými cíli, ale také cílovou skupinou. Berk a Rossi (1999) poznamenávají, že zatímco typickým uživatelem výsledků evaluačních šetření jsou rozhodovatelé, v případě základního výzkumu jsou to většinou ostatní výzkumníci a vědecká komunita.

Evaluační šetření vždy berou v potaz konkrétní kontext, ve kterém jsou prováděna. Podobně také jejich výsledky jsou platné vždy a pouze v daném kontextu; naproti tomu základní výzkum se snaží o formulaci obecně platných závěrů (či snad dokonce zákonitostí), které jsou na konkrétním aplikačním kontextu nezávislé. Weissová (1998) rozdíl mezi evaluačním šetřením a základním výzkumem vystihuje poukazem na způsob použití dosažených výsledků. Zatímco v případě základního výzkumu je ponecháno na přirozených a spontánních procesech, jak budou či nebudou výsledky použity, evaluační šetření je vždy iniciováno s konkrétním cílem.

V základním výzkumu se hypotézy odvíjejí od teorie a od stavu poznání, formuluje je obvykle samotný výzkumník a své výzkumné cíle a záměry koriguje na základě doporučení svých kolegů (či oponentů); samostatné uchopení zkoumaného problému je v základním výzkumu běžné. Naproti tomu v případě evaluačních šetření se jejich cíl odvíjí od zadání a od faktické situace, v níž se nacházejí jednotlivé zainteresované strany. Evaluátoři navrhnou design evaluačního šetření, konstruují příslušné instrumenty, analyzují získaná data a připravují závěrečné výstupy, nicméně každý z kroků vždy konzultují se zadavatelem dané evaluace, jenž by měl rovněž schvalovat každé významné rozhodnutí týkající se jejího provedení. Pokud by realizátor evaluačních šetření postupoval způsobem, jaký je běžný v základním výzkumu (tedy samostatně), porušil by tím oborové normy, které definují standardy dobré praxe.

Obvykle uváděná symbióza teorie a empirie (E. Babbie; F. Kerlinger, H. Lee) vedoucí k vzájemnému obohacování, kdy v jednom směru je základní výzkum odvozen z teorie a

v druhém, opačném, směru přispívá k jejímu obohacení, je při evaluačních šetřeních naplňována jen jednosměrně. Mohou vycházet z teorie (ačkoliv ani to není v praxi vždy pravidlem), není však jejich prvoplánovým cílem přispívat k jejímu budování, obohacení či k její verifikaci. Znalost teorie pochopitelně umožňuje lépe chápat případné vztahy mezi sledovanými jevy, dovoluje precizněji operacionálně definovat klíčové pojmy a snáze formulovat hypotézy, konstruovat modely a hledat příčiny. Davidsonová (2005) v této souvislosti poukazuje na tři metodologické úlohy, které odlišují evaluační šetření od základního výzkumu. Jde o zvážení důležitosti, určení věcného základu (*merit*) a syntézu získaných poznatků.

Použitý design evaluačních šetření je spoludeterminován praktickými okolnostmi (čas, rozpočet apod.) v podstatně větší míře než v základním výzkumu a kvůli ingerenci zainteresovaných stran je častokrát také kompromisem „politickým“. Při evaluačním šetření je poměrně běžný nesoulad v zájmech jednotlivých zainteresovaných stran a v cílech, které sledují, což v základním výzkumu zcela chybí (Kliner, Morgan 2000).

Evaluační šetření jsou prováděna účelově – jejich zamýšleným posláním je usnadnit rozhodovací proces (resp. umožnit, aby se rozhodovací proces mohl opřít o exaktní a empirické údaje a aby probíhal na věcném základě). Z tohoto, jistě neúplného, výčtu některých funkcí, které evaluační šetření plní, zřetelně vyplývá, že motivace k jejich provedení se značně liší od důvodů realizace základního výzkumu. V evaluačních šetřeních nejde o prvoplánový sběr či kumulaci poznatků kvůli uspokojení zvědavosti.

V neposlední řadě považujeme za důležité upozornit, že základní výzkum klade důraz na vnitřní validitu a upřednostňuje metodu před výsledkem. Naproti tomu aplikovaný charakter evaluačních šetření je podtržen důrazem na vnější validitu. Mnohem ochotněji než v základním výzkumu se u evaluačních šetření připouští kompromisní charakter designu – ten je determinován nejen principiálními otázkami metodologického charakteru, ale srovnatelnou vahou také praktickými okolnostmi (závažnost podporovaného rozhodnutí, časové a finanční limity atd.).

Tabulka č. 1 Srovnání vybraných parametrů základního výzkumu a evaluačních šetření

Základní výzkum	Evaluační šetření
Získávání nových poznatků; rozšiřování znalostí; rozvoj teorie	Získávání údajů pro rozhodování zainteresovaných stran
Snaží se dospět k obecně platným zákonitostem, tvrzením o vztazích mezi jevy (proměnnými); monotetičnost	Snaží se ověřit vlastnosti předmětu evaluace; idiografičnost
Klade důraz na zobecnitelnost závěrů na jiné populace	Klade důraz na užitečnost zjištění; získání prakticky významných poznatků
Přináší zjištění, která mají dlouhodobý dopad	Přináší zjištění, která mají krátkodobý dopad
Klade důraz na vnitřní validitu; upřednostňuje metodu před výsledkem	Klade důraz na vnější validitu; akceptuje, že volba metod je kompromisem
Vychází z jedné úrovně analýzy	Vychází z víceúrovňové analýzy

Zdroj: autor na základě pramenů citovaných v textu

Souhrn

Rekapitulace významných odlišností mezi základním výzkumem a evaluačním šetřením není samoúčelná. V konkrétní praxi totiž existují mnohé případy, kdy je pod hlavičkou evaluačního šetření realizován typ aktivit, které naplňují mnohé znaky základního výzkumu a neplní tak původní očekávání. Jak je patrné v kapitole sedmé, velmi často se evaluační šetření snaží o získávání nových poznatků, o rozšiřování obecných znalostí či o souhrn a systematizaci stávajících poznatků.

I přes uvedené rozdíly mezi základním výzkumem a evaluačními šetřeními lze však vyzorovat společný zájem o metodologii. Jak základní výzkum, tak také evaluační šetření totiž rozvíjejí metody používané k ověření definovaných cílů. V tomto ohledu platí, že metody sběru dat (pozorování, interview, dotazníková šetření), techniky výběru (pravděpodobnostní výběrové techniky, kvótní výběr, výběr úsudkem/záměrný výběr) i analytické postupy (testování hypotéz, analýza časových řad) využívané při zpracování shromážděných dat nejsou nijak odlišné, liší se jen způsoby jejich využití.

Srovnání evaluačních šetření a základního výzkumu je prospěšné také proto, že poukazuje na podobnosti a rozdíly v požadovaných schopnostech a kompetencích. Zdá se tedy, že provádění evaluačních šetření vyžaduje mnohem větší míru sepětí realizátora s danou konkrétní situací; zřetelná je rovněž vyšší míra provázanosti (či interakce) s ostatními

stranami zainteresovanými na daném evaluačním šetření. Ukazuje se rovněž, že realizátoři evaluačních šetření jsou méně citliví vůči metodám a analytickým postupům dominantně zaměřeným na dosahování či posilování vnitřní validity (méně statistické významnosti, více interpretačního významu). V neposlední řadě stojí za zmínku faktická absence zobecňujících závěrů, či snah o takovou formu zobecnění, která přesahuje rámec daného evaluačního šetření. I přes uvedené rozdíly platí, že požadavky kladené na realizátory nejsou natolik odlišné, aby vedly k absolutní dysfunkci evaluátorů a výzkumníků. Koneckonců první evaluační šetření v 60. letech v USA byla prováděna právě výzkumníky z oblasti základního výzkumu (House 1993, Suchman 1967, Alkin 2004).

3.1.2 Monitoring

Monitoring je systematický sběr dat dokumentující pokrok daného intervenčního programu či postup jednotlivých fází jeho implementace. Monitoring tedy sleduje průběžný vývoj (rozvoj) intervenčního programu s indikacemi o dosahování cílů, plnění plánu, čerpání rozpočtu a dodržování termínů. Jeho cílem je (1) průběžné zjišťování pokroku v realizaci intervenčního programu, (2) srovnání aktuálních informací s výchozím, předpokládaným plánem a (3) včasné upozornění na případné odchylky, jež následně umožňuje přijetí konkrétních opatření. Monitoring je primárně určen sponzorům daného intervenčního programu a jeho managementu, nicméně využití poznatků plynoucích z prováděného monitoringu je v mnoha případech rozsáhlejší. Monitoring je kontinuální, nebo periodicky opakovaný; v praxi se může vyskytovat i kombinace obou přístupů.

V tomto ohledu se ukazuje zásadní rozdíl mezi monitoringem a evaluačním šetřením: zatímco monitoring provází daný intervenční program po celou dobu jeho implementace (resp. po dobu jeho stanovené etapy), evaluační šetření se provádí jako nárazová, jednou či vícekrát opakovaná činnost. To vede k diametrálně odlišnému způsobu řešení monitorovacích a evaluačních zadání: monitorovací ukazatele se častokrát opírají o faktografické indikátory, jejichž zjištění či výpočet je (téměř) automatický a vyplývá z průběhu prováděných činností. Není tedy třeba koncipovat robustní metodiku sběru dat, jejich transformace a analýzy; rovněž nároky na získávání a zpracování takovýchto informací bývají relativně nízké.

Monitoring nemůže sám o sobě přinést informace či zmapovat kvalitu realizovaného intervenčního programu (či projektu), nemůže ani objasnit, proč daný intervenční program funguje, či proč ve svém fungování selhává. Tyto otázky jsou vyčleněny evaluacím, které však mohou základní poznatky čerpat právě z monitoringu. Monitoring tedy primárně slouží k identifikaci vznikajících problémů, které vyžadují jednak okamžitou nápravu a pravděpodobně (při opakovaném výskytu či při neúčinnosti zvolených korekcí) také navazující hloubkovou analýzu (evaluační šetření nebo audit). Je tedy zřejmé, že monitoring napomáhá rozhodovací činnosti na operativní úrovni a přináší podklady pro rozhodnutí o dílčích opatřeních. Pro zásadní rozhodnutí strategického charakteru však nejsou podklady z monitoringu dostatečné. V tomto ohledu můžeme tedy považovat monitoring za doplňkovou činnost při provádění komplexněji pojatých evaluačních šetření.

Zásadní rozdíl mezi monitoringem a evaluačním šetřením je třeba dále hledat v jejich účelu. Monitoring je prováděn s cílem ujistit se, že vše probíhá tak, jak má (tj. včas, v rámci rozpočtu atd.), zatímco evaluační šetření má za cíl umožnit dosažení cíle, změnit stávající praktiky, pokud nefungují nebo pokud jsou neefektivní, a poskytovat podporu pro vedení intervenčního programu.

Z hlediska konkrétních očekávání vyhovuje monitoring a evaluační šetření poněkud odlišným požadavkům. Zatímco monitoring si lze snadno představit jako víceméně automatizovaný systém průběžného či periodického ověřování sady předem definovaných indikátorů, evaluační šetření jsou založena na procesech, které jsou mnohem citlivější a vnímavější vůči danému předmětu evaluace a vůči konkrétnímu kontextu.

Tabulka č. 2 Srovnání vybraných parametrů monitoringu a evaluačních šetření

Monitoring	Evaluační šetření
Neobsahuje hodnotovou složku	Obsahuje jak věcný základ, tak také přínos
Nebere v potaz velikost odchylky či význam (a možné dopady) odchýleného parametru	Dopady odchylek na fungování programu a jeho schopnost plnit a dosahovat svých cílů je předmětem posouzení
Je průběžný	Je jednorázové
Vnitřní odpovědnost	Vnější odpovědnost
Slouží manažerům intervenčního programu	Slouží různým (mnohým) zainteresovaným stranám
Je spojen se slabým zájmem (rutinní řízení, indikace problémů)	Je spojen se silným zájmem
Operativní rozhodnutí	Strategické rozhodnutí

Zdroj: autor na základě pramenů citovaných v textu

Souhrn

Některé prameny zaměřují monitoring a evaluační šetření, resp. tyto dvě činnosti/oblasti nerozlišují s potřebnou přesností. Monitoring bývá některými autory a institucemi (např. General Audit Office) považován za součást evaluačních šetření, někteří dokonce monitoring charakterizují jako nejčastěji prováděný typ evaluačních šetření. Monitoring je rovněž popisován jako jedna z evaluačních technik, a dokonce i Rossi, Lipsey, Freeman (2004) používají pojmy evaluace procesu a monitoring programu jako synonyma. S ohledem ke skutečnosti, že monitoring bývá v mnoha případech používán samostatně, bez navazujících činností evaluačních šetření, považujeme jej za samostatnou oblast, jež evaluaci může pomoci, v některých případech může být zahrnuta do jejího rámce, přesto však zůstává autonomní sférou.

Oba typy činností považujeme za nutné oddělit s ohledem ke konkrétním procesům prováděným v rámci monitoringu a evaluačních šetření. Jak vyplynulo z uvedeného srovnání monitoringu a evaluačních šetření, jsou obě činnosti natolik rozdílné, že jejich směšování považujeme za neudržitelné.

3.1.3 Audit

Audit lze definovat jako systematický proces objektivního získávání a posuzování důkazů týkajících se ekonomických činností a událostí. Cílem tohoto procesu je nejprve zjistit stupeň souladu mezi tvrzeními o činnostech či událostech a stanovenými kritérii a poté sdělit výsledek zainteresovaným stranám (American Accounting Association 1973).

Cílem auditu je ujištění se o tom, že prováděné činnosti jsou v pořádku; audit nezjišťuje, které postupy fungují, a které ne, ale ověřuje, zdali se v daném případě postupuje správně, eventuálně podle nejlepších známých praktik (Dvořáček 2003). Evaluační šetření se naproti tomu více koncentrují na otázky, zdali jsou vykonávány správné činnosti (tj. ty, které vedou k cíli). Chelimská (1985) poznamenává, že cílem auditu (oproti evaluacím) není zjišťovat, kolik lidí se účastní daného intervenčního programu, ani ověřovat, jakým způsobem se na činnosti sledovaného intervenčního programu podílejí jednotlivé zainteresované strany či jaké přínosy jim intervenční program poskytuje. Těmito deskriptivními funkcemi se evaluační

šetření odlišují od auditu. Další principiální rozdíl spatřuje E. Chelimská v tom, že zatímco evaluační šetření zjišťují, které z pozorovaných změn jsou důsledkem působení intervenčního programu, a které nikoliv (kladou si tedy otázky o vztahu příčin a následků), audit se obdobnými kauzální hypotézami nikdy nezabývá.

V čem však lze spatřovat vysokou podobnost, to jsou otázky normativní povahy. V tomto ohledu mají evaluační šetření a audit mnoho společného: předmětem normativních otázek je, zda intervenční program funguje tak, jak bylo zamýšleno (Schwandt, Halpern 1988). Přestože evaluační šetření i audit k zodpovězení používají jiné metody, spějí v tomto ohledu ke stejnému cíli. Audit se obvykle zaměřuje na velmi úzký výsek činnosti posuzovaných programů, a to na jejich finanční stránku. Naproti tomu předmětem zájmu evaluačních šetření jsou výstupy a dopady sledovaných programů, což sice mnohdy zahrnuje i finanční aspekt jejich působení, nicméně ne v takovém detailu jako u auditu.

Audit se však může zaměřovat i na jiné než finanční aspekty. V takovém případě se hovoří o „věcném“ auditu či specifitěji o performance assessment (hodnocení výkonu). Příkladem takto pojatého auditu je audit operací, který je formální a systematickou revizí vykonávanou proto, aby bylo možné určit, do jaké míry jsou naplňovány stanovené cíle a které konkrétní podmínky mají být zlepšeny.

Tabulka č. 3 Srovnání vybraných parametrů auditu a evaluačních šetření

Audit	Evaluační šetření
Základními hodnotami jsou finanční atest, soulad s předepsanými pravidly a jistota	Základními hodnotami jsou účinnost, věcný základ a přínos
Nezkoumá fungování intervenčního programu	Vytváří představu, jak intervenční program funguje
Konfrontuje skutečný stav s předepsanou normou	Využívá kombinace mnoha pohledů na činnost a výsledky intervenčního programu
Stupeň standardizace je velmi vysoký	Stupeň standardizace je relativně nízký
Samostatnost auditora v rozhodování není vůbec možná	Samostatnost evaluátora v rozhodování je poměrně vysoká

Zdroj: autor na základě pramenů citovaných v textu

Audit tedy přináší systematický a exaktní přístup k poznávání a zlepšování efektivity při řízení rizik a procesu řízení. Finanční audit se zaměřuje na analýzu souladu s pravidly a

ustanoveními, zatímco audit výkonu (*performance*) se soustředí na zkoumání relevance, hospodárnosti, efektivity a účinnosti.

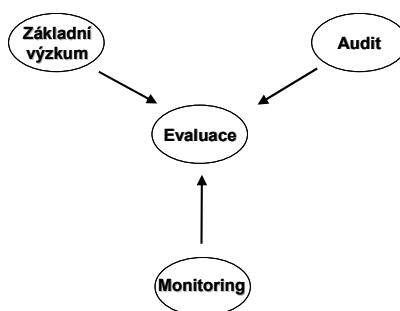
V konkrétní rovině sdílí evaluační šetření s auditem celou řadu principů a postupů. Předně je to systematický a empirický proces, ale není zde žádná ambice k zobecnění; nejde o sběr primárních dat. Dalším společným prvkem je důraz na objektivitu a nezávislost provádění obou činností a v neposlední řadě také samotný způsob vyhodnocení, při kterém se audit i evaluační šetření zaměřují na zjištění stupně souladu mezi skutečným stavem a stanovenými kritérii.

V některých konkrétních případech představuje audit jednu z metod použitých při evaluačním šetření intervenčního programu; zejména v evaluačních šetřeních zaměřených na zdravotnická zařízení a na poskytování zdravotní péče je audit poměrně běžnou součástí těchto činností.

3.1.4 Souhrn

Pomocí srovnání s ostatními podobnými činnostmi jsme poukázali na některé charakteristické rysy evaluačních šetření. Cílem uvedené analýzy není uměle oddělovat evaluační šetření od ostatních činností (jako např. od základního výzkumu, od monitoringu či od auditu), ale poukázat na některé specifické okolnosti použití evaluace v konkrétních aplikačních podmínkách. Kromě mnohých diferencí však tento rozboru poukazuje také mnohé podobnosti mezi sledovanými oblastmi a nastiňuje řadu aspektů, v nichž se jednotlivé činnosti vzájemně doplňují.

Obrázek č. 6 Vztah evaluačních šetření k ostatním disciplínám



Zdroj: autor

3.2 Procesní specifika realizace evaluačních šetření

Kromě toho, že se evaluační šetření vyznačují řadou specifických charakteristik, z nichž na některé jsme poukázali v předchozí kapitole, je i jejich provádění provázeno celou řadou odlišností, které se v jiných oblastech a disciplínách buď vůbec nevyskytují, nebo mají zásadně odlišný charakter. V následující části jsou prezentovány tři typy činností, které se významnou měrou podílejí na celkovém provedení evaluačních šetření, neboť každý z uvedených typů zásadně ovlivňuje hned několik etap jejich realizace. Pozornost je tedy věnována identifikaci a analýze zainteresovaných stran, přípravě a používání konceptuálních schémat a v neposlední řadě také metaevaluacím.

3.2.1 Zapojení zainteresovaných stran

Aplikovaný charakter evaluačních šetření podtrhuje jejich zakotvení v síti jednotlivců, skupin či organizací, kterých se dotýkají a jejichž zájmy může jejich realizace buď podpořit, nebo omezit, či dokonce ohrozit. Pojem *stakeholders* dnes již neoznačuje nestranné, vážené osoby držící kůl se sázkami (*stake*), jako tomu bývalo při soubojích středověkých rytířů (Patton 2008). V kontextu evaluačních šetření jsou takto označovány subjekty angažované v dané evaluační situaci, neboť mohou být prováděným evaluačním šetřením či jeho závěry přímo či nepřímo dotčeny (Fitzpatricková, Sanders a Worthen 2004;)¹¹. Podobně také Rossi, Lipsey a Freeman (2004) definují zainteresované strany jako „jednotlivce, skupiny či organizace, které mají významný zájem na tom, jak dobře intervenční program funguje, např. ti, kteří mají rozhodovací kompetence ve vztahu k programu, sponzoři a donátoři, administrátoři či pracovníci a klienti či zamýšlení beneficianti“.

Identifikace zainteresovaných stran¹² je tedy jednou ze základních (a raných) fází každého evaluačního šetření. Identifikaci zainteresovaných stran předpokládají také Standardy evaluace intervenčních programů, které vymezují tuto činnost hned v prvním ustanovení (cit. dle Sanders 1994).

¹¹ Tato definice je víceméně totožná se způsobem, jak novodobě pojem definoval R. E. Freeman (1984) ve své publikaci *Strategic Management. A Stakeholders Approach*.

¹² Plurál je zvolen záměrně. Jednak to odpovídá originálním pramenům v angličtině, jednak je tím neustále připomínáno, že zainteresovaných stran je při každém evaluačním šetření více.

Z uvedených definic jsou patrné tři základní pojmy, které vystihují roli zainteresovaných stran v evaluačním šetření, jde o zájem, angažovanost a interakci. V následujících odstavcích se k jednotlivým pojmům vrátíme s podrobnějším rozbohem.

Evaluační šetření přinášejí mezi zainteresované strany specifický prvek, který v analogické situaci základního výzkumu není přítomný, a to „zájem“. Jak je zdůrazňováno v mnoha pramenech, evaluační šetření jsou nadána silným vlivovým potenciálem, který je častokrát silnější než potenciál, jímž disponují ostatní typy šetření (Morris, Fitz-Gibbon, Herman 1987). Význam evaluačních šetření a jejich závěrů pro jednotlivé zainteresované strany přitom vyplývá ze samotného primárního účelu, pro který jsou tato šetření prováděna, tedy zejména zjištění účinnosti, odhad dopadů či ověření, zdali bylo dosaženo vytyčených cílů. Dále platí, že v aplikační praxi se od výsledků evaluačních šetření obvykle odvíjí celá řada klíčových rozhodnutí s tak vážnými důsledky, jako je např. budoucí podoba a uspořádání vzdělávacích kursů; rozdělení pracovních týmů na ty, které vyhovují stanoveným parametrům (a zaslouží si tak vyšší ohodnocení), a které nikoliv; pokračování či zastavení hodnoceného intervenčního programu atd.

Podle House (1980) reprezentuje zájem zobecněné prostředky k cílům, které jednotlivci či skupiny sledují. Předmět zájmu tedy zvyšuje či zlepšuje příležitosti k dosažení stanovených cílů. Jednotlivé zainteresované strany se pak zapojují do intervenčního programu spolu se svými zájmy a v souladu se svými možnostmi tyto zájmy prosazují. V některých případech se přitom dopouštějí omylů. Prosazují třeba taková opatření, která směřují proti jejich zájmům nebo volí takové varianty, které nejsou efektivní při dosahování daných cílů.

Zájmy je možné dělit na explicitní (deklarované, zjevné, vyjádřené, sdělované) a implicitní (neuvědomělé subjektu samotnému, neznámé pro okolí – skryté či zastřené, zamlčené). Z jiného pohledu lze rozlišovat zájmy vztahované k samotnému intervenčnímu programu (zájem na vyvolání změny v chování klientů, dosažení zlepšení v sociální situaci dané skupiny obyvatel, snížení nezaměstnanosti apod.) a zájmy týkající se prováděného evaluačního šetření (zájem na jasné identifikaci silných a slabých stránek, na rychlém provedení, na potvrzení dosaženého pokroku daného šetření apod.).

Evaluátoři nejenže usilují o takový návrh designu evaluačního šetření, který nediskriminuje žádnou ze zainteresovaných stran přímo zapojených do procesu plánování daného šetření, ale současně zvažují zájmy také dalších dotčených subjektů, které se dané evaluační situace účastní nepřímo (House, Howe 1999). Svě zájmy na výsledcích prováděného evaluačního šetření totiž mají jak sponzoři (kteří se chtějí ujistit o efektivitě jimi financovaného intervenčního programu) a manažeři intervenčního programu (kteří chtějí prokázat, že intervenční program dosahuje pod jejich vedením potřebné efektivitě), tak také pracovníci intervenčního programu (kteří prostřednictvím evaluačního šetření chtějí obhájit užitečnost a smysluplnost své pracovní pozice a ospravedlnit svůj vlastní výkon), a dokonce i případní beneficianti (kteří si např. nepřejí zrušení daného intervenčního programu). Zatímco sponzoři a manažeři intervenčního programu mohou též vystupovat v roli zadavatelů daného evaluačního šetření a mohou tak se konkrétní zájmy snadno prosadit prostřednictvím své angažovanosti při koncipování designu daného evaluačního šetření, ostatní zainteresované strany (pracovníci, beneficianti) tuto možnost nemají. Při provádění evaluačních šetření je tedy nutné, aby evaluátoři přijali taková opatření, která v první fázi povedou k sumarizaci deklarovaných zájmů (popř. dokonce i k odhalení či alespoň k odhadu zájmů skrytých). Ve fázi druhé je třeba tato opatření završit vypracováním takové strategie evaluačního šetření, která zajistí v maximální možné míře rezistenci klíčových poznatků a závěrů vůči případným vychýlením způsobeným různou intenzitou prosazování zájmů jednotlivými zainteresovanými stranami.

PŘÍPADOVÁ STUDIE:

V rámci probíhajícího věrnostního programu bylo provedeno formativně založené evaluační šetření zaměřené na evaluaci procesu. Na základě údajů z průběžného monitoringu programu se totiž ukázalo, že daný věrnostní program nenaplnuje cíle, které byly předem stanoveny (počty zákazníků zapojených v daném věrnostním programu, změna ve výši tržeb, vyzvednuté odměny za opakované nákupy a některé další). Primárním cílem evaluačního šetření bylo ověřit, zdali jsou jednotlivé činnosti nutné pro dosažení stanovených cílů vykonávány adekvátním způsobem. V rámci takto pojatého evaluačního šetření bylo identifikováno několik zainteresovaných stran (viz souhrnná tabulka) a následně byly identifikovány konkrétní zájmy, které jednotlivé zainteresované strany mají ve vztahu k věrnostnímu programu a k vybraným okolnostem jeho průběhu. Následující tabulka přináší přehled některých zájmů vybraných zainteresovaných stran (výčet není úplný):

Tabulka č. 4 Zájmy vybraných zainteresovaných stran

Zainteresovaná strana	Zájem
Sponzor (organizátor věrnostního programu; management obchodního řetězce)	<ul style="list-style-type: none"> - dosáhnout vyšších tržeb - udržet stávající zákazníky - získat nové zákazníky - stimulovat vyšší průměrnou útratu stávajících zákazníků / zvýšit share of wallet - zlepšit vnímání vlastní značky - minimalizovat náklady na realizaci programu (nízká cena incentive, barter, sdílení nákladů s dalšími partnery) - spokojenost klienta
Oddělení zodpovědné za věrnostní program (zaměstnanec sponzora; management věrnostního programu)	<ul style="list-style-type: none"> - vyšší míra zapojení oproti ostatním věrnostním programům - minimalizace negativních jevů provázejících věrnostní program - samoobslužnost programu (běží sám o sobě bez významného nasazení ostatních pracovníků – administrátorů) - snadná a fungující logistika (dostatek informačních materiálů, odměn apod.) - spokojenost sponzora
Klient (zákazník sponzora; nakupující; potenciální účastník věrnostního programu)	<ul style="list-style-type: none"> - maximalizovat získávanou hodnotu - jednoduché podmínky účasti v programu (minimalizace vstupních překážek) - zabezpečení anonymity (žádná osobní data)
Administrátor (prodejci organizátora; street-level; pracovníci; pokladní)	<ul style="list-style-type: none"> - jednoduchost administrace (aby věrnostní program nezdřoval od plnění ostatních pracovních povinností) - snadno zapamatovatelná pravidla - vyšší tržby (a z toho plynoucí vyšší osobní ohodnocení) - vyhnout se odmítnutí ze strany klienta
Vedoucí prodejny (zaměstnanec sponzora; přímý nadřízený administrátora)	<ul style="list-style-type: none"> - spokojenost sponzora

Zdroj: Evaluace věrnostního programu provedená autorem

Závěry a poučení:

a) Zadavatel evaluačního šetření (oddělení zodpovědné za věrnostní program) v první fázi vůbec nebral v potaz některé zainteresované strany (např. na přítomnost a roli vedoucích prodejen upozornilo až provedené evaluační šetření) a některé z různých zainteresovaných stran zadavatel ve svých původních aktivitách mezi sebou nerozlišoval (na relevanci odlišování vedoucích prodejen od administrátorů upozornilo rovněž až provedené evaluační šetření, v jehož rámci byla provedena podrobná analýza zájmů jednotlivých zainteresovaných stran).

b) Při návrhu věrnostního programu nebyly identifikovány zájmy některých zainteresovaných stran (prodejci) a nebyly alokovány prostředky pro uspokojení těchto zájmů (motivační program pro prodejce).

c) Jednotlivé zainteresované strany nebyly informovány o zájmech (a snahách) ostatních zainteresovaných stran.

Shrnutí případové studie:

Pečlivá identifikace jednotlivých zainteresovaných stran a důkladná analýza jejich zájmů umožnila poukázat na slabá místa v implementaci daného věrnostního programu a zvolit příslušná opatření k jejich nápravě. Případová studie dokumentuje smysluplnost Standardů, které přisuzují identifikaci zainteresovaných stran mimořádný význam. Z provedené případové studie je patrné, že to byla právě identifikace jednotlivých zainteresovaných stran (a jejich zájmů), co významným způsobem přispělo k využitelnosti provedeného evaluačního šetření a následném zlepšení daného věrnostního programu.

Konkrétní opatření, kterými lze uvedených cílů dosáhnout, nacházíme ve dvou rovinách, a to:

(1) v rovině výzkumného designu – zde jde především o získání adekvátního přístupu k zainteresovaným stranám; o volbu triangulačního designu, který umožňuje dosáhnout vyšší validity a s jehož pomocí lze snáze identifikovat případné rozpory v získávaných údajích;

(2) v rovině měření – instrumenty používané při výzkumu zainteresovaných stran by měly obsahovat takové prvky, které by umožnily efektivní obranu před zastíráním skutečných motivů a před předstíráním stavů, jež nekorrespondují se skutečností (Remr 2009a). Přístup založený na nepřímé formě usuzování podle sady indikátorů, který používá většina ostatních společenskovedních výzkumů, není v případě evaluačních šetření vždy adekvátní. Zvolená konceptualizace sledovaných jevů a konkrétní způsob jejich operacionalizace může totiž významným způsobem determinovat povahu získávaných výsledků a jejich validitu. Objektivně navíc dochází k variantnosti¹³ v definici způsobu měření jednotlivých jevů, což v případě evaluačního výzkumu může vyvolávat konflikt mezi jednotlivými aktéry a v konečném důsledku přispět k interpretační spornosti zjištěných výsledků (Bingham, Felbinger 2002)¹⁴.

Je třeba mít na paměti, že samotná identifikace jednotlivých zainteresovaných stran je jen prvním krokem k adekvátní analýze problematik zájmů v prováděných evaluačních šetřeních. Dále je nutné brát v potaz, že různé skupiny zainteresovaných stran mohou disponovat vlastními soubory nástrojů a strategií, pomocí kterých dosahují svých konkrétních zájmů týkajících se dalšího směřování posuzovaného intervenčního programu. Sponzoři a management (resp. zadavatelé evaluací) mohou např. své zájmy uplatňovat volbou či prosazením specifických témat vyjádřených v cílech, jež by dané evaluační šetření mělo naplňovat (Fitz-Gibbon, Morris 1987); tím mohou významným způsobem predeterminovat výsledek evaluačního šetření. K vychýlení může dojít také výběrem kritérií – tímto způsobem mohou evaluační šetření ovlivňovat zainteresované strany, které jsou podřízeny sponzorům a managementu. Nemají dostatek pravomocí k tomu, aby mohly na strategické úrovni definovat cíle (ať už intervenčního programu či prováděného evaluačního šetření), ale mohou v rámci

¹³ V aplikační praxi navíc může být sada indikátorů předmětem úvodního zadání garanta hodnoceného programu, a nikoliv předmětem teoreticko-empirické analýzy. Volba hodnocených parametrů a způsob jejich indikace tak může být účelová.

¹⁴ Snaha o dosažení nespornosti sledovaných jevů je patrná zejména v případě sumativního typu evaluací, které posuzují účinek zkoumaného programu.

stanovených cílů volit jednotlivé indikátory využívající přitom své informační a znalostní převahy taktické či operativní úrovně (Remr 2009b). V aplikační praxi dochází často k záměně výstupů a výsledků či k nahrazení indikátorů cílových snáze dosažitelnými procesními („pracuji 10 hodin denně“ namísto „vyrobím 200 ks“; „bylo proinvestováno 10 mld. Kč“ namísto „bylo postaveno 10 km nových dálnic“). V neposlední řadě je třeba uvést také možnost určení standardů požadovaných pro posouzení věcného základu a přínosu daného intervenčního programu: cílové hodnoty mohou být stanoveny buď neadekvátně nízko (vytvoření dvou nových pracovních míst v rámci programu rozvoje sociálních služeb v celkovém objemu podpory 85 mil. Kč ročně), nebo naopak záměrně příliš vysoko, aby jich nemohlo být dosaženo („uplatnění nejméně 30 patentů ročně po dobu následujících 15 let“).

Patton (2008) v této souvislosti rozlišuje tzv. silné a slabé zainteresované strany; odděluje tedy zainteresované strany, které mají k dispozici moc a nástroje, jak ovlivňovat daný intervenční program a příslušné evaluační šetření (silné zainteresované strany), od ostatních (slabých) zainteresovaných stran, které tyto možnosti nemají. Dále rozlišuje zájmy obou skupin zainteresovaných stran na silné zájmy, na nichž subjektům velmi záleží, a na zájmy slabé. V následující tabulce jsou shrnuty základní implikace pro volbu designu evaluačního šetření.

Tabulka č. 5 Matice zainteresovaných stran a jejich zájmů

	Slabé zainteresované strany	Silné zainteresované strany
Silný zájem	Design evaluace by měl podporovat a usilovat o jejich zahrnutí, zejména pokud jsou tyto zainteresované strany dotčeny výsledky evaluace (např. účastníci programu). Jejich zahrnutím dojde ke zvýšení heterogenity prováděné evaluace.	Tyto zainteresované strany mají vysoký potenciál jako primární skupina. Jsou to klíčové subjekty, které aktivně ovlivňují způsob použití evaluace, a to jak pro sebe sama, tak také pro jiné zainteresované strany.
Slabý zájem	Design evaluace by měl na tyto zainteresované strany pamatovat při komunikaci o výsledcích. Kontroverzní výsledky mohou tento „dav“ snadno přeměnit v početnou a aktivní zájmovou skupinu.	Evaluační design by měl kultivovat zájem těchto zainteresovaných stran a upozorňovat je na případy, kdy je jejich nezájem bariérou.

Zdroj: Patton 2008

Intervenční programy a jejich implementace, stejně tak jako samotná evaluační šetření, vytvářejí interakční prostor, v němž se střetávají jednotlivé zainteresované strany. Berk a

Rossi (1999) dokonce uvádějí, že pro téměř všechny problémy spojené s intervenčními programy platí, že jednotlivé zainteresované strany jsou ve vzájemném konfliktu – některým daný intervenční program vyhovuje, zatímco ostatním nikoliv. Taková situace vyvíjí na evaluátora tlak, neboť musí rozhodnout, které zájmy a v jaké míře při provádění evaluace zohlednit, popř. jak skloubit protichůdné zájmy jednotlivých zainteresovaných stran.

Nároky kladené na evaluátora jsou ještě dále komplikovány mnohými situačními okolnostmi. Některé zainteresované strany o sobě například nemusejí vědět, není mezi nimi žádná přímá interakce. Evaluační šetření tak může přinést jako jakýsi vedlejší produkt analýzu vztahů mezi jednotlivými zainteresovanými stranami. I za situace antagonisticky vymezených zainteresovaných stran je evaluátor nucen jednotlivé zainteresované strany identifikovat a zajistit jejich participaci.

Souhrn

Jak shrnuje House (1993): „Pojem *stakeholders* se stal pro evaluátory běžným pojmem a zahrnování zájmů těchto jednotlivců či skupin do procesu evaluace se stalo obecně akceptovanou praxí.“ Každé evaluační šetření je připravováno a realizováno v konkrétním uspořádání a jeho výsledky jsou tak platné jen v rámci daného situačního kontextu. Z tohoto pohledu je identifikace zúčastněných stran přirozenou součástí každého evaluačního šetření, což je třeba považovat za klíčový diferenciativní znak oproti základnímu výzkumu¹⁵ i oproti výzkumu prováděnému v aplikovaných oblastech¹⁶. Nejde přitom jen o normativní aspekt prováděných evaluačních šetření (např. zabezpečení rovného zacházení)¹⁷, ale také o jejich epistemologický základ, neboť bez znalosti všech zainteresovaných stran a bez důkladné analýzy jejich vztahů k předmětu evaluačního šetření by nebylo možné provést analýzu jejich zájmů, a potažmo by tak získaný výsledek nemohl být adekvátní a validní. Opodstatněná je v tomto ohledu otázka, kterou si klade E. Guba a Y. Lincolnová, zdali jsou evaluační šetření skutečně objektivní a nezávislá, nebo zda jsou souhrnem subjektivit jednotlivých zainteresovaných stran (Guba, Lincolnová 1989). Ukazuje se, že mnohost zájmů jednotlivých

¹⁵ Řada zainteresovaných stran není v případě základního výzkumu vůbec známa a zájmy zainteresovaných stran, které známy jsou, nebývají obvykle nijak zohledňovány.

¹⁶ V rámci aplikovaného výzkumu bývají obvykle zainteresované strany identifikovány, ale jen v některých případech jsou jejich zájmy fakticky zohledňovány při návrhu výzkumného designu.

¹⁷ Identifikace jednotlivých zainteresovaných stran a specifikace jejich zájmů ve vztahu k posuzovanému intervenčnímu programu je jedním ze základních pravidel definovaných Joint Committee on Standards.

zainteresovaných stran a případné tenze mezi nimi nemusejí být nutně negativním prvkem při realizaci evaluačních šetření, neboť nastíněná zájmová mnohost umožňuje provedení analýzy daného intervenčního programu z různých hledisek. To významně rozvíjí poznávací aspekt prováděných evaluačních šetření a současně také posiluje využitelnost získaných poznatků.

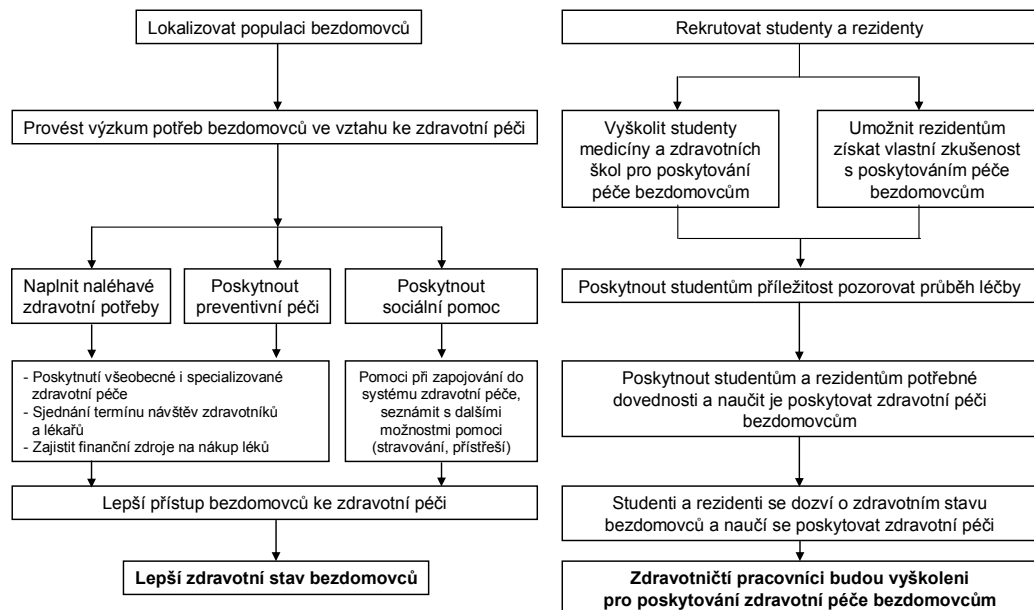
Identifikace zainteresovaných stran a pečlivá a podrobná analýza jejich zájmů představuje separátní fázi evaluačních šetření, kterou se odlišují od ostatních výzkumů. Chápeme ji však také jako specifický požadavek, který je evaluačním šetřením kladen na realizátory.

3.2.2 Konceptualizace a tvorba konceptuálních modelů

Zvýšený zájem o postupy konceptualizace v oblasti evaluací souvisí s posunem zájmu evaluátorů od vnějších projevů sledovaných intervenčních programů, tedy od zjišťování a ověřování, zda posuzované intervenční programy skutečně dosahují vytyčených cílů (tj. jsou efektivní), k vnitřním mechanismům, na jejichž základě intervenční programy působí. Primární snahou je tedy zjistit, jak jsou v rámci daného intervenčního programu vstupy transformovány na výstupy (Bouyssou, Marchant, Pilot, Perny, Tsoukiás, Vincke 1990). Tyto pokusy o otevření černé skříňky (*black-box*), za kterou jsou intervenční programy v rámci některých přístupů považovány, s sebou nesou také specifické metody, techniky a nástroje. Někteří evaluátoři poukazovali v této souvislosti na možnosti použití „pěšinkových“ diagramů (*path diagrams*), které zobrazují následnost jednotlivých kroků identifikovaných mezi samotnou intervencí (začátkem fungování intervenčního programu) a jejími výsledky (Weiss 1972). Suchman dokonce o několik let dříve doporučoval používání hierarchických diagramů zobrazujících zdroje (vstupy) a cíle (výstupy); jím navrhovaná schémata měla názorně dokumentovat vzájemné vztahy mezi příčinami a následky (Suchman 1967). V přibližně stejné době Stake (1967) publikoval svůj model potřeb, transakcí a výsledků intervenčních programů. I přes tyto pionýrské, a v té době poněkud ojedinělé, pokusy však nelze hovořit o skutečném rozmachu modelování procesů posuzovaných intervenčních programů až do přibližně 90. let 20. století, kdy nastal jejich opravdový rozkvět. V souvislosti s rozvojem evaluačních šetření se začaly objevovat požadavky na detailnější analýzu jednotlivých činností probíhajících v rámci hodnocených intervenčních programů, na zachycení jejich posloupnosti a vzájemné provázanosti. Postupně se tak začaly prosazovat dva typy konceptuálních modelů, a to explanační a implikační.

Klíčový představitel teorií vedeného přístupu k realizaci evaluačních šetření, Chen (2005), popisuje vytvářené explanační modely jako „specifikaci činností nutných k dosažení vytyčených cílů, dopadů, které mohou být dále anticipovány, a způsobů, jak lze těchto cílů a dopadů dosáhnout“. Přesun zájmu od strukturálního popisu intervenčního programu k jeho funkční explikaci je zde patrný na první pohled. Příkladem teoretického schématu je popis programu pro poskytování zdravotní péče bezdomovcům.

Obrázek č. 7 Ukázka explanačního modelu



Zdroj: Daponte 2008; upraveno autorem

V souvislosti s rozvojem evaluace výkonu (*performance evaluation*) v jehož rámci je obvyklé definovat jednotlivé etapy působení intervenčního programu, tzv. milníky¹⁸ a indikátory poukazující na splnění cílů v rámci každé jednotlivé etapy se při evaluačních šetřeních začaly používat implikační modely. Ty obvykle definují čtyři základní etapy, a to vstupy (*inputs*), které zahrnují soubor všech zdrojů nutných k fungování intervenčního programu. Etapa získávání a shromažďování vstupů je zakončena milníkem, který dokumentuje jejich dostupnost. Druhou etapou jsou činnosti (*activities*), které je nutné vykonat, aby z dostupných zdrojů bylo možné dosáhnout požadovaných cílů. Odpovídajícím milníkem činností, završujícím tuto etapu, je ověření jejich vykonání (tedy např. průběžný monitoring). Třetí etapou jsou výstupy (*outputs*), tedy přímé výsledky fungování posuzovaného intervenčního programu, které jsou očekávány nebo jsou dosaženy. Konečně čtvrtou etapu představují dopady (*impacts*), které daný intervenční program vyvolává¹⁹.

¹⁸ Jde vlastně o jakési rozhraní mezi dvěma klíčovými, po sobě následujícími etapami.

¹⁹ Někdy se namísto o dopadech (*impacts*) hovoří o dlouhodobých výsledcích (*long-term outcomes*), v takovém případě se pak v rámci třetí etapy hovoří o krátkodobých výsledcích (*short-term outcomes*).

Obrázek č. 8 Ukázka implikačního modelu

Cíle	Předpoklady	Cílová populace	Vstupy	Činnosti	Výstupy	Výsledky
Poskytnout bezdomovcům přístup ke zdravotní péči	Existence početné populace bezdomovců Obecná nedostupnost zdravotní péče pro mnohé bezdomovce Lepší přístup ke zdravotní péči přispěje k lepšímu zdravotnímu stavu bezdomovců	Bezdomovci v daném městě (čtvrti)	Pracovníci (lékaři, zdravotní sestry, studenti, dobrovolníci a další osoby zajišťující poskytování zdravotních služeb v terénu) Finanční zdroje (např. granty, dotace, veřejné rozpočty, výnosy ze soukromých sbírek apod.)	Lokalizovat populaci bezdomovců Provést zdravotní prohlídky bezdomovců Poskytnout preventivní péči Poskytnout okamžitou zdravotní péči Poskytnout informace o možnostech získání bezplatných léků	Počet bezdomovců, kterým je poskytnuta péče prostřednictvím daného programu Počet zdravotních prohlídek bezdomovců uskutečněných v rámci daného programu	Zlepšení zdravotního stavu bezdomovců
Vyškolit zdravotnický personál pro poskytování péče bezdomovcům	Zdravotní pracovníci se chtějí zapojit do programu pro poskytování zdravotní péče bezdomovcům Je třeba proškolit poskytovatele zdravotní péče v tématech, která s bezdomovectvím souvisejí	Studenti(-ky) medicíny a středních zdravotních škol Dobrovolníci se zájmem o poskytování péče bezdomovcům Ostatní poskytovatelé sociálních služeb; terénní pracovníci	Zdravotnické školy poskytující zázemí pro realizaci školení a poskytující studentskou pracovní sílu Odměny pro poskytovatele Materiálové příspěvky (zdravotnický materiál)	Nábor rezidentů a studentů Poskytnout rezidentům zkušenost s poskytováním zdravotní péče Poskytnout studentům možnost seznámit se s obsahem školení	Počet studentů, kteří jsou proškoleni	Zvýšený počet zdravotnických pracovníků vyškolených k poskytování péče bezdomovcům Počet absolventů, kteří využívají poznatky získané v rámci programu při své pracovní činnosti

Zdroj: Daponte 2008; upraveno autorem

Implikační modely umožňují strukturovaný popis jednotlivých prvků intervenčního programu vyskytujících se a působících v rámci dané evaluace a vyjasňují jednotlivé vstupy, činnosti a výstupy (spolu s výsledky a dopady), čímž vytvářejí jakýsi rámec pro provedení evaluačního šetření. Na základě sekvenčních schémat lze snáze identifikovat předpoklady, které nemusí být na první pohled zjevné (např. zamlčené předpoklady fungování intervenčního programu, skryté vstupy, latentní konflikt mezi jednotlivými vstupy apod.).

Obecně lze tedy shrnout, že implikační modely pomáhají identifikovat nerealistické kombinace zdrojů a výstupů, umožňují rozkrývat kritická místa jednotlivých intervenčních programů a významným způsobem přispívají k vyšší využitelnosti prováděných evaluací, a tím vedou ke zdokonalování samotného intervenčního programu.

Souhrn

Význam konceptuálních modelů spočívá kromě již uvedeného v tom, že představují rámec uvažování, ve kterém mohou jednotlivé zainteresované strany prezentovat své podněty a poznatky o působení posuzovaného intervenčního programu. V neposlední řadě platí, že pomocí uvedených modelů může být prezentována část výsledků evaluačních šetření.

Závěrem je třeba uvést, že některé evaluační přístupy nepovažují rozkrývání procesů probíhajících v rámci intervenčních programů za relevantní. V tomto ohledu by např. R. Tyler zcela jistě žádný takový model nevyužil, protože jeho evaluační přístup nevyžadoval znalost konkrétních mechanismů a způsobů, jak nové osnovy vedou k lepším výsledkům testovaných žáků. Jeho pozornost byla soustředěna na co nejpřesnější a co nejvíce validní ověření toho, zdali nové školní osnovy dosáhly definovaných cílů a přinesly žádoucí změnu v systému výuky, či nikoliv. Podobně také evaluace založené na experimentálním designu nevěnují tolik pozornosti způsobům, jak je výsledků dosahováno, ve srovnání s mírou pozornosti, která je věnována samotné randomizaci, volbě konkrétního výzkumného uspořádání a osvědčení platnosti získaných výsledků.

3.2.3 Metaevaluace

Již jsme uvedli, že jedním z cílů evaluací je poskytnout poučení z realizovaných aktivit (jakými mohou být např. implementace intervenčního programu či uskutečňování daného typu rozhodnutí). Otázkou však zůstává, nakolik se samotná oblast evaluací dokáže poučit ze svých vlastních aktivit, kroků a činností? Považujeme za samozřejmou součást rozvoje evaluačních schopností, že jednotlivé subjekty zapojené v evaluačních procesech mají vytvořeny mechanismy, které jim umožňují systematicky získávat a vyhodnocovat zpětnou vazbu o své vlastní činnosti.

Koncept metaevaluací byl poprvé popsán Scrivenem (1991), který definoval tuto činnost jako evaluaci provedené evaluace – či nepřímou také jako evaluaci evaluátorů. Metaevaluace tedy revidují provedené šetření, přičemž usilují o provedení evaluace samotných evaluátorů, neboť i oni jsou zainteresovanou skupinou svého druhu (Guba, Lincoln 1989) a je tak přirozené, že jsou podrobeni evaluaci stejně tak jako ostatní subjekty. Klíčovým motivem je v tomto ohledu snaha eliminovat subjektivní prvek v primární evaluaci a dosáhnout tak vyšší kredibility získaných závěrů.

Původní Scrivenovu představu o roli a významu metaevaluací později rozšířil zejména D. Stufflebeam, který chápe metaevaluaci jako soubor opatření sloužící k tomu, aby evaluace přinášely jasná zjištění a jednoznačné závěry (Stufflebeam, Shinkfield 2007). Z jeho pojetí je patrná snaha pojmut metaevaluace jako praktickou pomůcku při provádění vlastních evaluačních šetření, o tom koneckonců nejlépe svědčí jeho *Key Evaluation Checklist*, jenž může být považován za jakési vodítko, jak metaevaluace realizovat.

