



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Univerzita Hradec Králové
Ústav sociální práce

Evaluace sociálních programů

PhDr. Martin Smutek, Ph.D.

Recenzovali:

Mgr. Radka Janebová, Ph.D.

PhDr. Ondřej Štěch, Ph.D.

Publikace neprošla jazykovou úpravou.

Edice texty k sociální práci

Řada: Výzkumné metody v sociální práci - sv. 4

Studijní materiál vznikl za podpory projektu

Inovace studijních programů sociální politika a sociální práce na UHK s ohledem na potřeby trhu práce (CZ.1.07/2.2.00/28.0127), který je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

ISBN 978-80-7435-473-1

Obsah

1	Úvod	5
2	Základní pojmy evaluace sociálních programů	8
2.1	Co je evaluace?	8
2.2	Definice evaluace sociálních programů	11
2.3	Zaměření a cíle evaluace sociálních programů	12
2.4	Sociální a politický kontext evaluace	15
2.5	Pojem „stakeholder“ - „interesent“	16
3	Význam evaluace v kontextu rozvoje sociálního státu	18
3.1	Stručná historie evaluace	18
3.2	Vývoj přístupů k evaluaci sociálních programů	20
3.3	Současný vývoj programové evaluace	22
3.4	Rozvoj „Nového veřejného managementu“	23
4	Základní typy evaluace	26
4.1	Formativní a sumativní evaluace	26
4.1.1	Formativní evaluace	28
4.1.2	Sumativní evaluace	28
4.2	Evaluace procesu a evaluace dopadu	29
4.2.1	Evaluace procesu	30
4.2.2	Evaluace dopadu (impact, outcome evaluation)	31
4.3	Zvláštní případ evaluace: Evaluace bez zaměření na cíle	32
5	Role teorie v evaluaci	35
5.1	Teorie „o evaluaci“ a teorie „v rámci evaluace“	35
5.2	„Kořeny“, „kmen“ a „větve“ evaluačních teorií	37
5.2.1	Větev „Metody“	40
5.2.2	Větev „posuzování“/„ohodnocování“ (Valuing)	41
5.2.3	Větev „Využití“ (Use)	41
5.3	Závěrečná poznámka ke kapitole o stromu teoretických přístupů k evaluaci	42
6	Kvantitativní evaluace sociálních programů	43
6.1	Typy dat – kvantitativní a kvalitativní	43
6.2	Ztraceno v překladu: Podstata empirického sociálně-vědního výzkumu	44
6.3	Kvantitativní strategie	45
6.3.1	Experimentální design	45
6.3.2	Kvazi-experimentální design	47
7	Kvalitativní výzkumné nástroje v evaluaci	49

7.1	Kvalitativní strategie.....	49
7.1.1	Holistická perspektiva v kvalitativní evaluaci	50
7.1.2	Historie kvalitativních přístupů	51
7.1.3	Generalizace vs. extrapolace či transferabilita v kvalitativním výzkumu.....	52
7.1.4	Kreativita v kvalitativním výzkumu	53
7.1.5	Vědění versus tvořivost	54
7.1.6	Zásady kreativity v přípravě i realizaci kvalitativního výzkumu.....	54
7.2	Kombinace strategií – metodologický koktejl.....	56
7.3	Techniky sběru dat	58
8	Role hodnotitele v evaluaci	62
8.1	Role hodnotitele v designu a plánování nových programů.....	62
8.2	Externí a interní hodnotitel	63
8.3	Role „stakeholderů“	65
8.4	Vědecká důvěryhodnost a důvěryhodnost pro „stakeholdery“	66
9	Prezentace evaluačních zjištění a jejich praktické využití.....	68
9.1	Vyjasnit si účel studie: Typologie účelu výzkumu	68
9.2	Veřejně politické dokumenty.....	70
9.3	Typy použití evaluačních zjištění	71
9.4	Jak maximalizovat využití?	74
9.5	Závěrečná zpráva z evaluace	76
9.6	Titulní strana.....	78
10	Závěr	82
11	Rejstřík	88

1 Úvod

Cílem tohoto studijního materiálu je porozumění základnímu pojmosloví v problematice evaluace sociálních programů. Studijní předmět „Evaluace sociálních programů“ je zaměřen na teorii evaluace a tedy sumarizaci nejpodstatnějších teoretických přístupů, kdy je cílem porozumění teoretickým úskalím hodnotícího výzkumu, jeho typům a formám a jejich vazbě na základní výzkumné metody a techniky sociálních věd.

Student by měl být schopen na základě tohoto předmětu argumentačně legitimizovat (odůvodnit) potřebu provedení evaluace sociálního programu (politiky) a zdůvodnit výzkumné postupy a nástroje, kterými chce dojít k evaluačním zjištěním. Student by dále měl být schopen formulovat doporučení pro praxi vzešlá z hodnotícího výzkumu jak sumativního (převážně kvantitativního) typu, tak doporučení vzešlá z formativní (převážně kvalitativní) studie. Výstupem studia by mělo být komplexní porozumění studenta problematice evaluace v soudobé sociální realitě (sociálním státě), jelikož jde mj. i o jeden z nástrojů prosazování veřejného zájmu ve veřejné (sociální) politice.

Evaluace veřejných politik se stává v současnosti stále více a více potřebnou, a to ruku v ruce s tím, jak je vyžadována stále vyšší a vyšší odpovědnost realizátorů programů za jejich efektivnost, účelnost, celkově vzato za racionalitu takových opatření. Tlak na racionalitu opatření ve veřejné politice je spojen dále s manažerizací veřejné sféry, tedy s přebíráním nástrojů managementu sféry komerční. Tento text reflektuje manažerizační tendence ve formě tzv. Nového veřejného managementu, jehož důsledkem je zvýšený důraz na evaluaci sociálních programů. Manažerizačním trendům (což je kontext pro rozvoj teorie evaluace sociálních programů) je věnována celá jedna podkapitola této publikace, avšak na tomto místě chce autor podotknout, že tento text nemá být chápán jako příspěvek k manažerizačnímu trendu a prohloubení důrazu na ekonomickou efektivnost sociálních programů (jak bývá v tomto duchu evaluace často chápána). Tento studijní zdroj nemá ani čistě odsuzovat manažerizační tendence, tyto soudy necht' jsou na každém jednotlivém profesionálovi v praxi sociální sféry.

Znalosti získané v předmětu „Evaluace sociálních programů“ jsou využitelné v koncepční, plánovací a hodnotící rovině tvorby a analýzy veřejných (sociálních) politik jak na národní (mezinárodní) úrovni, tak na lokální (krajské, obecní) úrovni. Znalosti jsou tak prakticky využitelné jak v oblasti veřejné správy při koncipování jednotlivých tematických oblastí sociální politiky (rodinná politika, politika zaměstnanosti, vzdělávací

politika...), a také v procesech komunitního plánování sociálních služeb tak v oblasti neziskového sektoru v plánování a hodnocení neziskových sociálních programů (projektů).

Předmět „Evaluace sociálních programů“ vychází převážně z anglicko-jazyčných učebnic základů evaluace sociálních programů. V úvodu kurzu je probíráno základní pojmosloví, je vymezena oblast působení evaluačního výzkumu a je vysvětlen rozdíl oproti tzv. základnímu“ (akademickému) výzkumu.

Předmět se zabývá jak tzv. „sumativní“ tak i tzv. „formativní“ evaluací a uvádí oba tyto typy do kontextu relevantních výzkumných nástrojů, které při jejich realizaci lze využít. Je uveden cyklus (fáze) tvorby a realizace sociálního programu a vymezeny možné evaluační záměry v jednotlivých fázích tvorby a realizace sociálního programu (politiky).

V textu je systematicky použito pojmu „hodnotitel“ pro osoby, které provádějí evaluaci. Tento pojem umožňuje jednoslovně vyjádřit funkci takové osoby bez ohledu na to, zdali jde o experta z „vnějšku“, najatého k vedení evaluace, nebo o osobu „zevnitř“ organizace, která evaluaci vykonává v rámci svých pracovních povinností ve svém zaměstnání. Dále bude systematicky použit již zmíněný pojem „stakeholder“, jelikož bohužel není k dispozici zcela adekvátní český jednoslovný výraz pro celý takovýto komplex zainteresovaných aktérů (je možno využít v tomto duchu slova „interesent“, což je však opět cizí slovo).

V další části předmětu je věnována pozornost různým teoretickým přístupům k evaluaci (či evaluačnímu výzkumu) a k různým metodologickým směrům vycházejícím jak z pragmatismu, pozitivismu tak i z interpretativního výzkumného paradigmatu. V metodologické části kurzu je věnována pozornost jak běžným výzkumným nástrojům sociálních věd, tak i experimentálnímu (resp. kvazi experimentálnímu) výzkumnému designu. Závěr kurzu tvoří téma využití a využitelnosti zjištění z evaluace (evaluačního výzkumu) a dále prezentace podstatných náležitostí evaluační závěrečné zprávy a způsobů jejího rozšíření („diseminace“) mezi klíčové příjemce zjištění (tzv. stakeholdery).

V textu je nejprve zařazena kapitola o základních pojmech, které jsou s evaluačním výzkumem spojeny (kap. 2), dále pak stručné pojednání o historii evaluačního výzkumu (kap. 3), které uvede problematiku do kontextu rozvoje sociálního státu v průběhu 20. století. Obsáhlejší pozornost je pak věnována základním typům evaluačního výzkumu (kap. 4). To vše by mělo položit základy možnosti hlubšího pohledu na teoretická východiska programové evaluace (kap. 5), kde jsou na metafoře „stromu“ evaluačních

teorií diskutovány teoretické směry, které se primárně zaměřují na metody, ohodnocování či na využití evaluace. Dále se text zabývá možnostmi kvantitativní (kap. 6) a kvalitativní výzkumné strategie v programové evaluaci (kap. 7). Samostatná kapitola je věnována i roli hodnotitele a roli tzv. „stakeholderů“ (kap. 8). V závěrečné kapitole textu (kap. 9) je probrána problematika využití evaluačních závěrů a způsobů, jak toto využití maximalizovat.

2 Základní pojmy evaluace sociálních programů

Na různých definicích je možné přiblížit různá pojetí evaluace, z čehož následně můžeme vyvodit i důsledky uplatňování určitých vyhraněných přístupů. Než však k uvedení některých definic programové evaluace přistoupíme, je důležité uvést určitý „background“, tedy okolnosti, které tento typ výzkumu podkreslují a rozhodnout, co je evaluační výzkum a co výzkum „akademický“.

2.1 Co je evaluace?

To, co může posloužit k odlišení evaluační studie od jiných forem sociálního výzkumu, je jednoduše otázka smyslu a účelu evaluační studie. Rozdíl mezi tzv. „základním“ výzkumem (či „akademickým“ výzkumem) a evaluací byl Cronbachem a Suppesem (1969) popsán jako rozdíl mezi zkoumáním orientovaným na výsledek, porozumění jevu a zkoumáním orientovaným na budoucí rozhodování (ve veřejné politice). Tzv. „základní“ výzkum se zaměřuje na produkci znalosti a pravdy. Užitečná evaluace pak v první řadě podporuje skutečnou hmatatelnou akci. Pojem „evaluační výzkum“ autorů Rossiho a Freemana je možné chápat jako jakýsi hybrid, který má tendenci být spíše na znalost orientovaným výzkumem než výzkumem orientovaným na akci (dle Patton 1997:24).

Výzkumná strategie užitá v evaluačním výzkumu je určována prostřednictvím konkrétních, předem daných cílů evaluace. „Ty zpravidla nejsou orientovány na všeobecné vědecko-teoretické zájmy“ (Fiala, Schubert 2000:124). Dá se tedy říci, že cílem sociálně-vědního „akademického“ výzkumu je lepší porozumění společnosti, zatímco hlavním cílem evaluační studie je orientace na poskytnutí „zpětné vazby“ reálně existujícím sociálním programům a v ideálním případě i využití těchto poznatků v praxi.

V případě akademického výzkumu jde o tvorbu logických, koherentních a empiricky testovatelných teorií. Podle postřehu Veselého (2007:24) se „...nová teorie nebo empirický poznatek většinou těší pozornosti kolegů vědců zcela bez ohledu na jejich využitelnost ve veřejné politice. Dokonce lze často říci, že pro mnohé vědce je hlavním kritériem kvality jejich práce nikoli praktičnost, ale netriviálnost a originalita výsledků...“.

Aby však zde uvedené pasáže nevyzněly jako jednostranná kritika „základního“ resp. „akademického“ výzkumu, je potřeba uvést, že akademický výzkum „...vytváří nezbytnou základnu poznatků pro výzkum, který je již přímo orientovaný na veřejnou politiku“

(Veselý 2007:27). Dále pak také záleží na národní kultuře sociálních věd, nakolik rigidně má tendenci odlišovat „základní“ poznatky od jejich aplikace. „Mnohé akademické americké časopisy například explicitně požadují po autorech článků, aby svoji stať končili praktickými veřejně-politickými doporučeními nebo důsledky (policy recommendations, resp. policy implications). Relevance získaných poznatků pro veřejnou politiku je zde také jedním z důležitých aspektů žádostí o finanční podporu výzkumu (granty)“ (Veselý 2007:27). Hlavní charakteristické rozdíly mezi „akademickým“ výzkumem a výzkumem orientovaným na veřejnou politiku, kde nalézá uplatnění evaluace sociálních programů, uvádí následující tabulka.