Metaevaluace jsou prováděny dvěma základními způsoby. Jednak jako sumativní evaluace uskutečněné po dokončení primárního evaluačního šetření; sumativní forma metaevaluací je obvykle prováděna třetí stranou za účelem posílení legitimacy výsledků metaevaluací či z důvodů snazší identifikace zdrojů chyb objevujících se v primárním evaluačním šetření. Tento přístup má jednoznačně pozitivní efekty pro budoucí evaluační projekty a je faktickým spojením mezi jinak oddělenými individuálními evaluačními šetřeními (Remr 2008a). Pozitivního efektu metaevaluací však lze dosáhnout také využitím formativního uspořádání, kdy evaluátor sám hodnotí své vlastní evaluační šetření. V tomto případě metaevaluace umožňuje lepší pochopení, jak různá rozhodnutí v rámci daného designu ovlivňují kvalitu

výsledků a jak vlastně tato rozhodnutí dopadají na celý evaluační projekt. To dává evaluátorům příležitost ke zlepšení vlastních závěrů, a to dokonce ještě předtím, než jsou výsledky evaluačního šetření prezentovány zadavateli.

Výzvy a příležitosti

V posledním období lze zaznamenat úsilí mnohých evaluátorů o standardizaci metaevaluačních postupů, a potažmo také o standardizaci výstupů z prováděných metaevaluací. Zřejmou výhodou anticipované standardizace je především snadná dostupnost analýzy kvality provedeného evaluačního šetření. V kontextu evaluací je takovýto doklad důležitý a užitečný, neboť zainteresované strany, které jsou výsledky evaluace negativně dotčeny, častokrát napadají (účelově) právě kvalitu provedení evaluace (Benson 2001; Uusikylä, Virtanen 2000.). Existence standardizovaných metaevaluačních postupů by takováto zpochybňování významně eliminovala.

Ohrožení a limity

Metodika metaevaluací není doposud ustálená. Stále existuje řada fází procesu evaluací, které nejsou dotčeny žádným indikátorem v kontrolních seznamech, jež se pro provedení metaevaluací aktuálně využívají a jsou publikovány v odborné literatuře.

Je nutné, aby údajům metaevaluací i jejich samotnému konceptu rozuměli samotní zadavatelé a vyžadovali jejich provedení (Ciarlo 1981). Optimálním řešením je tzv. one-currency metaevaluace, tedy univerzální a obecně akceptovaný standard, podle něhož by byly metaevaluace prováděny. Rovněž je třeba dlouhodobě poukazovat na přidanou hodnotu pramenící z informací poskytovaných metaevaluacemi a poskytnout návod, jak přesně metaevaluace používat v každodenním rozhodování.

Specifickou bariérou rozvoje metaevaluací mohou představovat samotní evaluátoři, neboť již nyní je patrné, že zdaleka ne všichni jsou ochotní otevřít černou skříňku, kterou evaluační procesy představují.

3.2.4 Souhrn

Tři procesní specifika, která jsou charakteristická pro provádění evaluačních šetření zdaleka nevyčerpávají celou šíři postupů a činností, které v oblasti evaluací vznikly. Popsané oblasti však považujeme za jedny z nejvýznamnějších, neboť se k nim váže hned několik ustanovení Standardů. K identifikaci a zapojení zainteresovaných stran se váže mj. rozhodování o definici a specifikaci evaluační otázky, volba designu evaluačních šetření, příprava závěrečné zprávy z evaluačního šetření a několik dalších (dle Sanders 1994). Konceptuální modely obracejí pozornost k výběru informací a k popisu procesů probíhajících v posuzovaném intervenčním programu či jiném předmětu evaluace, čímž zásadně určují charakter celého procesu a přímo tak ovlivňují charakter a využitelnost získaných závěrů. Poslední z prezentovaných činností, metaevaluace, provází proces evaluačního šetření doslova od počátku až do konce. Ačkoliv metaevaluace nejsou nezbytnou součástí prováděných evaluačních šetření (v řadě případů se metaevaluace neprovádějí), přesto představují významný a mocný nástroj, který provádění evaluačních šetření odlišuje od ostatních aktivit.

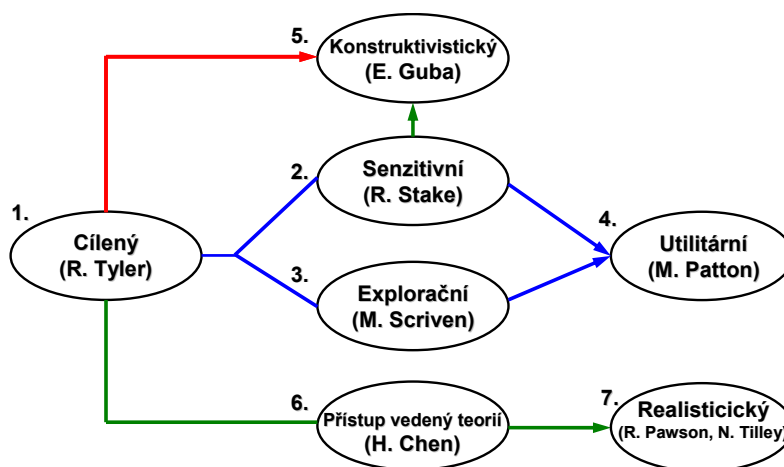
4. Analýza přístupů k realizaci evaluačních šetření

Čtvrtá kapitola je rozdělena do dvou podkapitol, kdy první se zaměřuje na teoretické možnosti a příležitosti, jak lze přistupovat k realizaci evaluačních šetření. V tomto ohledu poukazuje na obecně-metodologické předpoklady využití jednotlivých forem v závislosti na konkrétních podmínkách. Zatímco první podkapitola neopouští teoretickou rovinu, druhá z podkapitol přináší několik úhlů pohledu na reálně prováděná evaluační šetření. Taxonomie, o které se opírají jednotlivé typy evaluačních šetření prezentované v této části práce, vycházejí z aplikačních podmínek a odrážejí konkrétní okolnosti interakce evaluačních šetření s jejich sociálním kontextem. Pozornost je tak věnována typům evaluačních šetření diferencovaných dle účelu provedení, dle předmětu evaluace a dle vztahu zadavatele a realizátora.

4.1 Teoretické přístupy k realizaci evaluačních šetření

Pro teoretické přístupy k realizaci evaluačních šetření je v rámci chápání tohoto pojmu v naší práci charakteristické, že (1) vypovídají o pojetí externí validity, (2) odrážejí vztah šetření (údajů) a reality (konkrétní hodnocené situace), (3) vyjadřují/nastiňují základní metodologickou orientaci a v neposlední řadě také (4) definují ústřední zájem či hodnotu, která určuje celkové nasměrování prováděného evaluačního šetření (Alkin, Christie 2005).

Obrázek č. 9 Mapa jednotlivých evaluačních přístupů



Zdroj: autor

Na základě takto stanovených kritérií lze identifikovat celou řadu jednotlivých přístupů, které se v průběhu vývoje metodologie evaluačních šetření objevily. Z jejich dlouhého seznamu uvádíme celkem sedm přístupů, které reprezentují hlavní směry teoretického uvažování o způsobu realizace evaluačních šetření. Jsme si vědomi toho, že prezentovaný výčet není úplný; chybí zde mj. Fettermanův zplnomocňující přístup k provádění evaluačních šetření (*empowerment evaluation*), participativní a kolaborativní přístupy a mnohé další. Nicméně primárně sledovaný účel tento výčet beze zbytku naplňuje. Ukazuje totiž na pluralitu konkrétních forem jak mohou být evaluační šetření prováděna, přičemž sedm z uvedených přístupů současně odráží hlavní akcenty při provádění evaluačních šetření.

4.1.1 Cílené evaluace (goal-based evaluation; Ralph Tyler)

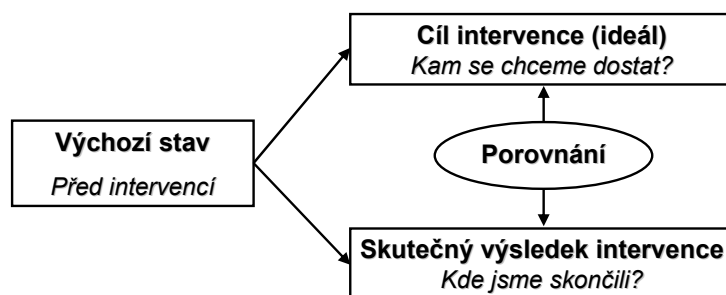
Tylerův přístup byl v daném historickém kontextu nevídaný. Podle historického exkursu Guby a Lincolnové (1989), zachycujícího vývoj evaluací, převládal v meziválečném období takový způsob evaluací, který využíval sumarizované výsledky standardizovaných znalostních, kognitivních či výkonnostních testů, jako jsou např. SRTA (Stone Reasoning Test in Arithmetic, jenž je používán od roku 1908) či SAB (Stanford Achievement Battery, vytvořený v roce 1922)²⁰. Tyto a mnohé podobné testy byly tehdy hojně používány, neboť bezvadně plnily svůj primární účel: individuální diagnostiku žáků. Kromě toho však výsledky těchto testů nacházely také sekundární využití, a to v sumarizované (agregované) podobě – na jejich základě byly porovnávány jednotlivé třídy v rámci školy, či dokonce celé školy v rámci dané oblasti či státu. R. Tyler však využití standardizovaných testů odmítl, a namísto toho vypracoval vlastní přístup k realizaci evaluačních šetření. Jeho záměrem totiž bylo vyhodnotit nově zaváděné osmileté osnovy a obával se, že diagnostické testy, primárně vytvořené pro měření konkrétních schopností jednotlivých žáků, nepřinesou potřebná data o jejich účinnosti. R. Tyler v této souvislosti poukazyval především na rozdílný předmět analýzy, tedy schopnosti jednotlivých žáků oproti konkrétně specifikovaným výsledkům používaných osnov. Při tomto srovnání by způsobnost používaných diagnostických testů vypovídat o výsledcích (efektivitě) nových osnov nebyla podle R. Tylera obhajitelná.

²⁰ V tomto období byl rozvoj evaluačních šetření svázán výlučně s oblastí školství a vzdělávání; ostatní typy evaluačních šetření, jako např. evaluace intervenčních programů, se objevily až v 60. letech.

Tylerův odkaz podstatně přesahuje rámec jednoho samotného přístupu k realizaci evaluačních šetření. Jeho hlavní doménou je obecný postup charakteristický tím, že k předem definovaným cílům (tj. očekávaným výstupům) jsou v rámci provádění evaluačního šetření přiřazovány skutečně dosahované a empiricky ověřitelné výsledky, přičemž oba typy informací jsou následně porovnány. Z tohoto důvodu je Tylerův přístup některými nazýván jako *objectives-oriented* (Alkin 2004), jiní o tomto přístupu píší jako o *goal-based evaluation* (Stufflebeam, Shinkfield 2007). Obě pojmenování však stejně dobře vystihují základní podstatu Tylerova přístupu a lze je používat jako synonyma.

Předem definované cíle intervenčního programu jsou tedy v rámci tohoto přístupu základním a jediným zdrojem standardů a kritérií. Aby bylo možné s cíli v rámci cíleného přístupu k evaluacím pracovat, měly by mít měřitelnou podobu – R. Tyler v tomto ohledu prosazoval takové cíle, které mají odraz v chování jednotlivců (srov. behaviorismus)²¹. Základní otázka cíleného přístupu k realizaci evaluačních šetření tedy zní: „Lze pozorovat očekávané prvky chování ve skupině účastníků intervenčního programu, či nikoliv?“

Obrázek č. 10 **Obecný princip cílených (goal-based) evaluací**



Zdroj: Patton 2008; upraveno autorem

Tylerův přístup k realizaci evaluačních šetření je založen na plném chápání komplexity evaluačního procesu, jehož cílem je dospět k závěru deklarujícímu míru, ve které posuzovaný program naplňuje své předem stanovené cíle. R. Tyler sám svá evaluační šetření prováděná v polovině 30. let v souvislosti s hodnocením nově zaváděných osmiletých osnov popsal jako posloupnost sedmi základních kroků takto (cit. dle Alkina 2004):

1) definování obecných cílů

²¹ V pedagogickém výzkumu mají takový charakter tzv. achievement testy.

- 2) kvantifikace cílů a jejich klasifikace (uspořádání cílů do menšího počtu základních skupinek)
- 3) nalezení specifických behaviorálních vzorců pro každou skupinku cílů
- 4) identifikace situací, ve kterých je možné zachytit (pozorovat) výskyt požadovaných behaviorálních vzorců; nalezení okolností, za nichž lze míru dosažení stanovených cílů jasně dokumentovat
- 5) výběr adekvátních metod získání potřebných údajů pro každou skupinku cílů; výběr či vývoj techniky měření stupně dosažení cílů
- 6) sběr dat včetně provedení předvýzkumu a volba optimální metody sběru dat
- 7) srovnání získaných údajů s předem definovanými cíli; stanovení způsobu využití získaných výsledků a nástin možností a omezení při jejich interpretaci

Tylerův přístup byl novátorský ještě v jednom ohledu: do té doby používané testování neposkytovalo žádnou zpětnou vazbu, která by mohla pomoci dosažené výsledky zlepšovat. Absence obecného standardu a akceptačních kritérií (vodítka, která by pomohla určit, jaké výsledky jsou „dobré“, a jaké nikoliv) neumožňovala naplnit bazální princip evaluace, tedy stanovit věcný základ (*merit*) a přínos (*worth*). Kdybychom i přes tyto výhrady připustili, že prováděné testování neslo jisté stopy evaluací, označili bychom je zpětně z dnešního pohledu jako evaluace sumativní (podrobněji o nich dále). Posuzované hodnoty totiž vytvářely pořadí jednotlivých tříd či škol a ukazovaly na úspěch a neúspěch jednotlivých subjektů. Naproti tomu evaluační šetření prováděná R. Tylerem přinášela zcela jiný typ údajů, zprostředkovávala totiž data poskytující detailně strukturované výsledky o diskrepancích mezi skutečnými a očekávanými či požadovanými cíli. Tím de facto provedená šetření poskytla návod, které z výsledků (tedy přesněji, které z behaviorálních vzorců dokumentujících určitou třídu cílů) je vhodné v příštím období zlepšovat. R. Tyler svým pojetím tedy umožnil, aby osmileté osnovy byly nejen porovnány oproti dosavadnímu kurikulu z hlediska rozvoje schopností a dovedností podrobnějších se žáků, ale současně aby se osmileté osnovy s intenzivní pomocí jím prováděných evaluačních šetření dále rozvíjely.

Tylerův odkaz je dodnes živým způsobem provádění evaluací. Jednak tím, že inspiroval mnohé pokračovatele k dalšímu rozvíjení cíleného přístupu, mezi všemi jmenujme např. Blouma et al. (1956) či Magera (1962), a dále tím, že umožnil jiným, aby vůči cílenému přístupu k evaluacím vymezili své vlastní přístupy (viz např. M. Scriven či E. Guba, o kterých je pojednáno ve větším detailu níže). Koneckonců současná podoba evaluačních šetření tak, jak je požadována Federální vládou USA (a do určité míry i Komisí EU) vychází do značné

míry právě z Tylerova cíleného přístupu k realizaci evaluačních šetření²². V obecnějších souvislostech lze hovořit o tzv. *objectives-oriented tradition*, která se objevuje i v jiných oborech: přestože nebyl doposud prokázán přímý vliv, nelze si nepovšimnout shody základních principů např. s dnes velmi rozšířeným proudem MBO (Management by Objectives) v manažerských studiích.

Nespornou výhodou Tylerova cíleného přístupu je jeho relativní jednoduchost a možnost snadného provedení, třeba i jen formou interně prováděných evaluačních šetření, bez pomoci speciálně školených evaluátorů. Za jeho silnou stránku lze považovat rovněž fakt, že tento přístup vede zúčastněné zainteresované strany (zejména pak manažery posuzovaných intervenčních programů) k zamyšlení nad jejich původními záměry a skutečně dosaženými cíli a podněcuje je k jejich precizaci a sdílení s ostatními zainteresovanými stranami. Někteří ze současných evaluátorů při hodnocení Tylerova odkazu poukazují na skutečnost, že jeho přístup vlastně není žádnou specifickou inspirací pro provádění evaluačních šetření, neboť jde o pouhou deskripci postrádající klíčové atributy samotné evaluace. Nejen R. Tylerovi, ale obecně celému cílenému přístupu je dále vytýkáno, že zcela odhlíží od kontextu, v němž jsou evaluační šetření realizována, a že nebere v potaz nezamýšlené dopady, ke kterým mohou posuzované intervenční programy vést. Cílenému přístupu k provádění evaluací je rovněž vytýkán relativně vysoký sklon k managerialismu (House 1993; Scriven 1993; Guba, Lincoln 1989) a sklon k obecně zaměřenému posuzování sledovaných intervenčních programů. Zastánci tzv. responzivního přístupu k evaluacím zase cílenému přístupu vyčítají, že jako formativní způsob evaluace vůbec nepočítá s individuálními hodnotami a hodnotovými systémy samotných klientů (v Tylerově případě žáků) a nebere v rámci posuzované situace ohled na jejich svobodné rozhodování a jednání (viz např. odlišné aspirace žáků škol v městských aglomeracích oproti aspiracím žáků na venkově).

Při hodnocení Tylerova přínosu nebyl doposud doceněn ještě jeden přínos. Tyler totiž svým přístupem přivedl do evaluací primární data a otevřel tak prostor pro pronikání společenskovědní metodologie (zaměřené na sběr a analýzu primárních dat) do oblasti evaluací. Přístup testování používaný do té doby se totiž zakládal výhradně na sekundární práci s jednotlivými testy (jejich vyplňování nebylo prováděno za účelem evaluací, ale byly evaluátory takto využívány).

²² R. Tyler se sám stal v 60. letech vedoucím NAEP (National Assessment of Educational Progress), což byl federální program pro sběr dat o výkonech žáků amerických škol.

Ilustrační příklad²³:

V rámci cíleného přístupu bude vypracován test počítačové gramotnosti, který bude po skončení kursu předložen každému z účastníků. Současně bude stanoveno (po dohodě s iniciátorem a s jeho souhlasem), že test se zaměří zejména na používání textového editoru a tabulkového procesoru a budou definovány konkrétní úlohy, které budou indikovat zvládnutí obsahu kursu. V neposlední řadě bude stanoveno, že pokud 80 procent účastníků kursu tento test zvládne (tj. dosáhne alespoň 75procentní úspěšnosti), bude možné považovat kurs za účinný. V takovém případě je pravděpodobné, že se jeho iniciátor rozhodne pro jeho rozšíření i do ostatních okresů.

4.1.2 Explorační přístup k provádění evaluačních šetření (goal-free evaluation; Michael Scriven)

Jak již bylo uvedeno, od dob R. Tylera byla významná část evaluačních šetření založena na porovnávání cílů vytyčených příslušnými zainteresovanými stranami před implementací programu se skutečně dosaženými výsledky. Tento způsob provádění evaluací, který Guba a Lincolnová (1989) nazývají „popisem“ a považují ho za tzv. druhou generaci evaluační metodologie, je využíván v řadě situací z rozličných důvodů dodnes a rozhodně ho nelze považovat ani za překonaný, ani za aplikačně omezený či nevhodný. Odborníci se však pochopitelně snaží navrhovat nové přístupy překonávající své ascendency v těch charakteristikách, jež vyvolávají největší kontroverze a jejichž výsledky zůstávají v mnohých ohledech za původním očekáváním. V tomto ohledu je tedy explorační přístup k realizaci evaluačních šetření třeba chápat jako Scrivenovu odpověď na cílený přístup, jako odpověď, jež přinesla nový pohled na způsob provádění evaluačních šetření.

Podstatou rozdílu mezi cíleným a exploračním přístupem je stupeň pozornosti věnovaný skutečným výstupům, výsledkům a dopadům posuzovaných programů. M. Scriven reagoval na obvyklý nedostatek cílených evaluací, které sice dokázaly přesně odpovědět na otázku, zdali bylo dosaženo vytyčených cílů a v jaké míře, nicméně vůbec nebraly v potaz

²³ Uvedený ilustrační příklad je uveden ve snaze o poskytnutí konkrétního obrazu shora uvedených charakteristik. Příklad je založen na předpokládaném provedení evaluačního šetření zaměřeného na posouzení efektivity rekvalifikačního kursu určeného nezaměstnaným osobám starším 50 let, který je zaměřen na osvojení práce s výpočetní technikou. Příklad je průběžný a opakovaně se k němu vracíme v každém z dále prezentovaných přístupů. Díky tomu si čtenář může vytvořit představu o odlišnostech, které mezi jednotlivými přístupy panují.

vyskytnuvší se nezamýšlené důsledky jejich působení²⁴. Explorační přístup k provádění evaluačních šetření tedy staví do pozice hodnocených prvků cíle samotné a odmítá jejich apriorní danost. Scriven (1973) zdůrazňoval, že evaluátor by dokonce cíle posuzovaného intervenčního programu neměl znát, aby jimi nebyl ovlivněn: „Goals bias the evaluator“²⁵. Toto je v kontextu Scrivenovy evaluační teorie tzv. kvalitativní význam objektivity; za objektivní je v tomto ohledu považován takový výsledek, který může prokázat, že není vychýlený²⁶. Evaluátor by tedy v souladu se základním principem tohoto přístupu měl věnovat pozornost **všem** výstupům, k nimž daný intervenční program směřuje, a to jak k výstupům zamýšleným, tak také nezamýšleným. Díky tomu mají evaluační šetření prováděná v souladu s principy tohoto přístupu mnohem lepší předpoklady zachytit i vedlejší a neočekávané důsledky posuzovaných intervenčních programů. Konkrétní evaluační praxe ukazuje, že explorační přístup k provedení evaluačních šetření dosahuje velmi dobrých výsledků zejména při identifikaci krátkodobých a střednědobých efektů (tj. výstupů a výsledků), nicméně pro hodnocení dopadů se výhody jejich aplikace výrazně snižují. Základní otázka exploračního přístupu k realizaci evaluačních šetření tedy zní: „Jaké jsou veškeré efekty dané intervence?“

M. Scriven ve svém exploračním přístupu poukazuje na potřebu vymanit se z rámce, který soustřeďuje pozornost realizátora evaluačních šetření na jednotlivé dílčí oblasti, a doporučuje provádět evaluační šetření takovým způsobem, při kterém by samotné hodnocení nebylo nijak svázáno s původními programovými cíli. Taková představa má navíc oporu v samotné normativní rovině evaluací; apriorní definice cílů, jejichž dosažení má dané evaluační šetření potvrdit či vyvrátit, může totiž vést k jednostranné akcentaci zájmů některých konkrétních zainteresovaných stran (např. iniciátorů intervenčního programu nebo jeho manažerů), což by bylo v rozporu nejen s požadavky definovanými Joint Committee on Standards (cit. dle Sanders 1994), ale i v rozporu s obecnými zásadami nestrannosti. M. Scriven sám zdůrazňoval, že explorační přístup k provádění evaluačních šetření je z tohoto důvodu spravedlivějším přístupem (Scriven 1993).

²⁴ Podoba evaluačních šetření byla v 60. letech ovlivněna zákonnými pravidly, která přesně specifikovala jejich podobu. Specifikace evaluačních šetření vycházely, s ohledem k účelu využití (např. rozhodnutí o pokračování či zastavení finanční podpory konkrétních škol z veřejných zdrojů), z Tylerovy metodiky.

²⁵ M. Scriven v této souvislosti někdy píše o tzv. bias-free evaluation.

²⁶ Kromě této kvalitativní objektivity se rozlišuje také objektivita kvantitativní, která ztotožňuje objektivitu s kvantifikací a spoléhá na shodu intersubjektivit (House 1980). Kvantitativní pojetí objektivity však M. Scriven zásadně odmítá.

Explorační přístup k provádění evaluačních šetření předpokládá konsenzus mezi všemi zainteresovanými stranami na jednotlivých kritériích posuzování a na relevanci zjištěných důsledků, a to jak zamýšlených a očekávaných, tak také těch nezamýšlených.

Z předchozího je patrné, že pokud je evaluační šetření realizováno v souladu s principy exploračního přístupu, měla by být minimalizována interakce evaluátora s ostatními zainteresovanými stranami (tedy zejména s manažery daného intervenčního programu a s jeho pracovníky, resp. se zadavatelem daného evaluačního šetření).

Explorační přístup k realizaci evaluačních šetření nevyklučuje při své aplikaci simultánní využití přístupu cíleného; z řady konkrétních případových studií je skutečně patrná snadná možnost oba přístupy kombinovat. Výhodou takovéto kombinace obou přístupů je maximalizace užitku z prováděného evaluačního šetření.

Explorační přístup k realizaci evaluačních šetření se však potýká s kritikou poukazující na jeho insuficienci při identifikaci pramenů pro stanovení kritérií dané evaluace. Jak nalézt tato kritéria? Jak provést operacionalizaci, když nejsou zadané (resp. evaluátorovi nejsou známé) cíle posuzovaného intervenčního programu?

Uvedení exploračního přístupu k evaluacím není ve vývoji evaluační metodologie ničím nahodilým; rozvinutí tohoto přístupu jasně vychází z diskursu 60. a 70. let, jenž byl veden na téma účelu a poslání evaluace jako takové. Explorační přístup k realizaci evaluačních šetření je koncipován jako *imanentní* součást intervenčních programů, jako průvodce zainteresovaných stran napomáhající v jejich rozhodovací činnosti v celém průběhu rozvoje a v celé šíři implementace posuzovaných intervenčních programů. M. Scriven svým přístupem za celý evaluační obor deklaroval jasně ambici zapojit evaluaci do procesu přípravy a implementace intervenčních programů v celé jejich šíři, čímž se pokusil marginalizovat tehdejší uzance, notně podporované cíleným přístupem, volající po provedení evaluačního šetření až v pokročilých fázích implementace intervenčních programů.

Ilustrační příklad:

Evaluátor, který bude používat explorační přístup, nezná záměry iniciátora hodnocených kursů, a tak bude věnovat pozornost jakýmkoliv přínosům, které kurs jeho účastníkům mohl přinést. Může tedy dojít k závěru, že kurs sice nenaučil jednotlivé účastníky programovat

makra ve Visual Basic, nicméně zdokonalil uživatelské návyky jednotlivých účastníků. Současně bude díky exploračnímu přístupu odhaleno, že si účastníci kursu vyměňují mezi sebou své zkušenosti z průběhu pohovorů u potenciálních zaměstnavatelů a sdílí mezi sebou nabídky o volných pracovních pozicích, což jim napomáhá k rychlejšímu získání nových pracovních míst.

4.1.3 Responsivní evaluace (responsive evaluation; Robert Stake)

Také R. Stake ve svém responsivním přístupu k realizaci evaluačních šetření vyšel z původního Tylerova cíleného přístupu, nicméně původní koncept rozšířil o další okolnosti, které provázejí dosažení vytyčených cílů v kontextu daného intervenčního programu. R. Stake v tomto ohledu upozorňuje, že kromě samotných výsledků je třeba přihlížet také k obecnějším podmínkám a okolnostem, za nichž je daný intervenční program připravován a implementován. Podle R. Stakea je třeba dále věnovat pozornost klíčovým procesům, které zajišťují, že intervenční program skutečně směřuje k naplnění svých cílů, a v neposlední řadě je nutné obracet pozornost ke standardům, na jejichž základě je intervenční program (resp. jeho cíle) posuzován. Stake (1980) poukazuje na fakt, že řada do té doby prováděných evaluačních šetření zdůrazňovala tvrzení o cílech, akcentovala používání objektivních testů a poukazovala na standardy (ne)dodržované administrátory daného intervenčního programu. Jako protiklad k takto koncipovaným evaluačním šetřením, jejichž design R. Stake sám označoval jako předem definovaný (*pre-ordinate*), navrhl vlastní přístup k provádění evaluačních šetření. Typickým znakem nově navrženého přístupu je obrat pozornosti ke komplexitě a k obtížně standardizované deskripci intervenčního programu. Zatímco Scrivenova odpověď R. Tylerovi v podobě exploračního přístupu je ve své podstatě „a-metodická“, neboť neodkazuje k žádnému konkrétnímu postupu, ani k žádným konkrétním metodám či technikám, jež by evaluátoři měli při aplikaci exploračního přístupu využívat, responsivní přístup k realizaci evaluačních šetření zřetelně odkazuje ke kvalitativní metodologii, „řeč čísel je příliš zjednodušující na to, aby dokázala zachytit celkový dojem“ (Stake 2004).

Responsivita Stakeova přístupu se naplňuje ve dvou základních směrech. Zprv platí, že evaluátor je vnímavý vůči subjektům zapojeným do procesu evaluačních šetření, což ve

Stakeově pojetí znamená apel na zohlednění zájmů těch zainteresovaných stran, které jsou realizací evaluačního šetření nejvíce dotčeny. Responsivní evaluátoři tak věnují významnou pozornost zejména administrátorům intervenčních programů, jejich klientům a beneficiům, snaží se naladit na vlnu těchto zainteresovaných stran a nechávají si jakoby „vnutit“ jejich pohled, jejich vnímání a jejich argumenty (Greene, Abama 2001). Z druhého Responsivní přístup předpokládá vnímavost evaluátorů vůči požadavkům, otázkám a očekáváním, která od prováděných evaluačních šetření mají jejich zadavatelé. Zatímco první směr Stakeova responzivního přístupu k realizaci evaluačních šetření odkazuje k později se rozvíjícím konstruktivistickým evaluacím (*per parenthesis*, E. Guba coby reprezentant konstruktivistického směru v oblasti evaluací na R. Stakea přímo odkazuje a činí z responzivního přístupu k realizaci evaluačních šetření jeden ze dvou zásadních pilířů svého vlastního přístupu), druhý směr vede k přístupu M. Pattona, který obrací pozornost k takovému návrhu evaluačních šetření, který podporuje využitelnost jejich výsledků.

Evaluační šetření realizovaná v souladu s principy responzivního přístupu nezakládají legitimitu svých záměrů na pojmech a konstruktech obvyklých v kvantitativních společenských výzkumech, jako je např. validita, nýbrž ji odvozuje z faktu, že jsou dosažené závěry akceptovány konsenzem důležitých zainteresovaných stran. R. Stake se v tomto případě opírá o již zmíněný koncept kvantitativní objektivity, čímž staví Responsivní přístup k realizaci evaluačních šetření do kontrapozice jak vůči cílenému přístupu (což není žádným překvapením, neboť oba přístupy vycházejí z diametrálně odlišných předpokladů a směřují k zásadně odlišnému pojetí smyslu a účelu evaluačních šetření), tak také vůči přístupu exploračnímu (což nemusí být nutně na první pohled patrné, neboť oba přístupy sdílejí ideu „otevřenosti“ při evaluaci věcného základu a přínosu).

Typickou vlastností evaluačních šetření prováděných v duchu responzivního přístupu je, že nemají žádný pevný design. Ten se totiž vytváří postupně krok za krokem až v průběhu realizace daného evaluačního šetření vždy na základě poznatků a dílčích závěrů vyplývajících z předchozí fáze. Evaluátor se tedy snaží citlivě reagovat na postupně se rýsující kontury zájmů jednotlivých zainteresovaných stran a operativně jim přizpůsobuje design prováděného evaluačního šetření. Responsivní evaluace tak přesouvají těžiště evaluátorovy pozornosti od cílů intervenčních programů a jejich naplnění či nenaplnění k důkladnému a detailnímu popisu (některých) zainteresovaných stran a jejich zájmů. Responsivní evaluace tak v mnoha případech nespějí k jednoznačným odpovědím, které nabízí cílený přístup, zato však

podrobně mapují předpoklady pro jejich dosažení a rozkrývají případné bariéry, kvůli kterým nemohou být očekávané cíle naplněny.

V tomto ohledu nás nepřekvapí, že se Responsivní přístup k realizaci evaluačních šetření metodologicky opírá o případovou studii; evaluace každého jednotlivého intervenčního programu je samostatnou studií, jež poskytuje evaluátorovi značný prostor pro rozhodování o využití konkrétních informací a dat i o provedení případných výzkumů přinášejících dokumentaci nutnou pro samotnou evaluaci. R. Stake dále poukazuje na vysokou roli kontextu, ve kterém jsou daná evaluační šetření prováděna a ve kterém dané intervenční programy působí. V této souvislosti je však nutné upozornit, že R. Stake chápe jako kontext jednotlivé zainteresované strany, jejich pozice, role a zájmy. Z uvedeného je zřejmé, proč Responsivní přístup k realizaci evaluačních šetření klade takový důraz na analýzu forem vnímání intervenčních programů ze strany jejich beneficentů a dalších zainteresovaných stran.

Responsivní evaluace nebývají obvykle završeny nějakým ultimativním závěrem, který by dával jednoznačnou odpověď, zdali posuzovaný intervenční program vyvolává očekávané efekty, či nikoliv. Namísto toho jsou dosažené výsledky vztahovány k očekáváním a k hodnotám jednotlivých zainteresovaných stran, a tím poskytují podpůrné informace pro rozhodování o dalším směřování sledovaného intervenčního programu²⁷. V responsivním přístupu k evaluačním šetřením platí, že evaluátor je jakýmsi mediátorem zájmů jednotlivých zainteresovaných stran, nicméně není považován za specifickou zainteresovanou stranu. Evaluační šetření prováděná v souladu s principy responzivního přístupu tedy nejsou završena subjektivistickým závěrem samotného evaluátora.

Při aplikaci responzivního přístupu je evaluátorův záměr vsazen do kontextu vyjádřeného přínosem intervenčního programu tak, jak ho vnímají relevantní zainteresované strany. V tomto ohledu je Responsivní přístup velmi blízký přístupu konstruktivistickému; současně se však od něj významně odlišuje tím, že nepovažuje evaluátory za zainteresovanou stranu *sui generis*²⁸.

²⁷ Ve Stakeově pojetí je myšlenka formativní funkce evaluací velmi zřetelná.

²⁸ Konstruktivistický přístup je v detailu popsán níže. Konstruktivisté považují roli evaluátora za aktivní (evaluátor je zúčastněný), což je v případě responzivního přístupu naopak – evaluátor je „pasivním posluchačem“.

Pohledem výzkumné metodologie je tak Stakeův přístup *de facto* apelem na využívání smíšené metodologie. Jak sám uvádí, „oba způsoby myšlení mohou stát vedle sebe“ (Stake 2004). R. Stake sice v tomto případě nepíše o kvalitativní, resp. kvantitativní metodologii, nýbrž odkazuje k responsivnímu, resp. cílenému přístupu k provádění evaluačních šetření, nicméně kvalitativně-quantitativní diskurs je přesto v pozadí jasně patrný. R. Stake ve svém responsivním přístupu sice volá po takových postupech, které by při provádění evaluačních šetření dokázaly reflektovat komplexitu posuzovaných intervenčních programů, současně však zdůrazňuje, že koncept měření jednotlivých jevů nelze opustit. Uplatnění responsivního přístupu tak obvykle vede k méně objektivním závěrům než cílený přístup, na druhé straně jsou však takto prováděná evaluační šetření lépe přizpůsobena naturelu samotných intervenčních programů.

Ve Stakeově responsivním přístupu k realizaci evaluačních šetření je zřetelná inspirace přístupem M. Scrivena a jeho apelem na uskutečňování hodnotících soudů²⁹. R. Stake však ve svém přístupu usiluje o zjemnění Scrivenova ultimativního imperativu tím, že varuje před kontraproduktivní eskalací vztahů mezi evaluátorem a zainteresovanými stranami, k níž důsledné následování Scrivenovy výzvy nevyhnutelně vede. V neposlední řadě poukazuje na relativitu těchto závěrů danou mnohostí jednotlivých zainteresovaných stran, neboť co je nejlepší pro jednu stranu, nemusí být nejlepší pro ostatní.

Ilustrační příklad:

V rámci responsivního přístupu není zvolen žádný způsob formální examinační, ale evaluátor např. navodí určitou situaci, kdy sám (sociálně skrytý jinou rolí) požádá účastníka kursu o radu týkající se práce s PC. Pokud účastník kursu dokáže v dané situaci zareagovat a využije přitom poznatky, které tvořily páteř daného kursu, bude daný intervenční program považován za účinný. Responsivní evaluátor projedná předem se zadavatelem konkrétní akceptační kritéria, tj. okolnosti, za jakých budou účastníci prověřováni, a význam, jaký bude přikládán konkrétním dovednostem.

²⁹ Scriven (1967) píše: „Evaluators must judge.“

4.1.4 Utilizační přístup k realizaci evaluačních šetření (utilization-focused evaluation; Michael Quinn Patton)

Základní myšlenka tohoto přístupu není zcela originální, neboť již Suchman (1967) uvádí, že „úspěch evaluačního projektu je závislý na jeho užitečnosti pro zadavatele, který se snaží zlepšit poskytované služby“. E. Suchman dále shrnuje, že evaluátor musí mít neustále na paměti potenciální užitečnost svých závěrů. Také R. Stake věnoval celou jednu polovinu responsivity svého přístupu využitelnosti provedených evaluačních šetření a nabádal, aby evaluátoři byli vnímaví vůči požadavkům zainteresovaných stran, které budou výsledky provedených evaluačních šetření využívat. M. Patton však učinil z kritéria využitelnosti a užitečnosti provedených evaluačních šetření jádro svého přístupu, zasadil ho do náležitého kontextu a dovedl k uceleně koncipovanému způsobu, jak k evaluačním šetřením a k jejich realizaci přistupovat.

M. Patton ve svém pojetí předpokládá, že evaluační šetření jsou od svého počátku vedena v těsné interakci evaluátora s klíčovými zainteresovanými stranami, a to zejména s těmi, které budou na základě výsledků dokončeného evaluačního šetření činit konkrétní rozhodnutí. V jeho snaze nejde pochopitelně o prvoplánové naplňování (jakýchkoliv) potřeb a požadavků konkrétních uživatelů závěrečných zpráv z evaluačních šetření, ale jde o rozsáhlou změnu v designu evaluačních šetření a úpravě jednotlivých postupů a procedur, díky níž by výsledky evaluačních šetření byly nadány vysokým potenciálem využitelnosti v daných podmínkách. S tím rovněž souvisí způsobilost závěrů prováděných evaluačních šetření podporovat zainteresované strany s rozhodovacími pravomocemi ve vztahu k intervenčnímu prostředí, přinášet tak skutečnou a prakticky využitelnou zpětnou vazbu a podílet se na dalším směřování těchto intervenčních programů. M. Patton si v tomto ohledu uvědomuje, že smysluplnost provádění evaluačních šetření a využívání jejich poznatků při řízení intervenčních programů je do značné míry závislá na tom, zdali a v jaké míře zadavatelé těchto šetření a další zainteresované strany (např. manažeři posuzovaných intervenčních programů) akceptují výsledky provedené evaluace. Svým přístupem se tedy M. Patton snaží předejít tomu, aby výsledky provedených šetření byly manažery intervenčních programů odkládány pro nepochopení či pro jejich (domnělou) irelevanci nebo z důvodu nedostatečné spolehlivosti formulovaných závěrů.

Nejde však jen o důraz na apriorní identifikaci uživatelů výsledků evaluačních šetření coby specifického typu zainteresované strany; M. Patton současně zdůrazňuje roli aktuální situace a konkrétního způsobu, jakým jsou výsledky provedeného evaluačního šetření využívány. V tomto ohledu lze tedy hovořit o jakési duální specifičnosti: na jedné straně je třeba uvažovat o specifičnosti subjektu, tedy přesné identifikaci a vymezení subjektů, kterým jsou doručovány výsledky provedených evaluací. Na druhé straně jde o specifičnost použití, tedy stanovení podmínek, možností a omezení, za kterých budou výsledky evaluací využívány (obyčejně se povaha a typ rozhodnutí, která se od výsledků evaluace budou odvíjet, časový rámec daného rozhodnutí, dopady rozhodnutí učiněného na základě výsledků provedené evaluace apod. specifikuje v zadávací dokumentaci k provedení evaluačního šetření). Používané metody a techniky jsou pak zvoleny s ohledem k povaze zainteresovaných stran používajících výsledky evaluačních šetření a způsobu jejich použití (Patton 1987).

Utilizační přístup k provádění evaluačních šetření tedy z celého spektra zainteresovaných stran obrací svou pozornost k těm subjektům, které činí na základě provedených evaluačních šetření další rozhodnutí. To totiž významným způsobem determinuje způsob, jakým jsou daná evaluační šetření prováděna. V porovnání se Stakeovým responsivním přístupem, který jako by balancoval mezi odpovědností vůči zadavatelům evaluačních šetření a zájmy případných beneficentů či klientů daných intervenčních programů, M. Patton přenáší veškerý důraz na zájmy zadavatelů evaluačních šetření. Vysoká pozornost věnovaná využitelnosti výsledků prováděných evaluačních šetření nachází svůj odraz také v jejich metodice, ve které je prvořadá pozornost věnována externí validitě. Vzhledem k tomu se u M. Pattona setkáme s větším odstupem od provádění experimentů a používání experimentálního designu, než např. u R. Tylera, jehož cílený přístup je s experimentálním designem bytostně spjatý.

Důraz na zainteresované strany, který je kromě utilizačního přístupu k provádění evaluačních šetření charakteristický také pro Stakeův Responsivní přístup, vede k individualizovanému či snad až personifikovanému popisu evaluační situace.

Zatímco M. Scriven či R. Stake předkládají své přístupy jako obecný způsob provádění (těch správných) evaluačních šetření, M. Patton od začátku koncipuje svůj přístup jako alternativu³⁰. M. Scriven pochopitelně připouští, že jeho explorační přístup lze kombinovat

³⁰ Např. výchozí definici utilizačního přístupu formuluje jako doplněk k obecným definicím evaluace.

s přístupem cíleným, a dokonce k tomu sám nabádá, nicméně zároveň poukazuje na omezenou validitu evaluačních studií založených výlučně na cíleném přístupu. Podobně také R. Stake naznačuje, že jiné než responsivní evaluace nemohou poskytnout tak bohatý popis posuzovaného prostředí, jaký nabízí jím navrhovaný přístup. M. Patton v tomto ohledu říká, že každá evaluace je vlastně svým způsobem (a v jisté míře) utilizačně zaměřená; nakonec tuto vlastnost jasně vyžadují i standardy pro provádění evaluačních šetření definované Joint Committee on Standards (cit. dle Sanders 1994). Pokud však jde o celkový přístup k volbě uspořádání daného evaluačního šetření, je ponecháno na evaluátorovi samotném, zdali podřídí zájmům zadavatele celý design, či nikoliv, zdali zvolí utilizační přístup, či se rozhodne pro nějaký jiný.

Přirozenou konsekvencí utilizačního přístupu k provádění evaluačních šetření je normativní relativismus: prováděná evaluační šetření mohou jednostranně zohledňovat normy a hodnoty spojené s jednou jedinou zainteresovanou stranou, se zadavatelem. Zde je patrná další odchylka od Scrivenova přístupu, neboť jeho explorační přístup spoléhá v prvé řadě na individuální etiku samotného evaluátora a teprve v druhé řadě se otevírá všem ostatním zainteresovaným stranám. Utilizační přístup má tedy v tomto ohledu mnohem blíže ke Stakeově responsivnímu přístupu, který je rovněž založen na předpokládaném hodnotovém relativismu vycházejícím z hodnot jednotlivých zainteresovaných stran, s výjimkou evaluátora samotného.

Utilizační přístup není (tak jako řada jiných přístupů) založen na rozvinutí jiného konceptu či na jeho kritice. M. Patton jej opírá o své vlastní zkušenosti s realizací evaluačních šetření a o praktické zkušenosti mnohých dalších evaluátorů, kteří se při provádění evaluačních šetření potýkají s proměnlivým přijetím svých snah, vycházejícím z individuální podstaty jednotlivých zainteresovaných stran. Např. Cronbach poukazoval na skutečnost, že nic neovlivňuje využívání výsledků evaluačních šetření více než osobní faktor – zájem zadavatelů poučit se z provedeného evaluačního šetření a snaha evaluátora zaujmout své auditorium (Stufflebeam, Shinkfield 2007).

To, co je na jedné straně klíčovou výhodou utilizačního přístupu k provádění evaluačních šetření (tedy zapojení zainteresovaných stran do samotného středu provedení daného evaluačního šetření), je současně také jeho největší slabinou. Sebemenší změna ve struktuře zainteresovaných stran, byť by se jednalo o výměnu jednoho jediného pracovníka, totiž může

celý proces evaluačního šetření vrátit na samý začátek nebo ho fakticky zastavit či významně zdržet. Spoléhání se na hodnoty, zájmy a očekávání jednotlivých zainteresovaných stran sice přibližuje prováděná evaluační šetření skutečným situacím (a pravděpodobně tak posiluje jejich vnější validitu a zvyšuje pravděpodobnost jejich faktického využití), nicméně staví tím evaluátory do středu manipulativních záměrů jednotlivých zainteresovaných stran, které se mohou snažit prostřednictvím formulace svých vlastních požadavků a záměrů dosahovat svých partikulárních (a častokrát vychýlených) cílů.

Jiná problematičnost utilizačního přístupu k provádění evaluačních šetření pramení ze samotného způsobu využití jejich výsledků. Zdaleka ne vždy lze totiž M. Pattonem požadovaný způsob použití předem stanovit či odhadnout (některá jednorázová evaluační šetření jsou realizována po dobu několika (mnoha) měsíců, přičemž existují i průběžně realizovaná a pravidelně opakovaná evaluační šetření, jež probíhají třeba i několik let). Dále se v konkrétní aplikační praxi ukazuje, že způsoby využití výsledků provedených evaluačních šetření mohou být mnohačetné, a tak by stanovení adekvátní váhy (a potažmo i významu jednotlivých způsobů použití) mohlo být v řadě případů náročnější, než samotné provedení evaluačního šetření. Utilizační přístup tak vlastně vypovídá o snaze evaluátorů vyjít vstříc zadavatelům evaluačních šetření, kteří nedovedou zasadit získané závěry do rámce probíhajících, resp. navazujících procesů.

Ilustrační příklad:

Evaluátor využívající přístup, který je zaměřen na využitelnost provedeného evaluačního šetření, nejprve prodiskutuje se zadavatelem konkrétní okolnosti, za jakých bude s výsledky provedeného šetření nakládáno a jaký typ rozhodnutí bude na základě provedené evaluace učiněn. V daném případě evaluátor zvolí některý z kvantitativních přístupů, neboť výsledky evaluace budou podporovat zásadní rozhodnutí o dalším osudu daného intervenčního programu (rekvalifikačního kursu). Jeho jasně sumativní charakter tak povede k volbě podobného uspořádání jako v případě cíleného přístupu. Evaluátor však bude věnovat pozornost nejen výsledkům, kterých jednotliví účastníci v testu dosáhli, ale zaměří se také na úspěšnost účastníků daného kursu při hledání zaměstnání a bude ji porovnávat s úspěšností ostatních nezaměstnaných starších 50 let (zvolen bude nejspíše kvaziexperimentální design). Na základě srovnání dosažených výsledků doporučí daný intervenční program k jeho rozšíření či navrhne jeho zastavení.

4.1.5 Konstruktivistický přístup k realizaci evaluačních šetření (constructivist evaluation; Egon Guba, Yvonna Lincolnová)

Konstruktivistický přístup (samotnými autory označovaný také jako „evaluace čtvrté generace“ či původně publikovaný pod názvem „naturalistický přístup“) je reakcí na obecně převládající pozitivistický přístup k chápání reality jako objektivní danosti a k možnostem jejího zkoumání a hodnocení. Čelní představitelé konstruktivistického přístupu, E. Guba a Y. Lincolnová, se přibližně od počátku 80. let snažili rozprodit vývoj evaluačního oboru diskuzí o hodnotovém pluralismu, kritikou přílišné koncentrace evaluátorů na výzkumnou metodologii a poukazováním na nevhodně koncipovaná evaluační šetření, která častokrát zúžila posuzované situace natolik, že ztratily svou vnější validitu.

V rámci tohoto historického kontextu je třeba poukázat na skutečnost, že 80. léta znamenala oproti předchozím dvěma desetiletím zásadní redukci počtu evaluačních šetření požadovaných federální vládou od externích evaluátorů, což bylo způsobeno jednak rozpočtovými škrty Reaganovy administrativy (bylo méně intervenčních programů, jež by bylo třeba evaluovat), a dále také významným posílením role interních evaluátorů, či spíše interních evaluačních týmů, působících v rámci jednotlivých ministerstev, což je považováno za důsledek rostoucího významu evaluací³¹. Evaluátoři se tak více než dříve obraceli k principiálním otázkám podstaty evaluací, na rozdíl od praktických aspektů provádění evaluačních šetření, jež dominovaly v předchozím období.

E. Guba a Y. Lincolnová při formulaci svého přístupu vycházeli jednak ze Stakeova responzivního přístupu, ze kterého převzali tezi, že evaluačního šetření by se měl účastnit nejen evaluátor a zadavatelé, ale také všechny další zainteresované strany, kterých se prováděné evaluační šetření dotýká. Konstruktivisté v tomto ohledu dokonce přiznávají všem zainteresovaným stranám, které jsou zapojeny do prováděného evaluačního šetření, plně angažovanou účast zahrnující jak spolurozhodování v otázkách cílů daného šetření, tak také právo ovlivňovat jednotlivé činnosti směřující k závěrům daného evaluačního šetření. Druhým pilířem, o nějž se tento přístup opírá, je konstruktivistická metodologie založená na hermeneutické dialektice. Evaluační šetření založená na konstruktivistickém přístupu tak nejprve identifikují, sumarizují a analyzují jednotlivé mentální konstrukce, které jsou

³¹ E. House (1993) v této souvislosti píše: „Evaluace začaly být až příliš důležitou činností na to, aby byly ponechány jen v rukách akademiků.“

v druhém kroku podrobeny vzájemnému srovnání a konfrontaci. V tomto ohledu E. Guba s Y. Lincolnovou zpochybňují epistemologické možnosti nezávislého a nezúčastněného poznávání (objektivní) reality a kladou je do protikladu s transakčním subjektivismem, o nějž se opírá jimi zdůrazňovaný intersubjektivní (sdílený) charakter individuálních konstrukcí vnějšího světa.

Na rozdíl od některých jiných přístupů, např. utilizačního či exploračního, je konstruktivistický přístup jasně spjat s konkrétními metodami a technikami. Rovněž jeho paradigmatická vyhraněnost a jasné paradigmatické zakotvení z něj činí jistou výjimku mezi ostatními přístupy. Jako jeden z mála přístupů (pokud ne jako jediný) odmítá objektivistický pohled na evaluace a vyzdvihuje předpoklad o neexistenci objektivní pravdy a objektivní hodnoty. Žádná z konstrukcí jednotlivých zainteresovaných stran není podle E. Guby a Y. Lincolnové o nic pravdivější a opravdovější než jiná konstrukce kteréhokoliv z dalších účastníků prováděného evaluačního šetření.

Konstruktivistický přístup je ze své podstaty zaměřený na jednotlivé zainteresované strany. V tomto ohledu lze najít podobnost se Stakeovým responsivním přístupem (což je zřejmě s ohledem k již uvedené inspiraci, kterou konstruktivistický přístup našel právě v responsivních evaluacích) i s Pattonovými evaluacemi utilizačními. Všechny tyto přístupy totiž kladou do středu své pozornosti buď všechny, nebo alespoň některé zainteresované strany a věnují značnou pozornost analýze jejich role v posuzovaném intervenčním programu a sledování konkrétních zájmů, jimiž jsou tyto zainteresované strany vedeny. Naproti tomu explorační (M. Scriven) a cílené (R. Tyler) přístupy obracejí svou pozornost k cílům sledovaných intervenčních programů a k metodám, jak co nejlépe ověřit jejich dosažení. Zainteresované strany vystupují v takto pojatých evaluačních šetřeních jen jako jeden z činitelů, jako jakási částička, v níž se zrcadlí přínos posuzovaných intervenčních programů, nikoliv však jejich věcný základ.

Mnoho evaluátorů poukazuje na značnou obtížnost provedení evaluačních šetření způsobem, který by využíval konstruktivistický přístup, přesto se však lze s konkrétními příklady jejich praktického využití setkat. Významnou bariéru aplikačního rozvoje konstruktivistického přístupu k realizaci evaluačních šetření představuje nutný souhlas zainteresovaných stran s principiálním cílem takto prováděných šetření, jímž je konvergence protichůdných zájmů, vizí, zkrátka subjektivních konstrukcí a apriorní zájem na dosažení konsenzu. Pawson a Tilley

(1997) proto zařazují konstruktivistický přístup založený na negociaci mezi nositeli jednotlivých konstrukcí do „útluného předměstského světa, v němž se odehrávají setkání mezi řediteli škol, učiteli a rodiči žáků, při kterých marní čas diskuzemi o charakteru domácích úkolů pro pátou třídu“. Rovněž další z předpokladů, hermeneutický dialektický proces, činí z konstruktivisticky prováděných evaluačních šetření dílo, které je v rámci obvykle požadovaných časových rámců (řádově měsíce) v podstatě neuchopitelné a neproveditelné. Konstruktivistický přístup tak sice nezměnil příliš evaluační praxi a vžité způsoby provádění evaluačních šetření, nicméně významným způsobem ovlivnil evaluační teorii a jakési „základy evaluačního myšlení“. Je téměř nemožné najít monografii či učebnici evaluací vydanou v posledních pětadvaceti letech, která by na konstruktivistický přístup neodkazovala. Je třeba na tomto místě poukázat na skutečnost, že samotný přístup se stal inspirací pro mnohé další (např. participativní přístup) a má rovněž své přímé pokračovatele – např. Bhola (1998), Schwandt (1984) či Huebner, Betts (1999).

V neposlední řadě je třeba vyzdvihnout přínos řady dílčích konceptů a námětů, které se v evaluační praxi rychle rozšířily bez ohledu na konkrétně uplatňovaný přístup. V tomto směru především idea identifikace všech subjektů dotčených evaluačních šetření a jejich zahrnutí do procesu evaluačních šetření coby specifických zainteresovaných stran je dnes obecně akceptovaným doporučením (Lay, Papadopoulos 2007), které má dokonce svůj normativní rozměr, neboť je definováno v Program Evaluation Standards (cit. dle Sanders 1994) jako *sine qua non* každého korektně provedeného evaluačního šetření.

Konstruktivistický přístup jako celek je však mnohými evaluátory (např. Scriven 1991) považován za „radikální přístup v současné epistemologii a výzkumné metodologii“, který na vlně kritiky využitelnosti experimentů v evaluační praxi vylil s vaničkou i dítě (Pawson, Tilley 1997). Konstruktivistickému přístupu nelze nevytknout neopodstatněný imperativ parity zainteresovaných stran v kontextu zjevně disparitní reality, kterou navíc všechny ostatní přístupy k provádění evaluačních šetření plně akceptují a reflektují. Další z odlišností konstruktivistického přístupu, která omezuje jeho využitelnost, se projevuje v nemožnosti posoudit adekvátnost závěrů z hlediska zavedených pojmů, jako je validita a reliabilita. E. Guba s Y. Lincolnovou sice v tomto ohledu nabízejí alternativní konstrukty (např. zavádějí pojem kredibility coby náhrady za vnitřní validitu či dependabilitu jako alternativu reliability), což však problém jisté izolace konstruktivistického přístupu neřeší.

Ilustrační příklad:

Při uplatnění konstruktivistického přístupu pozve evaluátor na společný brainstorming vybrané účastníky daného kursu; nezaměstnané, kteří nebyli do kursu přijati; lektory; zástupce potenciálních zaměstnavatelů; přímého nadřízeného zaměstnance, který se v minulosti kursu zúčastnil a získal nové zaměstnání; pracovníky daného úřadu práce; zástupce iniciátora daného intervenčního programu a rodinné příslušníky účastníků kursu. V průběhu diskuzí, které bude evaluátor moderovat, se zaměří na očekávání, s nimiž jednotlivé zainteresované strany vstupovaly do dané situace (konfrontace subjektivit) a bude zjišťovat, která z těchto očekávání se naplnila, a která nikoliv. Následně bude hledat způsoby, jak dosáhnout uspokojení potřeb každého z účastníků, a tyto diskuze budou opakovány do doby, než zúčastněné strany dosáhnou konsenzu. Výsledkem bude návrh uspořádání kursu, který bude iniciátorem považován za užitečný, (potenciální) účastníci v něm budou spatřovat cestu k získání nového zaměstnání, pracovníci úřadů práce budou přesvědčeni, že absolventy takto koncipovaného kursu dokáží snáze umístit na trhu práce a zaměstnavatelé budou absolvování kursu vnímat jako doklad o schopnostech a dovednostech uchazečů o zaměstnání.

4.1.6 Přístup k provádění evaluačních šetření vedený teorií (theory-driven evaluation; Huey-Tsyh Chen, Peter Rossi)

Přístup k provádění evaluačních šetření, který je veden teorií, lze považovat za příspěvek k diskusi mezi evaluátory o možnostech a limitech uplatnění randomizovaných experimentů. Představitelé tohoto přístupu nerozšiřují řady radikálních kritiků, kteří odmítají jak samotné experimenty, tak také jakýkoliv jiný design empirické studie šité podle pozitivistického stříhu. Přístup k provádění evaluačních šetření, který je veden teorií, se v tomto ohledu distancuje jak od konstruktivistického přístupu E. Guby a Y. Lincolnové, tak také od jejich descendentů v podobě participativních evaluací a řady dalších přístupů (zejm. od přístupu založeného na spolupráci /*collaborative*/ a od Fettermanova zplnomocňujícího přístupu /*empowerment*/).

Na druhé straně však nelze teorií vedený přístup přiřadit ani k obhájcům a šířitelům experimentálního designu, jakými byli např. D. Campbell, J. Stanley nebo T. Cook. Čelný představitel tohoto přístupu, H. Chen, obrací pozornost evaluátorů k mechanismu fungování intervenčních programů, a vytváří tak nový metodologický rámec, v němž mohou být

prováděny jak (dílní) experimenty, tak také jiné typy šetření. Významnou změnu oproti ostatním přístupům představuje způsob využití takto získaných údajů. Ty jsou totiž vztahovány k teorii popsanému kontextu daného intervenčního programu a především pak slouží k vysvětlení způsobu, jak intervenční program vlastně funguje. Prvek vysvětlení je v evaluačních šetřeních naprosto ojedinělý. Z mnoha pramenů je patné, že smysl, cíle a poslání evaluací (a evaluačních šetření) jsou spatřovány zejména v popisu (deskripci) a v jisté omezené míře v exploraci, nicméně pokus o vysvětlení (např. způsobu fungování intervenčního programu) se v uchopitelné a systematicky pojaté podobě objevuje až u H. Chena. Rozvoj přístupu k provádění evaluačních šetření, který je veden teorií (tj. obecně představou o mechanismech fungování posuzovaného intervenčního programu) později vyústil v rozvinutí ještě jiného přístupu, a to realistického.

Chenův přístup k evaluacím v podstatě směřuje k eliminaci interpretační nejistoty (zejména) v případě přijetí nulové hypotézy o fungování intervenčního programu v podmínkách experimentálního designu. Pokud totiž není shledán statisticky významný rozdíl mezi testovací a kontrolní skupinou, nabízí se vždy několik zásadních otázek či pochybností: Znamená výsledek testu, že daný intervenční program skutečně nevyvolává očekávané efekty? Co když je neprůkazný výsledek způsoben navrženým designem? Co když jsou použity takové nástroje měření, které jen nejsou dostatečně citlivé, aby mohly způsobený efekt zachytit? Mohli bychom dosáhnout jiného výsledku, tj. prokázat funkčnost daného intervenčního programu, pokud bychom zvýšili počet případů/velikost testovací a kontrolní skupiny? H. Chen proto v této souvislosti nabízí řešení, které spočívá v doplnění poznatků získaných z evaluačních šetření založených na experimentálním (či kvaziexperimentálním) designu o nový druh poznatků a znalostí – o „teorii“ působení daného intervenčního programu (Donaldson 2007; Rogers 2008.).

Ústřední pojem – „teorie“ – tedy v podání svého autora znamená pravděpodobný a obhajitelný model očekávaného působení daného intervenčního programu (Chen, Rossi 1983). Chenova „teorie“ je jakousi představou o mechanismech, kterými by daný intervenční program mohl působit a které by mohly vyvolávat pozorovatelné výstupy. Kondicionál v poslední větě je zvolen záměrně, neboť H. Chen sám nechápal svou definici teorie deskriptivním způsobem, tedy jako konstatování faktického stavu, nýbrž preskriptivně, tedy jako vyjádření žádoucího, lepšího či optimálního stavu (Chen 1990). Chenovu teorii by tedy bylo možné označit jako soubor vzájemně provázaných hypotéz o vztazích mezi jednotlivými

událostmi vyskytujícími se v rámci prvků daného intervenčního programu. Vlastní proces evaluačního šetření je pak *de facto* verifikací tohoto souboru hypotéz a modifikace výchozí představy na základě empiricky získávaných údajů.

Přístup k provádění evaluačních šetření, který je veden teorií, je založen na třech elementárních principech. Tím prvním je jeho zaměření na změnu (*action oriented*); zahrnuje totiž detailní popis jednotlivých mechanismů, pomocí nichž intervenční programy mohou dosahovat svého zlepšení. V tomto ohledu nejde ani tak o popis výchozí struktury či o zachycení jednotlivých prvků samotných intervenčních programů jako při tvorbě konceptuálních modelů, ale spíše o zachycení dynamiky intervenčního programu. Druhý princip Chenova přístupu je založen na konceptualizaci samotné intervence a detailním rozboru použité implementační strategie; v tomto ohledu se teorií vedený přístup k provádění evaluačních šetření zaměřuje na analýzu konkrétně použitých nástrojů dané intervence a na intenzitu jejich použití. Třetím z principů tohoto přístupu je volba kritérií, na jejichž základě jsou posuzovány výsledky daného intervenčního programu. Zde H. Chen navazuje jak na R. Tylera, tak také na M. Scrivena, kteří v rámci svých přístupů věnovali volbě a adekvátnímu stanovení těchto kritérií značnou pozornost. H. Chen však volí svůj odlišný způsob jejich volby – navrhuje totiž, aby byla volba kritérií pro posouzení účinnosti daného intervenčního programu ponechána na samotných zainteresovaných stranách.³²

H. Chen v rámci svého přístupu definoval celkem šest oblastí³³, jimž je třeba při provádění evaluačních šetření věnovat pozornost. Jde (1) o samotnou povahu intervence (*treatment*), jejíž poznání a analýza je nezbytným východiskem pro zachycení změn vyvolaných daným intervenčním programem, (2) o implementační prostředí, které jasně odkazuje ke kontextu, v němž se posuzovaná intervence, resp. intervenční program odehrává, a (3) o výsledky, ke kterým daný intervenční program dospěl. Druhá trojice oblastí zahrnuje jednak (4) dopady, jejichž detailní analýzu H. Chen považuje za ústřední téma celého evaluačního šetření, (5) intervenující mechanismus, což je v podstatě kauzální vztah spojující povahu intervence (*treatment*) a její výsledek (*outcome*) a (6) zobecnění, v souvislosti s nímž H. Chen upozorňuje, že získané poznatky mohou být využívány i s časovým odstupem, pro jiné populace či v jiných organizacích.

³² Pokud tato výzva zní jako utopie, která nemůže vést v praxi k adekvátním závěrům, odkazujeme na pravidla stanovení indikátorů využívaných pro monitoring a evaluaci operačních programů Evropské unie.

³³ H. Chen je nazývá domény (*domains*).

Chenovým cílem, který anticipuje v rámci svého přístupu, je snaha o dosažení vyváženého poměru mezi vnitřní a vnější validitou. To je další z mnoha prvků, jimiž se Chenův přístup odlišuje od všech ostatních. Jiné přístupy, jak H. Chen sám poukazuje, totiž buď kladou důraz jen na jeden z obou typů validity, nebo si problém validity vůbec nepřipouštějí (Chen 1990). H. Chen v tomto ohledu odkazuje na vyhlášené spory W. Gosseta s R. Fischerem vedené o povaze a vhodnosti znáhodněných experimentů, přičemž Chen jednoznačně straní Gossetovi, který namísto experimentálního designu preferoval takový design, který je sice méně exaktní, zato však usiluje o snížení celkové chyby a snaží se zvolit takové uspořádání, jehož postupy se co nejvíce podobají reálným situacím zkoumaných subjektů (Alkin 2004).

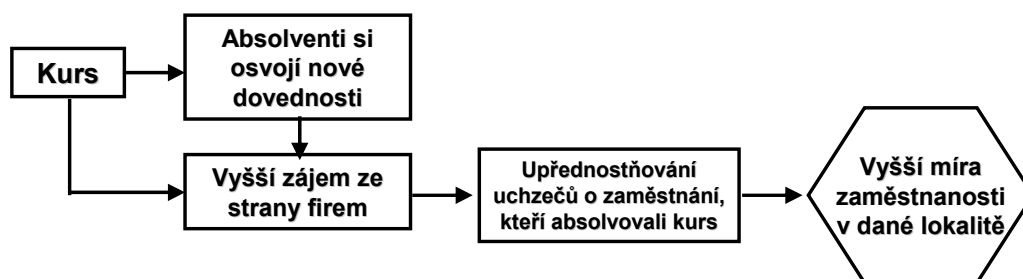
Nespornou výhodou Chenova evaluačního přístupu je, že konkrétní rozhodnutí, jež je při realizaci evaluačního šetření třeba učinit, jsou rámována teorií. Znamená to tedy, že charakter jednotlivých kauzálních vazeb, povaha sledovaných znaků (indikátorů) i samotná interpretace získaných výsledků má jasný a dokumentovaný základ, který usnadňuje implementaci dosažených výsledků (Molas-Gallard, Davies 2005).

Přístup k provádění evaluačních šetření, který jsou vedená teorií, lze tedy využívat zejména v takových případech, kdy není možné testovat vztahy mezi jednotlivými proměnnými přímo; v těchto situacích dokáže teorií vedený přístup zvýraznit výsledky sledovaných programů a zároveň také přesně specifikovat jednotlivé faktory, které k celkovému výsledku největší měrou přispívají. Daných závěrů je při využití tohoto přístupu dosaženo prostřednictvím deskripce mechanismu působení intervenčního programu a prostřednictvím specifikace a monitorování determinujících činitelů.

Ilustrační příklad:

Evaluátor provádějící evaluační šetření v duchu tohoto přístupu, si po úvodním seznámení s danou situací připraví „teorii“, jak daný kurs dosahuje očekávaných výsledků. Následně připraví takový design evaluace, který umožní vytvořenou teorii ověřit. Takový model může mít třeba podobu, která je patrná z následujícího obrázku:

Obrázek č. 11 Návrh konceptuálního schématu



Zdroj: autor

V průběhu evaluace bude výchozí model postupně upravován podle získávaných poznatků tak dlouho, až bude připravena finální podoba „ověřené“ teorie. Evaluátor bude během přípravy takové evaluace spolupracovat s jednotlivými zainteresovanými stranami, bude získávat konkrétní podkladové údaje o jednotlivých prvcích své teorie a bude přitom klást důraz na klíčové procesy. Ve své odpovědi zadavateli pak evaluátor zdůrazní kritické momenty, které se spolupodílely na úspěšnosti daného intervenčního programu.

4.1.7 Realistický přístup k provádění evaluačních šetření (realistic evaluation; Ray Pawson, Nick Tilley)

Přístup k provádění evaluačních šetření, který vychází z realismu, je podle svých autorů zlatým středem mezi dvěma proti sobě stojícími paradigmaty, pozitivismem a relativismem (Julnes, Mark, Henry 1998). Z výše uvedených přístupů k provádění evaluačních šetření je patrné, že oba z paradigmatických protipólů našly svůj odraz i v metodologii evaluačních šetření, neboť lze identifikovat jak přístupy odvolávající se k pozitivismu (zejména cílený přístup), tak také přístupy založené na relativistické tradici reprezentované konstruktivistickým přístupem E. Guby a Y. Lincolnové a jejich následovníky. Přístup k realizaci evaluačních šetření založený na realismu tedy jakoby obchází tyto tradiční póly, poukazuje na jejich neadekvátnost a zaměřuje svou pozornost na mechanismy vysvětlení sledovaných jevů.

R. Pawson a N. Tilley popisují svůj přístup k evaluačním šetřením na základě tří klíčových slov, která v sobě pojem „realistický“ obsahují. Jde o koncept skutečnosti (reality), realismu a realističnosti.

První pojem (*real*) odkazuje k předmětu evaluace. R. Pawson a N. Tilley v tomto ohledu jasně deklarují, že všechny jevy a fenomény jsou v rámci jejich realistického přístupu k provádění evaluačních šetření považovány za skutečné, za reálně existující. Reální jsou tedy nejen aktéři jednotlivých intervenčních programů, ale i samotné intervence a pochopitelně také jejich (dobré či nedostatečné) výsledky. Realistický přístup tedy odmítá zkoumat tyto jevy jako „narativní významy, které si vytvářejí jednotliví hráči“ (Pawson, Tilley 1997). Realistický přístup se přitom snaží o sociologické pochopení rovnováhy mezi zdroji a možnostmi, které jednotlivé zainteresované strany mají. Jejich interakce a vzájemná závislost je pochopitelně přítomná i ve skutečném světě, a proto je předmětem analýzy realistického přístupu (ibid.).

Druhé klíčové slovo (*realist*) se vztahuje k používané metodologii. Autoři realistického přístupu poukazují na mnohé chyby a nedokonalosti, které byly v různých fázích vývoje metodologie evaluačních šetření způsobeny přílišným zaměřením na technické detaily metod samotných, přičemž jejich podstata a obecný smysl se mnohdy vytratily. Svým přístupem se tak snaží vytvořit větší prostor pro „teorii“³⁴ a zdůrazňují přitom, že podstata vědeckého vysvětlení nemůže být svázána s žádnou jednotlivou metodou či způsobem měření. V minulosti se takovýto přístup projevoval příliš mechanickou aplikací experimentů a vedl k neuspokojivým výsledkům.

Třetím pojmovým prvkem (*realistic*) autoři apelují na samotný způsob provádění evaluačních šetření. Upozorňují, že tato šetření jsou jednou z forem aplikovaného výzkumu a jejich smyslem není formulovat obecné zákonitosti a spět k univerzálně platným závěrům, ale pomáhat v konkrétních situacích zainteresovaným stranám získat realistický obraz³⁵ posuzovaného intervenčního programu a realistické (tedy skutečnosti odpovídající) závěry. Aplikační možnosti evaluačních šetření jsou tedy dány schopností jednotlivých metod a technik přinášet informace o potvrzení či vyvrácení principů působení intervenčních programů pomocí generativní kauzality.

Východiskem realistického přístupu R. Pawsona a N. Tilleyho je realistická tradice filozofie vědy reprezentovaná zejména R. Bhaskarem, R. Harrém, I. Lakatosem a M. Hessem;

³⁴ Srovnávat smysl a význam pojmu s Chenovým přístupem; význam pojmu *teorie* a způsob jeho chápání je v realistickém přístupu velmi podobný.

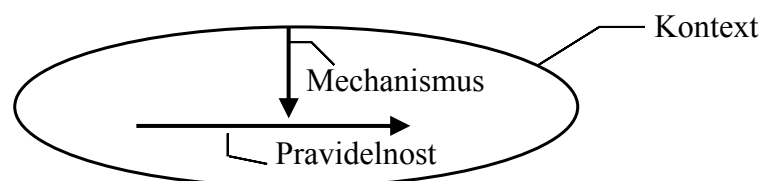
³⁵ Myslí se obraz, který není zahalen výzkumnou terminologií a je jen minimálně vychýlen samotnými metodami.

následování této tradice je užitečné chápat v obecnějším rámci jednoho z autorů (R. Pawsona), který se o tradici realismu opírá nejen v přístupu k evaluačním šetřením, ale rovněž ve výzkumné metodologii (viz zejména Pawson 1989) a v dalších aplikovaných oblastech (Pawson 2006).

Jádrem realistického přístupu je teze o neadekvátnosti experimentu jako metody zkoumání a odhalování kauzality. R. Pawson a N. Tilley přitom poukazují na odlišné použití experimentů ve společenských vědách oproti vědám přírodním a potažmo také na odlišné chápání vztahu mezi kauzalitou a metodou experimentu. Dokumentují přitom, že experimenty v přírodních vědách vycházejí z generativního chápání kauzality, tedy slouží k odhalování obecných mechanismů, jež se vyskytují v rámci daných vnějších podmínek, v rámci konkrétního kontextu. Naproti tomu společenské vědy žádné mechanismy pomocí experimentů neodhalují, kauzalitu považují za vztah, který je pouhou spojnicí ve vztahu mezi příčinou a následky. Společenské vědy tedy používají experimentální metody ke stanovení inkrementálních efektů, nicméně takové (*sukcesionistické*) chápání kauzality nevypovídá nic o transformativním potenciálu sledovaných jevů (Pawson, Tilley 1997). Namísto toho se obracejí k realismem popsané generativní kauzalitě.

Jádrem realistického přístupu jsou pravidelně opakované jevy, postupy či procesy, které vedou k daným výsledkům. Této pravidelnosti je dosaženo díky mechanismu, který působí v daném kontextu. Mechanismus lze přitom chápat a popisovat jak na mikroúrovni, která má individualistický charakter (zohledňuje např. izolaci jednotlivce, jeho subjektivní důvody apod.), tak také na makroúrovni, jež odkazuje ke společenským procesům a institucím (např. koheze či solidarita).

Obrázek č. 12 **Princip realistických evaluací**



Zdroj: Pawson, Tilley 1997

Autoři při explikaci svého přístupu používají příklad Durkheimovy teorie sebevraždy: pravidelné výkyvy či změny v míře sebevraždnosti zřetelně odpovídají *pravidelnosti*, která je naznačena v uvedeném modelu. Zjištěné míry sebevraždnosti se liší v závislosti na daném kontextu (katolické či protestantské vyznání) a jsou dány (vysvětleny) konkrétním mechanismem – v daném případě anomii (Pawson, Tilley 1997). Podobně také evaluátoři, kteří následují realistický přístup, se snaží vysvětlit zjištěnou pravidelnost (tj. účinnost daného intervenčního programu) pomocí konkrétního mechanismu, který působí v daném kontextu. Jeho roli nelze přitom v realistickém přístupu opomenout, neboť kontext přímo podmiňuje vztah mezi kauzálními jevy. Platí tedy, že jednání nemá smysl samo o sobě, ale pouze v daném kontextu. Podobně také stejná opatření (stejně intervenční programy) mohou produkovat různé závěry v závislosti na kontextu, ve kterém jsou aplikovány.

Platí, že kontext je základní a nepominutelnou součástí tohoto přístupu, nicméně se zdá, že jeho role je pasivní – zkrátka není zde místo pro skutečně dynamickou interakci intervenčních programů s jejich vlastním prostředím. Kontext je chápán jako invariantní danost, již je daný program determinován a v jejímž rámci přeměňuje klíčový mechanismus vstupy na výstupy (Henry, Julnes, Mark 1998). Efektivita sledovaného intervenčního programu je tak kontextem podmíněna; při změně vnějších podmínek však závěry provedené evaluace mohou přestat platit, měly by být ve změněném kontextu aktualizovány a evaluace by měla být provedena znovu. V tomto ohledu je dobře patrný rozdíl mezi realistickým přístupem a přístupem responsivním,

Je tedy zřejmé, že otázky „Co funguje?“ či „Funguje daný intervenční program?“ nejsou v rámci realistického přístupu vůbec relevantní; adekvátní dotazy by měly být formulovány např.: „Co funguje, pro koho a za jakých okolností?“ či „Jaké podmínky musí nastat, aby intervenční program mohl dosáhnout daných výsledků?“.

Ilustrační příklad:

Realistický evaluátor si nejprve stanoví kontext, v němž daný intervenční program působí (tj. zmapuje celkovou situaci na trhu práce v daném okrese, zjistí počty uchazečů o zaměstnání, počty volných pracovních míst, kvalifikační předpoklady uchazečů a nároky zaměstnavatelů; dále si vytvoří přehled o ostatních formách vzdělávání starších pracovníků – zohlední přitom činnost personálních agentur a dalších subjektů atd.). Dále zjistí, jaká je typická kariérní dráha absolventů daného kursu (za jak dlouho nacházejí pracovní místo, na jakou pracovní pozici

jsou přijímáni, na jaké další kursy se hlásí apod.). V neposlední řadě se realistický evaluátor pokusí definovat mechanismus, jakým posuzovaný kurs dosahuje v rámci daného kontextu svého cíle. Realistický evaluátor tak konstatuje, že vyšší znalosti práce s PC pomáhají administrativním pracovníkům zvládat efektivněji své pracovní úkoly, a proto jsou zaměstnavateli upřednostňováni při obsazování administrativních pozic. Kurs proto pomůže získat např. středoškolákům dodatečnou pracovní výhodu oproti ostatním uchazečům. Vzhledem k tomu evaluátor může navrhnout, aby posuzované kursy byly zaváděny v okresech, s vysokou poptávkou po pracovních pozicích v administrativě.

4.1.8 Souhrn

Každý z přístupů tedy působí jako jakýsi reflektor, který osvětluje jednu část daného intervenčního programu. V případě skutečně provedeného evaluačního šetření by se výsledky patrně nelišily tak výrazně, přesto by však byly jasně patrné.

Ze zkušeností a z dostupných podkladů se ukazuje, že žádný z existujících přístupů neposkytuje recept na všechny situace, ve kterých jsou evaluace prováděny. Existence většího počtu evaluačních přístupů je tak odrazem mnohosti hodnocených situací; pro každou situaci je vhodný jiný evaluační přístup. Současně to znamená, že v dané situaci lze využít několika různých přístupů a zohlednit tak jednotlivé parametry konkrétní aplikační situace. Tato variantnost v přístupech k realizaci evaluací na jedné straně představuje jistá rizika, neboť volba konkrétního přístupu může být v dané aplikační situaci provedena laickým subjektem (např. zadavatelem) bez adekvátní rozvahy nad implikacemi takovéto volby. Na druhé straně platí, že zmiňovaná mnohočetnost přístupů, jakkoliv činí oblast evaluací nepřehlednou, je smysluplná a v konečném důsledku užitečná, neboť přibližuje evaluační metodologii konkrétním aplikačním situacím.

Tabulka č. 6 Porovnání některých charakteristik jednotlivých teoretických přístupů k realizaci evaluačních šetření

	Cílený přístup (R. Tyler)	Explorační přístup (M. Scriven)	Responsivní přístup (R. Stake)	Utilizační přístup (M. Patton)	Konstruktivistický přístup (E. Guba, Y. Lincolnová)	Přístup vedený teorií (H. Chen)	Realistický přístup (R. Pawson, N. Tilley)
Definice	Kongruence mezi chováním a cíli	Srovnání výsledku s potřebami (<i>needs</i>)	Popis a posouzení hodnoty (<i>merit</i>)	Volba takového designu, který vede k využitelným poznatkům	Evaluační realita není objektivně daná, ale je mentální konstrukcí	Intervenční programy působí mechanismy, které jsou zachytitelné	Co funguje lépe, kde, pro koho, proč
Na co se zaměřuje	Cíle programu (jejich dosažení)	Posouzení efektů	Záměry; zamýšlené cíle	Splnění požadavků zadavatele	Dosažení konsenzu	Vytvoření představy o působení intervenčního programu	Popis mechanismu dosahování cíle v daném kontextu
Klíčová informace	Cíle a výstupy indikátory	Zamýšlené a nezamýšlené důsledky	Pozorování, záměry, smysl, standardy	Cokoliv využitelného zadavatelem	Konstrukce reality	Teorie působení intervence	Mechanismus, pravidelnost, kontext
Teoretický základ evaluace	Positivismus	Nespecifický	Senzitivní zaměření	Nespecifický	Fenomenologie	Vědecká metoda	Realismus
Metodologický akcent	Vnitřní validita	Nespecifický	Nespecifický	Nespecifický	Hledisko validity je odmítáno	Vnější validita	Vnější validita
Role evaluátora	Nezúčastněná, objektivní	Nezávislá autorita	Interaktivní účastník	Nezúčastněná; negociační	Je jednou ze zainteresovaných stran	Nezúčastněná	Nezúčastněná
Role kontextu	Bezvýznamná	Nespecifická	Subsidiární	Subsidiární	Daná vazbou s konstrukcemi jednotlivých zainteresovaných stran	Zásadní	Zásadní
Výhody	<ol style="list-style-type: none"> 1. Vysoká systematickosti 2. Stojí na tradičních vědeckých základech 3. Legitimizuje i jiné než konvenční metody sběru dat za předpokladu experimentálního designu 	<ol style="list-style-type: none"> 1. Intervenční program lze posuzovat i bez znalosti cílů (není vychýlený) 2. Upozorňuje na roli vedlejších a neočekávaných výsledků 3. Zůstává flexibilní při stanovení standardů 	<ol style="list-style-type: none"> 1. Zohledňuje rozdíly v hodnotách 2. Obrací pozornost k potřebám zainteresovaných stran 3. Za základ evaluací staví cíle jednotlivých zainteresovaných stran 	<ol style="list-style-type: none"> 1. Je otevřený jakékoliv metodě 2. Zdůrazňuje roli „využitelnosti“ 3. Lze ho snadno kombinovat s jinými přístupy 	<ol style="list-style-type: none"> 1. Důraz na zohlednění zájmů všech zainteresovaných stran 	<ol style="list-style-type: none"> 1. Pomáhá pochopit vnitřní mechanismus působení intervenčního programu 	<ol style="list-style-type: none"> 1. Umožňuje přizpůsobit intervenční program povaze zainteresovaných stran 2. Směřuje k praktické využitelnosti prováděných evaluací 3. Usiluje o zobecnění získaných poznatků
Nevýhody	<ol style="list-style-type: none"> 1. Nezahrnuje způsob definice standardů (cílů) 2. Příliš akcentuje pre-post design 3. Vede ke konvergentním efektům (předčasné závěry, odhlíží od dopadů) 	<ol style="list-style-type: none"> 1. Neposkytuje návod na provedení v praktických podmínkách 2. Předpokládá vysoce kompetentního evaluátora (příliš spoléhá na jeho osobnost) 3. Definice potřeb může být zkreslena samotným výzkumným procesem 	<ol style="list-style-type: none"> 1. Obtížná operacionalizace, chybějící metodika identifikace kritérií 2. Příliš mnoho parametrů, provádění evaluace není předem známo (smluvní obtíže) 3. Příliš neformální na to, aby přinášel obhajitelné výsledky 	<ol style="list-style-type: none"> 1. Obtížně obhajitelné při provedení sumativních evaluací 2. Příliš závislé na silné pozici evaluátora 3. Postup sám nestanovuje žádná faktická pravidla či doporučení 	<ol style="list-style-type: none"> 1. Obtížně proveditelné v reálných podmínkách 2. Zájmy různých zainteresovaných stran nemohou vždy konvergovat (dosahování shody nemusí být vždy žádoucí) 3. Vysoké riziko kooptace evaluátora některou z ostatních stran 	<ol style="list-style-type: none"> 1. Tím, že se zaměřuje na nepozorovatelné procesy, upíná se k opakování záměrů iniciátora intervenčního programu (popis mechanismu fungování) 	<ol style="list-style-type: none"> 1. Pracuje s nedokazatelnými předpoklady 2. Kontext chápe jako invariantní danost 3. Pro posuzování skutečných intervenčních programů je poněkud mechanistický

Zdroj: autor

4.2. Praktické přístupy k realizaci evaluačních šetření

Tato podkapitola obrací pozornost od teoretických přístupů, které jsou vymezeny na základě charakteristik spíše obecného a principiálního rázu, k typologiím evaluačních šetření, jež jsou naopak koncipovány s ohledem k praktickým okolnostem jejich provádění. Tyto přístupy jsou sledovány v celkem třech osách, které sledují účel provádění evaluačních šetření, předmět evaluace a charakter vztahů mezi klíčovými stranami zainteresovanými na provádění evaluačních šetření.

4.2.1 Dle účelu: formativní a sumativní

Rozdělení evaluačních šetření na formativní (*formative evaluation*) a sumativní (*summative evaluation*) je dílem M. Scrivena, který tuto dichotomii rozvinul na základě dlouhé a intenzivní disputace s L. Cronbachem. Jejich rozprava byla podnícena Tylerovými úvahami o obecném smyslu, účelu a poslání evaluačních šetření jako takových.

Formativní typ evaluačních šetření je spojován se snahou o zlepšování a rozvíjení daných intervenčních programů, politik a projektů³⁶. Výsledky takto pojatých evaluačních šetření tedy představují průběžnou zpětnou vazbu, která může být využívána pro korigování běžících programů nebo pro návrh nových opatření v dané oblasti. Výsledky formativních evaluačních šetření poskytují informace směřující k vylepšení stávajících intervenčních programů, přičemž otázka jejich samotné budoucí existence je automaticky předpokládána a v kontextu formativních evaluačních šetření není nijak zpochybňována (DePoy, Gilson 2003). Je tedy patrné, že formativní evaluační šetření provázejí běžnou činnost, rutinní působení daného intervenčního programu či jeho každodenní provoz a s ohledem ke své podstatě jsou obvykle prováděna opakovaně či, jak je zřejmé z řady konkrétních aplikací, dokonce průběžně.

Výsledky formativních evaluačních šetření slouží primárně takovým zainteresovaným stranám, jakými jsou manažeři intervenčních programů a jejich řadoví pracovníci, přičemž jinými zainteresovanými stranami (například sponzory či klienty) jsou výsledky formativních evaluačních šetření využívány naopak jen velmi zřídka (Fitz-Gibbon, Morris 1987). Formativní evaluační šetření tedy mají obvykle (nemusí to však platit vždy) interní povahu,

³⁶ M. Scriven spojoval formativní evaluační šetření primárně se způsobem implementace intervenčních programů.

což znamená, že iniciátorem, realizátorem i uživatelem výsledků takových šetření je stejný institucionální subjekt. V případě tohoto typu evaluačních šetření je pro evaluátory nutné zajistit u všech zúčastněných zainteresovaných stran akceptaci zásady, že žádný výsledek evaluace nebude posuzován ani jako dobrý, ani jako špatný. V této souvislosti je třeba připomenout primární funkci formativních evaluačních šetření, kterou je zprostředkování účinné zpětné vazby, jež může být využitelná při rozvoji daného intervenčního programu. Za takto stanovených okolností tedy nelze považovat případný výskyt značného množství námětů na zlepšení posuzovaného intervenčního programu za ukazatel jeho aktuálně nízké kvality či znak jeho nedokonalosti; formativní evaluační šetření rovněž nijak neposuzují užitečnost či smysluplnost jeho dalšího pokračování. Tyto principiální otázky jsou totiž vyhrazeny sumativním evaluačním šetřením.

Formativní evaluační šetření nelze ztotožňovat ani s brainstormingem, ani s exploračními nápady, jak daný intervenční program vylepšit; rovněž platí, že odkaz k potřebám jednotlivých cílových skupin, které zůstávají i během implementace daného intervenčního programu stále neuspokojeny, není dostatečným naplněním funkcí formativních evaluačních šetření. Jejich výsadní role spočívá v zohlednění primárních cílů posuzovaných intervenčních programů a v získávání takových údajů, které jsou přímo využitelné při jejich strategickém řízení.

V neposlední řadě je třeba poukázat na významné rozdíly mezi formativními evaluačními šetřeními a monitoringem (se kterým se v některých pramenech nesprávně zaměňují); klíčové rozdíly je třeba spatřovat ve vztahu obou činností k takovým tématům, jakými jsou (1) ověření správnosti definice cílové skupiny, (2) účinnost nástrojů stimulujících chování jednotlivců a (3) sledování výskytu nežádoucích efektů (Hong, Boden 2003). Zatímco monitoring je buď zcela pomíjí, nebo nedisponuje dostatečnými nástroji k jejich zachycení a analýze, formativní evaluační šetření jsou k jejich posouzení speciálně vytvořena.

Davidsonová (2005)³⁷ navazuje na základní typologii evaluačních šetření M. Scrivena a jednotlivé typy dále rozvíjí. Přichází s návrhem dvou typů formativních evaluačních šetření založených na rozdílných stavech jejich předmětu. Rozlišuje tak formativní evaluační šetření prováděná v rámci přípravy nových programů a postupy vytvořené a používané pro zlepšování již etablovaných a dlouhodobě existujících intervenčních programů. Role formativních evaluačních šetření je totiž v případě nových intervenčních programů zcela

³⁷ J. Davidsonová studovala problematiku evaluací pod vedením M. Scrivena.

specifická a zásadní; přináší všem zapojeným zainteresovaným stranám významné výhody a značnou přidanou hodnotu, neboť včasné odhalení případných nedostatků umožňuje jejich včasnou nápravu a minimalizaci negativních dopadů. Dále platí, že včasné provedená podrobná analýza nedostatků nově implementovaných intervenčních programů umožňuje hlubší a přesnější pochopení mechanismu jejich fungování v reálných podmínkách. Smysl a účel realizace formativních evaluačních šetření u nově implementovaných intervenčních programů lze tedy spatřovat zejména ve zdokonalování jejich věcného základu (*merit*). Naproti tomu význam formativních evaluačních šetření prováděných u již zavedených intervenčních programů spočívá v tom, že umožňují zodpovědným manažerům těchto intervenčních programů včasnou a relativně snadnou adaptaci na vnější (a v mnoha případech permanentně se proměňující) prostředí. Předpokládá se, že již zavedené a běžící intervenční programy jsou zbaveny postimplementačních obtíží, a tak se pozornost při jejich evaluaci může obrátit k interakci daných intervenčních programů s jejich bezprostředním okolím a zaměřit se na získávání takových podkladů, které umožňují efektivní změnu jednotlivých charakteristik v závislosti na změnách vnějších podmínek. Primárním cílem aplikace takto pojatých formativních evaluačních šetření tedy není vylepšování hodnocených intervenčních programů ve smyslu rozšiřování jejich věcného základu (*merit*), ale spíše posilování jejich přínosu (*worth*). Formativní evaluační šetření mají prospektivní charakter, neboť přinášejí podněty pro budoucí rozvoj posuzovaného intervenčního programu.

Za zmínku stojí také některé další specifické typy formativních evaluačních šetření, z nichž bychom rádi poukázali zejména na tzv. preformativní evaluační šetření (*pre-formative*) a na předběžná sumativní evaluační šetření (*early-warning summative*). V rámci preformativních evaluačních šetření je posuzován nejen samotný průběh intervenčních programů, ale zásadní pozornost je věnována zejména jejich iniciačním stimulům, tedy např. zadávací dokumentaci či objednávkám apod. V celé řadě situací jsou právě okolnosti vzniku dané intervence a původní motivace iniciátorů k zavedení daného intervenčního programu významným činitelem, který determinuje jeho základní parametry a má tak podstatný vliv na jeho samotné fungování i efektivitu. Rovněž je z aplikační praxe zřejmé, že v mnoha případech dochází k podcenění nesouladu mezi původními (ideálními) cíli, jež byly sledovány iniciátory dané intervence, a cíli, které se při přípravě intervenčního programu ukázaly jako skutečně dosažitelné (reálné).

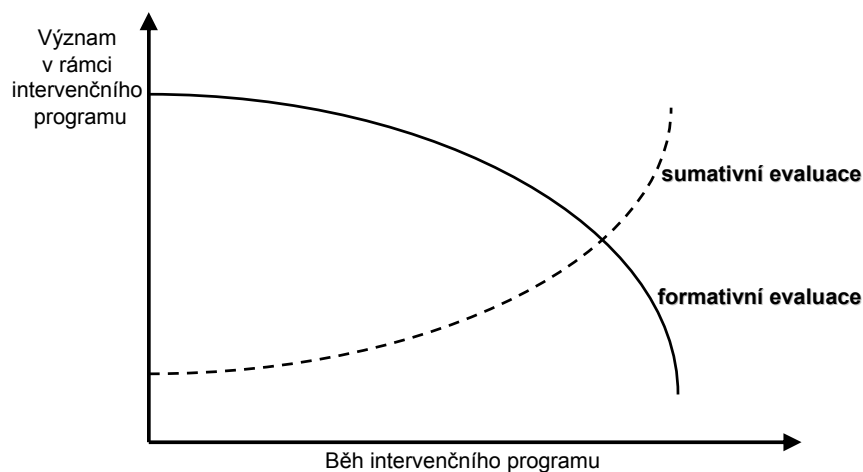
Předběžná sumativní evaluační šetření mohou být chápána jako „závěrečný test nanečisto“. Jejich základním principem je simulace sumativních podmínek, jež by měly v maximálně možné míře korespondovat s podmínkami v budoucnu prováděného sumativního evaluačního šetření. V praxi se této podobnosti dosahuje především aplikací identického uspořádání sběru vstupních dat, volbou shodných či analogických ukazatelů či použitím mechanismu zpracování a vyhodnocení údajů, jaký je plánován pro použití při sumativní evaluaci. Jinými slovy tedy platí, že předběžná sumativní evaluace je prováděna podle standardů sumativních evaluací, nicméně s jejich výsledky se důsledně zachází jako s výsledky evaluace formativní, to znamená, že z jejich výsledků není dovozována odpovědnost hodnocených subjektů a slouží výlučně jako nástroj rozvoje a zlepšení.

Sumativní evaluace jsou používány v situacích, kdy je třeba posoudit dosažené výsledky a rozhodnout o dalším pokračování, či zastavení hodnocených intervenčních programů. Cílem konkrétních evaluací v tomto případě není hledat návrhy na zlepšení, nýbrž konstatovat a osvědčit dosažení či nedosažení vytyčených cílů, popř. doporučit na základě příslušných zjištění další pokračování či zastavení posuzovaných intervenčních programů. Sumativní evaluace jsou prováděny buď po skončení daného intervenčního programu, nebo po přesně definované etapě (např. na konci zkušebního provozu, po roce působení apod.). Sumativní evaluace mají svou povahou velmi blízko k auditu, neboť se obě činnosti ohlížejí zpět do minulosti – jsou založeny na retrospektivní analýze fungování intervenčního programu ve vymezeném období. Sumativní evaluace i audit volají dotčené zainteresované skupiny k odpovědnosti, snaží se tedy vztahovat dosažené výsledky daného intervenčního programu ke konkrétním subjektům či okolnostem, které jejich dosažení usnadnily či podpořily, nebo naopak dosažení očekávaných cílů bránily či je zcela znemožnily. V neposlední řadě sumativní evaluace i audity jsou založeny na exaktních technikách sběru informací a metodách jejich analýzy. Sumativní evaluace jsou určeny (a častokrát také iniciovány) externími subjekty, které stojí mimo samotný intervenční program (obvykle jde o sponzory daného intervenčního programu či o jeho iniciátory); s ohledem ke způsobu využití jsou sumativní evaluace prováděny obvykle externími evaluátory.

Měli bychom rovněž uvést, že některé prameny poukazují na relativně okrajovou roli sumativních evaluací oproti evaluacím formativním. Například Stake (1975, 1982) uvádí, že

„není zcela obvyklé, aby evaluátor volil mezi červenou a modrou, aby činil závěry či aby poskytoval jeden jednotlivý ukazatel přínosu (*worth*)“. Podobně také Cronbach (1980) argumentuje vyšším důrazem na formativní (zlepšovací) smysl, který evaluace mají.

Obrázek č. 13 Význam formativních a sumativních evaluací během působení programu



Zdroj: Fitzpatrick et al. 2004

Základní představení obou typů evaluací diferencovaných účelem provedení nastínilo nejen klíčové rozdíly v jejich účelu, ale současně nastínilo některé zásadní rozdíly ve způsobu, jak jsou oba typy evaluací prováděny a jak odlišné požadavky na realizátory kladou. V této souvislosti bychom rádi poukázali na několik dalších charakteristik, ve kterých se formativní a sumativní evaluace odlišují. Na první pohled nemusí být patrné, že formativní evaluace jsou opakovaným, kontinuálním procesem, zatímco sumativní evaluace jsou většinou jednorázové; jejich opakování je sice principiálně možné, nicméně může být za daných situačních podmínek považováno za podezřelé. Další ze zásadních diferencí mezi formativními a sumativními evaluacemi je účel jejich provedení a záměr, který je jejich zadavateli sledován: v případě formativních evaluací je to vylepšení stávajícího intervenčního programu, zatímco v případě evaluací sumativních je to podpora při rozhodování o jeho další existenci, event. o stanovení odpovědnosti. Diference v záměrech obou typů evaluací se promítají do dalších okolností spojených s jejich prováděním v daných aplikačních podmínkách: zatímco formativní evaluace jsou zapojenými zainteresovanými stranami zodpovědnými za chod intervenčního programu většinou přijímány přátelsky a s ochotou participovat (McDavid,

Hawthorn 2006), v případě sumativních evaluací jednají dotčené zainteresované strany přesně opačně.

Dalším rozdílem není sice nutné věnovat na tomto místě detailní pozornost, nicméně uvádíme je pro úplnost alespoň v přehledové tabulce.

Tabulka č. 7 Rozdíly mezi formativní a sumativní evaluací

	Formativní	Sumativní
Cíl použití	Zlepšit daný intervenční program	Podpořit rozhodnutí o budoucnosti programu
Smysl provedení	Poskytnout zpětnou vazbu, aby mohl být intervenční program vylepšen	Poskytnout informace, na jejichž základě lze rozhodnout o pokračování, či zastavení intervenčního programu
Tematické vymezení evaluačních šetření	Vyjasnit cíle intervence, ověřit povahu implementace; identifikovat výstupy	Posoudit výstupní parametry; poukázat na problémy při implementaci
Typické otázky	Co funguje? Co je třeba zlepšit? Jak lze dosáhnout zlepšení?	Jaké jsou výsledky? Za jakých podmínek bylo výsledků dosaženo? Při jakých nákladech?
Kritéria pro volbu designu	Jaké informace jsou zapotřebí? Kdy?	Jaké důkazy jsou nutné pro jednotlivá rozhodnutí?
Intenzita zájmu hodnocených subjektů na výsledku evaluačních šetření	Nízká/středně vysoká	Velmi vysoká
Metodologie	Kvalitativní i kvantitativní	Výlučně kvantitativní
Frekvence sběru dat	Vysoká	Nízká (častokrát jen jednorázové)
Typická forma provedení	Interně	Externě
Role evaluátora	Interaktivní; je v těsném kontaktu s posuzovanými subjekty, zapojuje se do diskuzí nad výsledky evaluace	Nezávislá; evaluátor si zachovává odstup od hodnocených subjektů; výsledky tlumočí přímo zástupcům nadřazené úrovně
Příjemci výsledků evaluačních šetření	Manažeři, pracovníci, administrátoři intervenčního programu	Sponzoři, rozhodovatelé
Forma sdělení výsledků evaluačních šetření	Neformální	Formální

Zdroj: Fitzpatrick et al. 2004; Herman et al. 1987, upraveno a doplněno autorem

Souhrn

Rozlišování formativních a sumativních evaluací je specifickou dichotomií, která nemá žádný ekvivalent ani v základním výzkumu, ani v jiných oblastech výzkumu aplikovaného. Proto se nabízí otázka, zdali by podobná diferenciací nebyla relevantní i pro společenskovední výzkum. Každodenní evaluační praxe totiž poskytuje více než dostatek konkrétních podnětů dokumentujících, jak nebezpečné a eticky problematické může být, pokud jsou oba uvedené cíle zaměřovány, či spíše pokud dojde k proměně těchto cílů v průběhu realizace daného (evaluačního) projektu.

Tabulka č. 8 Vztah mezi formou provedení evaluace a způsobem využití jejích výsledků

		Způsob využití	
		Formativní	Sumativní
Forma provedení	Formativní	✓	Pole č. 1
	Sumativní	Pole č. 2	✓

Zdroj: autor

Pole č. 1 odráží situaci, kdy je evaluace provedena jako evaluace formativní, nicméně s jejími výsledky je nakládáno, jako by se jednalo o evaluaci sumativní. Nikdy nelze předem vyloučit, že taková situace během realizace evaluačního šetření a bezprostředně po jeho dokončení nenastane. Zdokumentované situace, ve kterých k podobnému „zneužití“ výsledků formativních evaluací došlo, naznačují, že se obvykle jedná o „názorový posun“ na straně zadavatele dané evaluace – získané závěry jsou použity v rozporu s jejich původním posláním. Musíme poukázat na neobhajitelnost tohoto počínání, a to jak z hlediska metodologického, tak také z hlediska etického. Inkompatibilita formativního záměru a sumativního resultátu spočívá zejména v tom, že výzkumné techniky prováděné při formativních evaluacích nemají dostatečnou vnitřní validitu pro účely, jež vyplývají ze sumativního určení. Koneckonců, rovněž externí validita spojovaná s formativními evaluacemi je z pohledu využití získaných závěrů pro sumativní účely (tj. pro vyvození odpovědnosti) insuficientní, neboť v čase, kdy jsou obvykle prováděny formativní evaluace (tj. v průběhu, či dokonce v raných fázích existence intervenčního programu), nelze očekávat, že se mohou naplno projevit výstupy a výsledky daného intervenčního programu. Etická problematičnost popisovaného nakládání s výsledky formativních evaluací vychází z porušení ustanovení Standardů (cit. dle Sanders 2004). Ovšem ke zneužití tohoto typu dochází

v naprosté většině případů ze strany zadavatelů evaluací, kteří nejsou žádným takovým kodexem vázáni. Způsob naložení s výsledky provedené evaluace je přitom plně v moci zadavatelů, kteří si častokrát nemusí diference mezi formativní a sumativní evaluací uvědomovat, popř. je zohlednit při využívání výsledků prováděných evaluačních šetření.

Pole č. 2 odráží situaci, kdy jsou výsledky sumativních evaluací používány, jako by se jednalo o výsledky evaluací formativních. V tomto případě se neadekvátně zvolená metodika projeví především nedostatečným rozsahem a podrobností informací, které mohou sumativní evaluace poskytnout. Jak jsme uvedli výše, sumativní přístupy kladou důraz na exaktnost, což se odráží jednak v relativně nízké otevřenosti při získávání informací (primární a jediný zájem je přece ověřit předem definované parametry), jednak nepoměrně vyšším formalismem, ať už při sběru dat, či ve fázi prezentace výstupů. V kontextu sumativních evaluací skutečně nelze očekávat takovou otevřenost zúčastněných subjektů, jaká panuje při formativních evaluacích. Z tohoto pohledu lze výsledky sumativních evaluací považovat za vychýlené, neboť dotazované subjekty mohou v obavě, že by jim poskytnuté informace mohly uškodit, řadu svých výpovědí zkreslit (viz rozsáhlá literatura dotýkající se problematiky vychýlení způsobených sociální desirabilitou). V neposlední řadě je třeba poukázat na časový rámeček sumativních evaluací, který je v naprosté většině aplikačních situací nekompatibilní s požadavky kladenými na formativní evaluace. Výsledky sumativních evaluací by zkrátka byly k dispozici příliš pozdě na to, aby mohly být reálně využity.

Někteří zadavatelé nerozlišují význam formativní a sumativní evaluace a požadují kombinaci obojího. Příčinou jejich postoje bývá nízké povědomí o nekompatibilitě účelů a forem a rovněž snaha ušetřit finanční prostředky (chtějí „vše v jednom“). Kombinace evaluačních typů však přináší některé problémy – závěrečné zprávy častokrát vznikají v procesu vzájemné spolupráce mezi realizátorem a zadavatelem. To je korektní a žádoucí v případě formativních šetření, nicméně pokud prováděné evaluační šetření obsahuje také část, která má jasně sumativní charakter (má např. zhodnotit míru zapojení a užitečnost konkrétních subjektů v minulém období), je takováto participace za hranicemi etické přijatelnosti.