Tabulka 1 - Rozdíly mezi „akademickým“ výzkumem a výzkumem orientovaným na veřejnou politiku (policy research, policy analysis)

	Akademický výzkum	Policy research a policy analysis
Hlavní cíl	Porozumění	Změna
Klient	Nespecifikovaný, pravda, veřejný zájem	Konkrétní klient, aktér politiky (nutná neustálá komunikace)
Způsob	Vysvětlení	Hodnocení
Předmět analýzy	Obor, disciplína (např. dopravní systém)	Problém (např. zahlcení osobní dopravou)
Sběr dat	Extenzivní fáze primárního sběru dat	Často pouze sekundární analýza dat
Výsledný produkt	Studie, odborný článek	Analýza, plán, memorandum

Zdroj: Veselý 2007:27

Patton (1986 dle Clarke 1999) rovněž poukazuje na odlišnost mezi evaluací a výzkumem. Ačkoli programoví evaluátoři používají čistě výzkumné metody ke sběru dat, jejich primární záměr je poskytnout informaci těm, kteří jsou odpovědní za tvorbu rozhodnutí o budoucím rozvoji programů. Proto je důraz kladen na nástroje, relevanci, praktičnost a pokrytí informačních potřeb specifických „rozhodovatelů“ (decision-makers). To je v protikladu k čistému výzkumu, kde je podle Pattona relativně velký důraz kladen na zobecnitelnost, kauzalitu a kredibilitu (důvěryhodnost) v rámci vědecké komunity.

Podle Marka (1996:230, vl. překl.) je pak v tomto duchu užitečné uvést, co evaluace programu není. Není to „... ,čistý“ nebo teoretický výzkum, který má za cíl rozvoj hypotéz a tvorbu teorií. Ačkoli je záměr ,čistého‘ výzkumu vytvářet znalost v našem vlastním zájmu, evaluační výzkum je vždy spojen s programem, který je poskytován občanům. Takovýto program může být již v činnosti, nebo může být teprve plánován.“

Podle Weissové (1997:516 in Clarke 1999:2, vl. překl.) lze evaluaci a její odlišnost od akademického výzkumu definovat následujícím způsobem: „Evaluace, na rozdíl od základního výzkumu, nemá za cíl hledání ‘pravdy’. Jejím záměrem je pomoci zlepšení programování a tvorby politiky“. Je to právě tato praktická orientace, která je jedním z hlavních odlišujících znaků evaluace od jiných výzkumných aktivit. Jako forma zkoumání je tedy evaluace orientována na akci. Evaluační výzkum se liší od jiných typů výzkumu spíše vzhledem k zamýšlenému záměru, než vzhledem k designu nebo použitým výzkumným metodám. Tento pohled na věc je v současnosti sdílen mnoha evaluačními teoretiky.

V diskusi o odlišnostech mezi evaluačním a akademickým výzkumem může být užitečné zmínit se o instituci, která je označována jako tzv. „think-tank“ (česky by se dalo říci přeneseně „mozkovna“). Jejich hlavní rolí je „...být určitým prostředníkem mezi akademickým výzkumem a světem politiky nebo obchodu. Koncept think-tanků pochází ze západní Evropy, kde jsou tyto instituce velmi rozšířené a některé z nich mají i nezanedbatelný vliv na veřejnou politiku v dané zemi nebo i v mezinárodním měřítku“ (Veselý, Nekola 2007:165). Rolí think-tanků je tedy zabývat se výzkumem politiky (policy analysis) a většinou se dotýkají témat ekonomie, veřejné správy, veřejné politiky a jsou zdrojem myšlenek, koncepcí a podněcovatelem debat v oblasti veřejné správy a přispívají tak k artikulaci veřejného zájmu. Je však třeba podotknout, že vzhledem k financování think-tanků, které je často zajišťováno soukromými nadacemi či mecenáši z oblasti průmyslu, není apriorním předpokladem jejich činnosti snaha o objektivitu. Často tak hájí zájmy svých sponzorů, donátorů. Přestože se v České republice tento koncept teprve rozvíjí, existuje několik institucí, které publikují politicky relevantní odborné dokumenty. Mezi významná analytická pracoviště v ČR pak podle Nekoly a Veselého (2007:165) patří Národní vzdělávací fond, Národní ústav odborného vzdělávání, Výzkumný ústav práce a sociálních věcí, CERGE-EI nebo Centrum pro sociální a ekonomické strategie (CESES). Explicitně jako think-tank se profiluje například Institut pro sociální a ekonomické analýzy (ISEA).“

V případě evaluace můžeme dále hovořit o formálním, systematickém zkoumání naplánované sociální intervence, což je úkol pro odborníky v oblasti sociálních věd, nebo můžeme hovořit o neformálním, subjektivním hodnocení v podstatě čehokoliv, což dělají v podstatě všichni lidé během svého běžného všedního dne. První typ evaluace je podle Guby a Lincolnové (1986 dle Clarke 1999) jasně odlišitelný od druhého jmenovaného díky

své podstatě a způsobu, jakým je veden. Formální evaluace je definována jako forma „disciplinovaného zkoumání“, která aplikuje vědecké procedury ke sběru a analýze informací o obsahu, struktuře a výsledcích programů, projektů a naplánovaných intervencí. Každodenní neformální evaluaci je pak pochopitelně možné označit za nesystematické, případně „laické“ hodnocení.

V tomto kontextu si můžete všimnout, že v této práci je systematicky používán výraz „evaluace“, protože české slovo „hodnocení“ má širší škálu kontextů, ve kterých je uplatňováno. Mj. český výraz „hodnocení“ evokuje i ono subjektivní laické hodnocení. Používání z angličtiny převzatého pojmu evaluace lze považovat za v současné době již pevně ustálené i v českých sociálních vědách a evokuje systematické hodnocení za využití relevantních sociálně-vědních výzkumných metod.

2.2 Definice evaluace sociálních programů

Není náhodou, že většina nejvlivnějších textů o základních teoriích, principech a technikách evaluace sociálních programů pochází z USA. Již zmiňovaný trend tzv. manažerismu zde nachází nejkvalitnější odezvu. Největší část z citovaných definic sociálních programů, přístupů k programové evaluaci a autorů obecně má tedy své kořeny právě v této zemi. Uveďme některé z možných definic: „Evaluace je typ výzkumu politik, který má lidem pomoci činit rozumná rozhodnutí o budoucím chodu programu... Evaluace se snaží být nezaujatá. Ve své podstatě se snaží reprezentovat celou škálu různých perspektiv těch, kteří jsou oporou programu“ (Weiss dle Alkin 2004:83, vl. překl.). „Programová evaluace je definována jako aplikace evaluačních přístupů, technik a znalostí k systematickému uchopení a rozvoji plánování, implementace a efektivity programů“ (Chen 2005:3, vl. překl.).

Neměli bychom rozhodně zapomenout na možná nejznámější autory na poli evaluace, a to díky jejich široce rozšířené učebnici „*Evaluation: A Systematic Approach*“, autory Rossiho a Freemana: „Evaluační výzkum je systematická aplikace procedur sociálního výzkumu za účelem hodnocení konceptualizace, hodnocení designu, hodnocení implementace a hodnocení nástrojů sociálně-intervenčních programů“ (1993:5, vl. překl.).

Podle Pattona (1986:14 in Clarke 1999:1, vl. překl.) je programová evaluace „...systematický sběr informací o aktivitách, charakteristikách a výsledcích programů, které jsou určeny konkrétním lidem k redukování nejasností, k zlepšení efektivity a k tvorbě

rozhodnutí s ohledem na to, co tyto programy dělají a co ovlivňují.“ Obecná Pattonova definice má tři navzájem vztážené komponenty: (1) systematický sběr informací o (2) potenciálně široké míře témat pro (3) mnoho možných použití. Odlišný přístup je reprezentován zmiňovanou široce používanou učebnicí Rossiho a Freemana, kteří definují evaluační výzkum jako systematickou aplikaci sociálně-vědních procedur při hodnocení programů sociální intervence. Avšak i Rossi s Freemanem sami zdůrazňují, že programová evaluace je i něco více než jen aplikace sociálně-vědních metod. Podle nich je to: „...rovněž politická a manažerská aktivita, vstup do komplexní mozaiky, ze které se vynořují politická rozhodnutí a možnosti alokace zdrojů pro plánování, design a implementaci programu určeného ke zlepšování životních podmínek lidí. V tomto smyslu evaluační výzkum rovněž potřebuje být viděn jako integrální součást sociální politiky...“ (1993:29, vl. překl.).

Evaluace je tedy typ výzkumu, který používá zavedené metody sociálně vědního výzkumu k evaluaci úspěšnosti (nebo neutrálně řečeno efektu) programu veřejné politiky, např. veřejných sociálních služeb (zde je použit pojem sociální služby ve velmi širokém slova smyslu - v pojetí, které odpovídá anglickému „welfare services“ a zahrnuje jak např. veřejné služby vzdělávání, tak např. veřejné zdravotnictví, tedy nejen sociální služby jako součást systému sociálního zabezpečení obyvatelstva).

V závěru této podkapitoly o definici programové evaluace by tedy mělo být zřejmé, že není možné poskytnout jednoduchou definitivní odpověď na otázku, co je evaluace. Jak podotýká Patton (1982:35 dle Clarke 1999:4), žádná jedno větná definice nebude stačit k plnému pokrytí praxe evaluace.

2.3 Zaměření a cíle evaluace sociálních programů

Pojmy „politika“, „program“ a „projekt“, jak již bylo zmíněno v úvodu, zde budou používány v podstatě vzájemně zaměnitelně, jelikož se na ně vztahují a aplikují stejné výzkumné principy. Např. i podle Chena (2005) je slovo „politika“ velmi blízké slovu „program“. Ačkoli se může politika jevit o mnoho rozsáhlejší než program (zde vzhledem k množství lidí, které ovlivňuje), principy, o kterých hovoří tento text, jsou relevantní i pro ni. To znamená, že ve zbytku knihy je fakticky možné chápat slovo program i jako politiku. Podle Nachmiase (1980) mohou být pojmy „projekt“, „program“ a „politika“

analyticky odlišeny podél kontinua specifičnosti, kde projekt bude nejspecifičtější soubor aktivit, konstruovaný k dosažení cíle, a politika naopak „nejrámovější“.

Celý tento text je o evaluaci sociálních programů. Ale co se tímto pojmem vlastně myslí? Jaké jsou charakteristiky a podstata sociálních programů, ke kterým se pak vztahuje evaluace? Programy, které hodnotitelé mohou hodnotit, mají rozdílné zaměření. Může jít o programy zdravotní politiky, programy vztahující se k politice zaměstnanosti nebo intervenční programy v podstatě v jakékoli oblasti sociálního státu. Mohou spadat do oblasti vzdělávání, kriminality nebo sociálního zabezpečení. Nicméně všechny takové programy mohou sdílet společné rysy, a to například ve snaze zlepšit lidské „well-being“, tedy jejich osobní bio-psycho-sociální „pohodu“. V celospolečenském významu pak mají tyto sociální programy zlepšit celkový „welfare“ obyvatelstva, tedy blahobyt obyvatelstva (odtud pak pojem „welfare state“, tedy doslovně „stát blahobytu“, resp. „sociální stát“, tzn. typ společnosti, který se stará o sociální blahobyt svého obyvatelstva). Je v tuto chvíli jedno zdali tak činí pomocí prevence nemocí, snižování míry chudoby či sociálního vyloučení nebo investicemi do vzdělávání. Pro zjednodušení je o všech těchto programech v této publikaci psáno jako o „intervenčních“ programech nebo, ještě jednodušeji prostě o sociálních programech.

Evaluace sociálních programů jsou pak vykonávány z mnoha rozličných důvodů (Rossi, Freeman 1993:13):

- mají posoudit „hodnotu“ běžících programů a odhadnout užitečnost pokusů, jak tyto programy zlepšit;
- mají stanovit užitečnost inovativních programů a iniciativ;
- mají zvýšit efektivitu programového managementu a administrace a vyhovět požadavkům na odpovědnost vůči programovým sponzorům, donátorům a jiným zainteresovaným aktérům (veřejnosti);
- evaluace rovněž mohou svou „troškou do mlýna“ přispívat k reálným a zároveň i metodologickým znalostem sociálních věd.

Jak cituje Matoušek Royseho a kol. (in Smutek 2005:10-11), slouží hodnocení programů k hledání odpovědí na následující otázky (byť by v zájmu evokování nejen kvantitativní, ale i kvalitativní analýzy mohly spíše začínat slůvky jako je „jak“, „proč“ a podobně):

- Pomáhá opravdu program klientům?
- Jsou klienti spokojeni se službami?
- Mění program situaci klientů a jejich život?
- Je zapotřebí poskytovat na program tolik peněz, kolik se poskytuje?
- Je nově zaváděný program lepší než program dřívější?
- Je možné určitý program zlepšit?
- Jsou pracovníci zaměstnaní v programu dobře využiti?