4.2.2 Dle předmětu: evaluace výsledku, procesu a výzkumu potřeb

Klasifikace evaluačních šetření založená na předmětu evaluace vychází z postupu intervenčního programu a z jednotlivých fází jeho rozvoje. V této práci jsou jednotlivé typy uváděny jakoby zpětně, od konečných fází zaváděných intervenčních programů směrem k ranějším etapám. Důvodem tohoto převrácení je význam, který je daným typům evaluací přiřkládán. Největší důležitosti se tedy těší evaluace výstupu (či přesněji výstupu, výsledku a dopadu, což je podrobněji vysvětleno dále), následuje evaluace procesu a relativně nejmenší váha je v oboru evaluací přiřkládána výzkumu potřeb. Pokud bychom však následovali chronologickou stopu, pak by tyto tři typy evaluací byly prezentovány v opačném pořadí.

4.2.2.1 Evaluace výsledku (*outcome evaluation*)

Evaluace výsledku lze považovat za tradiční typ evaluací; vůbec první evaluace prováděné R. Tylerem se totiž zaměřovaly právě na hodnocení výsledků nově zaváděných osmiletých osnov. Tento typ evaluačních šetření se tedy zaměřuje především na popis, exploraci a stanovení změn, které nastaly ve skupině beneficentů či jiných dotčených subjektů v důsledku působení daného intervenčního programu (Fitzpatrick, Sanders, Worthen 2004). Za hlavní účel evaluace výsledku někteří další autoři, např. Unrau, Gabor, Grinnell (2007), považují analýzu změny v chování cílové skupiny poté, co její zástupci byli podrobeni působení příslušného intervenčního programu.

Evaluace dopadu (*impact evaluation*) zahrnuje jednak měření pozitivních a negativních změn, které posuzovaný intervenční program přináší dotčeným subjektům, jednak stanovují, zdali jsou tyto změny skutečně způsobeny daným intervenčním programem (Datta, Grasso 1998). Evaluace dopadu mají *de facto* stejný cíl jako evaluace výsledku, odlišují se tím, že se zaměřují na dlouhodobě udržitelné efekty vyvolané daným intervenčním programem (Kirkpatrick, Clive 2007).

Přestože evaluace výsledku přímo vybízí k provedení sumativního evaluačního šetření, jehož cílem by bylo ověření, zdali daný intervenční program dosahuje vytyčených cílů, či nikoliv, lze evaluaci výsledku provádět také formativním způsobem (Schalock 2001). Formativní evaluace výsledku se přitom nesoustředí na identifikaci subjektů zodpovědných za to, že nebylo dosaženo stanovených cílů, nehledá „viníka“ neuspokojivých závěrů, ale pokouší se identifikovat hlavní bariéry bránící dosažení daných cílů a hledat způsoby jejich překonání.

Použití formativního způsobu evaluace výsledku předpokládá takovou povahu intervenčních programů, která dovoluje jejich opakování (např. etapové projekty).

4.2.2.2 *Evaluace procesu (process evaluation)*

Tento typ evaluace se zaměřuje na zachycení mechanismů působení sledovaných intervenčních programů. V rigidní podobě se procesní evaluace soustřeďují na soulad každodenních činností vykonávaných v rámci sledovaných intervenčních programů s předem definovanými postupy a pravidly (Baumgartl, Lorenzo, Mizikaci 2005). V méně rigidní podobě se procesní evaluace zaměřují také na odhalení možných příčin vyskytnuvších se problémů, úspěchů a změn plynoucích z běžného provozu daného intervenčního programu. Předmětem zájmu evaluací procesu tedy není intervenční program samotný a už vůbec ne jeho výsledky, nýbrž způsob, jak je výsledků dosahováno. Klíčové otázky evaluace procesu jsou: „Co se děje?“, „Jak se to děje?“, „Jak se zainteresované strany zapojují do intervenčního programu?“

V některých případech se o evaluaci procesu hovoří jako o tzv. procesní komponentě, čímž se posiluje přesvědčení, že tvoří – vedle analýzy výsledků – pouhou součást komplexní evaluace intervenčních programů. Nicméně je třeba dodat, že evaluace procesu bývá velmi často realizována jako samostatné šetření.

Podle některých evaluátorů zaměřených na evaluaci výsledku (jako např. Scriven) je třeba přistupovat k evaluaci procesu velmi uvážlivě, neboť tento typ evaluací nevede ze své podstaty k symetrickým výsledkům – tato metoda totiž v mnoha konkrétních situacích zvýrazňuje negativní jevy v průběhu realizace intervenčního programu, zatímco velké většině „standardně“ vykonávaných procesů nevěnuje téměř žádnou pozornost. Je proto smysluplné zaměřit se především na funkční prvky sledovaných procesů, jako např. na vnitřní a vnější komunikaci a vztah vůči klientům i vůči dalším zainteresovaným stranám.

Y. Unrauová poukazuje v této souvislosti na dva procesy, které zasluhují specifickou pozornost; jsou jimi proces poskytování služeb klientům (tj. na jedné straně činnost pracovníků daného programu a na straně druhé profil klientů a jejich spokojenost) a proces administrativní podpory, který pomáhá udržovat poskytování služeb zákazníkům³⁸ (Unrau et al. 2007).

³⁸ Její oblastí zájmu jsou především programy ve sféře sociální práce, nicméně platnost významu obou procesů je podle našeho přesvědčení univerzální.

Evaluace procesu bývá nejčastěji prováděna ve dvou podobách, a to jako evaluace kontextu (*context evaluation*) a dále jako evaluace samotného zavádění (*implementation evaluation*). Evaluace kontextu analyzuje, jak je průběh samotného intervenčního programu ovlivňován kontextem a bezprostředním okolím, naproti tomu evaluace samotného zavádění posuzuje průběh daného intervenčního programu či jednotlivých opatření (Bradley 2007). Evaluace samotného zavádění se velmi podobá monitoringu, neboť sleduje takové okolnosti, jakými jsou dodržování harmonogramu či rozpočtová kázeň. Přesto však není s monitoringem totožná, protože tato forma evaluace procesu postrádá sumativní ráz, který je monitoringu vlastní. Evaluace samotného zavádění má naopak ve velké většině formativní charakter, jelikož přichází s konkrétními náměty na zavedení účinných opatření, jež mají umožnit efektivnější implementaci bez toho, aby poháněla jednotlivé zainteresované strany k odpovědnosti.

Evaluaci procesu lze tedy považovat za jakousi extenzi monitoringu, neboť analyzuje, sumarizuje a rozvíjí údaje poskytované monitoringem. Od evaluace procesu je třeba odlišovat i formativní evaluaci. Evaluace procesu totiž může být prováděna nejen jako formativní, ale v jejím rámci lze anticipovat také sumativní cíle. Formativní evaluace procesu může přinést identifikaci nejslabších článků v celém procesu a nalezení způsobů, jak činnosti v těchto úsecích vylepšit (např. pomocí školení pro pracovníky intervenčního programu, zvýšením počtu pracovníků intervenčního programu, jiným rozdělením pracovních úkolů mezi pracovníky intervenčního programu atd.). Na druhé straně sumativní evaluace procesu kontroluje dodržování postupů. V takto pojaté evaluaci mohou být jednotliví pracovníci voláni k zodpovědnosti za jednání odchylovající se od stanovených postupů, či naopak mohou být odměňováni za jejich dodržování.

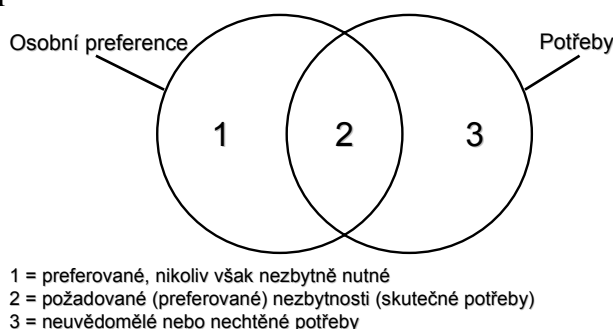
4.2.2.3 Výzkum potřeb (*needs assessment*)

Výzkum potřeb identifikuje a měří míru nenaplněných potřeb (požadavků) v rámci sledované cílové skupiny (Gaber 2000). B. Witkinová a J. Altschuld píší v tomto ohledu o rozdílech mezi tím, co je, a tím, co by mělo být (Witkin, Altschuld 1995), přičemž bezprostřední kontakt evaluátora se zainteresovanými stranami umožňuje odhadnout užitečnost konkrétních aspektů intervenčního programu (Posavac, Carey 2003). Reviere et al. (1996) definují výzkum potřeb jako způsob zjištění rozdílu mezi reálným a ideálním stavem reflektovaným pomocí konkrétních hodnot, jež jsou danými zainteresovanými stranami sdíleny. R. Reviere, ale i mnozí další, v této souvislosti zdůrazňují potenciální proměnlivost předmětu potřeby a

dále poukazují na mnohost variant, mezi nimiž lze volit při hledání způsobu uspokojení té které potřeby. V pozadí každého takového rozhodovacího procesu založeného na výzkumu potřeb je jednak existence samotné potřeby a současně také alespoň teoretická dostupnost způsobu jejího uspokojení. Výzkum potřeb je tedy posouzením stupně důležitosti a míry relevance jednotlivých problémů vnímaných konkrétními zainteresovanými stranami a hodnocením potenciálních dopadů, nákladů a proveditelnosti konkrétních činností, které mohou vzniklou potřebu uspokojit. McKillip (1987) definuje výzkum potřeb velmi přímočaře jako nástroj pro podporu rozhodování (např. o nabídce kursů celoživotního vzdělávání či o způsobu financování ústavu pro bezmocné adolescenty).

Klíčem k adekvátně provedenému výzkumu potřeb je jasné rozlišení mezi potřebami (*needs*) a osobními preferencemi (*wants*). V celé řadě případů totiž nemusí být skutečná potřeba tím, co příslušné zainteresované strany³⁹ samy chtějí, a způsob uspokojení potřeby může mít často takovou formu, kterou si tyto strany rozhodně nepřejí. Zásadní diference mezi potřebami a osobními preferencemi je třeba brát vždy v úvahu; platí totiž, že při absenci potřeby dochází k malfunkcím (např. intervenční program nemůže dosáhnout svých cílů), zatímco výsledkem chybějících osobních preferencí je „pouze“ nespokojenost některých beneficentů. Dále platí, že nízká míra saturace konkrétní potřeby nemusí být nutně vnímána samotnými beneficenty jako újma (Remr, Kotíková 2007). V tomto ohledu je jasně patrné sepětí potřeb s hodnotami – lidé s různými hodnotami mívají různé potřeby.

Obrázek č. 14 Srovnání vzájemného vztahu mezi potřebami a osobními preferencemi



Zdroj: autor

³⁹ Nejčastěji jde o beneficenty.

Nutnost pečlivého rozlišování mezi potřebami a osobními preferencemi je velmi dobře patrná v celé řadě konkrétních aplikačních situací; např. při výzkumech hodnocení výuky. Žáci obvykle dokáží velmi dobře definovat své osobní preference (*wants*) a v realizovaných výzkumech pak snadno deklarují svou míru spokojenosti s jejich naplněním. Na druhé straně však sami nemohou definovat kompletní portfolio svých potřeb spojených se vzděláváním a už vůbec nemohou identifikovat konkrétní řešení a posoudit jejich adekvátnost.

Při podrobnější analýze jednotlivých typů potřeb lze podle Altschulda a Witkinové (2005) rozlišovat tři základní dimenze potřeb:

a) Vědomé a nevědomé potřeby (*conscious needs* × *unconscious needs*)

Nevědomost je třeba vztahovat pouze k nositeli potřeby (potřebu si neuvědomuje subjekt, který je hodnocen, nicméně ostatní zainteresované strany si mohou být dané potřeby vědomy). Při identifikaci nevědomých potřeb je nutné zahrnout cíle samostatných cílových skupin (např. nejen jednotlivců dotčených daným programem, ale také expertů atd.).

b) Naplněné a nenaplněné potřeby (*met needs* × *unmet needs*)

Naprostá většina výzkumů potřeb (zejména těch, které se týkají nově zaváděných intervenčních programů) se zaměřuje na nenaplněné potřeby. Prostřednictvím řady technik a nástrojů se snaží takové potřeby identifikovat a zejména pak poukázat na jejich naléhavost. Nicméně i potřeby, které jsou aktuálně naplněné, vyžadují stejnou míru pozornosti. Jejich identifikace a popis totiž může pomoci při hledání případných variant jejich uspokojení (častokrát je v případě zavedených intervenčních programů kladena přílišná pozornost na způsob uspokojení dané potřeby, ovšem potřeba samotná ustupuje do pozadí). V mnoha případech se v důsledku změny vnějšího kontextu může stát, že původní potřeba zanikne, přesto však intervenční program koncipovaný pro její uspokojení přetrvává. Z podobných důvodů mohou být nelezeny nové a efektivnější způsoby, jak stávající potřeby uspokojovat.

c) Instrumentální a cílové potřeby (*treatment needs* × *outcomes needs*)

Rozlišení mezi instrumentálními (*treatment needs*) a cílovými potřebami (*outcomes needs*) má zásadní význam pro samotnou metodiku zkoumání potřeb. Instrumentální potřeby jsou totiž vždy spojeny se způsobem uspokojení potřeby a s konkrétními nástroji pro dosažení, naplnění či uspokojení dané potřeby. Instrumentální potřeby se tak přímo vztahují ke způsobu řešení vzniklých problémů. Naproti tomu v případě cílových potřeb jde o problém samotný;

příkladem cílové potřeby tak může být nízká míra nezaměstnanosti v daném kraji či regionu; instrumentální potřebou může v tomto případě být zavedení rekvalifikačních kursů pro nezaměstnané.

Davidsonová (2005) poukazuje na rozdílné účely výzkumů potřeb prováděných u programů nových a stávajících/zavedených. Výzkum potřeb se v případě zavedených programů využívá pro identifikaci sledovaných výsledků či dopadů a jejich relativního významu. Výzkum potřeb tak říká, na co se zaměřit, až budou posuzovány dopady programu. Provázání výzkumu potřeb a evaluace výsledku zakládá bytelnou konstrukci evaluace intervenčních programů, neboť ty jsou (nebo by měly být) koncipovány tak, aby uspokojovaly konkrétní potřebu či soubor potřeb. Do jaké míry tak skutečně činí, pak ověřuje evaluace výsledku. Jí se uzavírá logická smyčka poznání funkčnosti daného intervenčního programu.

V případě nově zaváděných programů je identifikace potřeby složitější o to, že řadu potřeb si lidé neuvědomují a nedovedou posoudit ani způsoby jejich saturace. Potřeby jsou totiž vždy závislé na daném kontextu a odvíjejí se od něj.

S konceptem „potřeby“ přišel Scriven, který jím v kontextu svého exploračního přístupu k provádění evaluačních šetření suploval apriorně definované cíle intervenčních programů. Potřeby tak ve Scrivenově přístupu poskytují evaluátorům objektivní základnu pro posouzení, zda bylo dosaženo výsledku, či nikoliv (House 1980).

Výzkum potřeb může mít jak formativní, tak také sumativní povahu (Soriano 1995). V případě formativního pojetí výzkumu potřeb je jeho smyslem nalezení způsobu, jak daný intervenční program nastavit a jak ho zavést, aby příslušné potřeby co nejlépe uspokojoval. Pokud má výzkum potřeb sumativní charakter, je většinou koncipován tak, aby odhalil existenci potřeby, případně jasně doložil, že daná potřeba neexistuje. Sumativní výzkum potřeb rovněž poskytuje údaje pro rozhodnutí, zdali má daný intervenční program vůbec začít. Např. v USA je výzkum potřeb od roku 2002 vyžadován u všech federálních intervenčních programů, neboť je součástí posouzení jejich efektivnosti. V tomto ohledu se striktně požaduje odkaz ke konkrétní potřebě (která v daný okamžik prokazatelně existuje) či k přesně vymezenému problému nebo zájmu, který plánovaný intervenční program bude naplňovat⁴⁰.

⁴⁰ Viz Office of Management and Budget, 2002, str. 6.

4.2.3 Dle vztahu zadavatele a realizátora: externí, interní a hybridní

Diferenciace na základě vztahu realizátora k předmětu evaluace rozlišující interní a externí formu, vychází z analogické dichotomie, jež je patrná ve velké části ostatních oblastí aplikovaného výzkumu. V kontextu evaluací je východiskem tohoto rozlišení fakt, že některé instituce využívají k provádění evaluací vlastní interní pracovníky.

Důvodů, proč organizace využívají vlastních interních evaluátorů namísto toho, aby realizovaly jednotlivé evaluační projekty na smluvním základě s externími realizátory evaluačních šetření, je hned několik. V první řadě platí, že provádění evaluačních šetření je v mnohých organizacích považováno za součást vlastního řízení, a tak vzniká potřeba po průběžné evaluaci jednotlivých činností (Russ-Eft, Preskill 2001). Navíc může být takovýto proces provádění evaluačních šetření jen těžko oddělitelný od ostatních procesů. Dále je patrné, že některé organizace provozují dokonce hned několik různých intervenčních programů, což opět vytváří předpoklady pro angažování interních evaluátorů⁴¹. House (1993) doplňuje, že některé veřejné organizace založily vlastní oddělení, která se evaluacemi systematicky zabývají. Ve všech uvedených případech se interní forma evaluace přímo nabízí, neboť umožňuje dané organizaci redukovat náklady na realizaci vysokého množství, relativně standardizovaných evaluačních projektů. Hlavní výhody interní formy realizací evaluací shrnuje Love (1991), který poukazuje především na přesnou a detailní znalost prostředí, cílů a záměrů – interní evaluátor je zkrátka součástí dané organizační struktury a kultury, a tak případné riziko plynoucí z nepochopení cílů zkoumaného evaluačního projektu je relativně nízké. Tato výhoda se dále promítá do volby adekvátnějších postupů a efektivnější komunikace získaných výsledků, neboť interní evaluátor je schopen lépe než externista určit, jaké informace, komu a v jakém čase poskytnout. Opomenout nelze ani dlouhodobější zájem interního evaluátora v porovnání s *ad-hoc* kontrahovaným externím subjektem.

Častým případem provádění interních evaluací jsou autoevaluační šetření, kdy je evaluace prováděna vlastními pracovníky a daná instituce provádí evaluaci sebe sama. Autoevaluace⁴² jsou z hlediska rolí jednotlivých zainteresovaných stran typem evaluačních šetření, při kterých splývá posuzovaný subjekt s rolí evaluátora.

⁴¹ V podmínkách České republiky je toto patrné mj. v případě Evropského strukturálního fondu.

⁴² S ohledem ke konkrétním podmínkám provádění autoevaluací v ČR používáme v této práci také pojem „vlastní hodnocení“.

Sommerland, Willoughby 1995) zmiňují několik základních charakteristik autoevaluací, jež jsme dále doplnili a rozšířili. V následujících přehledu se je snažíme stručně přiblížit.

a) Reflexivní praxe

Vlastní hodnocení je typ evaluační činnosti, který spoléhá na schopnost hodnocených subjektů provést reflexi sebe sama. Tato technika vede dotčené subjekty k tomu, aby věnovaly pozornost konkrétním aspektům své činnosti. Pokud je analýza realizována na institucionální (nikoliv individuální) hladině, je třeba zmíněnou sebereflexi chápat na úrovni celé organizace. V případě, že těžiště evaluačního šetření leží v individuální rovině⁴³, jsou poznatky získané jejich sebehodnocením zasazovány do kontextu ostatních osobních dat.

b) Průběžné zjišťování kvality

Subjekty provádějící vlastní hodnocení systematicky hledají poučení ve své vlastní činnosti a posuzují adekvátnost jednotlivých opatření vzhledem k cílům, kterých dosáhly. Smysl vlastního hodnocení z hlediska evaluační praxe je třeba hledat také ve stimulaci postupů a kroků, které umožní soustavně vyhledávat a vyhodnocovat zpětnou vazbu týkající se vykonávané činnosti; autoevaluace tak posilují odpovědnost za konkrétní kroky, které jsou na úrovni daných subjektů prováděny. Metoda vlastního sebehodnocení se využívá v situacích, kdy je dosažení definovaných cílů či ověření jiných cílů dané evaluace snadno identifikovatelné (a potenciálně lehce přezkoumatelné) a kdy není třeba používat komplexní mapu sofistikovaných indikátorů pro posouzení věcného základu a přínosu. Autoevaluace vycházejí z předpokladu, že zainteresované strany mají zájem na vlastním zlepšování. V neposlední řadě autoevaluace stimulují adekvátní řídicí procesy, neboť jejich provádění si vynucuje stanovení cílů natolik konkrétních, aby jejich dosahování mohlo být v procesu autoevaluace ověřováno.

c) Aplikační omezení

Zásadním nedostatkem autoevaluací je jejich subjektivní charakter, kvůli kterému jsou diskvalifikovány z některých aplikačních situací; autoevaluace nelze např. využít, pokud jsou sledovány sumativní cíle. Autoevaluace mají jednoznačně formativní charakter a jejich využitelnost je tak přímo závislá na důvěře mezi autoevaluátorem/posuzovaným intervenčním programem na jedné straně a zadavatelem takového typu evaluace na straně druhé. Jak již

⁴³ Sebehodnocení se může používat např. při evaluaci pracovníků (*personnel evaluation*).

bylo uvedeno v části věnované formativním a sumativním evaluacím, je eticky problematické využívat výsledky (formativně pojatých) autoevaluačních zpráv ke stanovení odpovědnosti (tj. k sumativním cílům).

Naproti tomu výhodou externích evaluátorů je potenciálně vyšší profesionalita a případná specializace na některé konkrétní typy evaluačních šetření. Pokud je tedy požadován specifický či méně standardní design evaluace, je pravděpodobnější, že externí evaluátor bude v tomto ohledu dosahovat vyšší excelence než evaluátor interní (Preskill, Torres 1999). Pochopitelným důvodem spolupráce s externími evaluátory je realizace tzv. sumativních evaluací; v tomto případě je účast interního evaluátora s ohledem na možný konflikt zájmů přinejmenším nevhodná a nežádoucí. Dále platí, že v rámci spolupráce s externími realizátory evaluačního šetření může zadavatel získat přístup ke specifickým evaluačním metodikám, což se v dlouhodobějším časovém horizontu může pozitivně odrazit v jejich diseminaci. V neposlední řadě je vhodné poukázat na přípravu zadávací dokumentace, v níž zadavatel popisuje kontext situace, ve které plánuje realizovat dané evaluační šetření, specifikuje své požadavky a uvádí způsob využití výsledků provedeného evaluačního šetření. Konkrétní zkušenosti ukazují, že činnosti spojené s přípravou zadávací dokumentace mají pozitivní vliv na následnou utilizaci výsledků evaluačního šetření.

V podmínkách ČR existuje několik různých oblastí, ve kterých jsou realizována evaluační šetření externí formou. Kromě řady jiných se jedná například o oblast poskytování rozvojové pomoci, jejíž účinnost je ověřována pomocí nezávisle prováděných evaluačních šetření⁴⁴, dále jde o sociálně-intervenční programy poskytované a organizované neziskovým sektorem, kdy jednotlivé subjekty (neziskové organizace) poptávají od externích poskytovatelů realizaci evaluačních šetření, ať už z důvodu korekce vlastní činnosti, podpory vlastního rozhodovacího procesu nebo s cílem nezávislé deklarace efektivity vykonávaných činností zahrnované do výročních zpráv. V neposlední řadě jsou pomocí externě prováděných evaluačních šetření ověřovány osvětové kampaně, jejichž cílem je zvýšení obecného povědomí o některých konkrétních sociálních problémech (kromě mnoha jiných např. domácí násilí, bezpečnost dětí na Internetu) či otázkách veřejného zdraví (rakovina kůže, preventivní péče o chrup a mnoho dalších) nebo podpora konkrétního chování obyvatel (nákup výrobků produkovaných v ČR, podpora ekologicky šetrných výrobků, třídění domovního odpadu atd.).

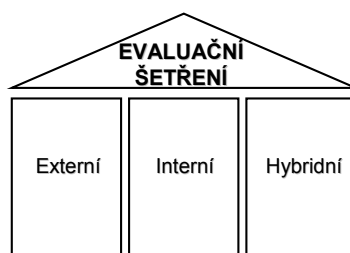
⁴⁴ Realizace takovýchto evaluačních šetření je založena na Pařížské deklaraci o efektivitě poskytované pomoci z roku 2005.

V praxi se v řadě případů nevyskytují takto krystalicky čisté, ideální typy interních či externích forem provádění evaluačních šetření. Standardně uváděnou dichotomii navrhuje tedy doplnit ještě o třetí typ, jakým mohou být evaluační šetření v aplikační praxi prováděny, a to o hybridní formu, která kombinuje prvky interně a externě prováděných evaluačních šetření. Přestože může existovat řada konkrétních variant a podob vzájemného provázání prvků interně a externě prováděných evaluačních šetření, v konkrétní aplikační praxi patří mezi nejvíce obvyklé takové uspořádání, při kterém externí subjekty⁴⁵ provádějí evaluační šetření v rámci interního evaluačního systému. V takovémto případě postupuje externí evaluátor podle metodiky připravené zadavatelem, což odlišuje hybridní formu od typicky externího způsobu provedení evaluačního šetření, kdyby to byl právě externí evaluátor, kdo by navrhl jeho metodologii. Hybridní evaluace tedy vychází z předpokladu, že jejich zadavatel disponuje alespoň částečnou interní kapacitou, jež se kromě přípravy metodiky podílí také na školení a koordinaci činnosti externích evaluátorů. Od interních evaluací se hybridní způsob odlišuje tím, že samotné provedení evaluačního šetření není uskutečněno interním týmem evaluátorů. Aniž bychom tyto okolnosti podrobovali hlubšímu zkoumání, platí, že mezi nejčastějšími důvody se vyskytují kapacitní potřeby (vysoký objem evaluačních požadavků v krátkém čase), etické požadavky (potřeba externího evaluátora, který není součástí sítě sociálních vztahů v rámci dané organizace) a kvalifikační předpoklady (vysoké nároky na odbornost hodnotitelů, jimiž daná organizace nedisponuje).

Využití hybridního způsobu provádění evaluačních šetření je široce uplatňováno zejména ve vědě, a to v podobě peer review či recenzního posuzování odborných článků, kdy jsou výsledky takto provedených evaluací zdrojem rad a námětů poskytované jednotlivým předkladatelům v naději na zdokonalení jejich statí (Hackett, Chubin 2003). Hybridní forma evaluací se používá také při evaluaci nabídek v rámci výběrových řízení nebo při posuzování grantových žádostí, kdy tato metoda napomáhá při rozhodovacích procesech o alokaci finančních zdrojů. S hybridní formou evaluačních šetření je možné se dále setkávat v rámci posuzování projektových žádostí o dotace poskytované v rámci operačních programů EU. Nicméně tento nástroj není úplným výčtem všech aplikačních oblastí, které hybridní formu evaluačních šetření využívají.

⁴⁵ Externími subjekty mohou v tomto kontextu být zejména jednotlivci – hodnotitelé; v praxi se lze setkat i s jinými pojmy, jež lze považovat za synonyma: posuzovatelé, experti.

Obrázek č. 15 **Formy provádění evaluačních šetření v aplikační praxi**



Zdroj: autor

Využívání hybridního způsobu provádění evaluačních šetření formou expertního posuzování se vyznačuje několika obecnými charakteristikami, a to především transparentností (Ross 1980), anonymitou (Kostoff 1997), efektivitou (Beemt, LePair 1991), flexibilitou (Ciechetti 1991) a reliabilitou (Howe 1982). V následující části je těmto charakteristikám věnována podrobnější pozornost.

Transparentnost

Hybridní způsob provádění evaluačních šetření formou expertního posuzování může být považován za transparentní proces, neboť v obvyklém případě jsou jednotlivé postupy definovány a určující pravidla jsou jasně stanovena. Zadavatel takovéto evaluace v zájmu zajištění konsistentního provádění jednotlivých evaluačních šetření disponuje uceleným a podrobným souborem norem a postupů, jimiž jsou externí hodnotitelé vázáni. V konkrétní aplikační praxi mohou mít tyto normy a postupy podobu příruček, manuálů či mohou být předávány externím hodnotitelům v podobě školení. Podobně přesně a podrobně jsou obvykle zadavatelem definována také jednotlivá kritéria a standardy, jejichž podstatným znakem je měřitelnost (resp. možnost ověření míry jejich naplnění) a universalita (resp. akceptace ze strany dané komunity a jejich legitimita). Platí přitom, že všechny zainteresované strany mohou k těmto podkladům získat přístup a mohou se seznamovat s jejich obsahem.

Anonymita

Obvyklý způsob provádění expertního posuzování vytěsňuje za hranici těchto transparentních procesů identitu konkrétních hodnotitelů. Tím jednak stěží jejich anonymitu, zbavuje zainteresované strany váhy interpersonálních vztahů, přirozené autority konkrétních subjektů a eliminuje tak významný prvek, který by za jiných okolností mohl významně vychýlit

způsob provedené evaluace. Současně tím nadřazuje metodiku evaluačních šetření nad individuální evaluační akty vykonávané jednotlivými hodnotiteli. Anonymizace posuzovaných objektů i samotných externích hodnotitelů je nástrojem k zajištění nezbytné neutrality a to jak jednotlivých aktérů, tak také celého procesu expertního posuzování i jeho výsledků. V této souvislosti je třeba upozornit na skutečnost, že v celé řadě aplikačních situací může být takovéto opatření neúčinné; navzdory tomu, že identita posuzovaného objektu není odhalena přímo, lze ji v některých případech úspěšně aproximovat na základě dostupných indicií.

Efektivita

V mnoha konkrétních aplikacích jde o posuzování již vytvořených děl či jiných artefaktů; ze strany subjektů posuzování tedy nepředstavuje tento proces evaluace žádnou dodatečnou zátěž. Pochopitelnou výjimku v tomto ohledu představuje jen evaluace projektů, které musejí být připraveny právě kvůli možnosti posouzení předkládaného záměru. Na druhé straně i úsilí vynakládané jednotlivými experty (studium podkladů, příprava posudku, cestovné) musí být obhajitelné v porovnání s významem této činnosti.

Flexibilita

Schopnost procesu expertního posuzování přizpůsobovat se konkrétním podmínkám či okolnostem dané aplikační situace činí z této formy evaluací universální techniku.

V konkrétních aplikacích se lze setkat s rozličnými typy objektů posuzování (tj. např. projekty, díla, výkony...), evaluační šetření mohou být prováděna s různým počtem etap (jednokolové či vícekolové), existují formy šetření, která jsou jednokriteriální či multikriteriální. Lze podle počtu expertů a podle způsobu jejich činnosti rozlišovat individuální posuzování prováděná izolovaně jedním či několika hodnotiteli nebo týmová posuzování, kdy skupina expertů provádí svou posuzovací činnost v kolegiu. Z hlediska závěrů, ke kterým celý proces expertního posuzování může dospět stojí za zmínku selekční, hierarchický či eliminační přístup. Všechny uvedené způsoby konkrétního designu expertního posuzování naznačují vysokou variabilitu této formy provádění evaluačních šetření a snadnou přizpůsobitelnost konkrétním aplikačním podmínkám.

Reliabilita

Spolehlivost expertního posuzování je námětem celé řady samostatných analytických úloh. Nejsou výjimečné případy, kdy je shoda mezi hodnotiteli nejen porovnávána, ale je také testována pomocí příslušných statistických testů (kromě Cohenova kappa, také např. Scottovo

pí či Krippendorffova alfa). Ačkoliv dosažení vysoké reliability není při expertním posuzování primárním cílem, neboť při jeho aplikaci jde zejména o získání obhajitelných stanovisek (Neuendorf 2002), přesto je reliability v rámci možností posilována vyšším počtem jednotlivých posouzení (tj. zapojením většího počtu hodnotitelů). Nízká reliability může být dokladem např. různě vysokých nároků jednotlivých hodnotitelů na úroveň dosažení klíčových charakteristik posuzovaného objektu nebo může vycházet z odlišných paradigmatických rámců použitých jednotlivými hodnotiteli. Rovněž je třeba mít na paměti, že míra znalostí a zkušeností s předmětem posouzení a se samotným postupem hodnocení může být mezi jednotlivými hodnotiteli velmi rozdílná, což může mít za následek nižší shodu mezi jednotlivými posuzovacími akty.

Při realizaci hybridní formy evaluačních šetření prováděné formou expertního posuzování je třeba zvážit také jednotlivé slabé stránky. Předně je nutné vzít v potaz, že se tento způsob provádění evaluačních šetření zaměřuje výlučně na dokumentované výsledky a umožňuje jen ve velmi omezené míře odhalit budoucí potenciál. Dále je třeba zohlednit odlišnosti mezi jednotlivými hodnotiteli vycházející z jejich vztahu k předmětu posouzení, ale také z jejich předchozích zkušeností s expertním posuzováním. Celý evaluační proces může dále selhávat na přílišném zaměření na nepodstatné aspekty daného objektu a na nedostatečné pozornosti, kterou externí hodnotitelé věnují věcnému základu. Poměrně častým nedostatkem je také nedostatečná argumentace hodnotících sentencí a nedostatečné zapojení odbornosti externího hodnotitele při samotném posuzování.

4.2.4 Souhrn

Každá z uvedených taxonomií má zásadní význam pro metodologii evaluačních šetření a každá z nich je námětem řady publikací a mnoha specificky zaměřených výzkumů, které jsou motivovány snahou o hlubší poznání klíčových charakteristik jednotlivých forem provedení evaluačních šetření. Z hlediska cílů naší práce mají zásadní význam typy evaluačních šetření diferencované dle vztahu realizátora k zadavateli. Na základě rozdělení evaluačních šetření na externí, interní a hybridní je v další, empirické, části věnována pozornost výsledkům výzkumných šetření, která mapovala konkrétní způsoby provádění jednotlivých typů evaluačních šetření v podmínkách České republiky.

5. Popis metodiky provedených výzkumů

Účelem této kapitoly je prezentovat obecný výzkumný rámec a popsat v detailu jednotlivé aspekty realizovaného výzkumného šetření, a to jak části desk research, tak také terénního výzkumu. Pozornost je věnována významným údajům o způsobu výběru a metodě sběru dat; nastíněn je rovněž způsob zpracování údajů a v neposlední řadě také hlavní limity, které omezují možnosti použití získaných závěrů.

5.1 Obecný výzkumný rámec

Vzhledem k zaměření práce, která se obrací ke konkrétním praktikám provádění evaluací v daných oblastech, považujeme za nezbytné dokumentovat vybrané jevy empirickými poznatky, jež doplňují základní teoretické vymezení jednotlivých přístupů.

Výzkum zaměřený na zjištění aktuální evaluační praxe v České republice sestává ze dvou hlavních výzkumných úloh. První z nich představuje desk research, který se opírá o obsahovou analýzu příslušných dokumentů a druhou výzkumnou úlohou je field research, v rámci kterého byla provedena série individuálních hloubkových rozhovorů se zadavateli a s realizátory evaluačních šetření.

A. Desk Research

V rámci daného výzkumu byla provedena obsahová analýza závěrečných zpráv z evaluačních šetření provedených v souvislosti s řízením operačních programů EU. Jednotlivé závěrečné zprávy jsou použity jako zdroj dat, který vypovídá o externím způsobu provádění evaluačních šetření. Dále byly analyzovány závěrečné zprávy shrnující výsledky z vlastního hodnocení škol (autoevaluace); s pomocí těchto zpráv jsou formulovány některé poznatky týkající se provádění evaluačních šetření interní formou. Samostatně byly rovněž analyzovány posudky oponentů a konzultantů bakalářských prací, neboť dokumentují stávající praxi hybridního způsobu provádění evaluačních šetření.

Záměrně byly vybrány oblasti, které nejsou vzájemně nijak provázány; každý ze tří analyzovaných případů se tak rekrutuje z odlišné oblasti.

Při návrhu jednotlivých indikátorů jsme měli na zřeteli potenciální porovnatelnost výsledků dosažených v rámci každého ze tří případů. Ačkoliv to není primárním cílem provedeného desk research, přesto je možné v některých ohledech vzájemná porovnání provádět.

B. Field Research

Terénní výzkum byl proveden v oblasti, která se, podobně jako první z případových studií, váže k oblasti řízení operačních programů EU. Oblast externě prováděných evaluačních šetření byla zvolena proto, aby bylo možné konfrontovat názory, postoje a motivy zadavatelů a realizátorů. Volba oblasti, která tématicky navazuje na některou z případových studií rovněž není nahodilá a je výsledkem snahy o vzájemné provázání skutečností zjištěných oběma typy výzkumů.

V rámci terénního výzkumu byli dotazováni jak zadavatelé evaluačních šetření z řad řídicích orgánů jednotlivých operačních programů, tak také realizátoři, kteří provedli v dané oblasti nejméně jedno evaluační šetření.

5.2 Stanovení cílových skupin

Desk Research

Nalezení relevantního případu, který by dokázal adekvátně reprezentovat stávající úroveň evaluační praxe při provádění jednotlivých typů evaluačních šetření v České republice, je úkol poměrně nesnadný, a to hned z několika důvodů. Především by zvolené případy měly být natolik běžné a frekventované, aby disponovaly obvyklým potenciálem k zobecnění.

Případové studie si obvykle takový cíl nekladou, nicméně, vzhledem k plánovanému použití takové studie (ilustrace nynější praxe při provádění evaluačních šetření v České republice) je alespoň teoreticky dosažitelná zobecnitelnost významnou součástí tohoto záměru.

Naší snahou bylo rovněž identifikovat oblasti, ve kterých jsou jednotlivé formy provádění evaluačních šetření trvalou či pravidelně opakovanou činností, a kdy je důvodné předpokládat existenci propracované metodiky a jasně dokumentovaného systému pro provádění takovéto formy evaluace.

Field Research

Pro účely terénního výzkumu byly stanoveny dvě cílové skupiny. První cílovou skupinu představují subjekty, které realizují evaluační šetření – tato cílová skupina je v některých částech textu označována jako strana nabídky. Kritérium provádění evaluačních šetření je přitom považováno za splněné, pokud daný subjekt veřejně deklaruje realizaci evaluačních šetření a/nebo pokud daný subjekt participoval na provedení nejméně jednoho evaluačního šetření. Hlavní předmět činnosti subjektu není přitom zohledňován. Definice této cílové skupiny však eliminuje fyzické osoby z důvodu posílení homogenity souboru realizátorů; explorační závěrečných zpráv navíc ukazuje, že případy, kdy je evaluace realizována jednotlivcem, jsou velmi výjimečné (v našem souboru závěrečných zpráv se vůbec nevyskytuje).⁴⁶ Do základního souboru jsou zařazovány subjekty bez ohledu na právní formu své činnosti; skupina realizátorů evaluací může proto zahrnovat jak podnikatelské subjekty, tak také subjekty nepodnikatelské. V rámci těchto subjektů je konkrétním respondentem vedoucí evaluačnického týmu a/nebo pracovník zodpovědný za provádění evaluačních šetření.

Druhou cílovou skupinu představuje strana poptávky, resp. zadavatelé evaluací. Vyloučení jsou interní zadavatelé, resp. zadavatelé interně prováděných evaluací. Do zkoumaného souboru tak mohou být zařazeny pouze subjekty, které zadaly provedení evaluačního šetření externím realizátorům. V rámci takovýchto subjektů je konkrétním respondentem pracovník evaluačního oddělení (optimálně jeho vedoucí či zástupce vedoucího). Do zkoumaného souboru jsou zařazovány výlučně veřejné instituce, neboť je naše pozornost zaměřena na evaluaci veřejných intervenčních programů.

5.3 Technika výběru

Desk research

V rámci desk research byly identifikovány pro každou případovou studii základní soubory dokumentů/dat, které byly následně analyzovány. V případě závěrečných zpráv z evaluačních šetření provedených externími realizátory je základním souborem sada dokumentů shromážděných Evropským sociálním fondem v ČR. V případě autoevaluačních zpráv jde o výsledky vlastního hodnocení škol v ČR, které jsou veřejně dostupné a v případě posudků bakalářských prací jde o posudky všech prací obhajovaných v letním semestru 2008 a v roce

⁴⁶ Fyzické osoby se v roli evaluátorů však angažují při hodnocení projektových žádostí.

2009. Vzhledem ke skutečnosti, že ve všech případech byly analýze podrobeny veškeré dokumenty, není problematika výběru v tomto ohledu relevantní.

Field research

V rámci této fáze byla s ohledem na maximalizaci pokrytí využita kombinace pěti různých přístupů k identifikaci relevantních subjektů. Konkrétně se jednalo o (1) prohledání obchodního rejstříku, (2) shromáždění evaluačních zpráv zadávaných orgány veřejné správy, (3) aplikace výběrové techniky snowball mezi realizátory, (4) získání kontaktů z databáze členů ČES a EES z ČR, (5) provedení soupisu referencí realizátorů směřujícího k doplnění seznamu zadavatelů.

Z takto shromážděných údajů byly připraveny dvě samostatné výběrové opory / seznamy subjektů představujících základní skupiny pro každou z cílových skupin. S ohledem na relativně nízký počet jednotlivých subjektů v obou výběrových oporách (řádově desítky subjektů) byly s žádostí o participaci kontaktovány všechny subjekty. Interview byla provedena se zástupci 5 operačních programů a s 11 realizátory.

5.4 Výzkumné instrumenty

Desk research

Záznamové archy byly připraveny v návaznosti na cíle výzkumu. V první fázi byly připraveny jen jejich návrhy, které byly podrobeny zkušebnímu záznamu. Poté byl seznam sledovaných indikátorů doplněn a upraven tak, aby nezvyšoval riziko nízké reliability dané rozdílným pojetím jednotlivých posuzovatelů a současně aby získávané údaje dosahovaly co nejvyšší validity.

K dosažení vytyčených cílů a k verifikaci formulovaných hypotéz byly jednotlivé indikátory, zvoleny tak, aby vystihovaly význam a podstatu každé z klíčových dimenzí. Indikátory byly doplněny o několik dalších proměnných klasifikačního charakteru (počet stran, typ autora apod.) a o další znaky, které usnadňují adekvátní charakterizaci. Zkrácený seznam jednotlivých zaznamenaných znaků je uveden v příloze.

Field research

Výzkumné instrumenty byly připraveny na základě pečlivého studia relevantní literatury, zahrnující jak samotný předmět zkoumání (tj. konkrétní typy prováděných evaluací), tak také metodiku analogicky koncipovaných výzkumů. Připravené výzkumné instrumenty však nevycházejí přímo z žádných jiných výzkumných instrumentů.

Výzkumné instrumenty pro obě cílové skupiny byly koncipovány jako polostrukturované, jelikož bylo nutné získat komplexní vhled do motivačních faktorů determinujících utváření designu evaluačních šetření⁴⁷.

5.5 Sběr dat

Desk Research

Výskyt stanovených charakteristik analyzovaných závěrečných zpráv byl zaznamenáván vždy třemi na sobě nezávislými hodnotiteli. Každý z hodnotitelů postupoval samostatně podle předem stanovených instrukcí; prováděl záznam sledovaných indikátorů, prováděl analýzu všech shromážděných zpráv a dokumentů a výsledky své analýzy průběžně zaznamenával do individuálního datového souboru. V další fázi byly individuální datové soubory převedeny do souhrnného datového souboru a současně byla v této fázi ověřována shoda výsledků individuálních hodnocení. V případě nesouladu faktografických údajů (jako např. počet stran, rok přípravy zprávy apod.) byla informace ověřena čtvrtým hodnotitelem přímo v originálním dokumentu (tj. v autoevaluační zprávě). V případě ostatních údajů byla data přetřansformována do nové podoby zohledňující míru shody a neshody mezi jednotlivými posuzovateli.

Field Research

Sběr dat byl realizován formou osobních interview⁴⁸ prováděných na počátku roku 2010.

⁴⁷ Pro argumentaci použité kombinace přístupů viz Patton 1990, s. 287 nebo Robson 2000, s. 92.

⁴⁸ Tato metoda sběru dat byla zvolena s ohledem ke třem skutečnostem: (1) osobní dotazování působí důvěryhodněji (Singleton, Straits 1999). Vzhledem ke skutečnosti, že řada získávaných údajů měla charakter citlivých informací týkajících se např. vzájemných vztahů mezi zadavateli a realizátory, je prvek kredibility prováděného výzkumného šetření významným faktorem. V této souvislosti je třeba dále doplnit, že jedním z témat dotazování jsou také otázky znalostního, či přinejmenším povědomostního charakteru, jež by při jiné metodě sběru dat nemusely přinést spolehlivé odpovědi. (2) Osobní dotazování poskytuje respondentovi prostor

Polostrukturovaný charakter interview poskytuje dostatečné možnosti získávat podrobnější informační podklady dokreslující a vysvětlující pozici sledovaných subjektů v rámci klíčových ukazatelů charakterizujících stávající situaci a používané praktiky.

Sběr dat byl proveden v několika krocích, a to (1) telefonický screening, (2) osobní dotazování, (3) telefonický follow-up.

Role autora

Autor této práce provedl osobně všechna face-to-face interview, proto případná vychýlení zapříčiněná způsobem vedení interview působí konstantně na všechny provedené rozhovory. Tím je sníženo riziko vychýlených dat oproti situacím, kdy sběr dat provádí větší počet tazatelů.

Během jednotlivých interview nenastaly žádné specifické okolnosti, které by mohly ohrozit věrohodnost získávaných údajů. Přestože autor práce/tazatel s mnohými z respondentů již v minulosti spolupracoval, nejsou známy žádné konkrétní okolnosti, jak by tato předchozí spolupráce mohla ovlivnit odpovědi poskytované respondenty.

5.6 Etické souvislosti sběru dat

Důvěrnost získávaných informací je zaručena; byla přijata opatření zajišťující ochranu respondentů před přímou i nepřímou identifikací jejich identity. Z tohoto důvodu není uveden seznam organizací, které jsou zahrnuty ve zkoumaném souboru.

Informace jsou publikovány výlučně v souhrnné podobě. Veškeré výstupy z analýz druhého, popř. třetího stupně třídění byly vždy kontrolovány z hlediska možné nepřímé identifikace konkrétních respondentů. Pokud bylo takovéto riziko identifikováno, byla analýza provedena

zeptat se kolegů, konzultovat, což u jiných metod sběru dat, jako zejm. v případě CATI, není vůbec možné a u dalších variant (kupř. samovyplňovací dotazníky) není takovýto proces pod kontrolou výzkumníka. Mnohých interview se účastnilo více osob na straně respondenta, což by jiná než osobní forma dotazování vůbec nedovolovala, nebo by významně snížila možnosti kontroly daného procesu. Metoda osobního dotazování rovněž umožňuje zpřesnit význam kladených otázek, což je v daném případě velmi relevantní argument. Přestože dotazování bylo prováděno s profesionály, kteří se systematicky zabývají realizací evaluačních šetření, dala se na základě předchozích analýz dokumentů (závěrečných zpráv z provedených evaluací) odůvodněně předpokládat nízká míra povědomí o terminologii vztahující se k jednotlivým metodám a evaluačním postupům či přinejmenším značná konfúze ve významu klíčových pojmů. (3) Osobní dotazování obvykle vede k vyšší návratnosti než jiné metody sběru dat, jako např. CATI či SAQ (Babbie 1995).

s upravenými daty (kategorie zahrnující větší počet případů), v krajním případě bylo od dané analýzy zcela upuštěno.

Před provedením interview byl celý výzkumný proces revidován z hlediska možných porušení etického kodexu AAPOR/WAPOR.

Respondenti byli informováni o pořízení audio-nahrávky z důvodu provádění následných analýz již při telefonickém screeningu a pak znovu ještě před samotným začátkem interview. Dotazovaní byli poučeni o svých právech týkajících se zajištění důvěrnosti jejich odpovědí a byl získán jejich informovaný souhlas.

6. Uplatnění evaluace v praxi

Tato kapitola přináší celkem tři případové studie dokumentující aktuální způsoby realizace evaluačních šetření v České republice. Volba případových studií vychází z typů evaluačních šetření definovaných v podkapitole 4.2.3, kde jsou rozlišena evaluační šetření prováděná externí, interní a hybridní formou. Každá z případových studií je prezentována v podobné struktuře a rovněž charakter sledovaných údajů v rámci možností zachován co nejpodobněji, aby byly vytvořeny podmínky pro formulaci společných závěrů.

6.1 Externí forma provádění evaluačních šetření

Externě prováděná evaluační šetření mají ve velké většině *ad-hoc* charakter. Značná pestrost jednotlivých oblastí, témat a očekávání spolu s jejich nízkým absolutním počtem⁴⁹, fakticky znemožňuje stanovení společného jmenovatele, na jehož základě by mohly být vybírány a vzájemně porovnávány konkrétní příklady dokumentující způsoby provádění tohoto typu evaluačních šetření v České republice. Jedna z mála oblastí, která umožňuje se uvedeným omezením vyhnout je evaluace operačních programů. Vzhledem k tomu⁵⁰ je případová studie vybrána právě z této sféry.

6.1.1 Kontext případové studie

Jednotlivé typy evaluačních šetření jsou zakotveny v nařízení Rady (ES) a dále specifikovány v dalších dokumentech⁵¹ upravujících pravidla strukturální a kohezní politiky Evropské unie.

⁴⁹ V rámci celé ČR se jedná o jednotky případů nanejvýš o několik desítek provedených evaluačních šetření, které odpovídají definičnímu vymezení používanému v této práci.

⁵⁰ Existují však i další důvody, na jejichž základě byla zvolena oblast evaluace operačních programů. Za zmínku stojí zejména jednoznačnost definice základního souboru, který v daném případě zahrnuje všechny doposud připravené, projednané a zveřejněné závěrečné zprávy z evaluačních šetření provedených v rámci specifikované oblasti; dále je to relativní dostatek zdrojových dat (tj. závěrečných zpráv z provedených evaluací) daný legislativním rámcem; přesné stanovení jednotlivých typů evaluačních šetření (*ex-ante*, *interim*, *ex-post*), které sjednocuje cíle a smysl jejich provádění a rovněž také možnost získání zadávací dokumentace k jednotlivým evaluačním šetřením, z níž může být patrné původní zadání. K těmto principiálním důvodům je třeba přidat ještě jeden praktický – výsledky evaluačních šetření operačních programů jsou veřejně dostupné; naproti tomu výsledky evaluačních šetření prováděných v jiných oblastech takovýto charakter vesměs nemají (a to i přesto, že jejich zadavatelem jsou z valné většiny veřejné organizace financované ze státního rozpočtu). Přístup ke konkrétním pramenům je tak v případě evaluací operačních programů podstatně snazší, a také případný výběr konkrétních dokumentů pro prováděnou analýzu je v rámci této oblasti transparentnější.