V plánovacích stádiích programů sociální intervence se evaluace zaměřují na stanovení rozsahu a závažnosti sociálních problémů, které vyžadují sociální intervenci. Zaměřují se také na design programů, které si kladou za cíl pokrytí těchto problémů. Ve stádiu řízení zavedených i nových programů pomáhají evaluace determinovat stupeň, ve kterém jsou programy efektivní - tzn. jak úspěšně dosahují své zamýšlené cílové populace a jak se starají o svěřené zdroje (veřejné finanční prostředky či dary, které jim například poskytnul sponzor).

Hodnotící výzkum je vhodný kdykoliv, kdy je plánován nějaký společenský zásah. Společenský zásah je proveden ve společenských souvislostech za účelem získání zamýšleného výsledku. Evaluační výzkum je pak tím procesem, kdy dochází k vyhodnocení, zda bylo dosaženo zamýšleného výsledku.

Témata evaluačního výzkumu jsou velmi různorodá. Klademe si např. otázky typu (Babbie 1998):

- Pomohly jednotlivé programy politiky zaměstnanosti vytvořit v regionu nová pracovní místa a snížit tak nezaměstnanost?
- Umožňují programy sociálních služeb poskytovaných v domácnostech seniorů oddálit či úplně eliminovat umístování seniorů do domovů pro seniory?
- Redukoval vzdělávací program o sexuálním zneužívání dětí tento typ sociální patologie?
- Přispělo zavedení školného na univerzitách k rychlejšímu rozvoji vysokých škol a ke zlepšení kvality studia? (minimálně tato diskuse, z výše uvedených, Českou republiku teprve čeká).

Podle Fialy a Schuberta (2000:83) můžeme dále pracovníě odlišit tři typy evaluací: (1) administrativní, (2) politickou a (3) vědeckou evaluaci.

(1) „Evaluace může probíhat čistě (vnitřně) administrativně, např. prostřednictvím představených nebo zvláštních revizních orgánů v rámci úřadu, který program realizoval. Ale i externí (vědecká) evaluace může být omezena pouze na interní administrativní kritéria, na jejichž základě jsou kladeny např. otázky, zda mohl být program realizován bez zaměstnání dodatečného personálu, nebo jestli výdaje odpovídaly finančnímu plánu apod.“

(2) „Politická evaluace se však musí řídit jinými kritérii a musí vypadat odlišně. Typickými otázkami politické evaluace je, zda program ovlivnil přesně tu cílovou skupinu, pro kterou byl koncipován, zda byl získán maximální počet potenciálních voličů a je-li výsledek spojován s příslušným politickým rozhodnutím.“

(3) „Také vědecká evaluace musí nejprve stanovit, zda má být hodnocen celý program, nástroje, jednotlivá opatření nebo různé implementační techniky, zda bude hodnocen celý politický proces, implementační proces nebo jen menší dílčí procesy, jestli má být měřen output, tzn. ‚aktivity vztahující se bezprostředně k aktérům, kteří program realizují‘, zda má být případně zjišťován impact, tzn. ‚změny, které lze sledovat u adresátů příslušné politiky‘, či zda je cílem analyzovat outcome, tedy ‚celkový účinek státních aktivit na veškeré účastníky příslušné politiky, tzn. zamýšlené, ale hlavně nezamýšlené účinky‘“ (Jann 1981:26 dle Fiala, Schubert 2000:83).

2.4 Sociální a politický kontext evaluace

Evaluační výzkum se nachází ve výrazném politickém kontextu, jelikož výsledky hodnotícího výzkumu se mohou stát „municí“, s níž oponenti programu mohou operovat, když se třeba snaží docílit zrušení programu. Výsledky programové evaluace tak mohou ovlivnit existenci lidí, jejich zaměstnání (které se váže na existující programy), výši jejich mezd (financování programu), mohou ovlivnit osudy klientů programu. Politika a věda jsou tedy integrálními součástmi programové evaluace. Podle Cronbacha a kol. (1980 dle Greene 1998) má hodnotitel politický vliv, i když na něj vůbec neaspiruje.

Prostřednictvím evaluačních studií stupují výzkumníci do politického procesu. Evaluační studie je pak možné chápat jako tzv. veřejně politické dokumenty, které mají vliv na tvorbu veřejné politiky. Jak uvádí Veselý (2007:14) „...existuje mnoho různých

podob, kterými lze ovlivňovat politiku prostřednictvím expertizy. Tyto způsoby lze nejlépe kategorizovat podle typů činnosti, kterou provádějí nebo které se účastní. Konkrétně jde o tyto aktivity:

- Deskripce a monitoring - sběr dat a informací o určité oblasti.
- Explanace - tvorba teorií a modelů o příčinách a důsledcích sociálních jevů.
- Komparace - porovnání stavu v určité oblasti se situací v jiné zemi.
- Evaluace - vyhodnocení naplnění cílů předchozích politik či programů.
- Tvorba doporučení - poskytnutí rady, jak postupovat při řešení určitého konkrétního problému.
- Legislativa - pomoc při tvorbě legislativy.
- Tvorba strategií a plánů - pomoc při vytváření strategických dokumentů.“

Evaluační výzkum se tak podle Earla Babbieho (1998) pohybuje v silném politickém a ideologickém rámci (tedy např. je hnán pouze dobovými trendy, jako je např. multikulturalismus, pozitivní diskriminace a podobně, které nemusí být součástí dominantního společenského světónázoru i v budoucnu) a byť jsou jeho výsledky dosahovány prostřednictvím vědeckých postupů (tedy v podstatě se snahou o objektivitu a neutralitu – alespoň v tzv. pozitivistické výzkumné tradici), přijímání výsledků různými zainteresovanými lidmi je většinou silně subjektivně zabarveno.

Na závěr této podkapitoly o politickém kontextu programové evaluace shrňme, že též známí autoři Rossi a Freeman (1993:406) rozpoznávají vysokou důležitost politického a sociálního kontextu a vyzývají hodnotitele ke kontinuálnímu hodnocení sociální ekologie prostoru. Tento prostor nazývají „arénou“, ve které evaluátoři chtě nechtě operují a vstupují do vztahu s tzv. „stakeholdery“.

2.5 Pojem „stakeholder“ - „interesent“

Ve všech evaluačních kontextech máme množství často soupeřících, potenciálních příjemců zjištění z hodnocení, tedy skupin a jednotlivců, kteří jsou se sociálním programem nějak spojeni a kterým se v žargonu programové evaluace říká „stakeholdeři“ (viz např. Chen 2005, Greene 1998 aj.). „Stakeholdeři“, které bychom jinak (ač také cizími slovy) mohli nazvat „interesenty“, popřípadě též „participanty“, pokrývají velkou šíři

subjektů od tvůrců politiky a financovatelů, přes administrátory sociálního programu a řadový personál, až k těm, kterým je program určen a mají z něj profitovat, případně ke všem občanům. Bohužel je obtížné nalézt v českém jazyce srozumitelné slovo, které by dokázalo nahradit cizí pojem „stakeholder“, proto bude v tomto textu i nadále využíváno právě tohoto cizího slova.

3 Význam evaluace v kontextu rozvoje sociálního státu

Evaluační výzkum je podle Earla Babbie (1998) pravděpodobně stejně starý jako všeobecný sociálně-vědní výzkum sám. Kdykoliv lidé provedli nějakou společenskou reformu, museli dávat pozor na její aktuální důsledky. V posledních letech rozsah hodnotícího výzkumu získává na popularitě a aktivuje mnoho nových přístupů, což se odráží i v mnoha nových knihách, specifických vysokoškolských kurzech a konečně i na realitě samotného života projektů, programů a politik, které se již nadále bez systematické evaluace neobejdou. Je tedy jasné, že sociální vědci ve funkci hodnotitelů budou nosit svou kůži na trh mnohem více než kdykoliv předtím.

Nehledě tedy na relativně nedávný vznik evaluace sociálních programů (jakožto samostatného pojmu, jakožto systematické vědecké aktivity a někde dokonce jako samostatné vědní disciplíny), není evaluace, jakožto lidská aktivita, nikterak nová. Evaluace sama o sobě není novým zájmem ani v rámci sociálních věd. Aktivity např. úředníků a jejich odpovědnost veřejnosti, to vše bývalo předmětem velkého zájmu společnosti v celé lidské historii.

Jak píše Suchman (dle Nachmias 1980:2, vl. překl.): „Lidská potřeba vědět je úzce spojena s přáním soudit. Přirozená zvědavost, která vede lidi k otázce ‚proč?‘, je rovněž základem k tomu, co člověka vede k objevování ‚Cui bono?‘ neboli ‚jak je to dobré?‘ či ‚kdo z toho má prospěch?‘.“ Dokonce lze říci, že i samotný proces učení se a socializace jedince se děje na základě toho, co by mohlo být nazváno „evaluační zpětnou vazbou“ (viz předchozí kapitola).

Počátky evaluace sociálních programů mohou být vystopovány do úplných počátků moderní vědy někdy okolo roku 1600. Jak Cronbach a jeho kolegové (1980 dle Rossi, Freeman 1993) naznačují, Thomas Hobbes a jeho současníci se před čtyřmi stoletími zabývali tvorbou numerických měř k hodnocení sociálních podmínek a identifikaci příčin úmrtnosti, nemocnosti a sociální disorganizace.

3.1 Stručná historie evaluace

Ačkoli tedy historické kořeny samotné evaluace sahají až do 17. století, systematický evaluační výzkum je podle Rossiho, Lipseyho a Freemana (2004:8) relativně moderní

historií 20. století. Aplikace sociálně-vědních výzkumných metod na evaluaci sociálních programů souvisí s rozvojem a vylepšováním samotných výzkumných metod stejně tak jako s ideologickými, politickými a demografickými změnami v období „moderny“.

Poprvé se systematická evaluace sociálních programů objevila v oblasti vzdělávání a veřejného zdraví. Před první světovou válkou se nejvýraznější snahy zaměřovaly na posouzení míry gramotnosti a na profesní přípravu na zaměstnání či na aktivity v oblasti veřejného zdraví, které měly za cíl snížit míru úmrtnosti na infekční nemoci atd. Ve 30. letech 20. století sociální vědci využívali převážně přísné vědecké metody v tradici osvícenské doby a pozitivistické metodologie sociálních věd.

Obzvláště během 30. let 20. století (v souvislosti s velkou hospodářskou krizí) se objevilo nebo výrazně narostlo mnoho problémů, jako je nezaměstnanost nebo bezdomovectví. Když hovoří Weinbach (2005) o situaci v tehdejších USA, shrnuje, že programy, které na tyto problémy reagovaly, bývají souhrnně označovány jako „New Deal“ (Nový úděl). Některé z těchto programů byly nepochybně úspěšné, jiné byly úspěšné méně. Celkově se však stále zvyšoval požadavek evaluace těchto nákladných aktivit, obzvláště ze strany kritiků těchto programů. To tedy výrazně přispělo k rozvoji programové evaluace.

Během 2. světové války se objevila potřeba evaluace vojenských programů, které se staly po boku evaluace sociálních programů nejfrekventovanějšími typy evaluace. Během 50. let 20. století pak podle Weinbacha (2005) dále roste význam programové evaluace a roste i množství provedených evaluací jak vládních, tak nevládních sociálních programů. Ti, kteří se evaluaci věnovali, pak rozvinuli postupně lepší a lepší evaluační metody.

V období po druhé světové válce se dále objevila spousta nových veřejných sociálních programů v oblasti bydlení, vzdělávání, profesní přípravy a prevence v oblasti zdraví. Podle Rossiho, Lipseyho a Freemana (2004:8-9) se dá říci, že do konce 50. let 20. století se evaluace stala běžnou součástí veřejných politik. Sociální vědci se angažovali při posuzování preventivních programů v oblasti kriminality a delikvence, v oblasti psychoterapeutické a psychofarmakologické léčby, v oblasti bydlení a dalších. Rozvoj metod sociálního výzkumu, hlavně rozsáhlých výběrových dotazníkových šetření a pokročilých statistických metod a zároveň rostoucí finanční podpora a administrativní know-how, umožnil v té době provádět rozsáhlé evaluační studie.

3.2 Vývoj přístupů k evaluaci sociálních programů

Jak referují Rossi, Lipsey a Freeman (2004:9) v 70. letech 20. století se evaluační výzkum v podstatě vydělil jako samostatná tématická oblast v rámci sociálních věd. První samostatný odborný časopis, věnující se evaluaci - *Evaluation Review* - vznikl v roce 1976 v nakladatelství SAGE Publications. Další časopisy, které se věnují evaluaci, vydává dodnes v největším objemu právě zmíněné nakladatelství SAGE Publications v USA.

V raných stádiích byla evaluace formována výhradně sociálními vědci. V pozdějších dobách začali podle Rossiho, Lipseyho a Freemana (2004) do evaluace stále více promlouvat příjemci zjištění z evaluací. S tím jak evaluace dospívala, došlo také k obratu ke kvalitativním výzkumným metodám.

Zapojení uživatelů (stakeholderů) do programové evaluace posunulo tuto oblast od čistě akademické sociální vědy směrem k politické a manažerské aktivitě a evaluace se tak stala podle Rossiho, Lipseyho a Freemana (2004) signifikantním vstupem do komplexní mozaiky, ze které nakonec vyvstávají politická rozhodnutí a zdroje pro sociální programy. Jako taková musí tedy být evaluace sociálních programů, jak již bylo řečeno v úvodních úvahách této práce, chápána jako integrální součást sociální politiky či obecněji veřejné politiky.

Podle Shadishe a kol. (1991 dle Matoušek 2007) lze v moderní programové evaluaci rozlišit stádium (1) naivního instrumentalismu, stádium (2) pragmatické a konečně stádium (3) syntetické.

1) Naivní instrumentalismus

Stádium naivního instrumentalismu bylo ovlivněno pozitivistickou a kvantitativně orientovanou metodologií sociálních věd. V 60. a 70. letech 20. století formulovali první významní teoretici výchozí principy hodnocení sociálních programů převážně se zaměřením standardní vědecké metody. Jejich teorie počítaly podle Shadishe a kol (1991 dle Matoušek 2007:139) „...s autoritou vědeckého důkazu jako s kritériem, kterému se budou podřizovat všichni - od klientů přes pracovníky sociálních agentur až po úředníky a politiky rozhodující o existenci sociálních programů.“ V tomto období se předpokládalo, že výsledky evaluace budou sloužit „...k rušení neefektivních služeb, k revizi služeb stávajících i ke konstrukci služeb nových, adekvátnějších“ (ibid.). „Ideálem výzkumného designu byl v té době experiment s kontrolní skupinou. Pokud není možné mít kontrolní

skupinu, je přinejmenším možné srovnávat různé skupiny, s nimiž chtějí pracovníci různých agentur dosáhnout stejného cíle“ (ibid.).

2) Pragmatická etapa

Větší pozornost věnovanou využití výsledků evaluací přinesl podle Shadishe a kol. (1991 dle Matoušek 2007:142) konec sedmdesátých let. V daném období (souvisí i s ekonomickou krizí poloviny 70. let, která vyvolala potřebu rozsáhlejších reforem sociálních politik v západním světě a 80. léta jsou následně ve znamení tzv. thatcherismu či reganády) bylo uznáno, že „...sociální programy mají velkou setrvačnost a ke změnám v nich dochází pomalu a postupně. Studium faktorů podporujících změny se zdálo být novou, nadějnou cestou k prosazení výsledků hodnotícího výzkumu. Tyto faktory bylo možné studovat od nejnižší, lokální úrovně až po úroveň nejvyšší, centrální“ (ibid.).

V evaluaci sociálních programů klesla vědecká pravda „...z pozice absolutní autority na jednu z využitelných pravd. Od výzkumníka se již nečeká formulování konečného hodnotového soudu - ten je ponechán na subjektech, jež mají na existenci programu přímý zájem, tedy na lidech, kteří program řídí a užívají (ibid.). Jednu z čelních představelek evaluační metodologie Carol Weiss přivedly tyto úvahy „...ke skepsi týkající se možnosti uzavírat hodnocení programu jednoznačným výrokiem o jeho kvalitě. Výzkumné nástroje neumožňují objektivní syntézu očekávání všech subjektů, pro něž má program význam. Hodnocení programu se skoro vždy provádí v poli protichůdných zájmů, které by však výzkumník měl umět reflektovat“ (Matoušek 2007:143). Evaluační výzkum v tomto spíše skeptickém pojetí však může politiky „osvěcovat“, resp. poskytovat jim „...širší orientaci v problémech a ukazovat ...aktuální záležitosti na pozadí obecnějších zákonitostí formulovaných teoretiky sociálního výzkumu“ (ibid.).

3) Syntetické stádium

Třetí stádium přístupů k hodnocení programů nazývá Shadish a kol. (1991 dle Matoušek 2007:144) stádiem integrativním. Shadish „...varuje před tím, aby se hodnocení opíralo o zájmy jediné skupiny, která má na program zájem, byť by šlo o skupinu s největší rozhodovací pravomocí. Kvalitní studie se dá udělat spíše než formou jednotného designu jako soustava po sobě následujících menších studií, které zpřesňují výzkumné otázky a zohledňují zájmy různých skupin“.

3.3 Současný vývoj programové evaluace

Evaluace sociálních programů je úzce spojena také s rozvojem výzkumu veřejného sektoru, z čehož povstaly dnes hojně používaná odvětví sociálních či politických věd zvaná analýza politiky (policy analysis). Jak již bylo naznačeno výše, počínaje vládou R. Reagana (a M. Thatcher ve Velké Británii) v 80. letech docházelo k výrazným škrtům v rozpočtech určených na veřejné sociální programy. Politický vývoj v posledních desetiletích v Evropě a USA je spojen s vyvažováním nákladů a výstupů sociálních programů.

Se vstupem do 21. století můžeme podle Rossiho, Lipseyho a Freemana (2004:15) říci, že „...většině národních politik západní společnosti dominuje fiskální konzervatismus, omezování odpovědnosti státu za sociální programy a skepse ohledně velkých sociálních programů (důchodové zabezpečení atd.)“. Tyto trendy mají na programovou evaluaci různé dopady. Na jedné straně změny v sociálních programech vyžadují evaluaci, jelikož je potřeba porozumět dopadům provedených změn. Na druhou stranu se odpovědnost za provádění rozsáhlých evaluací přesunula na stát (ibid.). Každopádně však podle Rossiho, Lipseyho a Freemana (2004) lze tvrdit, že všechny tyto okolnosti posilují roli evaluačního výzkumu v soudobých společnostech.

Evaluace orientované na teorii pak patří k nejnovějším přístupům. Metodologové dnes často zastávají názor, že dosavadní modely byly příliš zatíženy problémem „černé skříňky“ (Hendl 2005:298). Černá skříňka je „...prostor mezi vstupem a výstupem programu. Metodologové se dříve zaměřovali především na interní a externí validitu posouzení programu a na měření vstupů a výstupů černé skříňky. Vnitřku černé skříňky věnovali menší pozornost“ (ibid.). To vedlo k následujícím důsledkům (dle Hendl 2005:299): (1) Aplikace programu předpokládala ideální podmínky a absolutní racionalitu. (2) Jestliže se mlčky očekávaly ideální podmínky, pak se politický kontext považoval za rušivou proměnnou, kterou mají na svědomí politici, a evaluátor nepociťoval potřebu věnovat tomuto aspektu pozornost. (3) Protože vnitřní fungování programu a jeho teorie byly stranou zájmu, šlo při evaluaci pouze o zjištění, zda program funguje, nebo ne. Opomíjela se prokázaná skutečnost, že evaluace málokdy vede k zastavení programu. Zanedbávaly se otázky zlepšení programu nebo přípravy nových programů. Opomíjel se význam teorie programu.

Současná evaluace veřejných politik vychází z dřívějších evaluačních pokusů v sociálních vědách v několika důležitých aspektech. Její primární zájem je vysvětlovat a předpovídat, spoléhá na empirickou evidenci a analýzu, klade důraz na budování středně rozsáhlé teorie a je záměrem, aby byla užitečná nejen pro tvůrce politik. Základním záměrem je tedy moderní programová evaluace, jakožto vědecko-výzkumná aktivita.

Je uznávána myšlenka, že kořeny programové evaluace jsou pevně zapuštěny v rámci výdobytků sociálních věd. Všechny oblasti sociální vědy přispěly k rozvoji metod evaluačního výzkumu. Není proto podle Berka a Rossiho (1990 dle Clarke 1999) překvapující, že nejlepší evaluační výzkum a nejlepší evaluátoři stavějí na množství disciplín, používají eklektický repertoár konceptů a metod. K teoretickým a metodologickým aspektům v samostatných kapitolách této knihy.

3.4 Rozvoj „Nového veřejného managementu“

Výše v textu již byla naznačena souvislost prudšího rozvoje programové evaluace mj. i s rozvojem výzkumu veřejného sektoru, resp. veřejné správy (veřejné administrace). Zvyšující se výskyt slova „management“ v popisu řízení veřejných záležitostí je podle Colebatche (2005) možno vystopovat zpětně až k reformní agendě šedesátých a sedmdesátých let. Je třeba se soustředit na požadované výsledky a jejich dosažením mohou být pověřeni manažeři orientovaní na cíl. Poté následovala rétorika devadesátých let s jejím znovuoobjevením konceptu „governance“ (vládnutí), tvrdící, že veřejní činitelé musí dostat autonomii pro nalezení inovativních odpovědí na veřejné problémy a že musí „kormidlovat a ne veslovat“, tj. poskytnutí služby zajistit, ne ji nutně osobně poskytovat.

Pojmoslovnou stánku věci v kontextu americké společnosti uvádí např. Gunn (1987:33 dle Colebatch 2005:72), když tvrdí, že v padesátých letech se převažující zájem o zkoumání veřejného sektoru týkal institucionální reformy a používal se termín „veřejná správa“. V šedesátých letech vyvstal nový zájem o plánování: racionalisté prováděli „strategická rozhodnutí“ spojená s výrazem „veřejná politika“. V osmdesátých letech se projevila snaha snížit veřejné výdaje a přijmout metody soukromého sektoru a běžnějším se stal termín „veřejné řízení“ či „veřejný management“.

Tzv. Nový veřejný management (New Public management – zkráceně NPM) „...zaváděný od poloviny 60. let v USA a od poloviny 80. let v Evropě kladl důraz na ekonomičnost, účinnost a efektivnost nákladů a organizací, nástrojů a programů sociálního

státu, jakož i na zvýšení kvality jím poskytovaných služeb“ (Horák a Kulhavý 2008:7). Těchto cílů mělo být dosaženo „...zavedením nového způsobu řízení, tzv. nové vlády (new governance), která představovala tendenci uvolnit všudypřítomnou kontrolu, zflexibilnit řízení managementu, měření výkonů, kontraktování a privatizace“ (Horák a Kulhavý 2008:7).

Tento přístup podle Colebatche (2005:73) posiluje orientaci na cíl, „...ale jinak dává manažerům více autonomie a volnosti v tom, jak těchto cílů dosáhnout. Na různých místech to může znamenat různé věci, protože manažeři vypracují k dosažení cílů různé způsoby, avšak to, co politikou skutečně bude, vzejde nakonec z místních jednání.

Pollitt (2003) identifikoval následujících osm klíčových elementů Nového veřejného managementu:

- posun v zaměření managementu od vstupů a procesů směrem k výstupům a výsledkům;

- posun směrem k měření a kvantifikaci, hlavně skrze rozvoj indikátorů výkonnosti a systémů benchmarkingu (nástroj strategického managementu systematicky srovnávající výkon jedné organizace v porovnání s jinými);

- preference specializovanějších „plošších“ a autonomnějších organizačních struktur;

- nahrazení formálních hierarchických vztahů (jak mezi organizacemi tak uvnitř nich) horizontálními kontrakty;

- podstatně vyšší využití trhu a tržních mechanismů v poskytování veřejných služeb;

- důraz na kvalitu služeb a orientace na zákazníka;

- rozšiřování a „rozmazání“ hranic mezi veřejným sektorem, tržním sektorem a tzv. třetím sektorem;

- posun v hodnotových prioritách od universalismu, rovnosti, zabezpečení, směrem k efektivitě a individualismu.

Podobně shrnují klíčové prvky Nového veřejného managementu Foster a Plowden (1996:45-46 dle Colebatch 2005:73):

- oddělení nákupu veřejných služeb od jejich poskytování,

- sloužit raději zákazníkům - klientům (v managerizačním pojetí se však mění i slovník, využívající spíše pojmu platící zákazník než klient) než byrokratickým, politickým, nebo výrobním zájmům,

- dávat přednost tržním cenám před daněmi,
- jsou-li subvence nutné, poskytovat je přímo a transparentně,
- rozšířit konkurenci,
- decentralizovat nabídku,
- podpořit komunity, aby mohly nabízet služby,
- stanovit cíle volněji a kontrolovat raději výstupy než vstupy,
- dosáhnout deregulace,
- raději problémům předcházet pomocí plánování než je následně napravit.

Přístup zvaný NPM svým způsobem podle Colebatche (2005:73) posiluje orientaci na cíl „například tvrzením, že pověřením lídři by měli určit cíle a dohodnout s úřady jejich realizaci, ...ale jinak dává manažerům více autonomie a volnosti v tom, jak těchto cílů dosáhnout“. Zmíněný přístup reflektuje důležitost lokálních přístupů k tvorbě i administraci welfare (blahobytu).

Můžeme tedy vidět, jak tento druh manažerismu koresponduje s komponenty nových forem vládnutí (governance). Obojí identifikují jako důležité (a) oploštění státu na komplexní síť vzájemných sítí a partnerství, (b) procesy vyjednávání a kontraktualizace, které jdou příčně přes sektory a linie odpovědnosti, (c) techniky sběru informací, které dávají zpětnou vazbu o měřitelných výstupech a aktivitách a (d) důraz na stakeholding, tedy pozornost věnovanou aktérům, kteří mají zájem na fungování sociálních programů (dle Fitzpatrick 2005:163). Podle Colebatche (2005:74) však není vždy jasné, „...zda diskuse o novém veřejném managementu popisují, co je nebo čeho bylo dosaženo, nebo obhajují, co by mohlo nebo mělo být. Mnoho prací nerozlišuje jednoznačně mezi empirickým a normativním popisem světa“.