⁵¹ Nařízení Komise (ES, Euratom) č. 2343/2002 ze dne 23. prosince 2002 o rámcovém finančním nařízení pro subjekty uvedené v článku 185 nařízení Rady (ES, Euratom) č. 1605/2002, kterým se stanoví finanční nařízení o souhrnném rozpočtu Evropských společenství. Nařízení Komise (ES) č. 1685/2000 ze dne 28. července 2000, kterým se stanoví prováděcí pravidla k nařízení Rady (ES) č. 1260/1999 týkající se způsobilosti výdajů na

V současné době stojí za bližší specifikaci dva legislativní prameny, a to nařízení Rady (ES) č. 1260/1999, které v člancích 40 – 43 vyjmenovává požadované tři typy evaluačních šetření (ex-ante, mid-term a ex-post) a definuje jejich základní charakteristiky. Je třeba připomenout, že tento pramen upravoval obecná pravidla pomoci a fungování operačních programů v minulém programovacím období (6. rámcový program) a v současnosti již pozbyl své platnosti. Nicméně řada z jeho ustanovení je významná i dnes. Druhým z legislativních pramenů je nařízení Rady (ES) č. 1083/2006 o obecných ustanoveních o Evropském fondu pro regionální rozvoj, Evropském sociálním fondu a Fondu soudržnosti a o zrušení nařízení (ES) č. 1260/1999, které v člancích 47 – 49 stanovuje cíle a způsoby hodnocení operačních programů v programovacím období 2007 – 2013. Podobně jako v 6. rámcovém programu je i v současném období požadováno provedení ex-ante a ex-post evaluačních šetření, nicméně mid-term evaluace byla opuštěna a namísto ní je požadováno provádění on-going evaluací.

Z uvedených pramenů a z dalších relevantních dokumentů plyne, že ex-ante⁵² evaluační šetření jsou zaměřena na stanovení potenciálu daného operačního programu či na šetření spočívající v analýze jejich silných a slabých stránek, rozvojového potenciálu a identifikaci možných rizik či ohrožení. Ex-ante evaluační šetření se dále zaměřují na optimalizaci přidělování rozpočtových zdrojů a na zlepšování kvality programování.

Mid-term (in-itinere, interim) evaluace se provádějí v polovině realizace operačního programu a mají formativní charakter (jejich cílem je korigovat dosavadní postupy). Vyžaduje se, aby tento typ evaluačních šetření byl proveden nezávislým subjektem, a aby se jejich hlavní pozornost zaměřila na výstupy (a výsledky) daného operačního programu, na míru, v jaké byly dosaženy vytyčené cíle a na způsob využívání a čerpání finančních prostředků. Členské státy jsou tak zavázány provádět toto průběžné hodnocení v souvislosti s monitorováním operačních programů.

činnosti spolufinancované strukturálními fondy. Nařízení Komise (ES) č. 438/2001 ze dne 2. března 2001, kterým se stanoví prováděcí pravidla k nařízení Rady (ES) č. 1260/1999, pokud jde o řídicí a kontrolní systémy pro pomoc poskytovanou ze strukturálních fondů. Nařízení Komise (ES) č. 1386/2002 ze dne 29. července 2002, kterým se stanoví podrobná pravidla pro prováděcí nařízení Rady (ES) č. 1164/94, pokud jde o řídicí a kontrolní systémy pro pomoc poskytovanou z Fondu soudržnosti a o postupy provádění finančních oprav.

⁵² Oficiální překlad nařízení Rady (ES) popisuje ex-ante evaluace jako „předběžné hodnocení“. V tomto textu však s uvedeným překladem záměrně nepracujeme, neboť evokuje provizorní charakter takovýchto evaluací, jejich nedokončenost či neúplnost, což nevystihuje jejich podstatu.

Doporučuje se, aby při provádění mid-term evaluací byly využívány monitorovací zprávy, nicméně tento požadavek není vždy možné naplnit⁵³. Provedení mid-term evaluací nebylo v ČR vyžadováno vzhledem k faktu, že ČR vstoupila do Evropské unie až v samém závěru daného programovacího období (termín pro dokončení mid-term evaluací byl 31. prosince 2003).

Nařízení Rady (ES) v neposlední řadě také stanovuje povinnost realizovat ex-post evaluace zaměřené na zkoumání rozsahu využití zdrojů, účinnosti a efektivity programování fondů a socioekonomického dopadu. Ex-post evaluace se zaměřují na analýzu dopadů daného operačního programu, a proto se provádějí až s časovým odstupem⁵⁴. Nařízení Rady (ES)⁵⁵ explicitně požaduje, aby ex-post evaluace, kromě již uvedeného, věnovaly pozornost také hlavním faktorům, jež přispívají k úspěchu či selhání realizace, dosaženým cílům a výsledkům spolu s posouzením jejich udržitelnosti.

Provedení všech výše uvedených typů evaluačních šetření je požadováno v rámci každého z operačních programů, a to vždy podle předem specifikovaných požadavků a v přesně stanovených časových rámcích. Kromě těchto tří typů evaluačních šetření se lze setkat také s tzv. tematickými evaluačními šetřeními, která jsou prováděna podle konkrétních potřeb řídicích orgánů jednotlivých operačních programů a je na jejich uvážení, zdali vůbec, kdy, v jakém rozsahu a v jakých oblastech budou takováto evaluační šetření zadávat.

6.1.2 Popis případové studie

Pro naplnění cílů naší práce byly vybrány všechny závěrečné zprávy, které jsou veřejně dostupné na WWW stránkách Evropského sociálního fondu⁵⁶. K datu provedení této analýzy

⁵³ Obvyklou bariérou je v tomto ohledu nedostatečná kvalita monitorovacích dat (nevhodný zdroj dat), jejich neúplnost (řada ukazatelů není k dispozici) a nejednotný formát (změny v průběhu implementace).

⁵⁴ Dle původních předpokladů měly být ex-post evaluace operačního programu šestého rámcového programu dokončeny do konce roku 2009; v současnosti (jaro 2010) se však s jejich realizací nezačalo a je velmi nepravděpodobné, že by byly dokončeny dříve, než v roce 2011.

⁵⁵ Ex-post evaluace jsou v českém překladu nazývány „dodatečným hodnocením“, což opět považujeme za neadekvátní a nesprávný překlad.

⁵⁶ www.esfcr.cz

(podzim 2009) tak bylo k dispozici celkem 22 závěrečných zpráv, které lze tematicky rozčlenit do čtyř základních skupin. První skupinu představují závěrečné zprávy z evaluačních šetření prováděných v rámci Operačního programu Lidské zdroje a zaměstnanost, pro který jsou k dispozici jak povinně prováděná šetření, tak také tematické evaluace. Z výstupů povinně prováděných evaluací je analyzována jedna závěrečná zpráva z ex-ante evaluace a dvě zprávy z on-going evaluací. V rámci tohoto operačního programu bylo provedeno i jedno tematické evaluační šetření. Celkem jsou tak v rámci tohoto operačního programu analyzovány čtyři závěrečné zprávy.

Druhým z operačních programů, ke kterému se vztahuje několik závěrečných zpráv je Operační program Rozvoj lidských zdrojů; v tomto případě je k dispozici závěrečná zpráva z ex-ante evaluačního šetření (včetně ex-ante evaluačního dodatku), poté řada tematických evaluačních šetření zahrnujících evaluaci projektů zaměřených na rovné příležitosti na trhu práce, evaluaci projektů zaměřených na problematiku romských komunit a několik dalších. V případě tohoto operačního programu nejsou k dispozici výsledky on-going evaluačního šetření s ohledem ke skutečnosti, že tento operační program byl vyhlášen v rámci minulého programovacího období, kdy on-going evaluace ještě nebyly definovány; k dispozici je však mid-term evaluační šetření. Celkem je tak k dispozici 9 závěrečných zpráv. Třetí a čtvrtá ze samostatných oblastí již nemá vazbu na konkrétní operační program; třetí oblast se týká evaluace Jednotného programového dokumentu pro cíl 3 regionu NUTS-2 hlavní město Praha⁵⁷, kde jsou k dispozici jak ex-ante evaluační šetření (včetně dodatku), tak také evaluace dosaženého pokroku a evaluace přínosu. Třetí skupina zahrnuje celkem 4 závěrečné zprávy. Čtvrtá oblast shrnuje evaluační šetření provedená v rámci iniciativy Equal, pro kterou jsou k dispozici závěrečné zprávy z jednotlivých etap průběžného hodnocení a samostatná evaluace principu mezinárodní spolupráce. Celkem je v rámci čtvrté skupiny zpracováno 5 závěrečných zpráv.

Rámcový pohled na počty závěrečných zpráv vycházející z typů evaluačních šetření provedených ve všech čtyřech skupinách poskytuje následující tabulka:

⁵⁷ Cíl 3 je referenčním dokumentem, na jehož základě byla poskytována podpora rozvoji lidských zdrojů na území hlavního města Prahy (NUTS-2) s využitím českých národních zdrojů a zdrojů Evropského sociálního fondu.

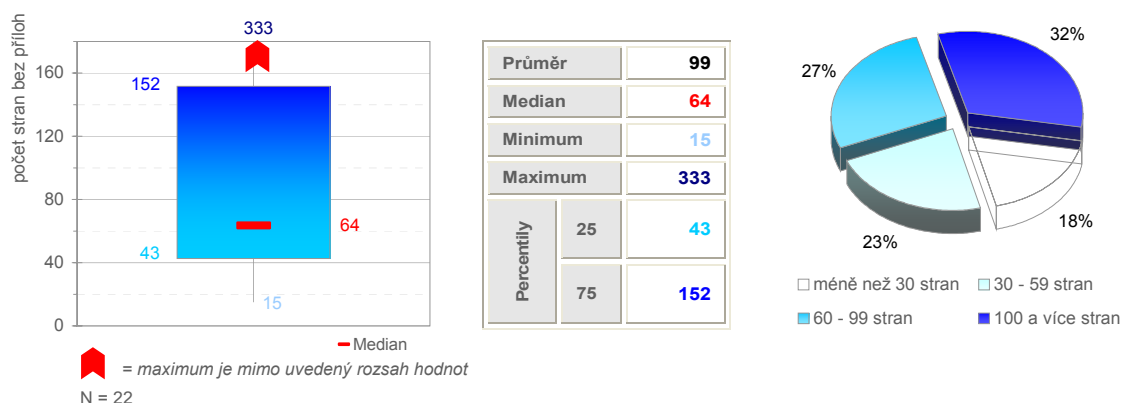
Tabulka č. 9 Počty posuzovaných závěrečných zpráv

Typ evaluačního šetření	Počet závěrečných zpráv
ex-ante	5 závěrečných zpráv
mid-term (2004 - 2006)/on-going (2007 - 2013)	8 závěrečných zpráv
ex-post	0 závěrečných zpráv
tematické evaluace	9 závěrečných zpráv
Celkem	22 závěrečných zpráv

Jak je na první pohled patrné, mezi posuzovanými závěrečnými zprávami nefigurují zprávy z ex-post evaluací, což je dáno faktickou absencí tohoto typu evaluací v ČR. Teprve v tomto časovém období (tj. jaro 2010) se totiž připravují ex-post evaluace operačních programů realizovaných ve zkráceném období 6. rámcového programu (2004 – 2006)⁵⁸.

Prvotní analýza rozsahu závěrečných zpráv ukazuje na vysoké rozpětí celkového počtu stran. V celkovém souboru všech 22 zpracovaných zpráv je téměř pětina kratší než 30 stran, přičemž nejkratší zpráva čítá 15 stránek.

Graf č. 1 Rozsah závěrečných zpráv (počet stran)



Na druhé straně má přibližně třetina zpráv více než 100 stran. Nejdelší zpráva obsahuje dokonce 333 stran, což prokazatelně vychyluje průměrný rozsah celého souboru analyzovaných zpráv (99 stran), neboť mediánová hodnota činí 64 stran.

⁵⁸ První výsledky těchto ex-post evaluací budou podle schváleného Evaluačního plánu Národního orgánu pro koordinaci k dispozici v druhé polovině roku 2012; ex-post evaluace aktuálních operačních programů pak budou podle současných předpokladů realizovány do konce roku 2015.

Celkový rozsah stran je třeba chápat jako prvotní předpoklad snadného používání výsledků provedených evaluačních šetření. Z hlediska snadné práce s jednotlivými zprávami je pochopitelně třeba vzít v potaz také jejich strukturu a konkrétní formu prezentace klíčových zjištění, přesto platí, že rozsáhlé zprávy přesahující 100 stran⁵⁹ budoucí utilizaci spíše komplikují, než aby ji usnadňovaly.

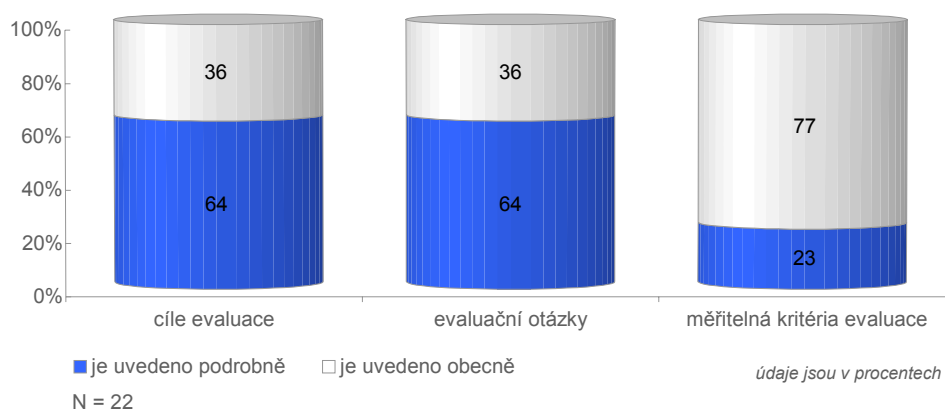
6.1.3 Výsledky

A. Orientace na cíl

Cíle evaluace spolu s formulací evaluačních otázek jsou páteří každého evaluačního šetření, neboť určují jeho obsahové zaměření a směřují rovněž ke stanovení konkrétního designu a k volbě metod sběru potřebných údajů. V mnoha případech představují jednotlivé okruhy cílů také osnovu pro přípravu závěrečné zprávy, které jsou uspořádány tak, aby umožnily snadné nalezení odpovědí na jednotlivé evaluační otázky.

Přesné stanovení cíle prováděného evaluačního šetření je tak klíčovým předpokladem jeho úspěšného provedení a dokončení. Obecně formulované cíle mohou komplikovat samotnou realizaci evaluačních šetření, neboť mohou vyvolat nesoulad v pochopení jeho cílů mezi zadavatelem a realizátorem. Z analyzovaných závěrečných zpráv je patrné, že naprostá většina (64 %) evaluačních šetření je založena na podrobně formulovaných cílech. Současně platí, že dvě třetiny ze závěrečných zpráv obsahují kromě detailně specifikovaných cílů také přesnou specifikaci evaluačních otázek.

Graf č. 2 Specifikace zaměření evaluačního šetření



⁵⁹ Údaj vyplývá z provedených hloubkových interview se zadavateli – blíže o tom viz kapitola sedmá.

Stanovení evaluační otázky však nemusí být výpovědí o přístupu realizátora a o jeho odborných předpokladech pro provádění evaluačních šetření, neboť evaluační otázky jsou v mnoha případech součástí zadávací dokumentace.

Z provedených analýz je patrné, že podrobně formulované cíle se ve větší míře vyskytují v závěrečných zprávách z evaluačních šetření zaměřených na výsledky operačních programů či dílčích aktivit prováděných v jejich rámci, což je dáno jejich převážně sumativním charakterem. Rovněž v rozsáhlejších závěrečných zprávách (o více než 60 stranách) je patrná vyšší míra pozornosti věnovaná detailnímu vymezení cílů prováděného evaluačního šetření a pečlivé formulaci evaluačních otázek.

V případě zbývající jedné třetiny závěrečných zpráv jsou cíle a evaluační otázky vyjádřeny jen povrchně. Je však třeba upozornit, že jejich obecné uvedení v závěrečných zprávách (popř. méně specifický způsob jejich prezentace) nemusí znamenat, že si evaluační tým jednotlivé cíle či evaluační otázky nestanovil podrobnějším způsobem⁶⁰. V souboru analyzovaných závěrečných zpráv se nevyskytl ani jeden případ, kdy cíle či evaluační otázky v publikovaném textu zprávy zcela absentovaly.

Pokud jde o podrobnost definování měřitelných kritérií prováděného hodnocení, jsou totiž výsledky analyzovaných závěrečných zpráv o poznání méně přesvědčivé. Měřitelná kritéria jsou podrobně uvedena v přibližně jedné čtvrtině závěrečných zpráv, zatímco převažující většina (77 %) uvádí měřitelná kritéria jen obecným způsobem. Jasně se tak naplňují některé charakteristiky evaluačních šetření, o nichž bylo v obecné rovině pojednáno v kapitole třetí⁶¹; důraz na užitečnost zjištění a nižší pozornost věnovaná technickým podrobnostem realizace evaluačních šetření je tedy třeba chápat jako doklad obvyklého zaměření evaluačních šetření na samotné výsledky spíše než na jako důkaz jejich nízké exaktnosti či nedostatečného konceptuálního ukotvení.

Lze tedy shrnout, že externě prováděná evaluační šetření důsledně naplňují zadání; vedou k vytyčeným cílům, naplňují svůj primární účel a v naprosté většině umožňují i dalším uživatelům ověřit, jak byly jednotlivé cíle naplněny.

⁶⁰ Vzhledem ke skutečnosti, že jde ve všech případech o dokončené a přijaté výstupy, je zřejmé, že zmíněné obtíže buď v procesu provádění jednotlivých evaluačních šetření vůbec nenastaly nebo byly uspokojivě vyřešeny.

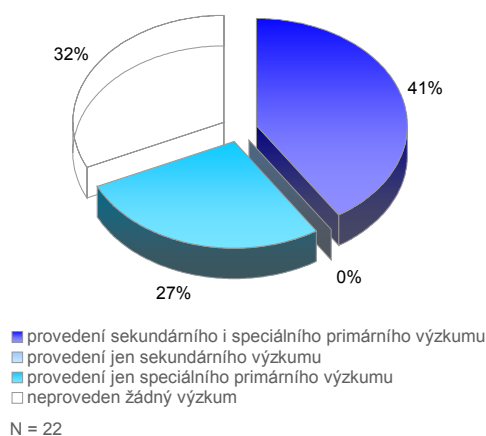
⁶¹ Viz část srovnávající evaluační šetření se základním výzkumem.

B. Exaktnost

Platí, že 95 % analyzovaných závěrečných zpráv (tj. všechny s výjimkou jediné) cituje starší dokumentaci, data či statistiky. V těchto zprávách jsou tedy jednotlivé záměry formulovány na základě empirických dat, ke kterým realizátoři evaluačních šetření získali přístup; uvedené údaje interpretují a dovozují z nich závěry pro vlastní evaluační úlohy. Je třeba upozornit, že citace empirických údajů nemusí automaticky předpokládat jejich analýzu, nemusí tedy být výsledkem analytické činnosti. I v této podobě je však patrná snaha jednotlivých evaluátorů dosahovat vnější validity svých závěrů používáním údajů popisujících kontext posuzovaných operačních programů a uváděním dat, která dokumentují dosažené výsledky či mapují výchozí situaci. Provedení sekundární analýzy dat, resp. specifikace postupu, který předpokládá (vyžaduje) provádění některých operací se získanými datovými údaji, je patrné v přibližně dvou pětinách závěrečných zpráv.

Zajímavé výsledky přináší analýza způsobů získávání empirických dat, o které se opírají o jednotlivé závěry specifikované v posuzovaných zprávách. V rámci přibližně jedné čtvrtiny (27 %) evaluačních šetření byl realizován primární výzkum, jehož výsledky představovaly jediný zdroj empirických údajů; v těchto případech není simultánní provedení sekundární analýzy zřejmé.

Graf č. 3 Provedení analýzy sekundárních dat a primárního výzkumu



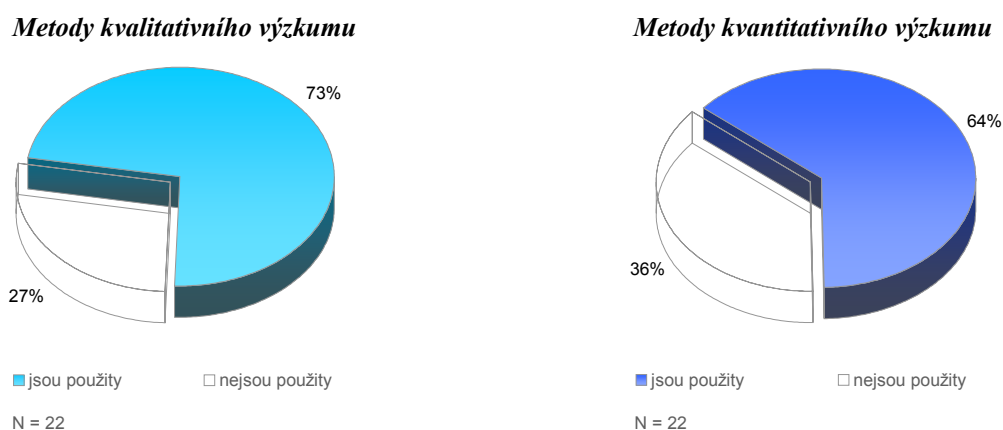
Dále platí, že přibližně třetina (32 %) závěrečných zpráv se opírá výlučně o citace sekundárních údajů, neboť v rámci realizace těchto evaluačních šetření nebyla provedena ani sekundární analýza, ani nebyl uskutečněn primární výzkum. Z podrobnější analýzy je v této souvislosti patrné, že se jedná zejména o výsledky ex-ante evaluací, které s ohledem ke svým specifickým cílům používají odlišné metody práce s informacemi a postupy jejich

vyhodnocení. V souboru analyzovaných podkladů zcela chybí závěrečné zprávy, při jejichž přípravě by byla provedena výlučně sekundární analýza. Ve dvou pětinách závěrečných zpráv je doloženo jak provedení analýzy sekundárních dat, tak také realizace primárního výzkumu.

Významnou měrou se na využívání sekundárních údajů podílejí data pocházející z monitoringu, která bývají externím realizátorům evaluačních šetření zpřístupňována, a to jak při přípravě průběžných evaluačních zpráv (on-going evaluací), tak také řady dalších evaluačních šetření (evaluačních šetření zaměřených na výsledky i některých tematických evaluací).

Z údajů uváděných v závěrečných zprávách je patrné, že zastoupení metod kvalitativního výzkumu dosahuje přibližně tří čtvrtinového podílu (73 %), zatímco podíl evaluačních šetření využívajících metody kvantitativního výzkumu je přibližně dvoutřetinový (64 %); vzhledem ke skutečnosti, že rozdíl 9 % představují dvě závěrečné zprávy, lze konstatovat, že metody kvalitativního i kvantitativního výzkumu jsou při provádění evaluačních šetření využívány přibližně stejnou měrou.

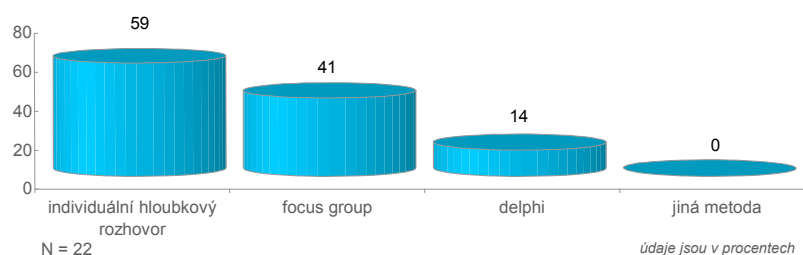
Graf č. 4 Využití metod kvalitativního a kvantitativního výzkumu



Dále je zřejmé, že 3 % provedených evaluačních šetření byly oba přístupy kombinovány, což není v kontextu evaluačních šetření žádný překvapující výsledek, neboť na tento jev poukazují mnohé prameny (Fitz-Gibbon, Morris 1987, Davidson 2005, Greene 2007, Mark, Shotland 1987).

V souvislosti s popisem používání empirických údajů stojí za zmínku také konkrétní metody, na jejichž aplikaci lze ze závěrečných zpráv usoudit. Ve třech pětinách evaluačních šetření byly podle specifikací v závěrečných zprávách prováděny individuální hloubkové rozhovory, což je hodnota, která činí z této metody nejvíce rozšířenou techniku, kterou realizátoři evaluačních šetření používají. Extenzivní používání této metody je dáno jednak nízkou finanční, časovou a organizační náročností a navíc je umocněno faktem, že realizátoři evaluačních šetření ve velké části konkrétních aplikačních situací disponují adekvátní výběrovou oporou (seznamy recipientů, databáze žadatelů apod.), která použití této metody usnadňuje. Za povšimnutí stojí rovněž poměrně vysoký podíl evaluačních šetření využívajících focus groups (41 %). Platí přitom, že nasazení této metody je patrné průřezově ve všech typech evaluačních šetření, tedy ex-ante, mid-term i při tématických evaluačních šetřeních.

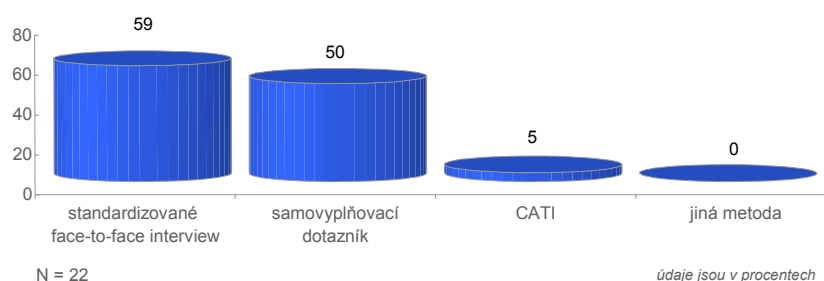
Graf č. 5 Konkrétní použité metody kvalitativního výzkumu



Pozn. Dopočet do 100 % představují závěrečné zprávy, ve kterých není daná metoda sběru dat zmíněna.

Analýza metod kvantitativního výzkumu přináší očekávané výsledky; za zmínku stojí jen CATI, které je využíváno marginálně za specifických okolností⁶².

Graf č. 6 Konkrétní použité metody kvantitativního výzkumu



Pozn. Dopočet do 100 % představují závěrečné zprávy, ve kterých není daná metoda sběru dat zmíněna.

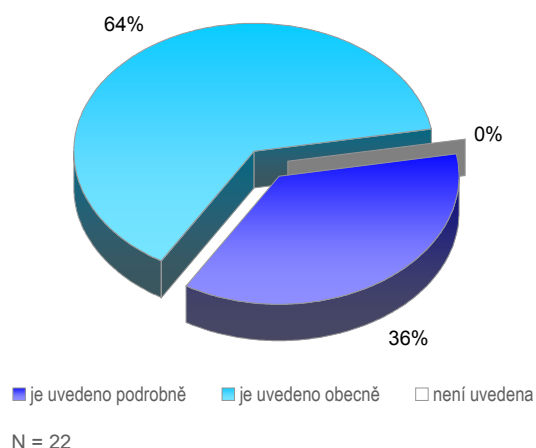
⁶² Jednotlivé subjekty, které realizují evaluační šetření, většinou nedisponují potřebným zázemím pro provádění kvantitativních výzkumů (nemají vlastní CATI studia, nedisponují tazatelskými sítěmi a nevlastní ani speciální softwarové nástroje pro správu a analýzu dat).

Z analýzy závěrečných zpráv je rovněž patrné, že kvantitativní výzkum je deklarován v závěrečných zprávách z evaluačních šetření, na kterých se podílel vyšší počet autorů, zatímco v případě použití metod kvalitativního výzkumu není tato diference patrná.

C. Evaluační přístup

V souvislosti s charakterizací principiálních odlišností, které vytvářejí jedinečnost evaluačních šetření a odlišují je od jiných činností jsme v kapitole 3.2.1 poukazovali, že specifickou úlohou evaluačních šetření je identifikace zainteresovaných stran a popis jejich zájmů. V tomto ohledu stojí za pozornost analýza vypovídající o míře, v jaké jsou ve vzorku analyzovaných závěrečných zpráv jednotlivé zainteresované strany skutečně identifikovány.

Graf č. 7 Identifikace jednotlivých zainteresovaných stran



Ukazuje se, že s pojmenováním a analýzou jednotlivých zainteresovaných stran se lze setkat ve všech zprávách, které byly zařazeny do zkoumaného souboru. Ve více než jedné třetině zpráv (36 %) je dokonce proveden jejich podrobný rozbor, zatímco zbývající zprávy (64 %) věnují této problematice spíše jen povrchní pozornost. S největší pečlivostí jsou zainteresované strany identifikovány ve zprávách z evaluačních šetření zaměřených na hodnocení výsledku, což je pochopitelné s ohledem k jejich cíli. Na druhé straně, podrobná identifikace zainteresovaných stran není patrná v žádné zprávě z ex-ante evaluací; v jejich případě jsou jednotlivé zainteresované strany popsány jen obecně. Z porovnání závěrečných zpráv připravených v rámci různých operačních programů je patrné, že podrobná identifikace a analýza zainteresovaných stran je v největší míře provedena v evaluačních šetřeních uskutečněných v rámci Operačního programu Rozvoje lidských zdrojů.

Identifikace zainteresovaných stran je nejen klíčovým prvkem evaluačních šetření, který umožňuje evaluátorovi orientaci v síti vztahů možných subjektů, jež mohou být dotčeny výsledkem prováděného evaluačního šetření, ale jde rovněž o naplnění Standardů provádění evaluačních šetření⁶³ a respektování rad uvedených v příručkách pro provádění evaluací operačních programů.

D. Doporučení

Samostatnou pozornost si zaslouží zjištění, že všechny (100 %) z analyzovaných zpráv obsahují návrh konkrétních kroků ke zlepšení v definovaných oblastech. Překvapující je v této souvislosti ta samozřejmost, s níž se v závěrečných zprávách doporučení objevují⁶⁴. Bez ohledu na typ prováděného evaluačního šetření, bez ohledu k jeho konkrétním cílům, smyslu a účelu jsou návrhy na zlepšení všudypřítomné.

V souvislosti s tímto poznatkem nelze opomenout diskusi mezi jednotlivými evaluátory, která se týká právě poskytování doporučení. Někteří z evaluátorů (Scriven 1991) totiž zastávají stanovisko, že evaluační šetření by měla být oproštěna od poskytování doporučení, jak reagovat na závěry provedené evaluace. Poukazují přitom na nestejnost rolí evaluátora a konzultanta, do níž může být realizátor evaluačního šetření vmanévrován, pokud bude kromě samotné evaluace navrhopat také další kroky. Rovněž se v této souvislosti poukazuje na ztrátu nezávislosti provedené evaluace i evaluátora samotného, neboť poskytování doporučených kroků nevyhnutelně vede ke stanovení priorit a k zaujetí stanoviska k předmětu evaluace, což je z hlediska profesních standardů evaluátora nepřijatelné. Evaluátor za dané situace přestává být tím, kdo pomáhá jiným v jejich rozhodovací činnosti a stává se sám rozhodovatelem.

Na druhé straně je z konkrétních závěrečných zpráv patrné, že poskytnutá doporučení nemají ve své většině charakter volby z možných alternativ, ale jde spíše o zprostředkování zkušeností jednotlivých zainteresovaných stran, s nimiž byli evaluátoři participující na jednotlivých projektech v kontaktu. Z kontextu je rovněž zřejmé, že tento prvek závěrečných zpráv je v téměř všech případech vynucen zadavateli, kdy návrh opatření tvoří součást samotné zadávací dokumentace a není-li uvedený požadavek odpovídajícím způsobem formalizován, bývá obvykle ventilován během procesu akceptace závěrečné zprávy a vpořádání jednotlivých připomínek.

⁶³ Ty však v českém evaluačním prostředí nejsou doposud zavedeny.

⁶⁴ Z povahy některých šetření lze o jejich relevanci přinejmenším pochybovat; v případě ex-post evaluace jde v první řadě o zhodnocení dopadu, tedy o posouzení dosavadního výkonu intervence prováděné v rámci daného operačního programu. Takováto závěrečná zpráva by tedy spíše než návrhu budoucího zlepšení posuzovaného intervenčního programu měla věnovat pozornost adekvátní evaluaci dosažených efektů a dostatečně argumentovat uvedená tvrzení o případných rezervách v jejich efektivitě.

6.1.4 Diskuse výsledků

Případová studie dokumentující stávající praxi v oblasti externě prováděných evaluačních šetření byla tedy vybrána z oblasti evaluace operačních programů, jež jsou vyhlašovány na základě cílů strukturální a kohezní politiky Evropské unie v časově specifikovaných rámcových programech. V této souvislosti je třeba poukázat na několik specifických okolností, které provádění tohoto typu evaluačních šetření provázejí; první se týká zobecnitelnosti získaných poznatků. Na základě analyzovaného vzorku závěrečných zpráv sice lze činit mnohé závěry týkající se zvyklostí a praktik při provádění evaluací operačních programů, nicméně jejich extrapolace mimo tuto oblast by mohla být zavádějící. Na evaluaci operačních programů je totiž nutné nahlížet jako na specifický evaluační systém, který je připraven tak, aby naplňoval poměrně konkrétní cíle, které jsou od něj očekávány, a aby napomáhal řídicímu orgánu při administraci operačního programu. Navzdory tomuto omezení je volba případové studie z oblasti evaluace operačních programů smysluplná, neboť celé této oblasti je přisuzován mimořádný význam, a to nejen v ČR, ale i v jiných evropských zemích.

Je třeba mít na paměti, že závěrečné zprávy jsou výsledkem procesu, při kterém zadavatelé specifikují své požadavky, připomínají dřívější verze závěrečných zpráv a větší či menší měrou se tak spolupodílejí na jejím obsahu a zaměření⁶⁵. Tato skutečnost je tedy významným činitelem, který je třeba zohlednit při interpretaci dosažených poznatků.

Přestože byla analýza provedena na všech zprávách dostupných z citovaného pramene, přesto je třeba poukázat na nízký absolutní počet případů (závěrečných zpráv). Robustnější datový základ by jednak pomohl zpřesnit celou řadu analýz a závěrů, a dále by umožnil formulovat některé nové hypotézy (např. o vztazích mezi charakteristikami závěrečných zpráv a složením analytických týmů; o vývoji některých ukazatelů v čase apod.).

6.1.5 Dílčí závěry

Na základě diskuse o přítomnosti doporučení v závěrečných zprávách je na místě věnovat tomuto jevu zvýšenou pozornost a je vhodné zaměřit další (analytické) úsilí na analýzu

⁶⁵ Např. absence konceptuálních modelů může být důsledkem nízkého povědomí o tomto nástroji mezi realizátory evaluačních šetření, ale stejně tak může být důsledkem rozhodnutí zadavatele o nezařazení tohoto výstupu do závěrečné zprávy.

charakteru doporučení, jejich vztahu k empirickým podkladům a rovněž k odhadu míry, ve které uvedená doporučení přispívají či napomáhají využitelnosti výsledků provedených evaluačních šetření.

Závěrečné zprávy obsahují ve značné míře veškeré podstatné náležitosti jako je zejména definice cílů, stanovené evaluační otázky apod. Je však třeba zároveň upozornit, že řada skutečností je popsána jen povrchně, což může být dáno snahou nezvyšovat rozsah závěrečných zpráv, na druhé straně tím klesají možnosti sekundárního využití připravených podkladů.

Samostatnou pozornost si zaslouhuje absence specifických evaluačnických nástrojů, jakými jsou např. konceptuální modely a rovněž je třeba konstatovat, že využití teoretických přístupů popsaných v podkapitole 4.1 této práce je nezřetelné. Chybí explicitní přihlášení se k nim a faktická podobnost v dílčích ohledech je spíše nahodilá. Ukazuje se tak, že praktické provádění evaluací je vedeno konkrétními okolnostmi, zatímco principy evaluační metodologie nejsou tak zřetelné.

6.2 Interní forma provádění evaluačních šetření

V oblasti interních evaluací je aktuální stav evaluační praxe prezentován na případové studii, která shrnuje některé poznatky o způsobu provádění vlastního hodnocení škol. Vzhledem k zákonem požadovanému provádění autoevaluací má takto zvolená případová studie jasně stanovený rámec možného zobecnění; souhrn poznatků z analyzovaného výběrového souboru tak může odrážet poměry v celé základní populaci (tj. v jednotlivých typech škol), a díky tomu je možné na základě analyzované skupiny případů usuzovat na obecné jevy provázející realizaci tohoto typu evaluačních šetření. Případová studie současně naplňuje podstatné znaky interně prováděných evaluací⁶⁶, neboť je prováděna vlastními pracovníky, kteří mají zájem na zlepšování vlastních činností a postupů.

6.2.1 Kontext případové studie

Vlastní hodnocení škol je v ČR vyžadováno zákonem č. 561/2004 Sb., o předškolním, základním, středním, vyšším odborném a jiném vzdělávání (školský zákon), který v §12 stanovuje náležitosti hodnocení škol, školských zařízení a vzdělávací soustavy. Současně také specifikuje konkrétní cíle, jež má vlastní hodnocení školy sledovat. V této souvislosti stojí za pozornost, že vlastní hodnocení školy je východiskem pro zpracování výroční zprávy o činnosti školy a rovněž slouží jako jeden z podkladů pro celkové hodnocení školy, které provádí Česká školní inspekce.

Daná oblast je kromě citovaného zákona dále upravena vyhláškou Ministerstva školství mládeže a tělovýchovy č. 15/2005 Sb., kterou se stanoví náležitosti dlouhodobých záměrů, výročních zpráv a vlastního hodnocení školy. Vyhláška mj. detailně specifikuje jednotlivé

⁶⁶ Je však třeba připomenout, že podle některých autorů se případy podobné vlastnímu hodnocení škol odlišují od obvyklé definice interních evaluací, neboť zadavatel zde není totožný s předmětem hodnocení. U interně prováděných evaluací totiž platí naprosté sjednocení všech tří elementů, tedy jak evaluátora a iniciátora, tak i předmětu evaluace. V případě vlastního hodnocení školy by tedy iniciátorem (či zadavatelem) evaluace musel být přímo ředitel školy, popř. jiný pracovník školy. V reálné situaci tomu však není, neboť iniciátorem vlastního hodnocení školy je MŠMT, které provádění tohoto typu evaluací nařizuje školám pomocí právního normativního aktu. Vzhledem ke skutečnosti, že v ostatních ohledech vlastní hodnocení škol vyhovuje definici interních evaluací, a s přihlédnutím ke skutečnosti, že i v daném případě MŠMT školám sice ukládá povinnost vlastní hodnocení provádět, ale primárním uživatelem výsledků vlastního hodnocení jsou samotní pracovníci školy, považujeme oblast za vyhovující cílům naší práce.

oblasti, kterých se vlastní hodnocení musí dotýkat. Platí tedy, že vlastní hodnocení školy se zaměřuje na:

- 1) cíle, které si škola stanovila,
- 2) posouzení, jakým způsobem škola vytyčené cíle plní,
- 3) přesnou identifikaci oblastí, ve kterých je třeba zlepšit úroveň vzdělávání spolu s návrhem příslušných opatření,
- 4) vyhodnocení účinnosti těchto opatření v následujícím období.

Kromě těchto obecných cílů anticipovaných vlastním hodnocením vyhláška dále specifikuje celkem šest hlavních oblastí, v nichž má být autoevaluace provedena. Jsou jimi:

- 1) podmínky ke vzdělávání,
- 2) průběh vzdělávání,
- 3) podpora školy žákům a spolupráce s rodiči,
- 4) výsledky vzdělávání,
- 5) kvalita řízení školy,
- 6) úroveň výsledků práce školy.

Zákon ani vyhláška nspecifikují, jak přesně má být autoevaluace provedena, nspecifikují podobu výsledků, nepředepisují žádné metody a ani nijak neupravují způsob projednání výsledků a jejich případnou implementaci. Školám je tak ponechán značný prostor k uplatnění vlastních metodik a k vlastnímu uchopení zadané povinnosti. Rovněž je třeba uvést, že zmíněná norma je zákonem imperfektním, neboť nestanoví žádnou sankci za nesplnění uložené povinnosti⁶⁷.

V návaznosti na zákonnou úpravu definovala Česká školní inspekce jednotlivé charakteristiky, podle nichž by vlastní hodnocení škol mělo být prováděno; vlastní hodnocení školy tedy (1) poskytuje zpětnou vazbu vzhledem k cílům školy, (2) vypovídá o silných a slabých stránkách školy, (3) vypovídá o účinnosti přijatých opatření, (4) slouží k tvorbě (korekcím) školního vzdělávacího plánu, (5) ovlivňuje DVPP, (6) ovlivňuje vzdělávací strategie školy, (7) vede k opatřením ředitele.

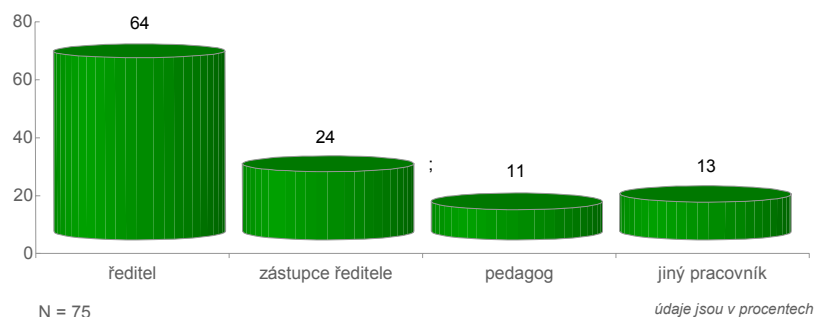
⁶⁷ Faktem zůstává, že díky soustavné kontrole ze strany České školní inspekce je tlak na dodržování zákonného ustanovení poměrně silný.

Provádění evaluací formou vlastního hodnocení vyžaduje náležité proškolení jednotlivých pracovníků, kteří sebehodnotící zprávy připravují; pro účely vlastní evaluace škol je připravena řada metodických pokynů, příruček a manuálů (Michek 2005, Michek 2006), které pomáhají adekvátním způsobem naplňovat vyhláškou definované požadavky. Rovněž jsou organizována speciální školení, jejichž cílem je jednak poskytnout jednotlivým pracovníkům škol již nabyté zkušenosti se způsobem provádění vlastního hodnocení, a dále ukázat způsoby, jak vlastní hodnocení využívat při rozvoji školy.

6.2.2 Popis případové studie

Největší podíl analyzovaných autoevaluačních zpráv byl připravován přímo řediteli škol (64 %) a přibližně jedna čtvrtina zpráv byla připravena zástupci ředitelů. Na přípravě autoevaluačních zpráv se autorsky dále podíleli v 11 procentech případů také ostatní pedagogičtí pracovníci a ve 13 procentech i ostatní zaměstnanci školy.

Graf č. 8 Funkce autora autoevaluační zprávy

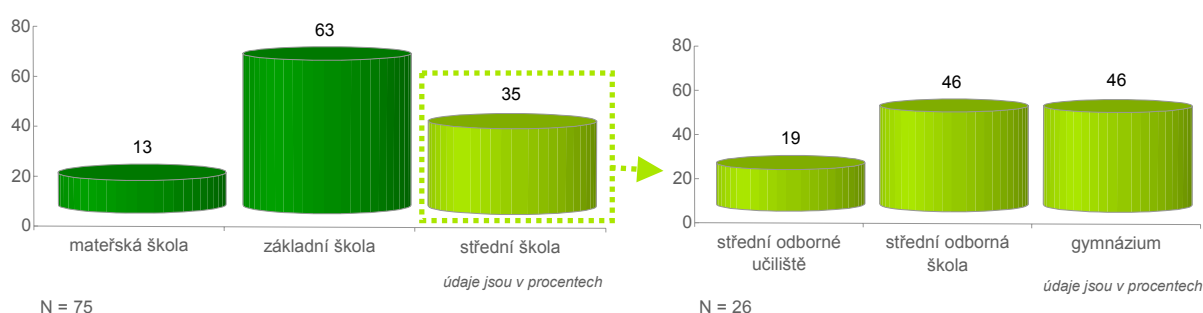


Pozn. Součet hodnot je vyšší než 100%, neboť na přípravě některých autoevaluačních zpráv se podílelo více autorů zastávajících různé funkce.

V této souvislosti zaujme skutečnost, že tři čtvrtiny (75 %) autoevaluačních zpráv bylo připraveno pouze jedním jediným autorem. Ačkoliv nelze vyloučit participaci ostatních členů pedagogického sboru jednotlivých škol a dalších pracovníků školy na samotném autoevaluačním procesu bez toho, že by byli zmíněni v závěrečné zprávě a přestože je pravděpodobné, že se ostatní pracovníci podíleli na autoevaluačních zprávách přípravou jednotlivých dílčích podkladů, lze v dané situaci vyjádřit pochyby o adekvátním zapojení všech zainteresovaných stran do procesu vlastního hodnocení.

Analyzovaný vzorek autoevaluačních zpráv zahrnoval závěry z vlastního hodnocení provedeného ve třech typech škol, a to v mateřských školách (13 %), na základních školách (63 %) a na školách středních (35 %). V rámci středních škol jsou rovnoměrně zastoupena gymnázia (coby reprezentanti poskytovatelů všeobecného vzdělání) a střední odborné školy. Podíl obou typů škol na sledovaném souboru činí 46 %. Zbývající část (19 %) připadá na střední odborná učiliště.

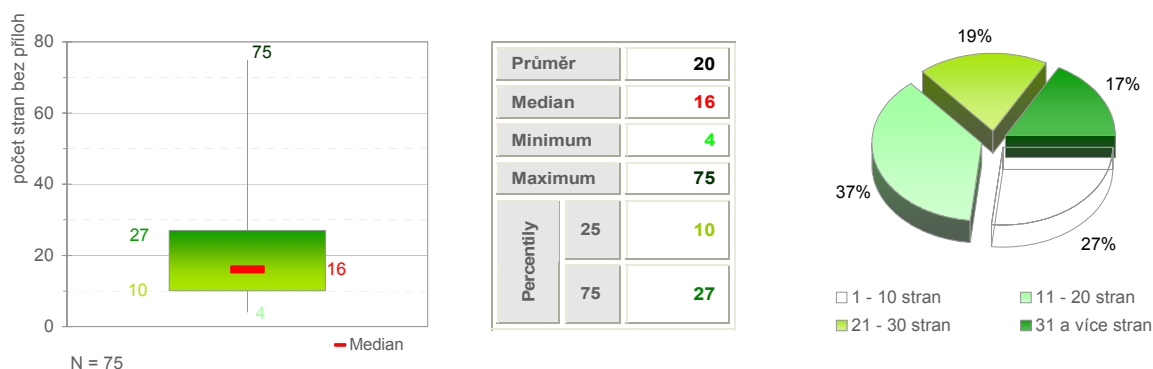
Graf č. 9 Typ subjektu



Pozn. Součet hodnot je vyšší než 100%, neboť některé školy mají duální charakter (typicky kombinují střední odborné učiliště a střední odbornou školu)

Bez zajímavosti není ani celkový rozsah zpráv, resp. rozpětí tohoto rozsahu; nejstručnější zpracovaná zpráva z vlastního hodnocení školy čítala 4 strany, zatímco druhý extrém představuje závěrečná zpráva o 75 stranách. Průměrný rozsah autoevaluační zprávy je 20 stran, přičemž mediánová hodnota činí 16 stran; z celkového počtu všech analyzovaných zpráv má 17 % rozsah vyšší než 30 stran.

Graf č. 10 Rozsah autoevaluačních zpráv (počet stran)



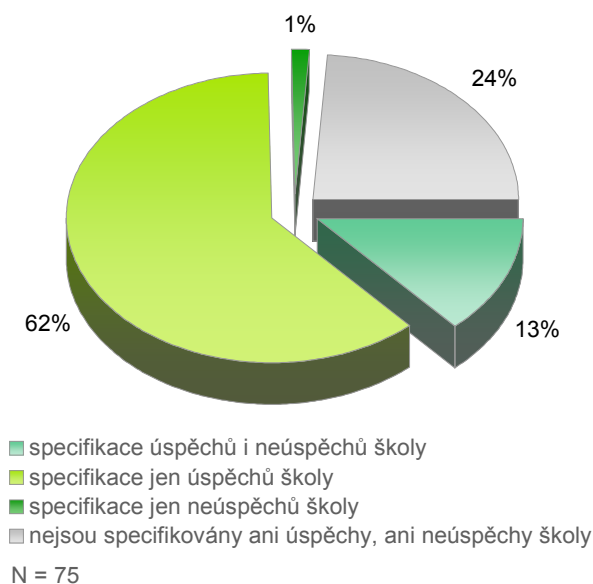
Tyto údaje zde nejsou uváděny samoučelně, neboť z dalších analýz je zřetelně patrná přímá úměra mezi rozsahem zpráv a adekvátností provedeného vlastního hodnocení včetně úplnosti pokrytí jednotlivých témat požadovaných vyhláškou Ministerstva školství mládeže a tělovýchovy.

6.2.3 Výsledky

A. Asymetrie vlastního hodnocení

Z konkrétních výsledků je patrné, že způsob, jakým jsou autoevaluace prováděny, může naplnit jejich poslání⁶⁸ ve značné míře, nikoliv však zcela. Některé z analyzovaných autoevaluačních zpráv totiž dokumentují faktickou absenci účinné zpětné vazby generované na úrovni jednotlivých subjektů. Konkrétně se ukázalo, že 62 % autoevaluačních zpráv obsahuje popis dosažených úspěchů, aniž by ovšem zmiňovaly neúspěšné aktivity škol.

Graf č. 11 Specifikace úspěchů a neúspěchů školy v daném období



Je přitom velmi nepravděpodobné, že by se neúspěšné činnosti tolika školám zcela vyhýbaly, resp. že by tak vysoký podíl škol provedl veškeré činnosti způsobem, který neumožňuje nalézt konkrétní poznatky pro jejich budoucí vylepšení. Z podrobnější analýzy navíc vyplývá,

⁶⁸ Např. identifikovat vhodné a méně vhodné praktiky v rámci jednotlivých činností vykonávaných danou vzdělávací institucí; směřovat ke zdokonalování školy v definovaných oblastech apod.

že mezi školami využívajícími zákonnou možnost provádět vlastní hodnocení jen jednou za dva roky⁶⁹, dosahuje podíl škol poukazujících výlučně na své úspěchy dokonce 70 %. Delší časový úsek přitom pravděpodobnost výskytu neúspěšných projektů pochopitelně zvyšuje. Lze tedy dovodit, že důvodem chybějící zmínky o takových činnostech může být nedostatečná schopnost identifikovat vlastní nedostatky nebo nedostatečná snaha hledat ve vlastních chybách poučení. Pravděpodobným vysvětlením je také obava ředitelů škol z případných následků, které by přiznání vyskytujících se chyb mohlo vyvolat. Ověření těchto hypotéz však není možné z dostupných údajů provést.

Uvedený poznatek je dále umocněn skutečností, že autoři přibližně jedné čtvrtiny autoevaluačních zpráv se jakékoliv specifikaci úspěchů a neúspěchů školy v posuzovaném období zcela vyhnuli – zprávy v tomto případě poskytují pouhý přehled vykonaných činností bez toho, že by byly adekvátně posouzeny. Tento přístup je patrný zejména v autoevaluačních zprávách připravených jedním jediným autorem (obvykle ředitelem školy). V této souvislosti stojí za připomenutí, že mezi důležité zásady evaluačního procesu patří právě schopnost uvědomovat si nedostatky (Bečvářová 2004).

Význam uvedeného zjištění je v zásadě dvojitý: zaprvé vypovídá o způsobu zpracování vlastního hodnocení a o konkrétních obtížích, které tento proces dotčeným subjektům přináší. V tomto ohledu lze výsledky využít při návrhu konkrétních opatření, která by pomohla jednotlivým školám provést vlastní hodnocení lépe a vytěžit z něj větší rozsah poznatků pro další činnosti. Patrně by bylo na místě dodatečně objasnit zainteresovaným stranám (zejména pak ředitelům škol a pedagogům) samotný smysl a účel provádění vlastního hodnocení. Zadruhé lze na uvedené zjištění pohlížet jako na indikátor evaluačních schopností, který vypovídá o nevhodném způsobu provádění autoevaluačních procesů; ty by totiž za běžných okolností měly vést k získání adekvátní zpětné vazby využitelné při dalším rozvoji.