4 Základní typy evaluace

Koncepty, jako jsou formativní evaluace, monitoring procesu a evaluace procesu jsou ve stávající literatuře často míchány dohromady, většinou pod pojem evaluace procesu. Nicméně ve skutečnosti jsou tyto evaluační přístupy odlišné ve smyslu odlišných premis, na kterých stavějí, odlišného typu dat, která sbírají, odlišného způsobu sběru dat a také odlišného způsobu, jak jsou informace zužitkovány. Tato kapitola se snaží uvést některé základní typy evaluace. Smyslem je zabránit výraznému zmatení pojmů a předejít tak situaci, kdy například „stakeholdeři“ budou vyžadovat jednu věc, jeden typ evaluačních zjištění, a hodnotitelé jim, na základě komunikačního nedorozumění o typu prováděné evaluace, budou poskytovat zcela jiný typ informací.

4.1 Formativní a sumativní evaluace

Michael Scriven bývá často považován za zakladatele jednoho ze základních rozlišení typů evaluace, když používá pojmy formativní a sumativní evaluace. Původně tyto pojmy použil k popisu dvou různých přístupů k evaluaci učebních plánů na školách. Tato dichotomie však byla posléze aplikována obecně na evaluaci sociálních programů.

Podle Scrivena (dle Clarke 1999:6-7) je formativní evaluace uskutečňována s cílem poskytnout zpětnou vazbu lidem, kteří se snaží něco zlepšit. Poukazuje to na určitý typ evaluačních aktivit, jejichž hlavním cílem je podpořit proces zlepšování sociálního programu. Následkem toho je ve formativní studii kladen důraz na identifikaci silných stránek a slabých stránek programu nebo intervence. V podstatě tento typ evaluace zahrnuje aktivity, kdy hodnotitel sbírá data o designu programu a jeho implementaci. Je kladen důraz na zkušenosti plánovačů programu, praktiků a participantů na programu. Hlavně je cílem zjistit, zdali je potřeba učinit nějaké změny pro zlepšení programu. Formativní evaluace tedy slouží ke zlepšení objektu evaluace „...pomocí zkoumání fungování politiky nebo programu, kvality implementace a organizačního kontextu...“ (Veselý, Paterová, Nekola 2007:82).

U tzv. „sumativní“ evaluace je primárním zájmem determinovat celkovou efektivitu (effectiveness) nebo dopad (impact) programu. Závěrem je doporučující aspekt, tedy zdali by program měl nebo neměl pokračovat. Jak vysledoval Patton (1986:66 dle Clarke

1999:7-8), sumativní evaluace mají tendenci být orientovány na závěr (conclusion-oriented), zatímco formativní evaluace mají tendenci být orientovány na akci (action-oriented). Souhrnná, resp. sumativní evaluace tak slouží k celkovému posouzení efektů politiky nebo programu.

S cílem podchytit základní elementy těchto dvou formálních přístupů k evaluaci je Herman a kol. (1987 dle Clarke 1999) prezentovali jako dva ideální typy situované na opačných koncích kontinua. Hlavní rysy formativní a sumativní dichotomie shrnuje následující tabulka.

Tabulka 2 - Formativní a sumativní evaluace – srovnání

	Formativní evaluace	Sumativní evaluace
Cíloví příjemci informací	programoví manažeři/ praktici	tvůrci politiky (policy-makers), financovatelé, veřejnost
Zaměření sběru dat	vyjasnění cílů, podstata implementace, identifikace výsledků (outcomes)	implementace, měření výsledků (outcomes)
Role hodnotitele	interaktivní	nezávislá
Metodologie	kvantitativní a kvalitativní (důraz na kvalitativní)	důraz na kvantitativní
Četnost procesu sběru dat	kontinuální monitoring	omezená
Procedury podávání zpráv	neformální, pomocí diskusních skupin a meetingů	formální zprávy
Četnost podávání zpráv	během fáze pozorování/ studia programu	po dokončení evaluace

Zdroj: Clarke 1999:8, adaptováno z Herman a kol 1987:26, vl. překl.

Na základě tohoto srovnání můžeme zdůraznit některé z hlavních odlišností mezi strategií formativní a strategií sumativní evaluace. Jak ukazuje tabulka, jsou zde jasně rozpoznatelné odlišnosti. Ty jsou obzvláště vidět v oblasti, jako je celková metodologická orientace, volba výzkumných metod, frekvence, s jakou jsou data sbírána, příležitosti k podávání zpráv o zjištěních z výzkumu a podstata vztahu mezi hodnotitelem a těmi, kdož jsou zaangażováni v programových aktivitách.

4.1.1 Formativní evaluace

V případě formativní evaluace jsou cílovou skupinou příjemců informací ti, kdo mají odpovědnost za plánování, řízení a realizaci opatření (intervence do životní situace klientů). Formativní hodnotitel pracuje úzce s programovým personálem, nerozvíjí jen porozumění tomu, jak je určitý program interpretován a implementován, ale rovněž je v pozici, kde poskytuje zpětnou vazbu o problémech a postupu práce a to vždy, kdy je potřeba, tedy kontinuálně. Funkce poskytovatele informací se stává hlavním rysem formativní evaluace. Hodnotitel při formativní evaluaci využívá jak kvalitativních, tak i kvantitativních metod a zaměřuje se na programové procesy. Podle Clarke (1999) je právě tento důraz na proces tím, co odlišuje formativní přístup od přístupu sumativního.

Myšlenkou stojící za takovouto formativní strategií je podle Pattona (1986 dle Clarke 1999) popsat, jak program aktuálně operuje, zvláště během svých raných fází. Úkolem hodnotitele, který pracuje v tomto módu, je zjistit, proč se věci dějí tak jak se dějí, jak se jednotlivé části programu k sobě navzájem hodí a jak program vnímají lidé. Primární záměr formativní evaluace je identifikovat směry, kterými by mohl být sociální program zlepšen. Proto je zpočátku důležité, že hodnotitel vyjasňuje zájmy, obsah a strukturu programu. Tyto kroky nejen, že identifikují různé komponenty programu, ale rovněž rozvíjejí porozumění tomu, jaká jsou očekávání od těchto různých elementů u různých skupin jedinců (manažeři, řadoví pracovníci, klienti...).

Jakmile jsou jednou programové mechanismy identifikovány, hodnotitel může rozkrýt rozsah, v jakém jsou dosahovány programové cíle. Pokud jsou určité aktivity shledány úspěšnějšími než jiné, mělo by to být oznámeno programovému personálu a program pak může být podle toho upraven. Díky své podstatě je formativní evaluační proces, který ve velké míře využívá hlavně kvalitativní techniky sběru dat, náročný na čas. Evaluátor musí udržovat úzký kontakt s programem, souvisle sbírat data a pravidelně poskytovat informace všem, kteří je mohou využít.

4.1.2 Sumativní evaluace

Jak bylo uvedeno výše, sumativní evaluace se primárně zaměřuje na determinaci efektivity opatření nebo naplánované intervence. Takovéto evaluace jsou obvykle objednány „tvůrci“ politik a financovateli, aby jim evaluace pomohla při rozhodování o budoucnosti určitého programu. Díky zaměření na zjištění rozsahu, v jakém program

dosahuje stanovených cílů, dominují kvantitativní metody a techniky sběru dat. Potřeba determinovat, zdali program pracuje nebo ne, znamená podle Clarka (1999), že sumativní hodnotitelé mají tendenci upřednostňovat experimentální a kvazi-experimentální výzkumný design (k podstatě experimentu a kvazi-experimentu viz metodologická kapitola).

Volba výzkumného designu a technik sběru dat nejsou jedinými oblastmi, kde je možné nalézt rozdíly mezi formativní a sumativní evaluační strategií. Výzkumníci, kteří se angažují v sumativní studii, přejímají spíše nezávislé a odloučené role v porovnání s formativními hodnotiteli, kteří častěji interagují s programovým personálem. V případě formativní studie je běžnou praxí hodnotitelů poskytovat informace a rady programovému personálu v pravidelných intervalech během monitorovací fáze, zatímco v čistě sumativní evaluaci nabývá zpětná vazba obvykle formu jedné, formálně sepsané zprávy, která je prezentována po dokončení evaluace. Jak vyzoroval Scriven (dle Clarke 1999), zatímco sumativní hodnotitel podává zprávu o programu (on a programme), formativní hodnotitel podává zprávu samotnému programu (to a programme).

Jaké jsou komparativní výhody a nevýhody obou zmíněných přístupů? Má formativní a sumativní evaluace stejnou důležitost? Mohou být kombinovány v rámci jediného výzkumného designu?

Když přijde na volbu mezi formativní a sumativní evaluací, budou podle Clarka (1999) někteří teoretici preferovat vždy buď jeden, nebo druhý přístup. Navíc sám Scriven, který s dichotomií formativní/sumativní přišel, přikládal sumativní evaluaci něco jako výsadní status. V jeho pohledu je primárním záměrem formativní studie před-připravit program pro sumativní hodnocení (dle Clarke 1999).

4.2 Evaluace procesu a evaluace dopadu

V evaluačním výzkumu jsou často odlišovány dva typy evaluace: (1) evaluaci procesu a (2) evaluaci dopadu.

Jak píše Nachmias (1980), evaluace procesu se zajímá o rozsah, v jakém je určitá politika (nebo program) implementována podle svých stanovaných propozic, zatímco evaluace dopadu se zajímá o přezkoušení rozsahu, kterým politika způsobuje změnu

v zamýšleném směru. Jinak řečeno je evaluace dopadu studií rozsahu, v jakém byly cíle dosaženy, zatímco evaluace procesu je studií toho, „jak“ byl program implementován.

V předchozí podkapitole jsme jasně rozlišili rozdíly mezi formativním a sumativním přístupem, nicméně je důležité nenakládat s dichotomií formativní/sumativní jako se synonymem k rozdílu mezi evaluací procesu (process evaluation) a evaluací výsledku (outcome evaluation). Patton (1982 dle Clarke 1999) podotýká, že díky nedostatku preciznosti v používání této terminologie, bývá formativní evaluace viděna jako výhradně na proces orientovaná (process-oriented) a sumativní evaluace jako jednoduše výsledkově orientovaná (outcome-oriented) aktivita. Scriven (1996 dle Clarke 1999) pak docela kategoricky tvrdí, že formativní evaluace nejsou druhem evaluace procesu a sumativní evaluace nejsou výhradně spojeny s výsledky. Ve skutečnosti se podle tohoto autora oba přístupy mohou zaměřovat jak na proces, tak na výsledek. Jinými slovy, lze takto uvažovat i o provedení kvantitativní (sumativní) evaluace, která poslouží jako nástroj pro dlouhodobé formování (proces implementace změn) organizace a rozvoj organizace. Obráceně také lze uvažovat o situaci, kdy převážně kvalitativní evaluace procesu implementace změn v organizaci je využito za účelem prezentace zjištění, čeho organizace docílila, tedy za účelem sumarizace výsledků.

Například Rossi a Freeman (1993) podotýkají, že evaluace dopadu (impact evaluations), které mají poskytovat odhad očekávaných výsledků navrhované intervence, mohou být vedeny jako součást celkové formativní evaluační strategie.

4.2.1 Evaluace procesu

Evaluace procesu je primárně zaměřená na rozsah, ve kterém byla určitá politika nebo program implementována podle svých proklamovaných osnov nebo záměru. To tedy zahrnuje periodické hodnocení fungování programu, s cílem detekovat odchylky od původních cílů, plánů a procedur. Podle Nachmiase (1980) tento formát často zahrnuje například audity finančních transakcí programu, účtovacího systému a přesnosti záznamů.

Význam evaluace procesu plyne podle Nachmiase (ibid.) z toho faktu, že původní programové cíle mohou být „až drasticky“ pozměněny (reformulovány), nebo dokonce negovány v průběhu implementace programu.

Tím, že se zaměřuje na změnu a na kontext implementace programu, nám může evaluace procesu ukázat, proč program dosáhl či nedosáhl určitých cílů a proč a jak by měly cíle být definovány a redefinovány během jejich vlastní implementace.

Jádrem přístupu k evaluaci procesu je poznámka, že implementace je souvislý sled procesů změn, který je iniciován už ve stádiu formulace programu. Pressman a Wildavsky (dle Palumbo, Sharp 1980, vl. překl.) definovali implementaci jako „...proces interakce mezi nastavením cílů a aktivitami vedenými k jejich dosažení.“ Podobně Berman (ibid.) zdůrazňuje, že implementace zahrnuje vzájemnou adaptaci mezi programem a organizačním nastavením pro implementaci. Žádná z těchto definic neříká, že cíle politik vždy budou implementovány způsobem, kterým byly původně pojímány. Obě definice indikují, že jsou cíle obecně adaptovány nebo změněny právě během implementace programu. Vhodnými otázkami pak může být: Jak mnoho změn můžeme během implementace očekávat? Dojde ke změnám v původních cílech, změnám v implementující organizaci a (pokud je politika úspěšná) ke změnám v cílové skupině, na kterou se program zaměřuje?

Proto například můžeme očekávat, že původní cíl programu se během implementace změní na specifitější a konkrétnější cíle. To bývá nazýváno procesem vyjasnění cílů (goal clarification). Vyjasnění se vyskytuje spolu s tím, jak implementující organizace přizpůsobuje program do jejích standardních operačních postupů a rovněž v odpovědi na tlak vytvářený na implementující organizaci různými vlivnými skupinami, mezi něž patří představitelé místních samospráv, tvůrci legislativy, neziskové organizace, klienti – obecně tedy „stakeholdeři“.

Proces proměny cílů může obsahovat i možnost substituce cílů jinými cíli, které jsou v protikladu s těmi originálně uvažovanými. Vyskytují se i změny způsobené nevhodnou nebo neadekvátní implementací. To vše bývá nazýváno procesem rozkladu cílů (goal dissolution).

4.2.2 Evaluace dopadu (impact, outcome evaluation)

S evaluací dopadu se budou pojit jistá omezení a potíže. První je podle Palumba a Sharpové (1980) v zásadě v tom, že evaluace dopadu by neměla být prováděna za absence evaluace procesu. S vedením evaluace dopadu totiž vyvstává potřeba jasně definovaných cílů evaluovaného programu a navíc ještě existence kritérií úspěchu

a neúspěchu. Otázkou tedy je, jak vést evaluaci dopadu programu, který má vágní nebo navzájem protichůdné programové cíle? Slovy zmíněných autorů, evaluace dopadu „...volá po existenci operacionalizovaně definovaných cílů politiky se specifikací kritérií úspěchu“ (Palumbo, Sharp 1980:289, vl. překl.). Na tuto problematickou oblast by mohla dát odpověď evaluace procesu, protože když sledujeme proces implementace, můžeme vidět specifické významy, které jsou dávány cílům programu různými skupinami zainteresovaných aktérů a jak se tyto subjektivní významy postupně mění v čase s tím, jak je sociální program postupně implementován.

Za druhé je problém s interpretací výsledku evaluace dopadu. Evaluace dopadu, která říká, že stanovených cílů nebylo dosaženo, nám neříká, zdali je selhání způsobeno chybnou teorií (chybnými předpoklady), na které je program vystavěn, nebo „slabou“, resp. nedostatečnou implementací. Podle Palumba a Sharpové (1980) může být pouze pomocí evaluace procesu podána odpověď na podobnou otázku, tedy zdali je selhání na straně východisek programu, nebo na straně „slabé“ implementace. „Bojujeme zde za to, aby evaluace dopadu nebyla vedena bez doprovodné evaluace procesu. Evaluace dopadu je statickou analýzou, analogickou k momentce získané v určitém okamžiku v čase. Reprezentuje situaci v tomto okamžiku, ale pokud je obrázek získán v čase, kdy se okolnosti rychle mění, pak dává falešný obraz o tom, co se děje nebo o tom, jaký bude konečný dopad programu. Evaluace procesu je na druhou stranu dynamickou analýzou, která umožňuje hodnotiteli sledovat změny tak, jak se postupně vyskytují a identifikovat bod, kdy se program stává institucionalizovaným a změny ustávají. V tomto čase by teprve měla být vedena evaluace dopadu, protože pouze pak může být měřen finální dopad“ (Palumbo, Sharp 1980:289-290, vl. překl.).

Když to shrneme, evaluace dopadu může ukázat, zdali program uspěl nebo selhal, ale není nastavena tak, aby nám ukázala „proč“ program selhal (nebo zdali program uspěl pouze proto, že programové cíle byly v průběhu implementace modifikovány).

4.3 Zvláštní případ evaluace: Evaluace bez zaměření na cíle

V klasickém modelu evaluace je zaměřena pozornost na zamýšlené služby a výsledky programu – na jeho cíle. Nicméně hodnotitel může podle Scrivena (dle Patton 2002:169) získat velmi zajímavé výsledky, když započne svou terénní práci bez znalosti cílů programu nebo alespoň bez zjevné snahy nastavit design evaluační studie za účelem

posouzení naplnění cílů programu. Michael Scriven (1972 dle Patton 2002) poprvé představil ideu „bezcílné“ evaluace (goal-free). Znamená pro něj provedení terénní fáze sběru informací o širokém okruhu aktuálních výsledků a dopadů programu a pak komparování zjištěných výsledků s aktuálními potřebami participantů programu (klientů). Hodnotitel se snaží co nejvíce potlačit rétoriku zaměřenou na cíle programu, v rozhovorech s programovým personálem nejsou řešena témata cílů programu, nejsou studovány dokumenty či brožury, ve kterých jsou cíle uvedeny, jsou studovány jen pozorovatelné dopady a dokumentovatelné efekty ve vztahu k potřebám participantů (klientů). Podle Scrivena (dle Patton 2002:169) existují čtyři základní důvody pro provedení „bezcílné“ evaluace:

1. zmírnit riziko prostého přímého studia stanovených cílů programu a tím i riziko nezachycení důležitých nezamýšlených výsledků;
2. zmírnit negativní konotace, které jsou spojeny s odhalením nezamýšlených účinků...;
3. eliminovat zkreslení, se kterým hodnotitel vstupuje do evaluace tím, že zná cíle programu;
4. posílit nezávislost hodnotitele tím, že se potlačí jeho závislost na znalosti cílů programu, které by mohly omezovat nezávislost hodnotitele v kvalitativním výzkumu.

Scrivenovými vlastními slovy (1972:2 dle Patton 2002:170) pak lze „bezcílnou“ evaluaci charakterizovat následujícím způsobem: „Jeví se mi to ve stručnosti tak, že zvažování a evaluace cílů programu je nadbytečný a možná i kontaminující krok... Čím méně hodnotitel slyšel o cílech projektu, tím méně se u něj rozvine tunelové vidění a tím více pozornosti může věnovat hledání skutečných efektů (spíše než jen kontrolovat údajné efekty)“.

„Bezcílná“ evaluace ve svém hledání aktuálních efektů zapojuje induktivní a holistickou strategii a vyvarovává se tak logicko-deduktivním omezením, která jsou vlastní běžnému, na cíle orientovanému přístupu k evaluaci. „Bezcílná“ evaluace je radikálním odklonem od běžného tradičního smýšlení o evaluaci. Nicméně náznaky podobného přístupu, čili volání po odklonu od přehnaného zdůrazňování cílů vidí např. Patton (2002:170) i u některých prominentních autorů jako je např. Peter H. Rossi (Rossi, Williams 1972:18), který tvrdil, že „...sociální program (a ostatně jakýkoli program), který nemá jasně specifikované cíle nemůže být vyhodnocen bez specifikování nějakých měřitelných cílů. Toto tvrzení je zřejmým truismem“. Carol Weiss (Rossi, Williams

1972:24-26) zdůrazňuje klíčovost cílů v evaluaci, když tvrdí: „Cíle musí být jasné, takže evaluátor ví na co se zaměřit... Tím začíná dlouhý a často bolestný proces dolování jasného, specifického a měřitelného cíle z lidí“.

Obzvláště kompatibilní s „bezcílnou“ evaluací je pak podle Pattona (2002:170) kvalitativní zkoumání, jelikož vyžaduje zachycení přímé zkušenosti participantů programu z jejich vlastní subjektivní pozice a v jejich přirozeném jazyce a pojmosloví. Hodnotitel tak může být otevřen jakýmkoliv údajům, které vzejdou ze zkoumaného fenoménu (programu). Je ale důležité podotknout, že „bezcílná“ evaluace může využívat podle Scrivena (dle Patton 2002:170) jak kvalitativní tak kvantitativní metody. Scriven dále uvádí, že „bezcílná“ evaluace může být provedena paralelně k tradiční, na cíle orientované evaluaci, ovšem jinými hodnotiteli.... Jde o maximalizaci silných stránek a minimalizaci slabých stránek obou přístupů.

5 Role teorie v evaluaci

Cílem této kapitoly je uvést přehled teorií o evaluaci a teorií v rámci evaluace formou tzv. „stromu“ evaluačních přístupů (teorií), která se stává z metodologicky orientované větve, arbitrárně orientované větve a pragmaticky orientované větve a u všech uvést hlavní představitele jednotlivých myšlenkových a metodologických škol.

V obecném smyslu hraje teorie v evaluaci sociálních programů mnoho rolí a je v evaluačním procesu „vitálním elementem“. Například může hodnotitelům poskytnout podklady pro výběr konkrétních výzkumných metod a metodologických přístupů, a to v různých myslitelných evaluačních kontextech. Teorie může rovněž pomoci zaměřit evaluaci správným směrem tím, že vede hodnotitele k hlavním a klíčovým otázkám a problémům. Celé téma teorie evaluace je v literatuře často přehlíženo. Mnohé texty o evaluaci ukazují programovou evaluaci jako ateoretickou, na metody orientovanou záležitost.

5.1 Teorie „o evaluaci“ a teorie „v rámci evaluace“

Když odkrýváme vztah mezi teorií a evaluací, je důležité rozlišit mezi teorií o evaluaci (about evaluation) a teorií v rámci evaluace (in evaluation). Prvním jmenovaným případem se zabývají Shadish a kol. (1991 dle Clarke 1999) v jejich přehledu některých hlavních teorií, které podkreslují praktickou programovou evaluaci.

Podle těchto autorů nám evaluační teorie říká kde, kdy a proč by měly některé výzkumné metody být aplikovány a jiné ne. Podotýká, v jaké posloupnosti by měly být metody aplikovány, naznačuje různé cestičky, kterými mohou být metody kombinovány, typy otázek, na které se dá více či méně dobře odpovědět pomocí určitých metod a také jaké výhody můžeme očekávat od určitých metod ve srovnání s jinými. Evaluační teorie jsou něco jako „válečné strategie a taktika“ (Shadish a kol. 1991 dle Clarke 1999), metody jsou něco jako „válečné zbraně a logistika“ (ibid.).

Jasná a srozumitelná teorie evaluace by měla obsahovat pět hlavních komponent. Autoři (ibid.) je popisují jako sociální programování (social programming), konstrukce znalosti (knowledge construction), užití znalosti (knowledge use), ohodnocování (valuing) a praxe (practice).

V zásadě pojem evaluační teorie referuje o teorii, která je aplikována na aktuální evaluační praxi. Jinými slovy, je to teorie o tom, jak dělat evaluaci. V tom je rozdíl oproti tomu, když se zaměřujeme na programovou teorii v evaluaci (in evaluation), kde je důraz kladen na specifikování toho, jak se předpokládá, že určitý program nebo intervence funguje.

Carol Weissová (dle Chen 2005) byla jednou z prvních teoretiček evaluace, která vnesla otázku teoreticky zaměřené (theory-focused) programové evaluace. Přístup vedený teorií (theory-guided) je často citován také jako theory-driven přístup (Chen a Rossi 1981 dle Chen 2005) nebo theory-based evaluace (Weiss 1995 dle Chen 2005). V praxi není teorie stojící za programem vždy zcela explicitní. Není totiž ani zcela jasně artikulována těmi, kdo program plánovali a vytvářeli jeho design. Jeden z prvních úkolů, se kterým se tedy hodnotitel setkává, je identifikovat teoretické předpoklady, které podkreslují plánovanou intervenci. Program nemůže být dobře vyhodnocen, dokud neporozumíme jeho podkreslujícím teoretickým základům.

Tento důraz na teorii má jasné dopady na druh role, kterou pak hodnotitel zastává. V rámci přístupu vedeného teorií (theory-driven), hodnotitel aktivně hledá a konstruuje teoreticky ověřený model sociálního problému proto, aby mu to pomohlo specifikovat mechanismy, kterými může program způsobovat změnu v sociálním problému. Proto je rozvíjení programové teorie krokem směrem k vyjasnění vztahů mezi programovými aktivitami a jeho efekty. Podle Chena (2005) pomáhá používání teorie takovýmto způsobem strukturovat evaluaci.

Chen (ibid.) tedy obhajuje nahrazení přístupu orientovaného na metody (method-oriented) perspektivou theory-driven. Evaluační výzkum je tedy do značné míry snahou o testování teorií (theory-testing), ale jak hodnotitel dospěje k programové teorii? Weissová (1997 dle Clarke 1999) identifikuje čtyři základní zdroje informací, které pomáhají při vytváření a specifikaci programových teorií. Jsou jimi: (1) dokumenty, (2) lidé, (3) dřívější výzkumy a (4) logické zdůvodnění.

Za prvé může hodnotitel získat vhled do podstaty programu díky studiu dostupné dokumentace. Do této kategorie spadají vnitřní zprávy, zprávy ze schůzí, žádosti o financování a literatura šířená v organizaci v souvislosti s programem. Nicméně tyto materiály mají pouze omezenou užitečnost, jelikož programové teorie nejsou vždy uváděny zcela explicitně.

Mnoho hodnotitelů tedy započíná svou, na teorii založenou (theory-based), evaluaci na rozhovorech s administrátory programu, praktiky, rozhodovateli a plánovači. Předchozí výzkumná zjištění jsou třetím možným zdrojem informací, které jsou hodnotiteli běžně dostupné. Čtvrtý zdroj základních informací pro vytváření programové teorie nachází Weissová v logickém zdůvodňování. Prvek tohoto logického usuzování lze nalézt v ostatních třech zmiňovaných informačních zdrojích.