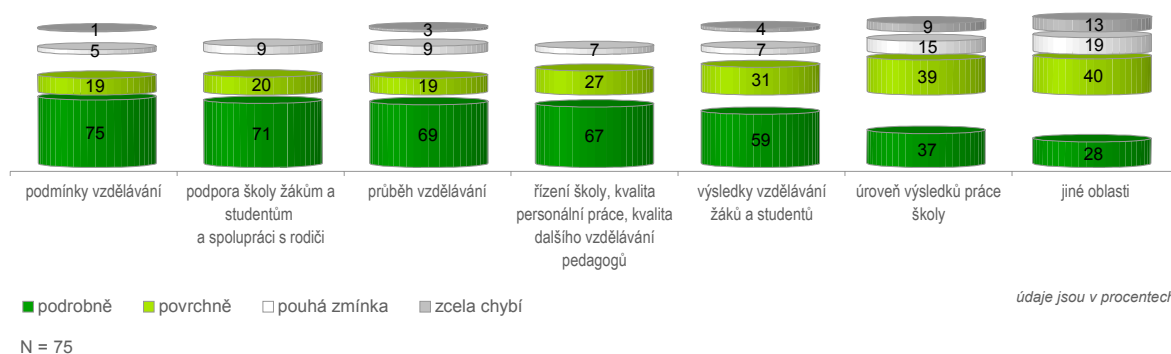
B. Podrobnost vlastního hodnocení

Již jsme uvedli, že vyhláška Ministerstva školství, mládeže a tělovýchovy 15/2005 Sb. školám předepisuje rámcovou strukturu vlastního hodnocení a specifikuje celkem šest oblastí, jichž se mají autoevaluace týkat. Z konkrétních dat je patrné, že se školy tímto ustanovením ve značné

⁶⁹ A nikoliv každoročně, jak je to běžné pro přípravu výroční zprávy.

míře řídí, neboť absolutní většina zpráv skutečně obsahuje přinejmenším zmínku o oblastech specifikovaných vyhláškou. Existuje sice i několik výjimek⁷⁰, nicméně je třeba uvést, že tyto absence nemají systematický charakter; v analyzovaném souboru se nevyskytuje žádná závěrečná zpráva, v níž by chyběla více než jedna oblast.

Graf č. 12 Obsah jednotlivých částí autoevaluační zprávy



Platí, že školy věnují největší pozornost podmínkám vzdělávání, které jsou ve třech čtvrtinách autoevaluačních zpráv obsaženy podrobně. Nicméně i dalším pěti vyhláškou požadovaným oblastem je věnována podrobná pozornost v nadpoloviční většině analyzovaných zpráv. Výjimku v tomto ohledu představuje oblast úrovně výsledků práce školy, jejímuž posouzení se podrobně věnovaly přibližně dvě pětiny škol; stejný podíl pak této oblasti věnoval povrchní pozornost.

Jako neopodstatněné se ukazují obavy, že se jednotlivé školy při svém vlastním hodnocení zaměřují výlučně na vyhláškou stanovené oblasti a zcela přitom opomíjejí ostatní aspekty života školy. Z konkrétních údajů vyplývá, že celkem 87 % škol ve svých autoevaluačních zprávách popisuje také jiné oblasti, než požaduje vyhláška; autoevaluaci v jiných oblastech věnují větší prostor střední odborné školy. V tomto ohledu se naplňuje předpoklad, že text vyhlášky stanovující jednotlivé oblasti provádění vlastního hodnocení je stimulem komplexního přístupu škol k provádění autoevaluačních procesů (Vašátková, Prášilová 2006), a nikoliv inhibítoem kreativního přístupu škol.

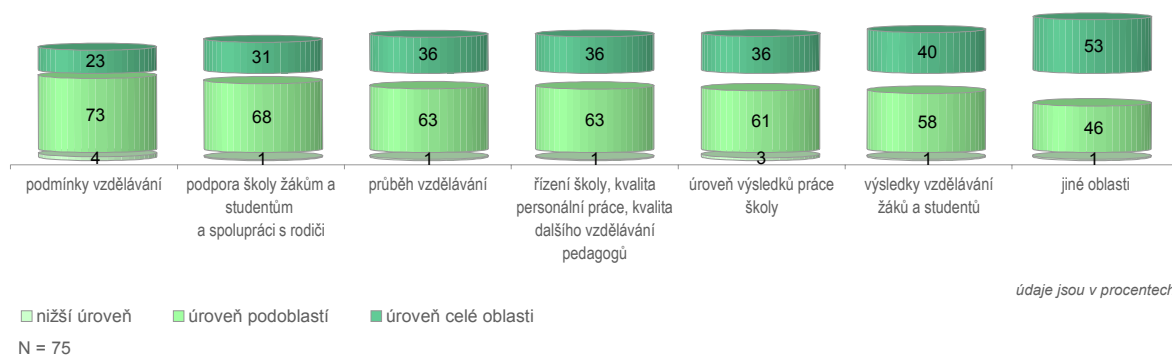
V autoevaluační zprávě může být samotná zmínka o specifikovaných oblastech sice dostatečným splněním požadavků daných vyhláškou, nicméně nevypovídá nic o

⁷⁰ Např. 9 % autoevaluačních zpráv neobsahuje žádné informace týkající se úrovně výsledků práce školy a 4 % zpráv vynechávají výsledky vzdělávání žáků a studentů.

smysluplnosti a podrobnosti zpětné vazby. Objevují se taktéž významné diference mezi jednotlivými oblastmi: např. podmínky ke vzdělávání jsou podrobně uváděny ve třech čtvrtinách sledovaných zpráv, zatímco úroveň výsledků práce školy je podrobně analyzována v 37 % zpráv. Významné rozdíly jsou patrné i mezi autoevaluačními zprávami z různých typů škol: větší pozornost je daným oblastem věnována středními školami ve srovnání se školami mateřskými a základními. Podobně také střední odborné školy věnují této části vlastního hodnocení významně větší pozornost než střední školy poskytující všeobecné vzdělání (gymnázia).

Volnost, která je školám ponechána při volbě konkrétního způsobu jak vlastní hodnocení v jednotlivých oblastech uskutečnit, jaké údaje použít a jak zajistit adekvátní přenos zpětné vazby na dotčené subjekty, se odráží ve značné diferenciaci způsobu provedení mezi jednotlivými školami. Je pochopitelné, že v tak široce definované oblasti, jakou jsou např. podmínky ke vzdělávání, existuje mnoho dílčích jevů iniciovaných různými zainteresovanými stranami, vyskytuje se řada činností s různým výsledkem a souhrnné posouzení má pak nutně jen orientační charakter. V téměř třech čtvrtinách autoevaluačních zpráv byla proto oblast podmínek vzdělávání pro účely vlastního hodnocení dále rozdělena na jednotlivé podoblasti a samotné hodnocení bylo provedeno teprve na takto konkretizované úrovni. V ostatních oblastech je takovýto přístup patrný v menší míře, nicméně stále platí, že většina z analyzovaných zpráv (58 – 68 %) dokumentuje provedení autoevaluační ve zpřesněné a konkretizované rovině podoblastí. Z dlouhodobějšího hlediska není v tomto směru patrný žádný trend; analýza založená na datu provedení neindikuje ani zlepšování, ani zhoršení v podílu škol, které si podoblasti stanovily.

Graf č. 13 Míra podrobnosti jednotlivých částí autoevaluační zprávy



V neposlední řadě je nutné zdůraznit, že volba konkrétních podoblastí nevychází ze skutečné situace v dané škole a nemusí tak vždy plně zohledňovat její specifické potřeby. Obyčejně jsou totiž jednotlivé podoblasti zvoleny na základě metodických doporučení a vzorových autoevaluačních zpráv, které jsou jednotlivým školám k dispozici. I přesto je však specifikace podoblastí významným signálem snahy o utilizaci výsledků prováděného vlastního hodnocení.

C. Evaluační přístup

Stanovení konkrétních podoblastí a provedení vlastního hodnocení na úrovni každé jednotlivé podoblasti má ještě další, neméně důležitý význam. Vymezení konkrétních podoblastí totiž usnadňuje stanovení jednotlivých indikátorů hodnocení. Každá dílčí podoblast tedy vede ke specifikaci konkrétních snadno měřitelných kritérií autoevaluace a potenciálně může vést také ke stanovení měřitelných standardů, jichž má být v posuzovaném období dosaženo. Ponechme nyní stranou, ve kterých konkrétních kritériích školy dosahují či nedosahují stanovených standardů, a obraťme pozornost k samotnému faktu jejich definice.

Graf č. 14 Specifikace kritérií vlastního hodnocení školy



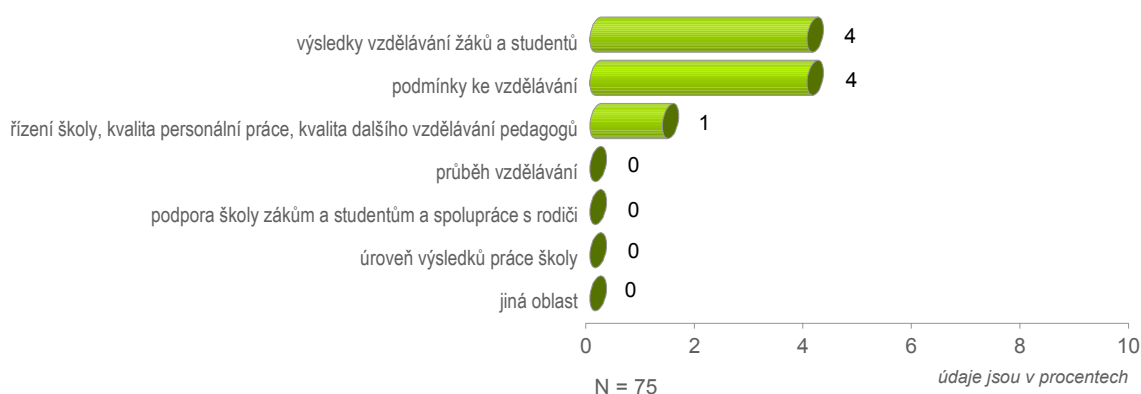
Obecně se z výsledků provedené analýzy ukazuje, že využívání vlastních kritérií při autoevaluaci není žádnou samozřejmostí. Pokud ponecháme stranou poněkud specifickou oblast výsledků vzdělávání žáků a studentů⁷¹, pak platí, že konkrétní kritéria si ve zbývajících oblastech stanovila přibližně jedna čtvrtina až jedna pětina škol v závislosti na konkrétní oblasti. Takový výsledek je z hlediska rozvoje evaluačních schopností a praktického

⁷¹ Tato oblast je specifická, neboť kritéria jsou zde dána z podstaty řady prováděných testů (zejm. Scio, Cermat).

provádění evaluace zcela zásadní. Nejde jen o samotný fakt absence kritérií, která jsou nezbytným prvkem evaluace, vlastní hodnocení nevyjímaje. Zjištěný stav totiž neumožňuje formalizované provedení evaluace a jednotlivé činnosti prováděné v tomto ohledu nemohou získat systematický, kontrolovaný charakter a jsou tak do značné míry otevřeny subjektivismu. V tomto ohledu nepřekvapí jeden ze závěrů České školní inspekce, která prováděla výzkumné šetření týkající se způsobu provádění vlastního hodnocení škol formou rozhovorů s představiteli vybraných škol⁷². V souhrnné zprávě je totiž charakterizována zjištěná praxe, kdy se „v mnohých školách vlastní hodnocení opíralo jen o pocity či názory ředitele vycházející z jeho bezprostřední zkušenosti, nikoli o exaktní zjišťování a vyhodnocování jevů v konkrétní oblasti prostřednictvím hodnotících nástrojů ...“ (Česká školní inspekce 2006). Expanze subjektivismu v autoevaluačních zprávách je přitom zřejmým důsledkem chybějícího evaluačního rámce založeného na specifikaci kritérií.⁷³

Pokud jde o stanovení standardů, jejichž dosažení by vlastní hodnocení mělo osvědčovat, je třeba konstatovat, že v analyzovaných autoevaluačních zprávách jsou patrné jen jakési náznaky jejich použití. Standardy autoevaluace jsou definovány jen u některých z vyhláškou stanovených oblastí (a to konkrétně u výsledků vzdělávání žáků a studentů, u podmínek ke vzdělávání a u řízení školy, kvality personální práce a kvality dalšího vzdělávání pedagogů) a navíc jen ve velmi malém počtu autoevaluačních zpráv (4 %, resp. 1 %).

Graf č. 15 Vytyčení standardů vlastního hodnocení školy



⁷² Celkem bylo zkoumáno 1.117 škol různých stupňů, od mateřských až po střední.

⁷³ Předpoklad, že subjektivistické pojetí autoevaluačního procesu je výsledkem nedostatku empirických dat či nerozvinutých schopností empirická data využívat se nenaplnil; v 96 % analyzovaných autoevaluačních zpráv totiž byla empirická data uváděna a interpretována.

Situaci dokresluje další ze zjištění, z něhož je patrné, že jen 9 % autoevaluačních zpráv obsahuje definici cílů, které provedení vlastního hodnocení sledovalo; naprosté většině autoevaluací tak chybí pevný rámec či plán provedení.

O uvedené skutečnosti svědčí také další ze sledovaných indikátorů. Při analýze autoevaluačních zpráv bylo zjišťováno, zdali je v závěrečné zprávě zmínka o aplikaci některého ze standardizovaných nástrojů vlastního hodnocení. Specifická pozornost byla v této souvislosti věnována metodě CAF (Common Assessment Framework), která byla popsána v několika metodických příručkách⁷⁴, takže zde existuje předpoklad snadného nalezení potřebných informačních pramenů o možnostech jejího použití. Analogická situace se týká druhé ze sledovaných metodik, a to modelu excelence vytvořeného Evropskou nadací pro management kvality (EFQM).⁷⁵

Ze získaných údajů je zřejmé, že model EFQM nebyl zmíněn v žádné z analyzovaných autoevaluačních zpráv a model CAF byl zmíněn v jednom jediném případě. Máme za to, že uvedený poznatek je třeba vnímat v kontextu ostatních zjištění; samotné nevyužívání uvedených modelů by nemuselo nutně znamenat nějaký problém či nedostatek. Jejich použitelnost v konkrétních podmínkách dané školy může být problematická, zkušenosti s jejich využitím nemusí být hluboké a adekvátní, a jistě existuje i řada dalších argumentů, proč v daném případě tyto metodické pomůcky nepoužít. Nicméně v kontextu neexistence cílů vlastního hodnocení, které je patrné přibližně v devíti zprávách z deseti, a v situaci, kdy 96 % zpráv neobsahuje definici standardů hodnocení u některého z širokého spektra kritérií, je absence ucelené metodiky vlastního hodnocení dalším důvodem proč by vlastní hodnocení škol mohlo být považováno za pseudoevaluaci (Stufflebeam, Shinkfield 2007).

D. Exaktnost

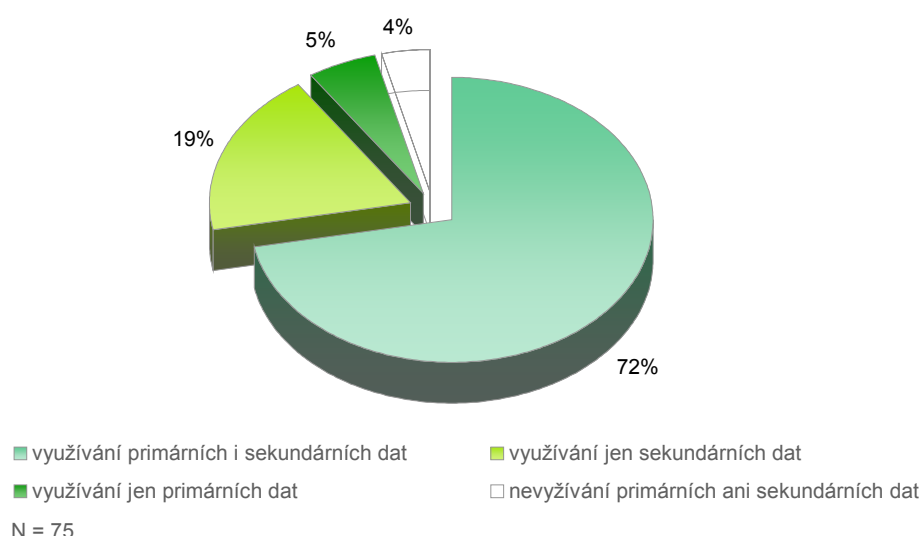
Zajímavé výsledky přinesla analýza zaměřená na používané zdroje informací a dat obsažených v autoevaluačních zprávách, resp. využívaných v procesu vlastního hodnocení. V případě externě prováděných evaluačních šetření je běžné, že se jejich realizátoři opírají o kombinaci sekundárních dat a výsledků primárních šetření. V jaké míře se však pracuje s empirickými údaji v interně prováděných evaluačních šetřeních, když výchozí zadání je

⁷⁴ Viz např. Michek 2005 nebo www.npj.cz.

⁷⁵ Dostupné zdroje jsou např. www.csg.cz; www.efqm.org.

velmi obecné a nedefinuje žádné konkrétní metodické požadavky? Téměř tři čtvrtiny (72 %) analyzovaných zpráv se opírají jak o sekundární data, tak také o výsledky primárních výzkumů prováděných v souvislosti s procesem vlastního hodnocení. Primární výzkumy jsou obvykle realizovány s cílem ověřit spokojenost žáků a rodičů s některými aspekty fungování školy. V případě těchto šetření stojí za zmínku snaha identifikovat jednotlivé zainteresované strany a zvolit takový výzkumný design, který by zohlednil specifika jednotlivých cílových skupin.

Graf č. 16 Využívání sekundárních dat a výsledků primárních výzkumů při autoevaluaci



Dále je patrné, že v přibližně jedné pětině autoevaluačních zpráv jsou sekundárně využívána dostupná data a dalších 5 % zpráv dokumentuje, že vlastní hodnocení školy provázal primární výzkum, který sloužil jako výlučný zdroj dat pro evaluaci vlastní činnosti. Naprostá většina autoevaluačních zpráv tedy dokumentuje jednotlivé skutečnosti s využitím empirických údajů, kterými školy buď již disponují, nebo které získaly v přímé souvislosti s prováděním vlastního hodnocení.

Mezi nejčastěji využívanými sekundárními zdroji dat figurují výsledky hodnocení žáků a jejich výkonů; jednotlivé školy dokládají úspěšnost svých vzdělávacích strategií na základě standardizovaného testování znalostí a výkonů žáků. Data pocházející z primárních výzkumů jsou ve velké většině údaji z výzkumů školního klimatu nebo z výzkumů hodnocení kvality výuky ze strany žáků a rodičů.

6.2.4 Diskuze výsledků

Shora uvedená zjištění dokumentující současnou praxi provádění vlastního hodnocení v mnoha ohledech korespondují s již zmíněnými výsledky výzkumného šetření České školní inspekce a některých dalších pramenů. Přesto je třeba poukázat na několik okolností, které je třeba vzít v potaz při posuzování validity prezentovaných závěrů:

Řada škol mohla provést důkladnou autoevaluaci, nicméně úroveň závěrečného výstupu o tom nemusí svědčit (a vice versa), a náš předpoklad kongruence mezi procesem a výstupem tak nemusí platit ve všech případech. Hlubková analýza aktuální praxe provádění vlastního hodnocení by vyžadovala šetření *in situ* (resp. provedení terénního šetření s kompetentními pracovníky jednotlivých škol). Zdá se tedy smysluplné oddělovat přípravu autoevaluačních zpráv od samotného procesu autoevaluace. Řada zpráv totiž vyvolává pochyby, zda byla skutečně připravena na základě řádně uskutečněného autoevaluačního procesu; oprávněnost takových pochyb přitom podporují i mnohá zjištění České školní inspekce (2006).

Skutečná situace v celém základním souboru (tj. ve všech školách v ČR) může být méně příznivá, než jak dokládají naše údaje. Je třeba mít na paměti, že v námi analyzovaném souboru autoevaluačních zpráv jsou jen výstupy, které byly zveřejněny. To může vyvolávat obavu, že při provedení stejné analýzy na úplném souboru (všech) autoevaluačních zpráv by mohly být v řadě ukazatelů hodnoty sledovaných parametrů horší.⁷⁶

6.2.5 Dílčí závěry

Autoevaluace lze využít k prokazování dosažených zlepšení, nicméně validní zdroj informací věcném základu a o přínosu ho lze považovat jen v omezené míře. Pomocí autoevaluací nelze odpovědět na otázky o dopadech; autoevaluace mohou být využity ke zlepšování vlastní činnosti a způsobu provádění jednotlivých procesů (formativní evaluace). Z hlediska obsahu lze vlastní hodnocení využívat při posuzování vlastního výkonu (na základě parametrů definovaných zadavatelem takové autoevaluace).

⁷⁶ Je třeba upozornit na zjištění České školní inspekce (2006), že přibližně jedna třetina škol žádnou autoevaluaci neprovádí, a to navzdory zákonem specifikované povinnosti.

Výsledky uvedené v této části poukazují na jeden z možných průvodních jevů, který se může vyskytnout v případě, že je evaluace (či autoevaluace jako v tomto případě) nařízena zákonem či jinak vyžadována externí autoritou. Existuje nebezpečí, že se povinné subjekty zaměří na plnění formálních náležitostí stanovených požadavků, ale příležitosti, které jim vlastní hodnocení poskytuje, nevyužijí.

Námi provedená analýza neumožňuje formulovat závěr, zdali je neadekvátní zpětná vazba odrazem nízkého povědomí o způsobu provedení vlastního hodnocení, či zdali je výsledkem obav kompetentních pracovníků škol z veřejného přiznání vlastních chyb. Obě uvedená zdůvodnění však mají úzkou souvislost s evaluační kulturou: první zdůvodnění vypovídá o nedostatečném rozvoji evaluačních schopností; zatímco druhé zdůvodnění vypovídá o neadekvátní percepci autoevaluací (a to jak samotnými autory zpráv, tak také jejich sociálním okolím)⁷⁷. Ať tak či onak, neschopnost či neochota provést reflexi vlastní činnosti znemožňuje budoucí zlepšování či případnou nápravu, což výrazně limituje možnost využití výsledků prováděného vlastního hodnocení.

⁷⁷ V tomto ohledu je poněkud zarážející fakt, že samotné autoevaluační zprávy jsou v takovém rozsahu veřejně přístupné.

6.3 Hybridní forma provádění evaluačních šetření

Případová studie, na níž je doložena konkrétní forma uplatnění evaluace v podobě hybridního evaluačního šetření, je z oblasti školství; dokumentuje, jakým konkrétním způsobem je využívána evaluace při posuzování bakalářských prací.

6.3.1 Popis případové studie

Konkrétní případ zahrnuje analýzu posudků konzultantů a oponentů bakalářských prací odevzdaných na jedné předem zvolené katedře v letech 2008 (letní semestr) a 2009 (zimní a letní semestr). Analyzovaný soubor zahrnuje všechny dostupné posudky k bakalářským pracím, jež byly v daném období odevzdány – celkem tak bylo zpracováno 148 posudků, přičemž 62 posudků je datováno k roku 2008 a zbývajících 86 pochází z roku 2009.

Pravidla zavedená na dané katedře vytvářejí dobré předpoklady pro aplikaci evaluace v této konkrétní oblasti, neboť jsou základní nároky kladené na bakalářské práce jasně stanoveny. Kromě toho existuje přesný výčet jednotlivých elementů, které musí každá bakalářská práce obsahovat. Tyto nároky jsou spolu s dalšími pravidly všem studentům známy, neboť jsou umístěny na veřejné části WWW stránek dané katedry. Z tohoto výčtu je tedy patrné, že bakalářská práce musí obsahovat⁷⁸:

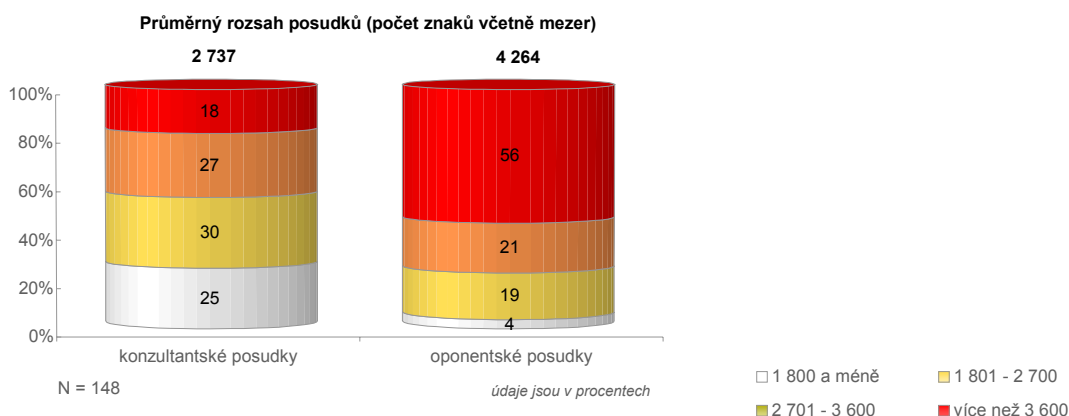
1. rozbor poznávacího problému na základě relevantní literatury,
2. formulaci cílů práce a výzkumné otázky,
3. popis použité metodiky,
4. argumentaci tvrzení, jež jsou v dané bakalářské práci předkládána,
5. závěry, ke kterým autor/ka dospěl/a.

Kromě těchto podmínek se dále požaduje minimální rozsah prací a splnění několika dalších formálních požadavků (podpis dohody o podmínkách užití, počet exemplářů atd.).

⁷⁸ Výčet je podobně jako později uvedený seznam součástí bakalářských prací doslovně převzat z WWW stránek dané katedry; bližší citace či odkaz zde není uveden, neboť chceme ponechat daný případ v anonymitě.

opponentskými posudky představují takto stručné výstupy 4 %; rovněž je patrné, že nadpoloviční většina opponentských posudků je delší než dvě normostrany, zatímco podíl konzultantských posudků tohoto rozsahu je přibližně třikrát nižší (činí 18 %).

Graf č. 18 Srovnání rozsahu posudků konzultantů a oponentů



6.3.2 Výsledky

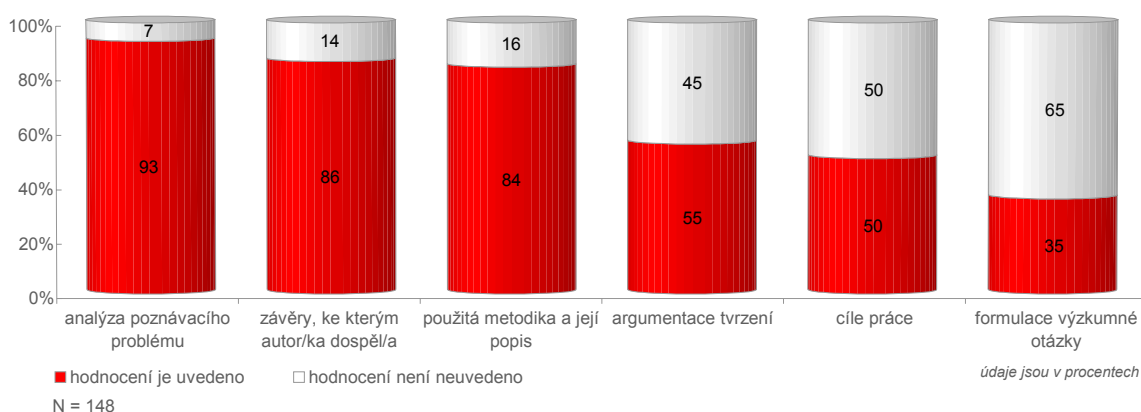
A. Orientace na cíl

Prvořadá pozornost se při analýze posudků zaměřila na zjišťování, zdali a v jaké míře je v analyzovaných posudcích skutečně zohledňováno naplnění obecných zásad přípravy bakalářských prací.

První z dovedností, kterou mají studenti svou bakalářskou prací prokázat, je schopnost jasně formulovat poznávací problém. Z výsledků provedené analýzy je patrné, že 93 % posudků tomuto kritériu skutečně věnuje pozornost⁷⁹. Pro ostatních 7 % posudků platí, že se jedná o texty kratší než jedna normostrana (mezi nejkratšími posudky se schopnosti studentů jasně formulovat poznávací problém vyhýbá 31 % posudků), a dále také o posudky k bakalářským pracím, jež mají charakter rozboru sociálního problému (v tomto případě není věnována pozornost formulaci poznávacího problému v 26 % posudků).

⁷⁹ V této souvislosti neobracíme pozornost ani k rozsahu, jaký je formulaci poznávacího problému v práci vyčleněn, a ani se zde nezabýváme otázkou výsledku hodnocení (tj. v jaké části posudků je konkrétní způsob formulace poznávacího problému hodnocen pozitivně, negativně či neutrálně)

Graf č. 19 Komplex zpětné vazby



Celkem 14 % posudků neposkytuje žádnou zpětnou vazbu, která by se týkala závěrů, ke kterým autor/ka bakalářské práce dospěl/a. V této souvislosti je třeba upozornit na skutečnost, že tento prvek absentuje významně častěji v posudcích konzultantů než oponentů: zatímco k závěrům bakalářské práce se nevyjadřuje 6 % oponentských posudků, mezi konzultantskými posudky je to 23 %. Podobně vysoký podíl posudků (16 %) zcela opomíjí zhodnotit použitou metodiku a její popis; platí přitom, že hodnocení této dovednosti nechybí v žádném posudku, v němž je hodnocena empirická práce založená na výzkumu.

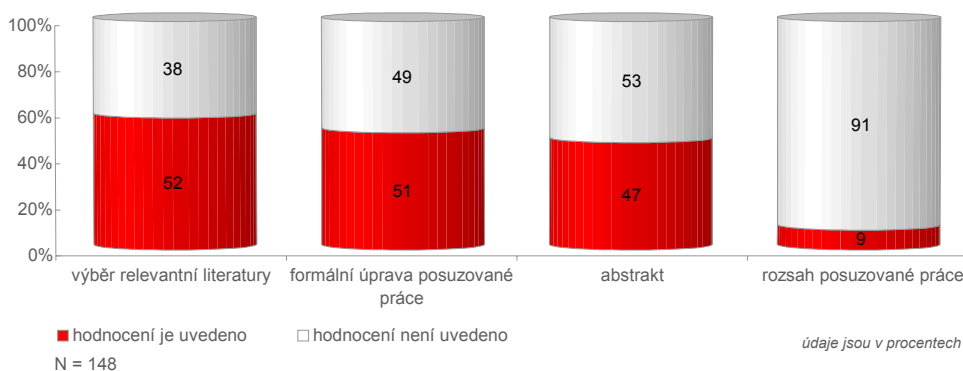
Dále stojí za pozornost, že polovina analyzovaných posudků se vůbec nevyjadřuje ke způsobu, jakým jsou v bakalářských pracích formulovány cíle a 65 % posudků v této souvislosti opomíjí poskytnout zpětnou vazbu týkající se formulace výzkumné otázky. Tento jev se vyskytuje zejména mezi posudky, které hodnotí bakalářské práce, jež mají charakter teoretické stati (kdy cíle takovýchto prací nejsou posouzeny ve třech čtvrtinách posudků). V této souvislosti stojí dále za zmínku, že v roce 2008 činil podíl posudků, jež se nevyjadřovaly k cílům práce 60 %, zatímco v roce 2008 to bylo 43 %.

Z uvedeného je tedy patrné, že významná část posudků, a to jak konzultantských, tak také oponentských, nekorresponduje se strukturou kritérií, podle nichž je vyžadováno, aby byly bakalářské práce připraveny. K podobným závěrům vede také analýza ostatních parametrů bakalářských prací; platí, že 91 % posudků se vůbec nevyjadřuje k celkovému rozsahu bakalářských prací⁸⁰; formální úpravy si všimá přibližně polovina analyzovaných posudků. Z konkrétních dat je v tomto případě patrný významný rozdíl mezi konzultantskými a

⁸⁰ Lze se domnívat, že toto kritérium je považováno za natolik samozřejmé, že je mu věnována jen okrajová pozornost.

opponentskými posudky, kdy otázky formální podoby bakalářské práce ponechává beze zmínky 61 % konzultantských posudků oproti 38 % posudků oponentských.

Graf č. 20 Komplement zpětné vazby



Výběr relevantní literatury představuje náplň přibližně tří pětiny (62 %) posudků; ostatní ponechávají tento aspekt bakalářských prací bez povšimnutí. Diference mezi konzultantskými a oponentskými posudky není v tomto ohledu významná, nicméně odlišnou skupinu posudků představují v tomto ohledu ty, které hodnotí empirické práce založené na výzkumu, tři pětiny takovýchto posudků (60 %) totiž nevěnují otázkám výběru relevantní literatury žádnou pozornost.

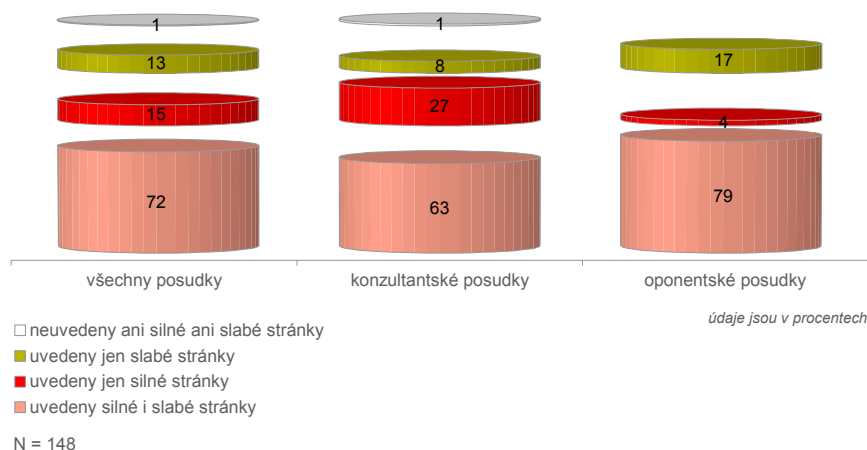
Zde uvedená absence hodnocení způsobu zpracování klíčových zadávacích podmínek může narušit zpětnovazební smysl a poslání posudků samotných a může ve svém důsledku přispět k rozvolnění pevných vazeb mezi vstupem (zadáním) a výstupem (finální podobou práce). V důsledku toho je využitelnost posudků jen selektivně omezená na některé aspekty předkládaných bakalářských prací a posudky tak neposkytují ucelenou evaluaci.

B. Symetrie expertního posuzování

Z provedené analýzy jednotlivých posudků je dále patrná relativně vysoká vyváženost silných a slabých stránek, na které dané posudky poukazují. Z konkrétních údajů je patrné, že přibližně sedm z deseti posudků obsahuje jak silné, tak také slabé stránky posuzovaných bakalářských prací. Vyšší podíl takovýchto posudků je možné vztáhnout k bakalářským pracím předkládaných v roce 2008, kdy bylo takto vyvážených posudků 77 %, naproti tomu

v roce 2009 činil jejich podíl 67 %. Nejvyšší podíl posudků s vyváženým poukazem na silné a slabé stránky je možné najít mezi rozsáhlejšími posudky, a dále mezi posudky, které hodnotí empirické práce založené na výzkumu, kde jejich podíl dosahuje 90 %.

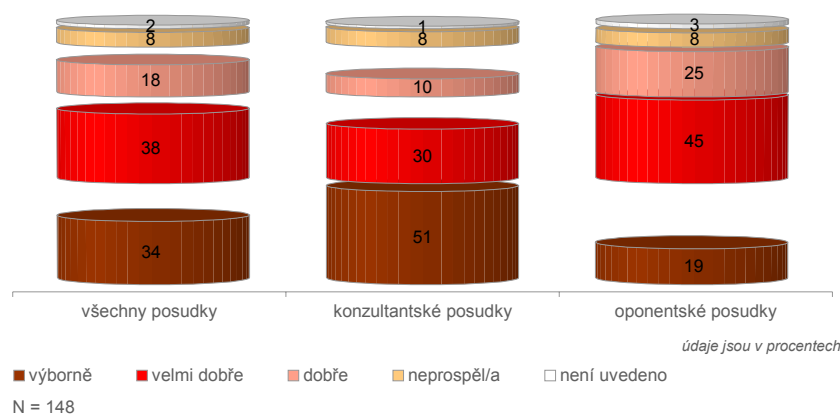
Graf č. 21 Silné a slabé stránky posuzovaných bakalářských prací



Ostatní posudky se dělí na dvě přibližně stejně velké skupiny, kdy první zahrnuje posudky vyzdvihující pouze silné stránky předkládaných prací (15 %), zatímco druhá skupina pak obsahuje posudky poukazující na výlučně slabé stránky (13 %). V tomto ohledu je patrný významný rozdíl mezi posudky konzultantů a oponentů, kdy oponentské posudky vykazují větší míru vyváženosti, neboť 79 % z nich obsahuje zmínku jak o silných, tak o slabých stránkách dané bakalářské práce. Naproti tomu mezi posudky konzultantů je podíl takto vyvážených posudků 63 %. Dále platí, že 27 % konzultantských posudků vyzdvihuje silné stránky bakalářských prací, ale nespecifikuje žádné slabiny; v případě oponentských posudků činí podíl takovýchto textů 4 %. Také je z dat zřejmé, že kratší posudky obsahují silné stránky ve větší míře než posudky delší, a to v poměru 23 : 4 %.

Z této části je patrné, že posuzovatelé zaujímají v podmínkách chybějící anonymity ve většině případů vyvážená stanoviska. Bylo by však třeba provést další, mnohem podrobnější analýzu, která by přinesla detailní informace o postojové vyváženosti jednotlivých posudků a případně ukázala na konkrétní oblasti inbalance v provedeném hodnocení.

Graf č. 22 Klasifikace bakalářské práce



C. Doporučení

Další z témat, jemuž byla v rámci analýzy posudků bakalářských prací věnována pozornost, se týká specifikace návrhů na případná zlepšení v jednotlivých oblastech. Z výše uvedených poznatků poukazujících na skutečnost, že mnoha oblastem analyzované posudky nevěnují žádnou pozornost je zřejmé, že v daných dimenzích nemohou tedy poskytovat ani případná doporučení. Následující údaje se tedy budou týkat vždy jen té podmnožiny posudků, která věnuje dané oblasti pozornost.

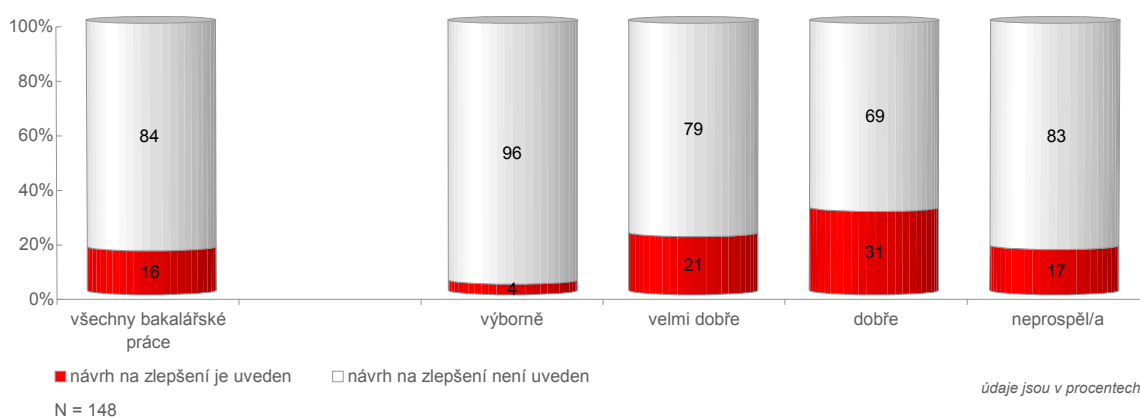
Obecně platí, že konkrétní náměty na zlepšení jsou patrné u posudků, které hodnotí teoretické statě; v tomto ohledu jsou patrné konkrétní náměty týkající se způsobu citací, zlepšení jazyka, stylistiky a pravopisu bakalářské práce, způsobu argumentace tvrzení a výběru relevantní literatury. Ve všech těchto ohledech přinášejí posudky k teoretickým statím konkrétní náměty na zlepšení v největší míře v porovnání s posudky ke všem ostatním typům bakalářských prací.

Dále je patrná souvislost specifikace konkrétních návrhů na zlepšení s celkovým rozsahem posudků, kdy delší posudky obsahují případné náměty na zlepšení ve větší míře než posudky kratší. Bez povšimnutí nelze ponechat diferenci mezi posudky konzultantů a oponentů, kdy oponentské posudky přinášejí konkrétní náměty ve významně větší míře než posudky konzultantů. Tento jev lze pozorovat konkrétně v případě hodnocení takových oblastí, jakými jsou návrh na zlepšení použité metodiky (11 : 32 %), návrh na zlepšení způsobu argumentace tvrzení (6 : 19 %); výběr relevantní literatury (1 : 20 %), zlepšení způsobu citací (1 : 19 %).

O významné formativní funkci posudků, jež spočívá v poskytnutí zpětné vazby využitelné při případném zlepšování dané práce (či spíše přístupu autora/autorky k přípravě práce srovnatelného významu), svědčí další ze zjištění. Ukazuje se totiž, že v případě šesti klíčových kritérií, která tvoří páteř posuzování jednotlivých bakalářských prací se zvyšuje podíl posudků, které poskytují návrhy na zlepšení dané práce ruku v ruce se zhoršující se klasifikací.

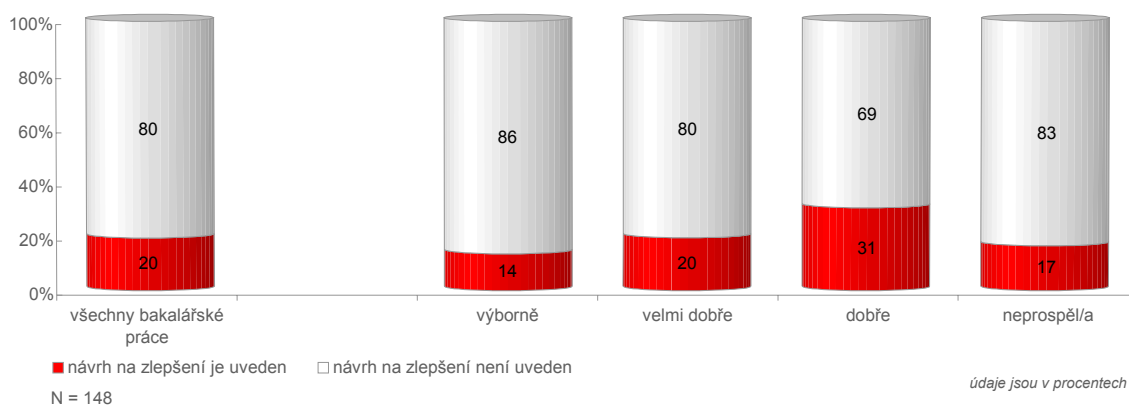
V případě analýzy poznávacího problému je patrné, že při „výborném“ hodnocení jsou autorovi/autorce nabízeny konkrétní návrhy na zlepšení v celkem 4 % posudků. Mezi bakalářskými pracemi, které jsou klasifikovány jako „velmi dobré“, činí podíl posudků s návrhem na zlepšení 21 % a u bakalářských prací hodnocených stupněm „dobře“ činí podíl posudků poskytujících návrh na vylepšení práce 31 %. Naznačené tendenci se vymykají práce, které daný posuzovatel považoval za nepřijatelné. V tomto případě není podíl posudků navrhujících konkrétní zlepšení ve způsobu analýzy poznávacího problému nejvyšší, nýbrž činí 17 %.

Graf č. 23 Návrh na zlepšení způsobu analýzy poznávacího problému



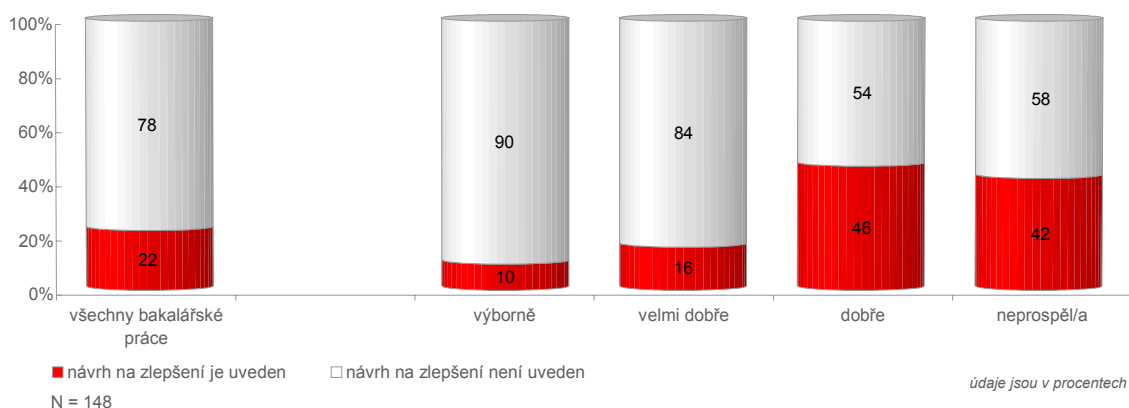
Podobný závěr vyplývá z analýzy navrhovaných zlepšení v druhé klíčové oblasti, kterou jsou závěry, k nimž autor/autorka dospěl/a. Tak jako v předešlém případě, také zde je patrné zvyšování podílu posudků navrhujících zlepšení v závislosti na stupni klasifikace.

Graf č. 24 Návrh na zlepšení závěrů, ke kterým autor/ka dospěl/a



U dalšího ze základních parametrů, popisu použité metodiky, je zřejmý rozdíl mezi posudky, jejichž autoři hodnotili bakalářské práce stupněm „výborně“ či „velmi dobře“ oproti posudkům s horší klasifikací. V případě druhé skupiny bakalářských prací jsou konkrétní návrhy na zlepšení uvedeny ve 46 % (dobře) resp. 42 % (neprospěl/a) posudků, zatímco v případě lépe hodnocených bakalářských prací je to 10 % (výborně) resp. 16 % (velmi dobře).

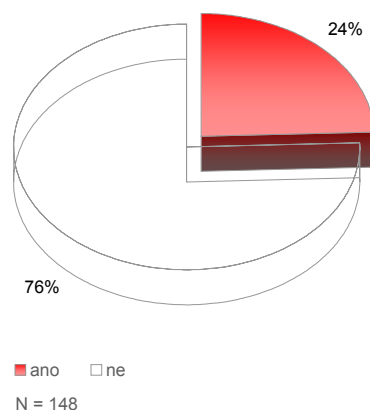
Graf č. 25 Návrh na zlepšení použité metodiky a jejího popisu



Obecně platí, že zdrojem doporučení jsou ve větší míře posudky oponentů, kteří v rámci všech šesti oblastí uvádějí návrhy na zlepšení ve větší míře než konzultantů.

Zvláštním typem doporučení je specifikace případných témat, která by mohla být diskutována při obhajobě práce. V tomto ohledu z provedené analýzy vyplývá, že v přibližně jednom posudku ze čtyř je podobný námět uveden. Významně odlišné jsou v tomto ohledu posudky delší než čtyři normostrany, v rámci kterých je specifikace námětu pro obhajobu uvedena ve dvou třetinách.

Graf č. 26 **Náměty na obhajobu**



Na základě získaných údajů lze tedy formulovat hypotézu, že posuzovatelé věnují pozornost těm oblastem (kritériím), ke kterým mají nějaký konkrétní návrh na zlepšení či nějakou věcnou připomínku, přičemž ostatní oblasti (kritéria) zcela opomíjejí.

6.3.3 Diskuse výsledků

Jistě mohou existovat i jiné případy expertního posuzování, které zohledňují základní evaluační pravidla ve větší míře, než v jaké jsou zohledněny ve zvolené oblasti.

Pravděpodobně hodnocení projektových žádostí o podporu z Fondů EU podávané v rámci některého z operačních programů by mohl posloužit jako příklad hybridní evaluace s lepšími výstupními parametry; podobně také peer review proces v některých vědeckých časopisech je v současné době upraven tak, tak aby zohledňoval základní evaluační principy.

Pro naši případovou studii však záměrně nebyla zvolena ani oblast evaluace projektů, ani peer review proces používaný některou z redakcí odborných časopisů. Evaluace projektů předkládaných v rámci výběrových řízení nebyla zvolena kvůli vysoké pravděpodobnosti možného zkreslení relevantní dokumentace, která provází realizaci jednotlivých fází výběrových řízení. Některá posouzení mohou být účelově vychýlena, což je okolnost, která by mohla mít dopad na charakter výsledků naší analýzy. Druhý důvod je praktického rázu a týká se dostupnosti jednotlivých dokumentů; některá výběrová řízení totiž mohou mít několikakolový průběh a získání všech potřebných údajů z jednotlivých fází takového procesu je nedosažitelnou metou. V neposlední řadě bylo naší snahou najít případovou studii v takové oblasti, v rámci které by byl předmětem evaluačního šetření skutečný výsledek, nikoliv jen projekt či nabídka (tj. příslib výsledku).

Případovou studii jsme nezvolili ani v oblasti peer review kvůli obvyklé vágnosti kritérií posuzování, jež by mohla zatemnit cíl, který je prezentací této případové studie anticipován.

Vzhledem ke skutečnosti, že předmětem posuzování jsou posudky, nikoliv však samotné bakalářské práce, nelze s určitostí rozhodnout, zdali chybějící oblasti hodnocení jednotlivých posudků jsou dány opomenutím ze strany posuzovatelů (konzultantů a oponentů) nebo zdali absentují i v samotných bakalářských pracích a hodnotitelé se tak vlastně nemají k čemu vyjádřit. Z uvedeného nástinu je dále patrné, že předmětem analýzy jsou výlučně jen posudky, nikoliv však samotné bakalářské práce. Součástí provedené (a zde prezentované) analýzy tedy nebyla ani adekvátnost navržené klasifikace vzhledem k věcnému základu dané bakalářské práce, ani shoda mezi konzultantskými a oponentskými posudky⁸¹, a dokonce ani diference v klasifikaci mezi jednotlivými hodnotiteli⁸².

Dále je třeba upozornit na problematičnost stanovení typu bakalářské práce – na základě analyzovaných posudků bylo možné toto jasně stanovit u přibližně poloviny posudků. Zapojení samotných bakalářských prací do procesu hodnocení by však velmi pomohlo při posilování reliability této proměnné.

Bylo by vhodné provést v některých případech metaevaluaci zohledňující nejen konzultantské a oponentské posudky, ale i proces, kterým bylo dosaženo bakalářské práce, postup výběru posuzovatelů (tedy oponentů) a proces hodnocení práce.

6.3.4 Dílčí závěry

Stávající podoba provádění evaluačních šetření hybridní formou ukazuje, že v některých případech založených na expertním posuzování nejsou evaluační šetření striktně vedena akceptačními kritérii. V dané případové studii je patrné, že autoři mnohých bakalářských prací nedostávají adekvátní zpětnou vazbu a že posuzovatelé se v mnoha případech zaměřují na jiná kritéria, než jsou ta, která byla zadána studentům. V praxi se pak posuzuje něco jiného, než co bylo zadáno. Z uvedených dat je patrné, že posudky obecně nemají pevnou strukturu a

⁸¹ Ačkoliv podobné úkoly bývají poměrně časté (viz např. Armstrong 1997; Cole, Dubin, Cole 1977, Garfield 1986; Marsh, Ball 1981)

⁸² Každá z těchto úloh je však na základě získaného datového materiálu proveditelná; necht' je tedy tento výčet považován za inspiraci pro další analýzy prováděné v rámci tohoto tématu

jen 10 % z nich se vyjadřuje ke všem požadovaným charakteristikám, které jsou od bakalářských prací očekávány.