Celkově vzato má programová teorie množství funkcí v rámci evaluačního procesu. Pomáhá identifikovat klíčové evaluační otázky, stanovit jaké výzkumné metody jsou nejvhodnější a vyjasňovat podstatu a obsah vztahu mezi programovými aktivitami a výsledky. V následující podkapitole se podíváme na pár základních evaluačních metodologií a na jejich východiska. Zajímavé vodítko k rozdělení hlavních evaluačních proudů poskytl Marvin C. Alkin a kol., když sestavili „strom evaluačních teorií“.

5.2 „Kořeny“, „kmen“ a „větve“ evaluačních teorií

Jak vyplývá z předchozího textu, mnozí teoretici založili své teorie na odlišných premisách, nicméně všechny evaluace mají kořeny v metodologii sociálně-vědního výzkumu a v sociálních vědách vůbec. Lépe porozumět evaluačním teoriím by podle Alkina (2004) mohlo pomoci sledovat cesty, jakými tyto evaluační „kořeny“ vyrostly v průběhu času do evaluačního stromu, kde různé větve reprezentují odlišná zaměření, odlišnou orientaci zastánců těchto teorií.

Všechny níže uvedené teoretické přístupy je možné využít při hodnocení sociálních programů. U každého přístupu je však akcentován jiný prvek. Jaká rizika takový důraz na určité prvky programové evaluace s sebou může přinést, to se snaží předjímat až závěr této práce.

Jak komentuje Alkin (2004) toto vysoké množství stávajících teoretických přístupů? „Jeví se mi to tak, že zde ve skutečnosti máme mnoho akceptovatelných způsobů, jak uchopit evaluaci. Někteří teoretici jsou přesvědčeni, že jejich přístup je tou pravou ‚Nirvánou‘. Já si to nemyslím. Určitou hodnotu má každý z těchto modelů. Při zkoumání pohledů různých teoretiků na programovou evaluaci čtenář často narazí na nejednu kapitolu, kdy se přistihne, že si v duchu říká: ‚ano, to dává smysl‘. Myslím si, že čtenáři a praktici udělají nejlépe, když budou sledovat určitý model, který jim dává největší smysl“ (Alkin 2004:x, vl. překl.).

Obecně vzato Alkin (ibid.) doporučuje, že je nejlepší mít základní definující strukturu, základní model nebo teorii, které se je dobré podržet, ale v zásadě by měl být hodnotitel potenciálně adaptabilní. Alkin zařazuje na „strom evaluačních teorií“ jen určitým způsobem vybrané osoby. Je jasné, že všichni, kteří o evaluaci píší, nemohou být považováni za ty, kdo rozvinuli unikátní evaluační teorii. I když i jejich příspěvek může být významný, nejsou teoretiky nebo tvůrci teorie. Alkin a Christina A. Christie vybírali ke kategorizování klíčové autory v oblasti evaluace čtyřmi způsoby:

1) Někteří autoři přispěli významně k základní metodologii výzkumu, která formuje důležité prameny pro další práci v oblasti evaluace. O těchto osobách referují jako o „metodoloziích“, kteří ovlivnili evaluaci.

2) Další kategorie autorů jsou analytici evaluačních otázek, problémů. Jejich práce, i když se třeba netýká konkrétně evaluačních modelů nebo teorie, významně přispěla k porozumění různým aspektům evaluace.

3) V další kategorii autorů identifikují evaluační interpretátory a učitele v oblasti sociálních věd a sociálně-vědní metodologie. Autoři v této kategorii mohou být rovněž prominentními analytiky evaluace (a mnoho z nich skutečně je), ale odlišují se tím, že píší učebnice evaluace, které, i když nerozvíjejí specifickou novou teorii, poskytují důležitý zdroj pro výuku evaluace sociálních programů a pomáhají interpretovat jemné nuance.

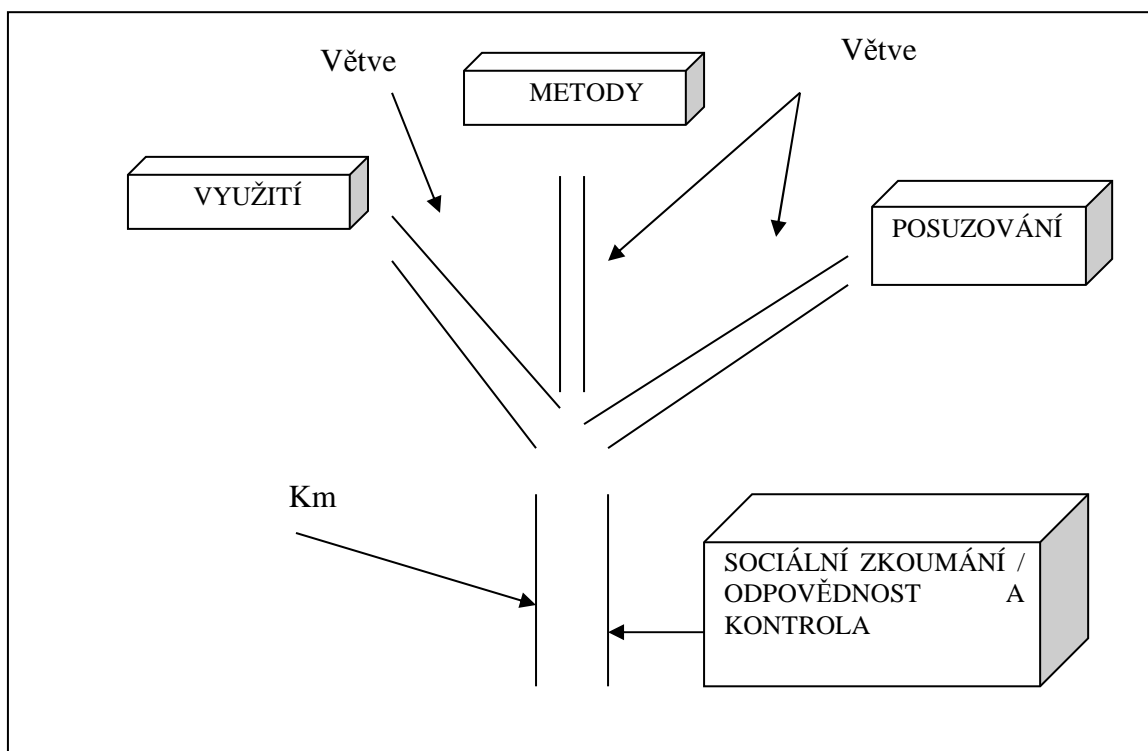
4) Dále jsou chápáni jakožto evaluační teoretici i autoři, kteří jsou spojeni pouze s určitou teoretickou evaluační pozicí. Mnoho z nich nemůže prezentovat ucelenou teorii, ale poskytli určitý specifický evaluační směr, který dále formuje programovou evaluaci.

Je třeba mít dále na paměti, že mnozí autoři podléhají silnému vnitřnímu vývoji v čase a jejich názory, které prezentovali v minulosti, již dnes neplatí a sami se k nim nehlásí, či je přímo negují. Pohledy různých teoretiků tedy nejsou fixovány v čase, jak by se mohlo zdát z čtení jejich publikované práce. Teoretici typicky mění své pohledy na věc v průběhu času a právě publikované práce často za těmito změnami zaostávají. To je podle Alkina (2004:7) zřetelné např. na práci Egona Guby, kdy Guba sám v reakci na kritický článek na jeho osobu článek sebereflexivně píše: „...v mnoha směrech nejsem ten Egon Guba, o kterém je řeč... Ten Egon Guba byl tím, který pracoval a psal před sedmi až deseti lety. Jestliže měl někdo problémy porozumět tomu, jaký byl tehdejší Egon Guba, sdílím s ním v současnosti stejný problém; v mé mysli se změnilo příliš mnoho věcí, takže je pro mne obtížné vyvolat zpět názory, které jsem tehdy měl.“

Přestože se smýšlení a názory konkrétních teoretiků v čase proměňují, přesto jejich pohled na věc (jedno zdali je nebo není stále aktuální) ovlivnil jiné teoretiky. Takže ačkoli se např. Guba podle Alkina (2004) zřekl sebe samého, jakým byl v minulosti, bylo jeho rané smýšlení velmi důležitou částí vlivu na jiné teoretiky evaluace. Různí teoretici stavěli na jeho myšlenkách (a pochopitelně i na myšlenkách jiných), bez čehož by dnes neměli myšlenky, jaké mají.

Alkin a Christieová (2004) tedy prezentují vztahy mezi teoriemi (resp. přístupy k evaluaci) ve formě evaluačního „stromu“ (viz obr. níže), kdy hlavní větve reprezentují jednu z těchto hlavních teorií: (1) větev využití, zužitkování (use), (2) větev metod (methods) a (3) větev posuzování/hodnocení (judging/valuing).

Alkin uznává, že tyto tři větve jsou „drastickým“ zjednodušením velmi komplexních vztahů, přesto je považuje za přínos pro utřídění a zvýraznění odlišností v teoriích evaluace.



Obrázek 1 - Strom evaluačních teorií

Zdroj: Alkin 2004:13, přeloženo a zjednodušeno autorem

Kmen stromu tvoří dvojí základy: odpovědnost a systematický sociální výzkum. Hlavní větev stromu (prostřední) je pokračováním kmene sociálního výzkumu. Je to evaluace vedená způsobem, který zdůrazňuje důležitost výzkumných metod. Tato větev je tedy nazvána „metody“. Další větev je nazvána posuzování či ohodnocování (valuing). Autoři, reprezentující tento přístup, zastávají názor, že přisuzovat údajům hodnotu je možná tou nejzásadnější komponentou práce hodnotitele. Třetí hlavní větev je nazývána „využití“ (use) a je charakteristická zájmem o to, jak budou evaluační informace použity a jak budou zaměřeny na ty, kdo budou dané informace používat.

Alkin (1972a in Alkin 2004) v textu, který definuje odpovědnost, hovoří o odpovědnosti k cílům, odpovědnosti k procesu a odpovědnosti k výsledku (outcome). Odpovědnost k cíli zkoumá, zdali byly stanoveny smysluplné a vhodné cíle. Rozhodovatelé a „horní linie“ managementu jsou odpovědnými jednotkami právě za tento druh odpovědnosti. Odpovědnost za proces reflektuje, zdali byly ustanoveny a implementovány vhodné procedury k dokončení těchto cílů. Odpovědnost tohoto druhu typicky nese management sociálního programu. Odpovědnost za výsledek (outcome) referuje o rozsahu, v jakém bylo dosaženo ustanovených cílů. Za výsledky jsou, podobně jako za proces, odpovědní manažeři sociálního programu.

5.2.1 Větev „Metody“

Většina teoretiků evaluace má na prvním místě metodologické zájmy a vidí výzkum jako podstatu vzniku programové evaluace vůbec. Nicméně jedna skupina teoretiků je obzvláště spjata s touto orientací. V sociálních vědách a psychologii tento důraz na výzkum závisí na dobře nastavených experimentálních studiích. Základními pracemi jsou pro tyto teorie rané práce Donalda Campbella (1957) a zvláště pak publikace Campbella a Stanleyho (1966), která definuje podmínky pro vhodný experimentální a kvazi-experimentální design (dle Alkin 2004).

Nicméně, než se budeme věnovat Campellovi a jeho následovníkům, je důležité zmínit ještě starší práci, která pomohla ustanovit pracovní rámec pro praxi evaluace, obzvláště v oblasti vzdělávání, práci Ralpa Tylera.

Alkin a Christieová (2004) vnímají práci Ralpa Tylera jako jeden z hlavních zdrojů pro moderní programovou evaluaci (pokud jsme schopni z dnešního pohledu přijmout 40. léta 20. století za moderní – zde použité slovo „moderní“ však se sociologickým pojetím

„moderny“ nemá nic společného). Záměrem evaluace je u Tylera validizovat hypotézy. Tylerův pohled na věc se stal známým jako úkolově-orientovaná evaluace (objectives-oriented nebo objectives-referenced). Tento přístup se zaměřuje na: (a) formulování tvrzení o vzdělávacích úkolech, (b) klasifikaci těchto úkolů (objectives) do hlavních typů, (c) definování a redefinování každého typu úkolů do konkrétního chování, (d) identifikování situací, ve kterých se od studentů dá očekávat, že budou vykazovat tyto typy chování, (e) selekci a zkoušení metod jak získat znalost vztaženou ke každému typu úkolů, (f) selekce úkolů na základě předchozích pokusů jiných metod pro další rozvoj a zlepšení a (g) nalezení smyslu pro interpretaci a použití těchto výsledků (Tyler 1942:498-500 dle Alkin 2004:18). Práce Ralpa Tylera, byť zaměřená úžeji na vzdělávací systém, ovlivnila řadu významných teoretiků evaluace a lze ji tedy chápat jako základ „větve“ evaluačních teorií zaměřených výrazně na sociálně-vědní metody a statistické procedury. V této linii pak pokračují další významní teoretici, mezi nimiž zaujímá přední postavení obzvláště Donald Campbell, jehož práci pak další autoři rozvíjejí.

5.2.2 Větev „posuzování“/„ohodnocování“ (Valuing)

Z kmene sociálního výzkumu vyrůstá větev teoretiků evaluace, kteří se zaměřují na ohodnocování (valuing) nebo na vytváření soudů (judgments). Teoretici na této větvi věří, že to, co odlišuje hodnotitele od jiných výzkumníků je, že hodnotitelé musí vytvářet hodnocení (value) na základě svých zjištění a v některých případech stanovit, které výsledky zkoumat. Tato větev evaluace má své základy u Michaela Scrivena.