Nedostatek přitom nespočívá v samotné posuzovací činnosti, ale v metodice hodnocení. Její absence znamená přesun odpovědnosti za úplné a správné provedení posouzení na jednotlivé posuzovatele, což může vést k imperfkcím při zpracování, neboť každý z posuzovatelů se zaměřuje na jiné aspekty posuzované bakalářské práce. Autoři posudků tak nutně volí jakýsi heuristický přístup při psaní svých posudků, kdy posuzují každou práci z její samotné podstaty. Kritéria pro posouzení tak formují individuálně na základě každé jednotlivé bakalářské práce.

Výsledek případové studie lze interpretovat také tak, že při chybějícím metodickém vedení mohou provedená evaluační šetření splnit značnou část očekávání (daný systém expertního posuzování funguje bez významnějších změn již řadu let), a ačkoliv neposkytují takovou přidanou hodnotu, jakou by mohly, svému účelu poslouží.

Zajímavou extenzí prezentované případové studie je zjištění, zdali autoři bakalářských prací pokračují ve stejném tématu i v diplomové práci. Tato informace by pomohla odpovědět na otázku, zdali je smysluplné, aby posudky obsahovaly náměty na dopracování, zlepšení či uváděly tipy pro další výzkum. Za situace, kdy existuje vysoký podíl studentů pokračujících ve stejném či alespoň velmi podobném nebo v navazujícím tématu, pak je pochopitelně specifikace takových námětů cenným příspěvkem, kterým mohou externí posuzovatelé významně posílit přidanou hodnotu, kterou autoři bakalářských prací v posudcích získávají.

6.4 Souhrn poznatků případových studií

Tři prezentované případové studie dokumentují současný stav ve vybraných oblastech využívajících evaluaci. Přestože možnosti zobecnění výše uvedených závěrů mohou být omezené a zcela jistě neodrážejí celou šíři a pestrost sociální reality (což ostatně nelze od případových studií ano očekávat), přesto vypovídají přesvědčivým způsobem o několika skutečnostech, které jsou relevantní ve vztahu k teoretickým poznatkům jež jsou v této práci analyzovány. Předně platí, že:

(1) formalizovaný způsob realizace evaluací má významná specifika oproti intuitivně prováděnému hodnocení. To s sebou nese potřebu tyto formalizované principy evaluací poznávat, hledat konkrétní způsoby formy jejich adekvátního uplatnění v aplikační praxi a také volit vhodné prostředky k jejich šíření. Konkrétně je z případných studií patrná nezbytnost evaluace všech, kteří formalizovaný způsob evaluace používají či provádějí.

(2) k uplatňování obecných evaluačních principů nedochází samovolně či spontánně. Evaluace je součástí komplexnějších procesů, a tak také její zavádění musí být stimulováno „shora“. Povinnost provádět evaluaci přitom není sama o sobě dostatečnou incentivou, neboť v rámci systémů vyžadujících konkrétní formu evaluační činnosti, může být takto zakotvená povinnost vykonávána jednotlivými subjekty pouze formálně, bez toho, že by zavedení evaluačního systému mohlo vést ke skutečným a žádoucím dopadům.

(3) korektní způsob provádění evaluace vyžaduje soustavnou zpětnou vazbu, která umožňuje jednotlivým evaluátorům lépe využívat principy evaluací ve prospěch zainteresovaných stran. Případové studie ukazují, že absence této zpětné vazby vede k nedokonalému provádění evaluačních šetření.

(4) konkrétní způsob uplatnění evaluace v odlišných podmínkách ukazuje na exaktní základ, o který se jednotlivá evaluační šetření opírají. Bez ohledu na charakter a míru diferencí jednotlivých kontextů, v nichž byl sledován způsob provádění evaluačních šetření je exaktní charakter evaluačních záměrů jasně zřetelný.

(5) specifickým typem užitečnosti či zvyšování přidané hodnoty evaluačních šetření je specifikace doporučení, jež se vine jak červená nit všemi sledovanými aplikačními oblastmi.

7. Realizace evaluačních šetření v praxi

Tato kapitola přináší některé z poznatků, které vyplynuly ze série hloubkových interview vedených se zástupci zadavatelů i realizátorů evaluačních šetření⁸³. Nejprve je věnována pozornost zadavatelům evaluačních šetření (straně poptávky), kdy je kromě základních popisných parametrů uvedena i obecná charakteristika této zainteresované strany.

Analogickým způsobem je zpracována informace o realizátorech evaluačních šetření, tedy o straně nabídky. Sedmou kapitolu uzavírá třetí část, která je dále strukturována do samostatných podkapitol zabývajících se konkrétními způsoby realizace evaluačních šetření, na které poukazují jak zadavatelé, tak také realizátoři.

7.1 Strana poptávky

V obecných souvislostech již byla strana poptávky v této práci mnohokrát zmíněna. Hlavními zadavateli evaluačních šetření v ČR jsou řídicí orgány jednotlivých operačních programů. V souladu s celkovým zaměřením naší práce, které je ve větším detailu specifikováno v kapitole páté, a s ohledem k oblasti, v níž byly prováděny i případové studie založené na analýze závěrečných zpráv z evaluačních šetření prováděných v souvislosti s administrací operačních fondů EU, omezíme se i na tomto místě pouze na tu část poptávky, která je s operačními fondy spjata.

Strana poptávky, reprezentující zadavatele evaluačních šetření, však nezahrnuje pouze řídicí orgány jednotlivých operačních programů. V případě některých typů evaluací (zejména ex-post) vystupuje v roli poptávky přímo Komise, která je zodpovědná za provedení tohoto typu evaluací. Dále je možné na straně poptávky identifikovat další orgány, které se podílejí na implementaci strukturální a kohezní politiky EU; patří sem zejména Národní orgán pro koordinaci, Evaluační jednotka a do určité míry i zprostředkující subjekty. Pokud bychom

⁸³ V úvodu je třeba připomenout, že vztahy mezi zadavateli (stranou poptávky) a realizátory (stranou nabídky) jsou determinovány stávající podobou obecně závazných předpisů. V tomto ohledu jsou vztahy upravovány především zákonem o veřejných zakázkách. Zadavatelé specifikují své požadavky v zadávací dokumentaci, jež tvoří přílohu k vyhlášeným výběrovým řízením; realizátoři připravují své nabídky, které jsou hodnoceny zadavateli na základě příslušného zákona. S vítězným uchazečem je za obvyklých podmínek uzavřena smlouva o realizaci evaluačního šetření.

stranu poptávky chápali nikoliv jako zadavatele evaluačních šetření, ale spíše jako uživatele výsledků evaluačních šetření, pak bychom na straně poptávky nemohli opomenout veřejnost. Následující tabulky přinášejí přehled klíčových zadavatelů evaluačních šetření:

Tabulka č. 10 Tematické Operační programy 2007 – 2013

Název OP	Řídící orgán	Charakteristika	Finanční prostředky
Integrovaný operační program	Ministerstvo pro místní rozvoj ČR	řešení regionálních problémů v oblastech infrastruktury pro veřejnou správu, veřejné služby a územní rozvoj	celkem 1, 58 mld. €
OP Podnikání a inovace	Ministerstvo průmyslu a obchodu ČR.	podpora rozvoje podnikatelského prostředí a podporu přenosu výsledků výzkumu a vývoje do podnikatelské praxe.	z fondů EU 3,04 mld. € , z českých veřejných zdrojů 0,54 mld. €.
OP Životní prostředí	Ministerstvo životního prostředí ČR.	zlepšování kvality životního prostředí a tím i zdraví obyvatelstva	z fondů EU 4,92 mld. € , z českých veřejných zdrojů 0,87 mld. €.
OP Doprava	Ministerstvo dopravy ČR	zkvalitnění infrastruktury a vzájemné propojení železniční, silniční a říční dopravy v rámci tzv. transevropských dopravních sítí (TEN-T)	z fondů EU 5,77 mld. € , z českých veřejných zdrojů 1,01 mld. €.
OP Vzdělávání pro konkurenceschopnost	Ministerstvo školství, mládeže a tělovýchovy ČR	zaměřený na zkvalitnění a modernizaci systémů počátečního, terciárního a dalšího vzdělávání, jejich propojení do komplexního systému celoživotního učení a zlepšení podmínek ve výzkumu a vývoji	z fondů EU 1,83 mld. € , z českých veřejných zdrojů 0,32 mld. €.
OP Výzkum a vývoj pro inovace	Ministerstvo školství, mládeže a tělovýchovy ČR	posilování výzkumného, vývojového a proinovačního potenciálu ČR	z fondů EU 2,07 mld. € , z českých veřejných zdrojů 0 mld. €.
OP Lidské zdroje a zaměstnanost	Ministerstvo práce a sociálních věcí ČR	snížování nezaměstnanosti prostřednictvím aktivní politiky na trhu práce, profesního vzdělávání, dále na začleňování sociálně vyloučených obyvatel zpět do společnosti, zvyšování kvality veřejné správy a mezinárodní spolupráci v uvedených oblastech	z fondů EU 1,84 mld. € , z českých veřejných zdrojů 0,32 mld. €.
OP Technická pomoc	Ministerstvo pro místní rozvoj ČR	podpora jednotného přístupu na národní úrovni pro zajištění aktivit efektivního řízení, kontroly, sledování a vyhodnocování realizace Národního strategického referenčního rámce (NSRR)	celkem 0,29 mld. € .

Zdroj: <http://www.strukturalni-fondy.cz/Programy-2007-2013>

Tabulka č. 11 Regionální Operační programy 2007 – 2013

Název OP	Řídící orgán	Charakteristika	Finanční prostředky
OP Praha Konkurenceschopnost	hl. m. Praha, Magistrát hl. m. Prahy	podpora investičních projektů v Praze	z fondů EU 234,94 mil. € , z českých 41,46 mil. €.
OP Praha Adaptabilita	hl. m. Praha, Magistrát hl. m. Prahy	podpora neinvestičních projektů (vzdělávání, sociální integrace, podpora zaměstnanosti a rozvoj lidských zdrojů ve výzkumu a vývoji)	z fondů EU 108,39 mil. € , z českých 19,13 mil. €.
ROP NUTS II Severozápad	Regionální rada regionu soudržnosti Severozápad.	Regionální operační programy se zaměřují na zlepšení dopravní dostupnosti a propojení daného regionu vč. modernizace prostředků veřejné dopravy, podporu rozvoje infrastruktury i služeb cestovního ruchu, přípravu menších podnikatelských ploch a zlepšování podmínek k životu v obcích a na venkově především prostřednictvím zkvalitnění vzdělávací, sociální a zdravotnické infrastruktury, odstraňování ekologických zátěží.	z fondů EU 745,91 mil. € , z českých 131,63 mil. €.
ROP NUTS II Severovýchod	Regionální rada regionu soudržnosti Severovýchod.		z fondů EU 656,46 mil. € , z českých 115,85 mil. €.
ROP NUTS II Střední Čechy	Regionální rada regionu soudržnosti Střední Čechy.		z fondů EU 559,08 mil. € , z českých 98,66 mil. €.
ROP NUTS II Jihozápad	Regionální rada regionu soudržnosti Jihozápad.		z fondů EU 619,65 mil. € , z českých 109,35 mil. €.
ROP NUTS II Jihovýchod	Regionální rada regionu soudržnosti Jihovýchod.		z fondů EU 704,45 mil. € , z českých 124,31 mil. €.
ROP NUTS II Moravskoslezsko	Regionální rada regionu soudržnosti Moravskoslezsko.		z fondů EU 716,09 mil. € , z českých 126,37 mil. €.
ROP NUTS II Střední Morava	Regionální rada regionu soudržnosti Střední Morava.		z fondů EU 657,39 mil. € , z českých 116,01 mil. €.

Zdroj: <http://www.strukturalni-fondy.cz/Programy-2007-2013>

Uvedený přehled může vyvolávat dojem těsného sepětí řídicích orgánů jednotlivých operačních programů s jednotlivými ministerstvy, nicméně konkrétní zkušenosti jednotlivých aktérů takovýto dojem nepotvrzují:

„Tam je častokrát skoro až rivalita mezi lidmi z operačního programu a z ministerstva; jsou to jako takové skupiny samy pro sebe.“ (R04)

Z hlediska dalšího rozvoje evaluačních šetření a z hlediska jejich možné diseminace je významné, do jaké míry se vyskytuje vzájemná interakce mezi současnými zadavateli evaluačních šetření se zadavateli potenciálními, tedy do jaké míry mohou subjekty, jež mají vlastní zkušenost s evaluačními šetřeními a jejich závěry možnost interagovat s ostatními subjekty, které by alespoň potenciálně mohli benefitovat z provádění evaluačních šetření.

„Ty evaluace maj ještě jeden efekt. Ta státní správa se díky tomu začíná učit spolu mluvit, komunikovat, domlouvat se na věcech, který musej dělat společně. Takže to přispívá ke zlepšení toho celkového institucionálního rámce tady u nás.“ (R07)

O smysluplnosti takovýchto procesů svědčí takové případy, jakými jsou zejména Irsko (McNamara, O'Hara, Sullivan 2009; Boyle 2005), Itálie (Leeuw 2002) či Španělsko (Bustelo 2006) a dokonce také Švýcarsko (Bussmann 2008). V těchto zemích se evaluační šetření, podobně jako v současnosti v České republice, začala realizovat v souvislosti s implementací strukturální a kohezní politiky EU, nicméně v současné době jsou evaluační šetření prováděna poměrně běžně i pro celou řadu ryze národních programů a politik.

„Tady je velký prostor pro osvětu zadavatelů, že je v jejich zájmu, aby evaluací procházely i ty národní a regionální programy.“ (R01)

„Ta problematika evaluací by neměla být spojována pouze s těmi Strukturálními fondy, což u nás je. Když se řekne evaluace, tak každý si k tomu vybaví jen tohle. Není nejmenší důvod, aby dotační tituly, které nejsou financovány z EU, nepodléhaly stejným pravidlům evaluace.“ (R06)

„Evaluace má u nás obrovský význam. Všechny ty naše struktury se totiž soustředí na tu administrativně-byrokratickou stránku, ale vůbec nesledují ty konce, unikají jim dopady. U nás se zajímáme o to, kolik kilometrů silnic se postaví, možná sem tam někoho napadne podívat se na to, kolik tam vlastně jede aut. Ale správně bysme se měli ptát, jak jsou spokojení lidi v těch autech. A tohle nás ta evaluace učí – všimnout si dopadů.“ (R02)

7.2 Strana nabídky

Adekvátní a nezkreslenou rekognoskaci strany nabídky lze považovat za samostatný výzkumný úkol s vlastním významem a přínosem, neboť fakticky neexistuje žádné sdružení⁸⁴ či asociace, která by sdružovala všechny či alespoň většinu relevantních subjektů. Navíc, jak bude patrné z provedené analýzy, povaha a zaměření jednotlivých subjektů je poměrně pestrá (společnosti se odlišují svým předmětem činnosti). Smyslem činností prováděných v rámci explorace je tedy identifikovat jednotlivé subjekty, které se prováděním evaluační činnosti systematicky zabývají.

Tabulka č. 12 Seznam realizátorů evaluačních šetření

Název subjektu	Hlavní předmět činnosti	Specifikace evaluační činnosti	Počet zaměstnanců
Akses, spol. s r.o	monitorování a evaluace projektů financovaných EU, řízení projektového cyklu	evaluace projektů PHARE, operační programy EU	1 – 5
ALVIT - INOVACE A VZDĚLÁVÁNÍ s.r.o.	problematika rovných příležitostí, semináře, vzdělávání	Evaluace přínosu pro cílové skupiny	1 – 5
AMITY Enterprises	Pomoc různorodým subjektům získat finanční prostředky z EU pro jejich projekty.	systematické zjišťování přínosu projektů (otázky dopadu, relevantnosti, udržitelnosti, efektivity a další využitelnosti projektu)	1 – 5
BERMAN GROUP	odborná technická asistence a konzultační služby v oblasti místního a regionálního ekonomického rozvoje	není k dispozici	neuv.
Centrum pro komunitní práci střední Morava (CpKP)	Regionální politika EU, regionální rozvoj ČR, podpora a vzdělávání neziskových organizací	Evaluace regionálních operačních programů	neuv.
Cross Czech	asistence při prosazování a následné realizaci projektu, sledování dotačních titulů a vypsaných programů v rámci EU a státních programů ČR	hodnocení a monitoring projektů podporujících regionální rozvoj	10 – 19
DAP Services	nabídka specializovaných služeb a nástrojů, které pomáhají zjistit a pochopit individuální a skupinové postoje	vnitřní evaluace škol	6 – 9
DHV CR, spol. s r.o.	multidisciplinární poradenství, realizace projektů v oblastech životního prostředí, infrastruktury, prostorového plánování a rozvoje lidských zdrojů	evaluace programů a projektů v oblasti implementace fondů EU v ČR, makroekonomické a sektorové analýzy a studie	50 – 99
Ernst & Young s. r. o.	odborné služby v oblasti finančního poradenství a auditu	Evaluace regionálních operačních programů a oblasti životního prostředí	250 – 499
EuroManagers	poradenské služby spojené s přípravou projektů financovaných programy a fondy EU	není k dispozici	10 – 19
EuroProfis	příprava, řízení a hodnocení projektů v oblasti strukturálních fondů	ex-ante, midterm, ex-post evaluace	6 – 9
EU-servis.cz	konzultační, koučinková a expertní činnost v oblasti realizace veřejných výdajových programů a projektů	evaluace programů spolufinancovaných z evropských fondů	6 – 9
Evasco	Evaluace a Assessment, studie a analýzy, školení a semináře, konzultační činnost	evaluace projektů v rámci implementace fondů EU	1 – 5
GLE o. p. s.	evaluace projektů a programů zaměřených na sociální integraci, řízení programů evropských strukturálních fondů	není k dispozici	1 – 5

pokrač.

⁸⁴ V ČR působí Česká evaluační společnost, která je začleněna v síti Evropských evaluačních společností. Aktuálně však sdružuje jen zlomek realizátorů evaluačních šetření.

Název subjektu	Hlavní předmět činnosti	Specifikace evaluační činnosti	Počet zaměstnanců
Cheval, s.r.o.	poradenství a vzdělávání v oblasti řízení firemních inovací, pomoc při získávání finančních prostředků z fondů EU	není k dispozici	1 – 5
Horizonty, o.s.	podpora profesionalizace neziskových organizací (formou školení a seminářů)	není k dispozici	1 – 5
IREAS centrum, s.r.o.	zpracování nezávislých odborných analýz a evaluací, vzdělávací a publikační činnost.	zpracování nezávislých odborných analýz a evaluací	neuv.
Jacobs Consultancy	konceptní a projektová příprava dopravních staveb	Evaluace programů rozvoje dopravy	6 – 9
Klíč o.s.	problematika školství	testování žáků, evaluace, autoevaluace škol	neuv.
Metodické a evaluační centrum, o.p.s.	metodická a odborná činnost v oblasti hodnocení škol, poradenství, konzultace	hodnocení škol	10 – 19
Národní monitorovací středisko pro drogy a drogové závislosti	monitoring situace v oblasti psychotropních látek	hodnocení preventivních programů, drogové péče a léčby	neuv.
Naviga 4	dotiční programy, komunikační strategie, evaluace a analýzy, vzdělávání	evaluace strukturálních fondů EU, vyhodnocení efektivity komunikačních aktivit	1 – 5
Navreme Boheme	konzultační služby, evaluace, výzkum, trénink	evaluace v rámci evropských projektů, veřejné správy a samosprávy, vzdělávacího systému a rozvoje lidských zdrojů	1 – 5
NTI-consulting	poradenství v oblasti HR, poradenství v nezaměstnanosti, vzdělávání	není k dispozici	1 – 5
Prev-Centrum o.s.	realizace primární a sekundární prevence a poskytování psychosociální pomoci osobám ohroženým závislostmi a dalšími společensky nežádoucími jevy	není k dispozici	10 – 19
PricewaterhouseCoopers Audit s. r. o.	poradenství, audit, daně, řízení rizik	není k dispozici	250 – 499
ProGrant	poradenství v oblasti čerpání finančních prostředků ze strukturálních fondů EU, zpracování rozvojových strategií a demografické studie	čerpání ze strukturálních fondů EU	1 – 5
Projekt Odyssea	realizace seminářů osobnostně sociálního rozvoje pro žáky a učitele	autoevaluace v osobnostně sociální výchově v rámci ŠVP	1 – 5
PVM Projekt s.r.o.	poskytování komplexních služeb v oblasti tvorby programů a projektů financovaných, resp. spolufinancovaných ze zdrojů EU	zhodnocení dosažených výstupů a výsledků, zpracování průběžných a závěrečných zpráv	6 – 9
REDECo s. r. o.	Strategické plánování a programování	Předběžné, průběžné a následné hodnocení, hodnocení žádostí o podporu	neuv.
RegioPartner	komplexní poradenské služby v oblasti socio-ekonomického rozvoje území	evaluace programů a procesů v jakémkoliv časovém okamžiku procesu/projektu	1 – 5
SANANIM, s.r.o.	služby v oblasti péče a léčby nealkoholových závislostí	evaluace výsledků léčby v terapeutické komunitě	10 – 19
Scotia s.r.o.	poradenství, lidské zdroje, diagnostika	evaluace zaměstnanců	1 – 5
Sociotrendy	výzkumné projekty, konzultace a dílčí výzkumné činnosti	není k dispozici	neuv.
SPF Group	příprava a realizace regionální politiky EU v ČR (technická asistence při tvorbě a řízení operačních programů, evaluační studie, příprava projektů)	není k dispozici	neuv.
Společně k bezpečí o. s.	řešení komunálních problémů	není k dispozici	1 – 5
Tima Liberec	příprava projektů a dotace, průzkumy veřejného mínění, evaluace, analýzy a strategie	evaluace programů a projektů	1 – 5
VeryVision	Projektové poradenství a dotační management, Analýzy, studie, akční a evaluační výzkum	vyhodnocení sociálních a jiných intervenčních programů	1 – 5

Zdroj: autor

Jednou z klíčových charakteristik, kterými lze vystihnout současnou stranu nabídky, je izolovanost jednotlivých evaluačních šetření a velmi omezené možnosti akumulace zkušeností a rozvoje evaluačních schopností. Jednotliví realizátoři totiž kromě samotné

evaluační činnosti vyvíjejí řadu jiných aktivit a podíl evaluační činnosti se podle odhadu kompetentních zástupců jednotlivých subjektů pohybuje přibližně mezi 5 – 20 %. Rovněž je ze získaných poznatků patrné, že v České republice doposud nepůsobí žádné pracoviště výlučně zaměřené na provádění evaluačních šetření. Ani v tomto ohledu se však daná situace příliš nevymyká zkušenostem z jiných zemí, kde bylo nebo stále ještě je nutné vytvářet evaluační kapacitu od základu (Mihalache 2009)

S ohledem k vysoké míře fragmentace strany nabídky, která se projevuje relativně vysokým počtem realizátorů (ve vztahu k počtu a k objemu zakázek), nelze prozatím hovořit o kumulaci zkušeností a poznatků⁸⁵. Podle Bieniase a Lewandawské (2009) bylo v České republice od roku 2004 do současnosti provedeno celkem 44 nařízených evaluačních šetření (tj. ex-ante, mid-term a on-going), což je počet, který přibližně odpovídá celkovému počtu realizátorů. Na základě vlastních údajů můžeme odhadnout, že počet tematických evaluací činí ve stejném období dalších cca 50 – 60 šetření, nicméně i tak je počet realizátorů ve vztahu k celkové velikosti a kapacitě tohoto odvětví příliš vysoký. Situaci dokládá i zjištění z obsahové analýzy vzorku závěrečných zpráv, kdy realizátoři uskutečnili jen výjimečně více než jedno evaluační šetření. To má pochopitelně dopady na odborné zázemí jednotlivých realizátorů:

„Ono je to pro ně těžký, protože oni v tom nepracují. My tomu sice rozumíme nejvíc, ale tu evaluaci si sami udělat nemůžeme. Všechno jim říct nejde a oni se to všechno z těch našich dokumentů nenaučí, no.“ (Z05)

„Evaluátoři neznají všechny ty souvislosti, a pak jsou ty výsledky hrozně obecný.“ (Z02)

⁸⁵ O případných úsporách z rozsahu ani nemluvě.

7.3 Způsob provádění evaluačních šetření

Získané informace o způsobu provádění evaluačních šetření jsou prezentovány formou tří zásadních dilemat, která se opakovaně vyskytovala v několika interview, a to jak na straně zadavatelů, tak také realizátorů. První dilema vypovídá o akcentu buď na obsahovou stránku prováděných evaluačních šetření nebo na metodologické aspekty. Druhé dilema zachycuje střet role administrátora operačního programu a zadavatele konkrétního typu evaluačních šetření. Konečně třetí z dilemat se obrací k samotné podstatě evaluačních šetření a všímá si hranice mezi evaluací a pseudoevaluací.

7.3.1 Dilema metody a výsledku

Současný přístup zadavatelů (strany poptávky), lze charakterizovat důrazem, který přikládají výsledku prováděného (zadaného) evaluačního šetření. Tento akcent není ostatně v kontextu evaluačních šetření ničím překvapivým; už v kapitole čtvrté, kde jsou specifikovány klíčové vlastnosti evaluačních šetření, jsme upozornili na důraz přikládaný konečnému výsledku, a to zejména ve srovnání se základním výzkumem. Praktické projevy této charakteristiky, s níž je možné se setkat i v každodenní aplikační praxi v České republice, však stojí za zvláštní pozornost.

Důraz na výsledek, který mimo jiné zatlačuje do pozadí metodologické aspekty provádění evaluačních šetření, je patrný už ze samotné zadávací dokumentace k mnohým evaluačním šetřením. Tyto dokumenty totiž jen zcela výjimečně specifikují metody, které by měly být pro řešení zadávaných úloh použity. Absence metodologicky relevantních kritérií poskytuje realizátorům evaluačních šetření na jedné straně značnou volnost, pokud jde o využití konkrétních metodologických postupů, otevírá tím prostor pro uplatnění jejich vlastní představy o vhodném designu daného evaluačního šetření a v konečném důsledku je faktorem stimuluje soutěž mezi jednotlivými realizátory. Na druhé straně však může představovat značné riziko, a to hned dvojího charakteru. První vyplývá z volby nevhodného přístupu k provedení evaluačního šetření a z výběru neadekvátních metod. Jak bylo patrné už v kapitole 4.1 této práce, může odlišně zvolený přístup vést k jiným závěrům a k následnému posunu v možnostech další práce s výsledky provedeného evaluačního šetření. Volba designu

evaluačního šetření však předpokládá jednak praktické zkušenosti s používáním jednotlivých technik a postupů, což lze v podmínkách České republiky naplnit jen částečně s ohledem ke skutečnostem uvedeným v kapitole 7.2. Dále předpokládá povědomí o jednotlivých metodách a technikách, a to jak na straně zadavatelů, tak také na straně realizátorů. Takovýto předpoklad není zcela samozřejmý, jak plyne z několika rozhovorů:

„Experimentální design? Hm, to by nešlo. My se musíme řídit hlavně zadáním ... a odpovídat na evaluační otázky ... na nějaké experimentování my nemáme čas.“ (R01)

„Logic models? Víte, v těchto zakázkách a výstupech evaluací se člověk nesmí dostat do moc filozofické roviny, protože to ten zadavatel nepotřebuje. On potřebuje znát ty konkrétní doporučení, závěry a tak. My sami jsme častokrát museli předělávat zprávu, protože jsme tu naši metodu popisovali moc podrobně.“ (R07)

„My si zas tak moc vymýšlet nemůžeme. Jednak na to není čas, a pak taky zadavatelé po nás chtějí to, co je v těch bruselských příručkách. ... No nevím, jak by se tvářili, kdybychom přišli s něčím úplně jiným.“ (R03)

Přesná specifikace jednotlivých výzkumných nástrojů není pochopitelně v rámci ToR možná, a navíc by byla i velmi kontraproduktivní, neboť by eliminovala významný prvek inovačního procesu. Nicméně specifikace alespoň minimálních standardů či požadavků kladených na způsob získání dat, na jejich validitu a reliabilitu by, zdá se, mohla předejít případným nedorozuměním týkajících se očekávané kvality dat.

„Chybí tady poučení zadavatelé. Oni nerozumí tomu, jak se ty evaluace vlastně provádějí. Oni vůbec nemají představu o tom, jaké náklady generuje takové provedení jedné fokusní skupiny.“ (R05)

Druhý typ rizika je dán nastavením konkrétních kritérií vyhodnocení nabídek realizátorů, jež jsou v mnoha případech definovány tak, že cena má významně vyšší váhu, než ostatní parametry navrhovaného evaluačního šetření. To může některé realizátory motivovat k navrhování takového designu evaluačních šetření, který v maximální možné míře spoléhá na již existující data získaná od zadavatele či dalších zdrojů.

„ ... tenhle systém vede k tomu, že vyhrávají nabídky, které vůbec nepočítají s žádnými výzkumy; celá ta jejich evaluace pak stojí jen na pár číslech ze Stat'áku. “ (R03)⁸⁶

V tomto ohledu je vhodné připomenout skutečně používané metody kvalitativních a kvantitativních výzkumů prováděných v rámci evaluačních šetření, jejichž závěrečné zprávy byly analyzovány v kapitole šesté. Provedený desk research prozatím neprokázal, že by realizátoři evaluací eliminovali provádění primárních výzkumů, nicméně s postupně se snižujícími rozpočty alokovanými na evaluační šetření a jistou benevolencí zadavatelů nelze tuto možnost zcela vyloučit. Jak k této situaci uvádějí sami realizátoři:

„Když těm úředníkům je jedno, jak k tomu dojdeme. Oni od nás potřebují nějaký závěr, potřebují, aby byl alespoň trochu nějak doložený a hlavně, abychom pod tím byli podepsaní. A mám pocit, že všechno ostatní jde tak trochu mimo ně. Prostě to neřeší. “ (R04)

To ostatně dokládají i některá interview provedená se samotnými zadavateli evaluačních šetření. Ukazuje se, že při výběru realizátora se navrženým metodám nevěnuje ve všech případech adekvátní pozornost:

„ ... pokud odpoví na zadané evaluační otázky, tak nám je jedno, jak k tomu chtějí dojít. “ (Z01)

„ ... no, my do těch jejich metod stejně nevidíme, ale když to tak navrhují a jsou to profesionálové, tak přeci vědí proč. “ (Z06)

Zdá se tedy, že z pohledu zadavatelů a realizátorů existují dva zásadní důvody, které vedou k současnému stavu: zaprvé je to obecně nízké povědomí odborné veřejnosti o metodologii evaluačních šetření. Tento závěr ostatně vyplývá i z řady dalších šetření prováděných mezi zadavateli a realizátory evaluačních šetření (Bienias, Lewandowska 2009; UNEP 2006, ESTEP 2007). Zadruhé je to konkrétní nastavení parametrů výběrových řízení, kdy v některých případech nekoresponduje stanovená váha výběrových kritérií s faktickým významem daného parametru a jeho potenciálním dopadem na výsledek poptávané zakázky.

⁸⁶ Podobné výhrady se opakovaly také v několika dalších interview; uvedená citace je v tomto ohledu zastupuje.

V této souvislosti považujeme za nutné zmínit některé implikace vyplývající se současného uspořádání.

a) vzniklá situace tak, poněkud paradoxně, může vést k tomu, že zadavatelé budou zapojeni do procesu rozhodování o volbě designu evaluačního šetření více, než pokud by základní parametry metod byly specifikovány už v zadávací dokumentaci. Větší spektrum navrhovaných řešení a přístupů nutně vyvolá dodatečné nároky na jejich posouzení v rámci jednotlivých výběrových řízení.

b) kombinace vysoké váhy ceny (v řadě konkrétních případů je to 60 – 70 %) při rozhodování o výběru a absence minimálních metodologických standardů může vést k volbě suboptimální varianty či nabídky. V konkrétních situacích pak může docházet ke srovnávání neporovnatelných variant.

c) apel zadavatelů na erudovanost a etickou zodpovědnost by dosahoval významně vyšší efektivity v podmínkách akreditovaného výkonu evaluačnické činnosti či v podmínkách jakéhosi povinného členství v profesní komoře. Za stávající situace však není účinný.

7.3.2 Dilema administrátora programu a zadavatele evaluačních šetření

Současnou praxi při provádění evaluačních šetření doprovází v některých případech rolový konflikt na straně zadavatelů. V této souvislosti je třeba připomenout, že typickými zadavateli evaluačních šetření prováděných v rámci evaluace operačních programů jsou samotné řídicí orgány, tedy subjekty, jež jsou primárně zodpovědné za administraci daného operačního programu. Situace, kdy zadavatelem evaluačních šetření je administrátor programu, jsou sice v praxi poměrně běžné a samy o sobě ještě případný rolový konflikt nezakládají, nicméně významným činitelem v takovýchto situacích je typ prováděného evaluačního šetření a konkrétní způsoby použití jeho výsledků. V případě formativních evaluací, jejichž cílem je zlepšování daného programu (ať už operačního či intervenčního), je zadání ze strany administrátora obvykle vedeno jeho snahou o další rozvoj programu a vychází z potřeby získat zpětnou vazbu k doposud prováděným činnostem. Inicivace takového evaluačního šetření ze strany administrátora programu je tedy v zásadě přirozenou součástí jeho činnosti. V daném případě se předpokládá, že výsledky provedeného evaluačního šetření slouží administrátorovi, který na jejich základě provádí jednotlivé korekce své vlastní činnosti nebo

procesů, jež jsou v rámci daného programu vykonávány. Z již uvedeného rozboru jednotlivých typů prováděných evaluací je patrné, že v případě ex-ante evaluací se sumativní prvek téměř nevyskytuje nebo je natolik okrajový, že neutváří charakter těchto šetření.

Ovšem v situaci, kdy administrátor daného programu vystupuje v roli zadavatele sumativního evaluačního šetření, jehož cílem je posoudit efektivitu daného programu, je riziko konfliktu obou rolí, resp. konfliktu zájmů vysoké. Tato situace panuje při on-going evaluacích, jež se podle nařízení Rady (ES) provádějí v souvislosti s monitorováním operačních programů, zejména pokud toto monitorování odhalí významný odklon od původně stanovených cílů nebo pokud jsou podány návrhy na revizi operačního programu. Výsledky on-going evaluací tak přímo či nepřímo vypovídají o činnosti řídicího orgánu operačního programu v uplynulém období, a on-going evaluace tak svým způsobem hodnotí jeho činnost.

„Setkáváme se i tím, že ti zadavatelé nechtějí slyšet, že jsou některé věci špatně nastavené, že nejsou dobré, že procesy neprobíhají tak, jak mají. A ten zadavatel, když je tam finální zpráva, která o tom mluví, tak o tom nechce slyšet, dohaduje se s námi, že to tak není. Přestože to z těch fokusních skupin jasně zní. A je to případ hlavně těch formálních evaluací, tam je největší zájem ze strany zadavatelů na tom, aby ta evaluace řekla, že je všechno plus minus v pořádku.“ (R04)

Sumativní prvek obsažený v tomto typu evaluace podtrhuje fakt, že závěry on-going evaluací jsou předkládány monitorovacímu výboru a následně také Komisi. Řídící orgány jsou za dané situace tedy jen jedním z několika uživatelů výsledků provedeného evaluačního šetření; superiorita ostatních uživatelů pak z řídicích orgánů činí zainteresovanou stranu se silným zájmem⁸⁷ na konkrétních výsledcích provedeného evaluačního šetření.

„... no, třeba zrovna posledně jsme tu zprávu museli předělávat snad pětkrát. Nakonec jsme v ní mohli nechat jenom takový ty obecný formulace o tom, že se všechno plní a že není žádný problém. Ostatní muselo jít z té zprávy pryč.“ ... „no, já je chápu, že to maj těžký, ale k čemu to pak je.“ (R07)

⁸⁷ Viz strana 48 této práce.

„Setkáváme se hlavně u on-going evaluací s tím, že se zadavatel snažil zmírňovat ty evaluační výroky.“ (R03)

Situaci dále navíc eskaluje požadavek na zveřejnění závěrečných zpráv z takovýchto evaluačních šetření:

„... my vůbec nesouhlasíme s tím, aby se ty zprávy musely dávat na web. I když se ta evaluace udělá dobře, i když je pro nás pozitivní, tak je v těch zprávách stejně strašná spousta nebezpečných věcí, který se daj snadno zneužít. Pak stačí, aby to dostal do ruky nákej nedovzdělanej pisálek a je oheň na střeše.“ (Z03)

„Oni nám, řekli, že máme pravdu s těmi nedostatky, že o nich taky vědí. Ale ono to bylo moc citlivý zrovna s tímhle tím něco dělat. Tak jsme jim to dali jenom na zvláštním papíru bokem, ale do zprávy, co se pak pověsila na net se to nedostalo.“ (R09)

„Když je ta zpráva veřejná, tak je snaha od toho zadavatele, aby nebyla kontroverzní. A evaluátoři jsou pak pod obrovským tlakem, aby ty zprávy nebyly kontroverzní.“ (R07)

Lze vyjádřit hypotézu, že díky těsné návaznosti on-going evaluací na monitoring byl zodpovědností za provádění této novinky pověřen právě řídicí orgán operačního programu. Aktuální praxe však ukazuje, že pokud by zadavatelem tohoto typu evaluačních šetření byl jiný subjekt, nemusel by podobný konflikt zájmů vznikat.

Podobná situace je navíc patrná i v jiných oblastech, v nichž se evaluační šetření provádějí. V tomto ohledu bychom rádi poukázali na analogickou situaci v oblasti vlastního hodnocení školy, kde jde o střet formativního poslání autoevaluační činnosti se sumativní utilizací ze strany České školní inspekce coby nadřízeného orgánu, který autoevaluační zprávy používá k externímu hodnocení. Na asymetrii, která tento typ evaluačních šetření provází jsme již poukázali výše (viz str. 129).

Pokud pomineme aspekt konfliktu rolí či zájmů jednotlivých zainteresovaných stran, pak lze na situaci týkající se adekvátního využívání výsledků z provedených evaluačních šetření pohlížet také optikou evaluační kultury a míry jejího rozvoje.

„Já si myslím, že je to taky o celkovém klimatu. Ta evaluace neslouží k tomu, aby někoho ponižovala nebo dehonestovala nebo tak něco, ale je to nástroj k tomu, jak se posunout dál, jak se zlepšit a tak. A když se v těch zprávách něco kritizuje, tak to není nějaká námitka ad hominem, ale je to podnět pro ten program. Problém je, že ty zadavatelé to berou všechno strašně osobně.“ (R08)

Uvedený případ je možné komentovat ještě z pozice realistického přístupu k provádění evaluačních šetření, který prosazují R. Pawson a N. Tilley. Jak je patrné z detailního popisu tohoto přístupu, které je uvedené v kapitole šesté, zásadním prvkem při evaluacích implementovaných programů sehrává kontext. Postup a pravidla, která jsou funkční a vedou přímo k požadovaným výstupům mohou fungovat jinak, pokud začnou být vykonávána v jiném kontextu. Z postoje zadavatelů je velmi dobře patrné, že v situaci, kdyby závěrečné zprávy neměly veřejný charakter, nemusel by být tlak na jejich pozitivní vyznění tak silný, jakým je v podmínkách nařízené publicity. Změna kontextu (tj. veřejný či neveřejný charakter výsledků evaluace) pak signifikantním způsobem ovlivňuje samotný mechanismus utilizace výsledků evaluačních šetření.

7.3.3 Dilema evaluace a pseudoevaluace

Poslední z významných dilemat se týká samotného věcného základu evaluačních šetření, která jsou v České republice aktuálně prováděná. Za pseudoevaluaci lze označit takový typ činnosti, který sice evaluaci připomíná, nicméně nenaplnuje některý z principiálních znaků (Stufflebeam 2007).

„Většina těch evaluací se nezadáva až tak proto, že by někdo měl pocit, že potřebuje ten program zlepšit nebo že by se chtěl dozvědět ten skutečný stav, ale zadávají se proto, že je to v podmínkách Evropské unie, že tento typ evaluací se musí provést. Mnohdy ty kompetentní nevědí, co s tím. Oni mají třeba zadáno, že musí provést nějakou evaluaci, tak ji zadaj, ale když jsou výsledky na stole, tak ani nevědí co s tím.“ (R08)

Důvody proč některé z prováděných šetření mají pseudoevaluační charakter mohou být v zásadě dvojího charakteru. První skupina důvodů je dána obranou zájmů některé zainteresované strany participující za zadání a dotčené výsledky provedeného evaluačního

šetření. Další z důvodů vycházejí z interakce provedeného evaluačního šetření a jeho výsledků s povahou rozhodovacích procesů, v jejichž rámci jsou dané evaluace prováděny (Armstrong, Wells, Woolford 2002). V rámci provedeného šetření tuto skutečnost potvrzují samotní aktéři:

„Tu evaluaci udělali opravdu moc dobře. Opravdu. Ale v podstatě to nemělo výsledek. Takový ty nesmysly, jako je zdvojený VaVpI⁸⁸ s OPP⁸⁹ – to všichni zpracovatelé, všichni evaluátoři říkali, že je to nesmysl ... rozdělení Doprava⁹⁰ a SFŽP⁹¹ – to samý. Všichni dopředu to říkali, všichni to věděli, že je to nesmysl, ale nic se nestalo. Vůbec nic. Politická úroveň si to prostě převálcovala a dneska to ty samý politici kritizují.“ (Z04)

Souvislost mezi klamným charakterem evaluace a nízkou utilizací je v tomto ohledu jasně patrná. Z dlouhodobého hlediska je třeba hledat takový rámec pro provádění evaluačních šetření, který dokáže podobným tlakům a tendencím čelit. Ani v tomto ohledu nelze hovořit o specifické charakteristice evaluačních šetření prováděných v České republice, neboť podobný přístup lze zaznamenat i v jiných zemích (Hansson 1997, Pollitt 1998).

⁸⁸ VaVpI = Operační program Věda a výzkum pro inovace

⁸⁹ OPP⁸⁹ = Operační program Podnikání a inovace

⁹⁰ Jde o Operační program Doprava

⁹¹ SFŽP = Státní fond životního prostředí

8. Závěry

Na základě výsledků provedených výzkumů a s podporou teoreticky zakotvených přístupů lze tedy identifikovat celkem pět oblastí, které vytvářejí současný charakter evaluačních šetření prováděných v České republice. Závěrečná část této práce je věnována jejich stručnému shrnutí⁹².

(1) mimořádný důraz kladený na doporučení vycházející z provedené evaluace

Tento rys je jasně patrný ve všech částech prováděného výzkumu. V každé ze tří oblastí, v nichž byly analyzovány závěrečné výstupy z provedených evaluačních šetření, i v hloubkových rozhovorech prováděných se zadavateli i s realizátory se vyskytuje mimořádná pozornost věnovaná doporučením. Realizátoři evaluačních šetření spatřují v poskytování doporučení jednu ze svých klíčových kompetencí a významnou případnou hodnotu, kterou mohou poskytnout. Analogicky i zadavatelé očekávají, že posouzení jejich minulé činnosti vyústí v nástin budoucích kroků, na které by se měli zaměřit. Ukazuje se tak, že doporučení jsou významným katalyzátorem utilizace evaluačních šetření a jejich výsledků.

Zjištěný závěr mj. obohacuje evaluační metodologii o další prvek, který odlišuje evaluační šetření od auditu a otevírá prostor k dalšímu rozvoji a precizaci postupů, které posilují prospektivní charakter evaluačních šetření.

(2) nedostatek zpětné vazby k provedeným evaluačním šetřením

Ukazuje se, že současnému způsobu provádění evaluačních šetření v ČR chybí systematicky vykonávaný proces získávání zpětné vazby ke své vlastní činnosti. V žádné ze sledovaných oblastí není prováděna metaevaluace. Za jistou výjimku lze považovat pouze oblast vlastního hodnocení škol, které je pod dohledem nadřízeného orgánu. Ani v tomto případě však nejsou metaevaluace prováděny systematicky se zřetelem k jejich formativní funkci umožňující budoucí zlepšování realizátorů evaluačních šetření. Zkušenosti z jiných zemí s rozvinutější evaluační kulturou přitom jasně poukazují na smysluplnost a užitečnost takovéto činnosti.

Ukázalo se, že námi aplikovaný postup při analýze závěrečných zpráv z evaluačních šetření umožňuje naplnit podstatnou část funkcí metaevaluace. Zdá se tedy, že dosažené výsledky mohou posloužit jako základ budoucí systematicky prováděné metaevaluace v jednotlivých oblastech.

⁹² Podrobné analýzy, jež jsou základem těchto závěrů lze najít v šesté a v sedmé kapitole.

(3) exaktní základ prováděných evaluačních šetření

Získané poznatky jasně dokumentují, že aktuálně uplatňovaný způsob provádění evaluačních šetření je v podstatné míře založen na exaktních údajích a přesně stanovených důkazech. Značná pozornost, která je v jednotlivých výstupech (i také v samotné činnosti mnohých evaluátorů) věnována získání, analýze a interpretaci empirických údajů, ať už číselných či jiných, je patrná napříč všemi oblastmi, na něž se soustředilo naše výzkumné úsilí. V tomto ohledu je třeba připomenout, že naprostá většina posuzovaných výstupů pocházela z evaluačních šetření, která neměla ryzí sumativní charakter, takže nutnost provádění konkrétních důkazů nebyla typem prováděných šetření *a priori* dána. Uvedené zjištění zakládá v kombinaci s již zmíněným důrazem na doporučení specifický charakter evaluačních závěrů.

(4) klamný charakter vynucených evaluačních šetření

Problém pseudoevaluací není specifikem, které by bylo omezeno výlučně jen na provádění evaluací v rámci zkoumaného prostředí. Rozsáhlá literatura, která se danou problematikou zabývá, je jasným důkazem o palčivosti takovýchto situací, kdy cíle realizovaných evaluačních šetření nesměřují k nestrannému a nevychýlenému posouzení věcného základu a přínosu. Na základě našich výzkumů je souvislost pseudoevaluací s konfliktem rolí či zájmů jasně patrná. Zdá se, že eliminace klamného charakteru evaluačních šetření je možná např. změnou v zapojení jednotlivých zainteresovaných stran. Na druhé straně nelze uvažovat o zmírnění požadavků na provádění evaluačních šetření, neboť význam systematicky budovaného prostředí stimulačního a poptávajícího výsledky evaluačních šetření je nezastupitelný.

(5) vyhýbavý přístup ke specifickým evaluačním technikám a nástrojům

Nevyužívání specifických metod, technik a nástrojů, které evaluační metodologie nabízí realizátorům jednotlivých šetření považujeme za průvodní jev relativně krátké historie provádění evaluačních šetření. Z analýzy situace v odlišných kontextech je zřejmé, že aplikační dovednosti realizátorů evaluačních šetření nejsou rozvinuty tolik, aby se staly samozřejmou součástí výsledků evaluačních šetření. DO té doby je třeba počítat s tím, že charakter evaluačních šetření bude silně determinován profesní skladbou evaluátorů. Význam našich zjištění v tomto ohledu spatřujeme v zachycení počátečního stavu, který umožní případná budoucí srovnání.

* * *

Přílohy

EVALUACE OPERAČNÍCH PROGRAMŮ SEZNAM SLEDOVANÝCH CHARAKTERISTIK

IDENTIFIKAČNÍ ÚDAJE

1) Jaký operační program?

- Lidské zdroje a zaměstnanost (LZZ)	1
- Rozvoj lidských zdrojů (RLZ)	2
- JPD 3	3
- EQUAL	4
- Ostatní	5

2) O jaký typ evaluace se jedná?

- ex-ante	1
- interim/mid-term	2
- on-going	3
- ex-post	4
- tematická evaluace	5
- jiný typ	6
- není zřejmé	7

3) Rok zpracování evaluační zprávy?

- rok _____

4) Počet členů evaluačního týmu?

- _____ osob

FORMÁLNÍ PODOBA ZÁVĚREČNÉ ZPRÁVY

5) Rozsah závěrečné zprávy bez příloh?

_____ stran

6) Je v závěrečné zprávě souhrn hlavních závěrů?

- ano	1
- ne	2

OBSAH ZÁVĚREČNÉ ZPRÁVY

7) Vypovídá závěrečná zpráva o silných a slabých stránkách posuzovaného programu/projektu?

- ano	1
- ne	2

- 8) Obsahuje závěrečná zpráva výčet činností prováděných v daném (posuzovaném) období?**
- ano 1
 - ne 2
- 9) Obsahuje závěrečná zpráva návrh konkrétních kroků ke zlepšení v definovaných oblastech?**
- ano 1
 - ne 2
- 10) Vypovídá závěrečná zpráva o účinnosti posuzovaných opatření/programu?**
- ano 1
 - ne 2
- 11) Navrhuje závěrečná zpráva zrušení, pokračování či změnu posuzovaného programu?**
- ano 1
 - ne 2
- 12) Zkoumá závěrečná zpráva efektivitu prostředků vynaložených v rámci posuzovaného programu?**
- ano 1
 - ne 2
- 13) Posuzuje závěrečná zpráva vhodnost aktivit vykonaných v rámci programu vzhledem k potřebám cílových skupin?**
- ano 1
 - ne 2
- 14) Posuzuje závěrečná zpráva soulad mezi plánovanými a uskutečněnými aktivitami?**
- ano 1
 - ne 2

TYP EVALUACE

- 15) Byla v rámci evaluace provedena Need Assessment (analýza potřeb) cílových skupin?**
- ano, velmi podrobně 1
 - ano, ale jen povrchně 2
 - ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy . 3
 - ne, (není o tom zmínka) 4
- 16) Byla v rámci evaluace provedena analýza procesů?**
- ano, velmi podrobně 1
 - ano, ale jen povrchně 2
 - ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy . 3
 - ne, (není o tom zmínka) 4

17) Byla v rámci evaluace provedena analýza výstupů?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4

18) Byla v rámci evaluace provedena analýza výsledků?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4

19) Byla v rámci evaluace provedena analýza dopadů?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4

METODIKA EVALUACE

20) Jsou stanoveny cíle evaluace?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4

21) Je uvedena evaluační otázka?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4

22) Obsahuje závěrečná zpráva popis metodiky provedení evaluace?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4

23) Jsou definována měřitelná kritéria hodnocení?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4

24) Jsou stanoveny standardy (referenční hodnoty), vůči kterým se poměřuje zjištěný stav?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4

25) Jsou nastíněny další možné směry rozvoje programu/posuzované oblasti?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4
26) Je charakterizován samotný program, jeho principy, mechanismy působení apod.?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4
27) Je uveden konkrétní přínos programu pro jednotlivé cílové skupiny?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4
28) Jsou identifikovány jednotlivé zainteresované strany?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, v závěrečné zprávě jsou však jen zmíněny výsledky bez podrobnější analýzy .	3
- ne, (není o tom zmínka)	4
29) Byla použita starší dokumentace, data, statistiky?	
- ano	1
- ne	2
30) Byl v rámci evaluace proveden sekundární výzkum?	
- ano	1
- ne	2
31) Byla v rámci evaluace provedena speciální primární výzkumná šetření?	
- ano	1
- ne	2
32) Které metody kvalitativního výzkumu byly využity?	
a) focus groups	1
b) individuální hloubkové rozhovory	2
c) delphi	3
d) jiná metoda: _____	4
33) Které metody kvantitativního výzkumu byly využity?	
a) fa-to-face interview	1
b) samovyplňovací dotazníky	2
c) CATI	3
d) jiná metoda: _____	4

34) Je uvedena SWOT analýza?

- ano, je prezentována v závěrečné zprávě 1
- ne, ale píše se o ní 2
- ne, ani zmínka 3

35) Je uvedena CBA (cost-benefit analysis) nebo CEA (cost-effectiveness analysis)?

- ano 1
- ne 2

36) Je uveden LOGIC MODEL?

- ano 1
- ne 2

* * *

A U T O E V A L U A Č N Í Z P R Á V Y SEZNAM SLEDOVANÝCH CHARAKTERISTIK
--

IDENTIFIKAČNÍ ÚDAJE

1) Typ školy

- MŠ	1
- ZŠ (1. stupeň)	2
- ZŠ (2. stupeň)	3
- SŠ	4

2) Pokud se jedná o SŠ, jaký je druh školy?

- SOU	1
- SOŠ	2
- gymnázium	3
- jiný typ SŠ	4

3) Jaké konkrétní období je předmětem AEZ?

- rok(y) _____

4) Jak dlouhé období autoevaluační zpráva postihuje (rok nebo dva)?

- jeden rok	1
- dva roky	2
- delší období	3

5) Funkce autora autoevaluační zprávy:

- ředitel školy	1
- zástupce ředitele školy	2
- pedagog	3
- jiný pracovník (výchovný poradce, psycholog, ...)	4

6) Počet pracovníků připravujících autoevaluační zprávu:

_____ pracovníků

7) Je v autoevaluační zprávě zmíněna spolupráce při její přípravě s externím odborníkem?

- ano	1
- ne	2

FORMÁLNÍ PODOBA AUTOEVALUAČNÍ ZPRÁVY

8) Rozsah autoevaluační zprávy?

_____ stran

OBSAH AEZ

9) Rozsah autoevaluační zprávy: obsahuje autoevaluační zpráva část věnující se podmínkám ke vzdělávání?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, oblast je však pouze zmíněna	3
- ne, zcela chybí (ani zmínka)	4
10) Na jaké úrovni je provedena autoevaluace v části věnující se podmínkám ke vzdělávání?	
- na úrovni celé oblasti	1
- na úrovni konkrétních podoblastí	2
- na úrovni nižší než je podoblast	3
- vůbec	4
11) Rozsah autoevaluační zprávy: obsahuje autoevaluační zpráva část věnující se průběhu vzdělávání?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, oblast je však pouze zmíněna	3
- ne, zcela chybí (ani zmínka)	4
12) Na jaké úrovni je provedena autoevaluace v části věnující se průběhu vzdělávání?	
- na úrovni celé oblasti	1
- na úrovni konkrétních podoblastí	2
- na úrovni nižší než je podoblast	3
- vůbec	4
13) Rozsah autoevaluační zprávy: obsahuje autoevaluační zpráva část věnující se podpoře školy žákům a studentům a spolupráci s rodiči?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, oblast je však pouze zmíněna	3
- ne, zcela chybí (ani zmínka)	4
14) Na jaké úrovni je provedena autoevaluace v části věnující se podpoře školy žákům a studentům a spolupráci s rodiči?	
- na úrovni celé oblasti	1
- na úrovni konkrétních podoblastí	2
- na úrovni nižší než je podoblast	3
- vůbec	4
15) Rozsah autoevaluační zprávy: obsahuje autoevaluační zpráva část věnující se výsledkům vzdělávání žáků a studentů?	
- ano, velmi podrobně	1
- ano, ale jen povrchně	2
- ano, oblast je však pouze zmíněna	3
- ne, zcela chybí (ani zmínka)	4

- 16) Na jaké úrovni je provedena autoevaluace v části věnující se výsledkům vzdělávání žáků a studentů?**
- na úrovni celé oblasti 1
 - na úrovni konkrétních podoblastí 2
 - na úrovni nižší než je podoblast 3
 - vůbec 4
- 17) Rozsah autoevaluační zprávy: obsahuje autoevaluační zpráva část věnující se řízení školy, kvalitě personální práce, kvalitě dalšího vzdělávání pedagogických pracovníků?**
- ano, velmi podrobně 1
 - ano, ale jen povrchně 2
 - ano, oblast je však pouze zmíněna 3
 - ne, zcela chybí (ani zmínka) 4
- 18) Na jaké úrovni je provedena autoevaluace v části věnující se řízení školy, kvalitě personální práce, kvalitě dalšího vzdělávání pedagogických pracovníků?**
- na úrovni celé oblasti 1
 - na úrovni konkrétních podoblastí 2
 - na úrovni nižší než je podoblast 3
 - vůbec 4
- 19) Rozsah autoevaluační zprávy: obsahuje autoevaluační zpráva část věnující se úrovni výsledků práce školy, zejména vzhledem k podmínkám vzdělávání a ekonomickým zdrojům?**
- ano, velmi podrobně 1
 - ano, ale jen povrchně 2
 - ano, oblast je však pouze zmíněna 3
 - ne, zcela chybí (ani zmínka) 4
- 20) Na jaké úrovni je provedena autoevaluace v části věnující se úrovni výsledků práce školy, zejména vzhledem k podmínkám vzdělávání a ekonomickým zdrojům?**
- na úrovni celé oblasti 1
 - na úrovni konkrétních podoblastí 2
 - na úrovni nižší než je podoblast 3
 - vůbec 4
- 21) Rozsah autoevaluační zprávy: obsahuje autoevaluační zpráva část věnující se jiným oblastem (vlastním programům, akcím, projektům apod.)?**
- ano, velmi podrobně 1
 - ano, ale jen povrchně 2
 - ano, oblast je však pouze zmíněna 3
 - ne, zcela chybí (ani zmínka) 4
- 22) Na jaké úrovni je provedena autoevaluace v části věnující se jiným oblastem (vlastním programům, akcím, projektům apod.)?**
- na úrovni celé oblasti 1
 - na úrovni konkrétních podoblastí 2
 - na úrovni nižší než je podoblast 3
 - vůbec 4

23) Vypovídá autoevaluační zpráva o silných a slabých stránkách školy?	
- ano	1
- ne	2
24) Obsahuje autoevaluační zpráva výčet činností prováděných v minulém období?	
- ano	1
- ne	2
25) Vyjmenovává autoevaluační zpráva úspěchy školy dosažené v minulém období?	
- ano	1
- ne	2
26) Vyjmenovává autoevaluační zpráva neúspěchy školy minulého období?	
- ano	1
- ne	2

PŘÍPRAVA AUTOEVALUAČNÍ ZPRÁVY /METODIKA AUTOEVALUACE

27) Jsou stanoveny cíle autoevaluace?	
- ano	1
- ne	2
28) Obsahuje autoevaluační zpráva popis metodiky provedení auto-evaluace? Je uvedena metodika zajišťování, jak škola dosáhla vytyčených standardů?	
- ano	1
- ne	2
29) Ve kterých oblastech jsou definována vlastní a podrobnější kritéria hodnocení?	
a) podmínky ke vzdělávání	1
b) průběh vzdělávání	2
c) podpora školy žákům a studentům a spolupráce s rodiči.....	3
d) výsledky vzdělávání žáků a studentů	4
e) řízení školy, kvalita personální práce, kvalita dalšího vzdělávání pedagogických pracovníků	5
f) úroveň výsledků práce školy, zejména vzhledem k podmínkám vzdělávání a ekonomickým zdrojům	6
g) jiná oblast (vlastní programy, akce, projekty, ...)	7
30) Ve kterých oblastech si škola vytyčila standard, se kterým se poměruje (srovnání s referenčními hodnotami)?	
a) podmínky ke vzdělávání	1
b) průběh vzdělávání	2
c) podpora školy žákům a studentům a spolupráce s rodiči	3
d) výsledky vzdělávání žáků a studentů	4
e) řízení školy, kvalita personální práce, kvalita dalšího vzdělávání pedagogických pracovníků	5
f) úroveň výsledků práce školy, zejména vzhledem k podmínkám vzdělávání a ekonomickým zdrojům	6
g) jiná oblast (vlastní programy, akce, projekty, ...)	7

- 31) Byla pro potřeby autoevaluační zprávy použita starší dokumentace školy?**
 - ano 1
 - ne 2
- 32) Byla pro potřeby autoevaluační zprávy provedena speciální výzkumná šetření?**
 - ano 1
 - ne 2
- 33) Kdo byl zdrojem dat (respondentem) v případě prováděných výzkumných šetření?**
 a) žáci 1
 b) učitelé 2
 c) rodiče 3
 d) zřizovatel školy 4
 e) někdo jiný 5
- 34) Byl při přípravě autoevaluační zprávy (při provádění autoevaluace) použit model EFQM?**
 - ano 1
 - ne 2
- 35) Byl při přípravě autoevaluační zprávy (při provádění autoevaluace) použit model CAF?**
 - ano 1
 - ne 2
- 36) Byla při přípravě autoevaluační zprávy (při provádění autoevaluace) použita SWOT (TOWS) analýza?**
 - ano 1
 - ne 2
- 37) Vyplývají z provedené autoevaluace návrhy na zlepšení?**
 - ano 1
 - ne 2

* * *

POSUDKY BAKALÁŘSKÝCH PRACÍ SEZNAM SLEDOVANÝCH CHARAKTERISTIK

1. Rok

- 2008	1
- 2009	2

2. Typ posudku

- konzultantský	1
- oponentský	2

3. Kód autora posudku

- _____

4. Rozsah posudku - počet znaků včetně mezer

- _____

5. Typ posuzované práce:

- teoretická stat'	1
- rozbor sociálního problému	2
- analýza problému sociální politiky	3
- návrh sociálně politického opatření	4
- empirická práce založená na výzkumu	5
- z posudku nevyplývá; nelze stanovit	6

6. Je v posudku uvedeno hodnocení formální úpravy posuzované práce?

- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5

7. Je v posudku uveden návrh na zlepšení formální úpravy posuzované práce?

- ano	1
- ne.....	2

8. Je v posudku uvedeno hodnocení způsobu citací v posuzované práci?

- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5

9. Je v posudku uveden návrh na zlepšení způsobu citací v posuzované práci?

- ano	1
- ne	2

10. Je v posudku uvedeno hodnocení poznámkového aparátu v posuzované práci?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
11. Je v posudku uveden návrh na zlepšení poznámkového aparátu v posuzované práci?	
- ano	1
- ne	2
12. Je v posudku uvedeno hodnocení jazyka, stylistiky a/nebo pravopisu posuzované práce?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
13. Je v posudku uveden návrh na zlepšení jazyka, stylistiky a/nebo pravopisu posuzované práce?	
- ano	1
- ne	2
14. Je v posudku uvedeno hodnocení rozsahu posuzované práce?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
15. Je v posudku uveden návrh na zlepšení rozsahu posuzované práce?	
- ano	1
- ne	2
16. Je v posudku uvedeno hodnocení analýzy poznávacího problému?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
17. Je v posudku uveden návrh na zlepšení způsobu analýzy poznávacího problému?	
- ano	1
- ne	2

18. Je v posudku uvedeno hodnocení výběru relevantní literatury?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
19. Je v posudku uveden návrh na zlepšení výběru relevantní literatury?	
- ano	1
- ne	2
20. Je v posudku uvedeno hodnocení cílů práce?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
21. Je v posudku uveden návrh na zlepšení způsobu formulace cílů práce?	
- ano	1
- ne	2
22. Je v posudku uvedeno hodnocení formulace výzkumné otázky?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
23. Je v posudku uveden návrh na zlepšení způsobu formulace výzkumné otázky?	
- ano	1
- ne	2
24. Je v posudku uvedeno hodnocení použité metodiky a jejího popisu?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
25. Je v posudku uveden návrh na zlepšení použité metodiky a jejího popisu?	
- ano	1
- ne	2
26. Je v posudku uvedeno hodnocení argumentace tvrzení?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5

27. Je v posudku uveden návrh na zlepšení způsobu argumentace tvrzení?	
- ano	1
- ne	2
28. Je v posudku uvedeno hodnocení závěrů, ke kterým autor/ka dospěl/a?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
29. Je v posudku uveden návrh na zlepšení závěrů, ke kterým autor/ka dospěl/a?	
- ano	1
- ne	2
30. Je v posudku uvedeno hodnocení abstraktu?	
- pochvala	1
- kritika	2
- pochvala i kritika současně	3
- neutrální zmínka	4
- žádné hodnocení	5
31. Je v posudku uveden návrh na zlepšení abstraktu?	
- ano	1
- ne	2
32. Vyzdvihuje posudek silné stránky bakalářské práce?	
- ano	1
- ne	2
33. Poukazuje posudek na slabé stránky, nedostatky bakalářské práce?	
- ano	1
- ne	2
34. Obsahuje posudek náměty na obhajobu (otázky k obhajobě)?	
- ano	1
- ne	2
35. Jak je práce klasifikována?	
- výborně (1)	1
- velmi dobře (2)	2
- dobře (3)	3
- neprospěl/a (4)	4
- není uvedeno	5

* * *

RÁMCOVÁ STRUKTURA INTERVIEW Z A D A V A T E L É
--

- 1) Jaká je role vaší organizace při řízení operačního programu?**
(kompetence, řešené úlohy, konkrétní předmět činnosti, vztah k nadřízeným orgánům, komunikace s ostatními řídicími orgány, komunikace s daným resortním ministerstvem, s NOK apod.)
- 2) Jaký je vztah organizace (odboru) k evaluacím?**
(zodpovědnost za konkrétní evaluační šetření, ...)
- 3) Jaká je Vaše role v rámci organizace?**
(struktura pracovní činnosti, zodpovědnost za konkrétní činnosti, doba zaměstnání, předchozí zkušenost s OP atd.)
- 4) Jaký je vztah vaší pracovní pozice k evaluacím?**
(zadávání externích evaluací, kontakt s evaluátory, návrh evaluačního systému, jak příprava podkladů pro rozhodnutí ostatních odborů; odhad podílu pracovní doby věnované evaluacím)
- 5) Jaký je pracovní tým, který se zabývá evaluacemi?**
(počet pracovníků, délka praxe, vystudované obory; realizované činnosti v souvislosti s evaluacemi; provádění evaluací interní x externí formou)
- 6) Realizované typy evaluací za posledních cca 5 let:**
(počty a typy konkrétních projektů – ex-ante, on-going, ex-post, tematické; konkrétní náplň tematických evaluací; formativní a sumativní evaluace, evaluace procesů, výsledků a dopadů, provádění metaevaluací, podíl evaluací OP oproti evaluacím v jiných oblastech)
- 7) Proces přípravy zadání evaluace**
(iniciace, příprava ToR, vyhlášení, komunikace s uchazeči; doba procesu, obvyklé problémy, osvědčené praktiky)
- 8) Proces realizace evaluace a akceptace závěrečné zprávy**
(doba procesu, poskytování podkladových údajů a informací, nejčastější nedostatky, hodnocení realizátorů a jejich činnosti, identifikace jednotlivých oblastí pro možná zlepšení, ...)
- 9) Metodika prováděných evaluací**
(požadované/preferované metody provedení evaluace; orientace v metodologických otázkách, využívání kursů, příruček, konzultací)
- 10) Hodnocení vztahu k realizátorům**
(odhad jejich klíčových kompetencí, hlavní nedostatky; odhad jejich dovedností v oblastech jako návrh designu, používání evaluačních metod, příprava závěrečných zpráv, komunikační dovednosti při prezentaci výsledků z provedených evaluací, znalost metod, další kompetence apod.)
- 11) Charakteristiky závěrečných výstupů z provedených evaluací**
(spokojenost, konkrétní výtky a výhrady, způsoby další práce se závěrečnými zprávami, sdílení výsledků s dalšími odbory, implementace opatření, apod.)

RÁMCOVÁ STRUKTURA INTERVIEW R E A L I Z Á T O Ř I
--

- 1) Jaký je hlavní předmět činnosti vaší organizace?**
(konkrétní předmět činnosti, řešené úlohy, vztahy k ostatním subjektům: partnerství a aliance; zapojení v asociacích; apod.)
- 2) Jaký je vztah organizace k evaluacím?**
(prováděné typy evaluačních šetření, oblasti provádění evaluací; podíl realizace evaluačních šetření na celkové činnosti)
- 3) Jaká je Vaše role v rámci organizace?**
(struktura pracovní činnosti, zodpovědnost za konkrétní činnosti, doba zaměstnání, předchozí zkušenost s evaluacemi atd.)
- 4) Jaký je vztah vaší pracovní pozice k evaluacím?**
(realizace evaluací, řízení evaluačních projektů, kontakt se zadavateli, návrh designu evaluačního šetření, příprava závěrečné zprávy; provádění sběru dat, vedení interview, analytické činnosti apod.; podíl pracovní doby věnovaný evaluacím)
- 5) Jaký je pracovní tým, který se zabývá evaluacemi?**
(počet pracovníků, délka praxe, vystudované obory; konkrétní realizované činnosti v souvislosti s evaluacemi; působení pracovníků v roli externích hodnotitelů apod.)
- 6) Realizované typy evaluací za posledních cca 5 let:**
(podíl evaluací OP oproti evaluacím v jiných oblastech; konkrétní obor a oblasti provádění evaluací, počty a typy konkrétních projektů – ex-ante, on-going, ex-post, tematické; konkrétní náplň tematických evaluací; formativní a sumativní evaluace, evaluace procesů, výsledků a dopadů, provádění metaevaluací)
- 7) Proces přípravy evaluačního šetření**
(příprava projektu šetření, komunikace s zadavatelem; doba procesu, obvyklé problémy, osvědčené praktiky; partnerství; vytváření ad-hoc týmů)
- 8) Proces realizace evaluace a akceptace závěrečné zprávy**
(doba procesu; hodnocení vztahu k zadavatelům; možnosti posílení utilizace apod.)
- 9) Metodika prováděných evaluací**
(požadované/preferované metody provedení evaluace; orientace v metodologických otázkách, vlastní zázemí pro provádění výzkumů, využívání konkrétních technik (CBA, ROI, NPV, BSC, ...; volba designu evaluačních šetření, SW vybavení apod.)
- 10) Hodnocení vztahu k zadavatelům**
(odhad jejich klíčových kompetencí, hlavní nedostatky; odhad jejich dovedností v oblastech jako příprava ToR, používání výsledků z provedených evaluací, znalost metod, další kompetence apod.)
- 11) Charakteristiky závěrečných výstupů z provedených evaluací**
(spokojenost, konkrétní výtky a výhrady, způsoby další práce se závěrečnými zprávami, sdílení výsledků s dalšími odbory, implementace opatření, apod.)

Seznam použité literatury

- Ajzen I., Fishbein M., 1980. **Understanding Attitudes and Predicting Social Behavior.** London, Prentice-Hall
- Alkin M. C., 2004. **Evaluation Roots.** London, SAGE Publication
- Alkin M. C., 1973. **Evaluation Theory Development.** In B.R. Worthen & J.R. Sanders (Eds.), **Educational Evaluation: Theory and Practice.** (pp. 150 – 155). Belmont, California: Wadsworth Publishing Company, Inc.
- Alkin M. C., Christie C. A., 2005. **Theorists' Models in Action.** San Francisco, Jossey-Bass
- Altschuld J. W., Witkin B. R., 2000. **From Needs Assessment to Action.** London, SAGE Publishing
- Alvarez R. M., Brehm J., 2002. **Hard Choices, Easy Answers.** New Jersey, Princeton University Press
- American Accounting Association, 1973. **Studies in Accounting Research. A Statement of Basic Auditing Concepts.** Sarasota
- Armstrong J.S., 1997. **Peer Review for Journals: Evidence on Quality Control, Fairness, and Innovation.** Science Engineering Ethics, 3, 63-84.
- Armstrong H.W., Wells P., Woolford J.L., 2002. **Evaluation, Governance and the Inclusion of the Third Sector in Regional Policy Partnerships.** Paper presented to the Regional Studies Association Conference *Evaluation and EU Regional Policy* Aix-en-Provence, 31 May to 1 June 2002
- Babbie E., 1995. **The Practice of Social Research.** London, Wadsworth Publishing Company
- Batini C., Scannapieco M., 2006. **Data Quality. Concepts, Methodologies and Techniques.** Berlin, Springer-Verlag
- Baumgartl B., Lorenzoni M., Mizikaci F., 2005. **Tools for Process Evaluation.** Wien, Navreme
- Bečvářová Z., 2004. *Evaluační procesy.* Dostupné z www.rvp.cz/clanek/36/11.
- Belson W. A., 1967. **Tape recording: its effect on accuracy of response in survey interview.** Journal of Marketing Research No. 4
- Benson A. P., Hinn D. M., Lloyd C. (eds.), 2001. **Visions of Quality: How Evaluators define, understand and represent program quality.** London, JAI
- Berk R. A., Rossi P. H., 1999. **Thinking About Program Evaluation 2.** London, SAGE Publishing

- Bhola H. S. 1998. **Program evaluation for program renewal: A study of the national literacy program in Namibia.** In Studies in Educational Evaluation No 4
- Bickman L. (ed.), 2000. **Research Design.** London, SAGE Publication
- Bickman L., (ed.) 1987. **Using Program Theory in Evaluation.** In: New Directions for Program Evaluation, No. 33. San Francisco, Jossey-Bass
- Biemer P. P., Lyberg L. E., 2003. **Introduction to Survey Quality.** New Jersey, Wiley
- Bienias S., Lewandowska I. (eds.), 2009. **Evaluation Systems in the Visegrad Member States.** Warsaw, Ministry of Regional Development
- Bingham R. D., Felbinger C. L., 2002. **Evaluation in Practice. A Methodological Approach.** New York, Seven Bridges Press
- Bloom B. S., Engelhart M. D., Furst E. J., Hill W. H., Krathwohl D. R., 1956. **Taxonomy of Educational Objectives: Handbook I. Cognitive Domain.** New York, David McKay
- Bloom H. S., Cordray D. S., Light R. J. (eds.), 1988. **Lessons from Selected Program and Policy Areas.** San Francisco, Jossey-Bass
- Boruch R. F. (ed.), 2005. **The Analysis of the American Academy of Political and Social Science.** London, SAGE Publishing
- Boruch R. F., 1972. **Ralations among Statistical Methods for Assuring Confidentiality of Social Research Data.** Social Science Research No 1
- Boruch R. F., 1997. **Randomized Experiments for Planning and Evaluation.** London, SAGE Publishing
- Bouyssou D., Marchant T., Pilot M., Perny P., Tsoukiás A., Vincke P., 1990. **Evaluation and Decision Models.** Massachusetts, Kluwer Academic Publishers
- Boyle R., 2005. **Evaluation Capacity Development in the Republic of Ireland.** ECD Working Paper Series No. 14
- Bradley C. J. (ed.), 2007. **Process Use in Theory, Research, and Practice.** San Francisco, Jossey-Bass
- Brinkerhoff R. O., Dressler D. E., 1990. **Productivity Measurement. A Guide for Manager and Evaluators.** London, SAGE Publishing
- Brown S. R., Melamed L. E., 1990. **Experimental Design and Analysis.** London, SAGE Publishing
- Bussmann W., 2008. **The Emergence of Evaluation in Switzerland.** Evaluation 2008, 14, 499-506
- Bustelo M., 2006. **The Potential Role of Standards and Guidelines in the Development of an Evaluation Culture in Spain.** Evaluation 2006, 12, 437-453
- Byrn D., 2002. **Interpreting Quantitative Data.** London, SAGE Publication

- Campbell D. T., 1999. **Experimenting Society**. In Campbell D. T., Russo M. J., 1999. *Social Experimentation*. Thousand Oaks, SAGE Publication
- Campbell D. T., Russo M. J., 2001. **Social Measurement**. London, SAGE Publishing
- Campbell D. T., Stanley J. C., 1963. **Experimental and Quasi-Experimental Designs for Research**. Chicago, Rand McNally
- Caracelli V. J., Preskill H. (eds.), 2000. **The Expanding Scope of Evaluation Use**. San Francisco, Jossey-Bass
- Cicchetti D.V., 1991. **The Reliability of Peer Review for Manuscript and Grand Submissions: A Cross-Disciplinary Investigation**. *Behavioral and Brain Sciences* 14, 119–186.
- Chapman R. A., 2003. **Etika ve veřejné službě pro nové tisíciletí**. Praha, SLON
- Chelimsky E., 1997. **The Coming Transformation in Evaluation**. In Chelimsky E., Shadish W. R. (eds.), 1997. *Evaluation for the 21st Century*. London, SAGE Publishing
- Chelimsky E., Shadish W. R. (eds.), 1997. **Evaluation for the 21st Century**. London, SAGE Publishing
- Chen H-T., 1990. **Theory-Driven Evaluations**. London, SAGE Publishing
- Chen H-T., 2005. **Practical Program Evaluation**. London, SAGE Publishing
- Chen H-T., Rossi P. H., 1983. **Evaluating with Sense: The Theory-driven Approach**. In: *Evaluation Review* No. 7
- Christie Ch. A. (ed.), 2003. **The Practice-Theory Relationship in Evaluation**. San Francisco, Jossey-Bass
- Ciarlo J. A., 1981. **Utilizing Evaluation. Concepts and Measurement Techniques**. London, SAGE Publications
- Clarke A., 1999. **Evaluation Research. An Introduction to Principles, Methods and Practice**. London, SAGE Publications
- Cochran W. G., Cox G. M., 1957. **Experimental Designs**. New Jersey, John Wiley & Sons
- Cohen J. 1994. **The Earth is Round ($p < .05$)**. *American Psychologist* No 49
- Cole S., Rubin L., Cole J.R., 1997. Peer Review and the Support of Science. *Scient. Am.*, 237, 34–41.
- Colton D., Covert R. W., 2007. **Designing and Constructing Instruments for Social Research and Evaluation**. San Francisco, Jossey-Bass
- Combs W. L., Falletta S. V., 2000. **The Targeted Evaluation Process**. Alexandria, ASTD
- Cook T. D., Campbell D. T., 1979. **Quasi-Experimentation. Design & Analysis Issues for Field Settings**. Boston, Houghton/Mifflin

- Cook T. D., Reichardt C. S., 1979. **Qualitative and quantitative methods in evaluation research.** London, SAGE Publications
- Coryn C. L. S., Scriven M. (eds.), 2008. **Reforming the Evaluation of Research.** San Francisco, Jossey-Bass
- Cousins J. B., Whitmore E., 2007. **Framing Participatory Evaluation.** In New Directions for Evaluation No 114, San Francisco, Jossey-Bass
- Creswell J. W., Clark V. L., 2007. **Designing and Conducting. Mixed Methods Research.** California, SAGE Publications
- Cronbach L., 1980. **Toward Reform of Program Evaluation.** San Francisco, Jossey-Bass
- Česká školní inspekce, 2006. **Souhrnné poznatky o vlastním hodnocení škol ve školním roce 2005/2006.** Zpráva z výzkumu, Praha
- Daponte B. O., 2008. **Evaluation Essentials.** San Francisco, Jossey-Bass
- Datta L-E., Grasso P. G., (eds.), 1998. **Evaluating Tax Expenditures: Tools and Techniques for Assessing Outcomes.** San Francisco, Jossey-Bass
- Davidson E. J., 2005. **Evaluation Methodology Basics.** London, SAGE Publication
- DePoy E., French G. S., 2003. **Evaluation Practice.** Toronto, Thomson
- Dinero T. E., 1999. **Selected Practical Problems in Health and Social Research.** London, The Haworth Press
- Donaldson S. I., 2007. **Program Theory-Driven Evaluation Science.** New York, LEA
- Donaldson S. I., Christie C. A., Mark M. M. (eds.), 2009. **What Counts as Credible Evidence in Applied Research and Evaluation Practice?** London, SAGE Publishing
- Drulák P. et al., 2008. **Jak zkoumat politiku.** Praha, Portál
- Dvořáček J., 2003. **Interní audit a kontrola.** Praha, C. H. BECK
- Eagly A. H., Chaiken S., 1993. **The Psychology of Attitudes.** Belmont, Thomson / Wadsworth
- EStep, 2008. **Final Report on the Framework to Analyse the Development of Evaluation Capacity in the EU Member States.** Zpráva z výzkumu
- Fajkus B., 2005. **Filosofie a metodologie vědy.** Praha, Academia
- Fetterman D. M., Kaftarian S. J., Wandersman A., (eds.) 1996. **Empowerment Evaluation.** London, Sage Publications
- Fields D. L., 2002. **Taking the Measure of Work.** London, SAGE Publishing
- Fiske S. T., Taylor S. E., 1991. **Social Cognition.** New York, McGraw-Hill

- Fitz-Gibbon C. T., Morris L. L., 1987. **How to Design a Program Evaluation**. London, SAGE Publishing
- Fitzpatrick J. L., Morris M. (eds.), 1999. **Current and Emerging Ethical Challenges in Evaluation**. San Francisco, Jossey-Bass
- Fitzpatrick J. L., Sanders J. R., Worthen B. R., 2004. **Program Evaluation**. Boston, Pearson Education
- Fournier, 2005. **Evaluation**. In. Mathison S., 2005. **Encyclopedia of Evaluation**. London, SAGE Publishing
- Fox J. A., Tracy P. E., 1986. **Randomized Response. A Method for Sensitive Surveys**. London, SAGE Publishing
- Frechtling J. A., 2007. **Logic Modeling Methods in Program Evaluation**. Singapore, John Wiley & Sons
- Freeman R. E., 1984. **Strategic Management. A Stakeholders Approach**. Boston, Pitman
- Gaber J., 2000. **Meta-needs Assessment**. Evaluation and Program Planning No 23
- Garfield E., 1986. **Refreeing and Peer Review: Part 1, Opinion and Conjecture on the Effectiveness of Refereeing**. *Curr. Contents*, 31, 3–11.
- Gay G., Bennington T. L., (eds.) 1999. **Information Technologies in Evaluation (Social, Moral, Epistemological and Practical Implications)**. San Francisco, Jossey-Bass Publishers
- Ginsberg L. H., 2001. **Social Work Evaluation**. Boston, Allyn & Bacon
- Gliner J. A., Morgan G. A., 2000. **Research Methods in Applied Settings (An Integrated Approach to Design and Analysis)**. New Jersey, LEA Publishing
- Gredler M. E., 1996. **Program Evaluation**. New Jersey, Prentice Hall
- Greene J. C., 2007. **Mixed Methods in Social Inquiry**. San Francisco, Jossey-Bass
- Greene J. C., Abma T. A. (editors), 2001. **Responsive Evaluation**. San Francisco, Jossey-Bass
- Greene J. C., Caracelli V. J. (editors), 1997. **Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms**. San Francisco, Jossey-Bass
- Guba E. G., Lincoln Y. S., 1989. **Fourth Generation Evaluation**. London, SAGE Publishing
- Gupta K., 2007. **A Practical Guide to Needs Assessment**. San Francisco, Pfeiffer
- Hall D., Hall I., 2004. **Evaluation and social research – introducing small-scale practice**. New York, Palgrave McMillan
- Hansson F., 1997. **Evaluation Traditions in Denmark: Critical Comments and Perspectives**. *Evaluation* 1997, 3, 85-96

- Harrison M. I., 2005. **Diagnosing Organizations (Methods, Models and Processes)**. London, SAGE Publishing
- Hedrick T. E., Bickman L., Rog D. J., 1993. **Applied Research Design (A Practical Guide)**. London, SAGE Publication
- Hendl J., 2005. **Kvalitativní výzkum. Základní metody a aplikace**. Praha, Portál
- Henry G. T., Julnes G., Mark M. M. (eds.), 1998. **Realist Evaluation: An Emerging Theory in Support of Practice**. San Francisco, Jossey-Bass
- Hersen M., Balow D. H., 1976. **Single-case Experimental Designs: Strategies for Studying Behavioral Change**. New York, Pergamon Press
- Hong D. H., Boden M., 2003. **R&D Programme Evaluation – Theory and Practice**. Hants, Ashgate
- House E. R. (ed.), 1983. **Philosophy of evaluation**. In: New Directions for Program Evaluation, No. 19. San Francisco, Jossey-Bass
- House E. R., 1980. **Evaluating with Validity**. London, SAGE Publishing
- House E. R., 1993. **Professional Evaluation. Social Impact and Political Consequences**. London, SAGE Publishing
- House E. R., Howe K. R., 1999. **Values in Evaluation and Social Research**. London, SAGE Publishing
- Howe M.J.A., 1982. **Peer Review: Improve or be Rejected**. Behavioral and Brain Sciences 5, 218-219.
- Huebner A.J., Betts S.C., 1999. **Examining Fourth Generation Evaluation: Application to Positive Youth Development**. Evaluation 1999, 5, 340-358
- Hyatt J., Simons H., 1999. **Cultural Codes Who Holds the Key?: The Concept and Conduct of Evaluation in Central and Eastern Europe**. Evaluation 1999, 5, 23-41
- Hyman H. H. et al., 1954. **Interviewing in social research**. Chicago, University of Chicago Press
- Isaac S., Michael W. B., 1997. **Handbook in Research and Evaluation**. California, Edits
- Israel M., Hay I., 2006. **Research Ethics for Social Scientists**. London, SAGE Publishing
- Julnes G., Mark M.M., Henry G.T., 1998. **Promoting Realism in Evaluation: Realistic Evaluation and the Broader Context**. Evaluation 1998, 4, 483-504.
- Juran L. M., Gryna F. M., 1980. **Quality Planning and Analysis**. New York, McGraw-Hill
- Kane M., Trochim W. M. K., 2007. **Concept Mapping for Planning and Evaluation**. London, SAGE Publication
- Katz D., 1942. **Do interviewers bias poll results?** Public Opinion Quarterly No. 6

- Kazdin A. E., 1982. **Single-case Experimental Designs**. In Kendall P. C., Butcher J. N. (eds.) *Handbook of Research Methods in Clinical Psychology*. New York, John Wiley
- Kerlinger F. N., Lee H. B., 2000. **Foundations of Behavioral Research**. California, Wadsworth Inc.
- Khakee A., 2003. **The Emerging Gap between Evaluation Research and Practice**. *Evaluation* 2003, 9, 340-352
- Kimmel A. J., 1988. **Ethics and Values in Applied Social Research**. London, SAGE Publishing
- Kimmel A. J., 2007. **Ethical Issues in Behavioral Research**. Malden, Blackwell Publishing
- Kirkpatrick C., Clive G. (eds.), 2007. **Impact Assessment and Sustainable Development**. Cheltenham, EE Publishing
- Kluckhohn C. K., 1951. **Values and Value Orientations in the Theory of Action**. In Parsons T. and Shils E. A. (eds.), *Toward a General Theory of Action*. Cambridge, Harvard University Press.
- Kostelecký T., Patočková V., 2006. **Fungování národních, regionálních a lokálních vlád – problém měření výkonu vlád (“government performance”)**. In *Sociologický časopis* 42
- Kostoff R.N., 1997. **The Principles and Practices of Peer Review**. In: *Science and Engineering Ethics* 3, 1, 19–34.
- Krasner L., Houts A. C., 1984. **A Study of the “Value” Systems of Behavioral Scientists**. *American Psychologist* No 39
- Kushner S., Norris N., 2007. **Dilemmas of Engagement: Evaluation and the New Public Management**. Oxford, JAI Press
- Langbein L., Felbinger C. L., 2006. **Public Program Evaluation. A Statistical Guide**. New York, M. E. Sharp
- Laudan L., 1984. **Science and Values**. California, University of California Press
- Lay M., Papadopoulos I., 2007. **An Exploration of Fourth Generation Evaluation in Practice**. *Evaluation* 2007, 13, 495-504
- Leeuw F.L., 2002. **Evaluation in Europe 2000: Challenges to a Growth Industry**. *Evaluation* 2002, 8, 5-12
- Levin H. M., McEwan P. J., 2001. **Cost-Effectiveness Analysis 2nd Edition**. London, SAGE Publishing
- Levin I. P., 1999. **Relating Statistics and Experimental Design**. London, SAGE Publishing
- Lipsey M. W., 1990. **Design Sensitivity**. London, Sage Publications
- Love A. J., 1991. **Internal Evaluation. Building Organizations from Within**. London, SAGE Publishing

- Mager R. F., 1962. **Preparing Instructional Objectives**. Palo Alto, Fearon
- Mark M. M., Shotland R. L., (eds.), 1987. **Multiple Methods in Program Evaluation**. San Francisco, Jossey-Bass
- Marsh H.W., Ball S., 1981. **The Interjudgmental Reliability of Reviews for the Journal of Educational Psychology**. *J.Educ. Psychol.*, 73, 872–880.
- Mathison S., 2005. **Encyclopedia of Evaluation**. London, SAGE Publishing
- McDavid J. C., Hawthorn L. R. L., 2006. **Program Evaluation & Performance Measurement**. London, SAGE Publishing
- McKillip J., 1987. **Need Analysis. Tools for the Human Services and Education**. London, SAGE Publishing
- McNamara G., O'Hara J., Boyle R., Sullivan C., 2009. **Developing a Culture of Evaluation in the Irish Public Sector: The Case of Education**. *Evaluation* 2009, 15, 101-112
- Mertens D. M., 2005. **Research and Evaluation in Education and Psychology**. London, SAGE Publishing
- Mihalache R., 2009. **Evaluation Networking in Romania**. *Evaluation* 2009, 15, 473-482
- Michek S., 2005. **Vlastní hodnocení školy**. Dostupné z: www.nuov.cz
- Michek S. a kol., 2006. **Příručka pro sebehodnocení poskytovatelů odborného vzdělávání**. Dostupné z: www.nuov.cz.
- Mohan R., Bernstein D. J., Whitsett M. D. (eds.), 2002. **Responding to Sponsors and Stakeholders in Complex Evaluation Environments**. San Francisco, Jossey-Bass
- Mohr L. B., 1995. **Impact Analysis for Program Evaluation 2**. London, SAGE Publishing
- Molas-Gallard J., Davies A., 2005. **Toward Theory-Led Evaluation. The Experience of European Science, Technology, and Innovation Policies**. *American Journal of Evaluation*, 2, 1-18
- Morris L. L., Fitz-Gibbon C. T., Herman J. L., 1987. **Evaluator's Handbook**. London, SAGE Publishing
- Morris L. L., Fitz-Gibbon C. T., King J. A., 1987. **How to Assess Program Implementation**. London, SAGE Publishing
- Morris L. L., Fitz-Gibbon C. T., Lindheim E., 1987. **How to Measure Performance and Use Tests**. London, SAGE Publishing
- Morris M., 2008. **Evaluation Ethics for Best Practice**. New York, The New Guilford
- Myers R. J., 1957. **Accuracy of age reporting in the 1950 United States Census**. *Journal of the American Statistical Association* No. 49

- Neubert S., 2000. **Social Impact Analysis of Poverty Alleviation Programmes and Projects**. London, Frank Cass
- Neuendorf K.A., 2002. **The Content Analysis Guidebook**. Thousand Oaks, CA, SAGE Publishing
- Nowak S., 1975. **Metodologie sociologických výzkumů**. Praha, Svoboda Praha
- Orr L. L., 1999. **Social Experiments**. London, SAGE Publishing
- Owen J. M., 2006. **Program Evaluation. Forms and Approaches**. New York, The Guilford Press
- Parker R. N., Hudley C., (eds.), 2006. **Pitfalls and Pratfalls: Null and Negative Findings in Evaluating Interventions**. San Francisco, Jossey-Bass
- Parry H. J., Crossley H. M., 1950-1951. **Validity of responses to survey questions**. Public Opinion Quarterly No. 14
- Patton M. Q., 1982. **Practical Evaluation**. London, SAGE Publishing
- Patton M. Q., 1987. **How to Use Qualitative Methods in Evaluation**. London, SAGE Publishing
- Patton M. Q., 2008. **Utilization-Focused Evaluation**. London, SAGE Publishing
- Pawson R., 1989. **A Measure for Measures. A Manifesto for Empirical Sociology**. London, Routledge
- Pawson R., 2006. **Evidence-based Policy. A Realist Perspective**. London, SAGE Publications
- Pawson R., Tilley N., 1997. **Realistic Evaluation**. London, SAGE Publications
- Pollitt Ch., 1998. **Evaluation in Europe: Boom or Bubble?** Evaluation 1998, 4, 214-224
- Posavac E. J., Carey R. G., 2003. **Program Evaluation. Methods and Case Studies**. New Jersey, Pearson Education Inc.
- Preskill H., Russ-Eft D., 2005. **Building Evaluation Capacity**. London, SAGE Publishing
- Preskill H., Torres R. T., 1999. **Evaluative Inquiry for Learning in Organizations**. London, SAGE Publishing
- Presman J. L., Wildawsky A., 1984. **Implementation**. Berkeley, University of California Press
- Rektořík J., Šelešovský J. a kolektiv, 2003. **Kontrolní systémy veřejné správy a veřejného sektoru**. Praha, Ekopress
- Remr J., Kvapilová E., 1998. **Důchodové zabezpečení v Čechách a na Slovensku**. In. Potůček M., Radičová I. (eds.), Sociální politika v Čechách a na Slovensku po roce 1998. Praha, Nakladatelství Karolinum

- Remr J., 2005. **Základní rozhodovací dichotomie při poskytování rodinné pomoci seniorům.** Sborník II. mezinárodní konference Problematika – Generace 50 Plus, 69-73
- Remr J., Kotíková J., 2007. **Podpora zaměstnávání starších osob.** VÚPSV Praha
- Remr J., 2008a **Meta-Evaluations. Useful Inspiration from Evaluation Research Colleagues.** Paper presented to the Conference *Quality Criteria in Survey Research*, Cadenabbia, July 2008
- Remr, J., 2008b. **Lazarsfeld's contribution to studying attitudes in evaluation research.** In *Advanced Lazarsfeldian methodology*. Praha, Karolinum
- Remr J., 2009a. **Social Desirability Response Bias: How Far It Can Go?** Paper presented to the Conference *Public Opinion and Survey Research in a Changing World*, Lausanne, September 2009
- Remr J., 2009b. **Možnosti a limity využití principu meta-evaluací ve společenském výzkumu.** Paper presented to the PhD Conference, FSV UK Praha
- Research Governance Framework, 2005. Dostupné z www.dh.goc.uk.
- Reviere R., Berkowitz S., Carter C. C., Ferguson C. G., 1996. **Needs Assessments: A Creative and Practical Guide for Social Scientists.** Washington, Taylor & Francis
- Rice S. A., 1929. **Contagious Bias in the Interview.** *American Journal of Sociology* No. 35
- Robson C., 2000. **Small-Scale Evaluation.** London, SAGE Publishing
- Rodríguez-Campos L., 2005. **Collaborative Evaluations.** Florida, Llumina Press
- Roe A., 1961. **The Psychology of the Scientist.** *Science* No 134
- Rogers P. J., Hacsí T. A., Petrosino A., Huebner T. A., 2000. **Program Theory in Evaluation: Challenges and Opportunities.** San Francisco, Jossey-Bass Publishers
- Rogers P.J., 2008. **Using Programme Theory to Evaluate Complicated and Complex Aspects of Intervention.** *Evaluation* 2008, 14, 29-48
- Rokeach 1972
- Rosenthal R., 1966. **Experimenter Effects on Behavioral Research.** New York, Appleton-Century-Crofts
- Ross P.F., 1980. **The Sciences Self-Management: Manuscript Refereeing, Peer Review and Goals in Science.** Lincoln, MA, The Ross Company
- Rossi P. H., Lipsey M. W., Freeman H. E., 2004. **Evaluation. A Systematic Approach.** London, SAGE Publications
- Russ-Eft D., Bober M. J., de la Teja I., Foxon M. J., Koszalka T. A., 2007. **Evaluators Competencies.** San Francisco, Jossey-Bass Publishers

- Russ-Eft D., Preskill H., 2001. **Evaluation in Organizations**. New York, Basic Books
- Sanders R. J., 1994. **The Program Evaluation Standards 2**. London, SAGE Publishing
- Saxe L., Fine M., 1981. **Social Experiments. Methods for Design and Evaluation**. London, SAGE Publications
- Schalock R. L., 2001. **Outcome-Based Evaluation**. New York, Kluwer Academic Press
- Schuman H., 1966. **The random probe: a technique for evaluating the validity of closed questions**. American Sociological Review No. 31
- Schwandt T. A., 1984. **An Examination of Alternative Models for Sociobehavioral Inquiry**. Indiana, IN
- Schwandt T. A., Halpern E. S., 1988. **Linking Auditing and Metaevaluation**. London, SAGE Publishing
- Scriven M., 1967. **The Methodology of Evaluation**. In Stake R. E. (ed.) Curriculum evaluation. American Educational Research Association Monograph Series on Evaluation No 1. Chicago, Rand McNally
- Scriven M., 1973. **The Methodology of Evaluation**. In Worthen B. R., Sanders J. R. (eds.) Educational Evaluation: Theory and Practice. Belmont, Wadsworth
- Scriven M., 1986. **New Frontiers of Evaluation**. In Evaluation Practice No 7
- Scriven M., 1991. **Evaluation Thesaurus**. London, SAGE Publishing
- Scriven M., 1993. **Hard-Won Lessons in Program Evaluation**. San Francisco, Jossey-Bass
- Scriven M., 2001. **An overview of evaluation theories**. In Evaluation Journal of Australasia
- Shadish W. R., Cook T. D., Campbell D. T., 2002. **Experimental and Quasi-Experimental Designs**. Boston, Houghton Mifflin Company
- Shadish W. R., Cook T. D., Leviton L. C., 1991. **Foundations of Program Evaluation**. London, SAGE Publishing
- Shadish W. R., Newman D. L., Scheirer M. A., Wye Ch., (eds.), 1995. **Guiding Principles for Evaluators**. San Francisco, Jossey-Bass
- Shapiro S, Eberhardt J.C., 1947. **Interview differences in an intensive interview survey**. International Journal of Opinion and Attitude Research No 1
- Shaughnessy J. J., Zeichmeister E. B., 1990. **Research Methods in Psychology**. New York, McGraw-Hill
- Shnoes C. J. Murphy-Berman V., Chambers J. M., 2000. **Empowerment Evaluation Applied: Experiences, Analysis and Recommendations from a Case Study**. American Journal of Evaluation No 21
- Shrader-Frechette K., 1994. **Ethics of Scientific Research**. Maryland, Rowman & Littlefield

- Sieber J. E., 1992. **Planning Ethically Responsible Research**. London, SAGE Publishing
- Singleton, Jr. R. A., Straits B. C., Straits M. M., 1993. **Approaches to Social Research**. Oxford, Oxford University Press
- Smith N. L., Brandon P. R., 2008. **Fundamental Issues in Evaluation**. New York, The Guilford Press
- Sommerland E., Willoughby S., 1995. **Self-evaluation and the Practitioner Evaluator**. *Evaluation* 1995, 1, 105-118.
- Soriano F. I., 1995. **Conducting Needs Assessments. A Multidisciplinary Approach**. London, SAGE Publishing
- Spaulding D. T., 2008. **Program Evaluation in Practice**. San Francisco, Jossey-Bass
- Spector P. E., 1981. **Research Designs**. London, SAGE Publishing
- Stake R. E., 1967. The Countenance of educational evaluation. In *Teachers College Record* No 68
- Stake R. E., 1975. **Program Evaluation, Particularly Responsive Evaluation**. Kalamazoo, Western Michigan University Evaluation Center
- Stake R. E., 1980. **Program Evaluation, Particularly Responsive Evaluation**. In Dockrell W. B., Hamilton D. (eds.) *Rethinking Educational Research*. London, Hodder and Stoughton
- Stake R. E., 2004. **Standards-Based & Responsive Evaluation**. London, SAGE Publishing
- Stake R. E., 2006. **Multiple Case Study Analysis**. New York, The Guilford Press
- Stake R. E., Trumbull D. 1982. **Naturalistic Generalization**. In *Review Journal of Philosophy and Social Science* No 7
- Stecher B. M., Davis W. A., 1987. **How to Focus an Evaluation**. London, SAGE Publishing
- Stufflebeam D. L. (ed.), 2001. **Evaluation Models**. San Francisco, Jossey-Bass
- Stufflebeam D. L., Shinkfield A. J., 2007. **Evaluation, Theory, Models & Applications**. San Francisco, Jossey-Bass
- Suchman E. A., 1967. **Evaluative Research: Principles and Practice in Public Service and Social Action Programs**. New York, Russell Sage
- Tilley N., 2000. Realistic Evaluation: An overview. Dostupné z: <http://www.danskevalueringsselskab.dk/pdf/Nick%20Tilley.pdf>
- Tukey J. W., 1991. **The Philosophy of Multiple Comparisons**. In *Statistical Science* No 6
- UNEP, 2006. **Czech Republic: Integrated Planning and Assessment of National Development Plan**. Výzkumná zpráva

- Unrau Y. A., Gabor P. A., Grinnell R. M., 2007. **Evaluation in Social Work The Art and Science of Practice.** Oxford, Oxford University Press
- Uusikylä P., Virtanen P., 2000. **Meta-Evaluation as a Tool for Learning: A Case Study of the European Structural Fund Evaluations in Finland.** *Evaluation* 2000, 6, 50-65
- Valente T.W., 2002. **Evaluating Health Promotion Programs.** Oxford, Oxford University Press
- Van den Beemt F., Le Pair C., 1991. **Grading the Gain: Consistent Evaluation of Research Proposals.** *Research Evaluation*, 1, 3–10.
- Van Der Knaap P., 2000. **Performance Management and Policy Evaluation in the Netherlands: Towards an Integrated Approach.** *Evaluation* 2000, 6, 335-350.
- Vaš'átková J., 2006. **Úvod do autoevaluace školy.** Olomouc, Univerzita Palackého
- Vaš'átková J., Prášilová M., 2006. **Realizace autoevaluačních procesů ve školách.** In KOCUROVÁ, M. (ed.) *Současné metodologické přístupy a strategie pedagogického výzkumu sborník anotací 14. konference ČAPV*, 2006.
- Veselý A., Nekola M. (eds.), 2007. **Analýza a tvorba veřejných politik.** Praha, Slon
- Wadsworth Y., 1993. **“What is Participatory Action Research?”** Melbourne, Action Research Issues Association
- Warwick D. P., Kelman H. C., 1973. **Ethical Issues in Social Intervention.** In Zaltman G. (ed.), 1973. *Processes and Phenomena of Social Change.* New York, John Wiley
- Weber M, 1998. *Metodologie, sociologie a politika.* Praha, OIKOYMENH
- Weinbach R.W., 2005. **Evaluating Social Work Services and Programs.** Boston, MA, Allyn and Bacon
- Weiss C. H, 1972. **Evaluation Research.** New Jersey, Prentice-Hall
- Weiss C. H, 1987. **The Circuitry of Enlightenment.** Knowledge: Creation, Diffusion, Utilisations No 8
- Weiss C. H, 1998. **Evaluation.** New Jersey, Prentice Hall
- Wholey J. S., Hatry H. P., Newcomer KE. (eds.), 2004. **Handbook of Practical Program Evaluation.** San Francisco, Jossey-Bass
- Winkler J., 2002. **Implementace veřejných programů a přístupy k jejich hodnocení.** Brno, ESF MU
- Witkin B. R., Altschuld J. W., 1995. **Planning and Conducting Needs Assessments.** London, SAGE Publishing
- Wyatt K. L., Philips C. C., 2009. **The Logic Model Guidebook.** London, SAGE Publishing
- Yin R. K., 2003. **Case Study Research. Design and Methods.** London, SAGE Publications

Dokumenty

Government Performance Recording Act, 1993

Nařízení Rady (ES) č. 1260/1999

Nařízení Rady (ES) č. 1083/2006

Zákon č. 561/2004 Sb., o předškolním, základním, středním, vyšším odborném a jiném vzdělávání (školský zákon)

Vyhláška Ministerstva školství, mládeže a tělovýchovy č. 15/2005 Sb.

Evaluační plán NSRR pro období 2007 – 2013