5.2.3 Větev „Využití“ (Use)

Větev orientovaná na využití zjištění z evaluace (use) má své základy v „teoriích orientovaných na rozhodnutí“ (decision-oriented theories). Tato kategorie teorií je spojena s designem evaluací, které jsou zaměřeny na informování o výsledcích, což je využitelné při dalším rozhodování. Evaluační výsledky tak mají přímý dopad na programové rozhodování a organizační změnu.

5.3 Závěrečná poznámka ke kapitole o stromu teoretických přístupů k evaluaci

Jak vidno z předchozího textu, mnoho současných autorů se nutně musí zabývat otázkou „využití“ evaluace a otázkou důvěryhodnosti zainteresovaných aktérů k této evaluaci. Vystává otázka, zdali je tedy možné sledovat pouze vědeckou přesnost evaluace. Dokonce i autoři, kteří jsou stále zařazováni na výše uvedenou větev „metod“, která zdůrazňuje právě vědeckou přesnost, jsou nuceni stále více prostoru ve svých přístupech věnovat otázce „využití“ evaluace a přizpůsobování designu evaluace těm, kteří jsou nazýváni „stakeholdery“.

Myslím si tedy, že můžeme v současné době sledovat určitý odklon od větve „metod“ a stejně tak od větve „posuzování“ směrem k větvi „využití“. Tento trend může však do značné míry vést k riziku ztráty vědecké objektivity (pokud však kdy vůbec existovala). Existuje riziko ztráty autonomie evaluátora, který se tak může stát přeneseně za výsledky samotného programu odpovědným zmíněným „stakeholderům“ a z lidsky pochopitelných důvodů se snažit prezentovat program v lepším světle, silou interpretace zlepšovat programové výsledky. Je jedno, zdali tak učiní vědomě či podvědomě, samotný fakt narušení autonomie a vědecké přesnosti zde figuruje v obou případech.

6 Kvantitativní evaluace sociálních programů

Cílem této kapitoly je popsat hlavní smysl zapojení kvantitativní výzkumné strategie do hodnocení sociálních programů. Dále je cílem seznámit se se základními pojmy jako je experiment a kvaziexperiment a základními technikami sběru dat. Kapitola však není zamýšlena jako čistě metodologická, na to jsou výzkumné metody popsány jen velmi stručně a publikace o teorii evaluace sociálních programů nemá suplovat metodologické učebnice v sociálních vědách. Spíše se jedná o základní vymezení dvou velkých paradigmat v sociálně-vědním výzkumu, ze kterých pak, ať už výhradně z jedné či ze vzájemné kombinace obou, vždy evaluace sociálních programů vychází.

6.1 Typy dat – kvantitativní a kvalitativní

Rozhodně je historicky důležité rozdělení na kvantitativní a kvalitativní data, byť tento dualismus přinesl v historii sociálních věd spoustu debat a sporů ohledně jejich využití (v souvislosti s kvalitativním výzkumem se v českém jazykovém prostředí často používají pojmy „informace“ nebo „údaje“ spíše než „data“, v následujícím textu však bude slovo „data“ někdy používáno i v souvislosti s kvalitativním výzkumem – v závislosti na tom, jaké pojmy původně uvádí různí citovaní autoři).

Kvantitativní metody v zásadě vytvářejí data, která mohou být prezentována v numerické podobě a mohou být předmětem různých typů statistických analýz, zatímco kvalitativní metody produkují údaje, které nejsou vhodné pro statistické vyhodnocení. Jak poznamenávají Miles a Huberman (1984 dle Clarke 1999) hlavní rozdíl mezi těmito dvěma přístupy je, že jeden pracuje s čísly, zatímco druhý je spojen se slovy.

Konvenční obrázek výzkumníka pracujícího v tradici kvantitativní je ten, že jde o odloučeného „vyšetřovatele“, který používá standardizované měřicí techniky, které jsou vytvořeny ke sběru objektivních fakt. Tato fakta jsou pak přístupná statistickým analýzám. V případě kvalitativního výzkumu je hlavním nástrojem sběru údajů samotný výzkumník a udržování odstupu od údajů není možné. Kvalitativní data jsou používána k získání detailů o subjektivních zkušenostech lidí, které je možné v kontextu evaluace sociálních programů označit např. jako „plánovače“ programu, manažery, stakeholdery a další.

6.2 Ztraceno v překladu: Podstata empirického sociálně-vědního výzkumu

Přední český sociolog Miroslav Petrusek (1993:76-77) poukazuje na „překladatelskou“ podstatu sociálně-vědního výzkumu. Překladatelská podstata empirického výzkumu má podle něj podobu, která se dá přehledně shrnout asi takto:

1) „P/z jazyk: přirozený jazyk komunikace se zadavatelem. Proměnou sociálních funkcí sociologie, která přestává být akademickou, či jen akademickou vědou, se mění i sféra mimovědní komunikace sociologa. Sociolog musí komunikovat s tím, kdo výzkum objednává a platí (zadavatel, klient).

2) T jazyk: teoretický jazyk sociologie jako vědy. Tento jazyk je sám základem intraparadigmatické plurality (existuje více než jeden T jazyk); uvnitř téhož paradigmatu funguje jako jazyk, v němž je formulována výchozí teorie, hypotézy a v němž jsou interpretována empirická data.

3) T/h jazyk: teoretický jazyk hypotéz, který se liší od jazyka čisté teorie tím, že je na jedné straně bližší přirozenému jazyku (hypotézy jsou bližší ‚zdravému rozumu‘ než ‚čistá teorie‘), ale na druhé straně je podřízen diktátu potenciálních ‚tvrdých dat‘, jež v duchu scientisticky koncipované sociologie jediná mohou verifikovat hypotézu. Jazyk hypotéz je v tomto smyslu značně ‚hybridní‘.

4) O jazyk: jazyk operacionalizačních procedur, přesněji jazyk, v němž jsou vyjádřeny operacionalizované hypotézy, jež mají být přímo verifikovány empirickým výzkumem. V podstatě jde o observační jazyk ve smyslu Carnapově.

5) P/r jazyk: přirozený jazyk komunikace s respondentem. V případě, že základní metodou sběru dat je dotazník, jde o jazyk dotazníku, v případě, že jde o rozhovor, jde o jazyk rozhovoru, který pak nutně zahrnuje i paralingvistické fenomény.

6) C jazyk: jazyk počítačových procedur (od computer).

7) S jazyk: matematicko-statisticky kondenzovaná informace, která ‚vyjde‘ z počítače. V podstatě jde o tabulky, grafy, diagramy, histogramy, korelační a jiné koeficienty, které jsou v následující fázi předmětem překladu do přirozeného jazyka.

8) T/i jazyk: teoretický jazyk interpretace dat, jež jsou vyjádřena v S jazyce. V této fázi ještě není ‚výstup‘ výzkumu vyjádřen ve ‚zcela‘ přirozeném jazyce, ale spíše v jazyce teoretickém.

9) P/z jazyk: přirozený jazyk výzkumné zprávy, tj. v podstatě opět jazyk komunikace se zadavatelem.“

Podle Petruska (1993:78) „...je zjevné, že nutných překladových procedur je v průběhu empirického výzkumu nejméně šest! Konkrétně: překlad z jazyka zadavatele do jazyka teorie (tento krok může variovat podle povahy výzkumu), překlad z jazyka teorie do jazyka výzkumu (operacionalizace), překlad z jazyka primárních dat do jazyka počítačových procedur (kódování atd.), překlad z jazyka počítače do jazyka matematicko-statistického výstupu z počítače, překlad matematicko-statistického výstupu do jazyka teorie (interpretace) a překlad do přirozeného jazyka závěrečné zprávy.“

Problém to představuje pro sociální vědy podle Petruska (1993:78) značný. „Pravděpodobnost, že výchozí informace zůstane po tolikanásobném překladu nedotčena, že se provedenými překlady dosáhne pouze odhalení latentní informace, je nepatrná. Jinými slovy – jazykový aspekt výstavby sociologie jako teoreticko-empirické vědy má podstatné důsledky nejen metodologické, ale i věcné: realita, která vstupuje do výzkumu, z něj vychází „nějak“ změněna. Avšak jen část, dokonce jen menší část těchto změn je reflektována.“ Sociální vědy jsou podle Petruska „dány jazykově“ a jinak ani býti dány nemohou (Petrušek 1993:79): „Situace v sociologii je podstatně složitější než ve vědách nonsociálních. Pro ty existuje pouze problém vztahu přirozeného jazyka a jazyka vědy, pro sociologa však existuje navíc ještě vztah jazyka sociologie a jazyka aktéra sociálních situací.“

6.3 Kvantitativní strategie

V případě kvantitativní metodologické strategie hraje významnou roli slovo experiment. V praxi sociálních věd jsou však s „klasickým“ experimentem nemalé problémy. Proto se často přistupuje k tzv. „kvazi-experimentu“, který se stal v rámci evaluace sociálních programů populárním postupem.

6.3.1 Experimentální design

Uskutečnění experimentu často brání etické důvody, shrnutelné pod hlavičku manipulace s lidmi proti jejich vůli. Úspěšnost experimentu předpokládá vytvoření

systému, který je (Žižlavský 2003:89): (1) jednoduchý, (2) uzavřený a (3) tudíž umělý. Zmíněné předpoklady pak nejlépe splňuje laboratorní prostředí. „Manipulujeme pouze s příčinným jevem, což znamená jeho vyvolání nebo změnu jeho intenzity. Potom sledujeme, zda vůbec nastane účinek, a když nastane, pak jej měříme. Jev - příčinu operacionalizujeme na *nezávislou proměnnou* a jev - účinek na *závislou proměnnou*“ (ibid.).

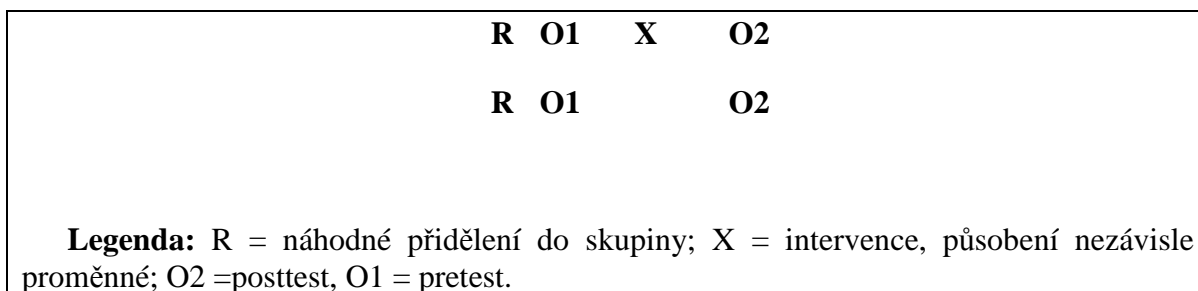
Řádný experiment lze rozpoznat podle následujících náležitostí (Shaugnessy, Zechmeister, 1990 dle Žižlavský 2003:89):

- Interní validita. Jsou eliminovány ostatní možné příčiny (vlivy) působící na výsledek experimentu (na závislou proměnnou).
- Reliabilita. Opakováním experimentu je dosaženo stejného výsledku (hodnoty závislé proměnné).
- Senzitivita. Experiment rozpozná efekt na závislou proměnnou, i když je nepatrný.
- Externí validita. Experimentální zjištění jsou zobecnitelná za hranice experimentálního prostředí.

Pokud chceme dosáhnout externí validity, musíme náhodně vybrat případy, které chceme zkoumat. Náhodně vybraný vzorek pak rozdělíme na experimentální skupinu a na kontrolní skupinu. „Experimentální skupina bude vystavena působení nezávislé proměnné, kdežto kontrolní skupina nebude. ...*Náhodným přidělením* zajistíme totožnost obou skupin, co se týče jejich struktury. Tímto opatřením zvýšíme interní validitu experimentu, protože po vystavení experimentální skupiny účinkům nezávislé proměnné by se měly obě skupiny lišit pouze hodnotami závislé proměnné, takže změna hodnoty závislé proměnné naměřená u experimentální skupiny je vysvětlitelná pouze vlivem nezávislé proměnné, a ne odlišným složením skupin“ (Žižlavský 2003:89).

Nejprve změříme hodnoty závislé proměnné u obou skupin před experimentem. Toto měření nazýváme „pretestem“ (ve vzorci, který je uveden níže je pretest označen jako O1). „Po pretestu vystavíme experimentální skupinu působení nezávislé proměnné (ve vzorci označeno písmenem X). Následně změříme hodnotu závislé proměnné u obou skupin - posttest (ve vzorci označen jako O2)“ (Žižlavský 2003:90).

Naměřená data dále „...statisticky analyzujeme, porovnáváme hodnoty nezávisle proměnné mezi skupinami v pretestu a také v posttestu. Pokud se v pretestu hodnotou závisle proměnné skupiny neliší a posttestu ano, potom můžeme konstatovat, že rozdíl je způsoben nezávisle proměnnou, tudíž, že mezi nezávisle proměnnou a závisle proměnnou existuje kauzální vztah“ (ibid.).



Obrázek 2 - Schéma experimentu

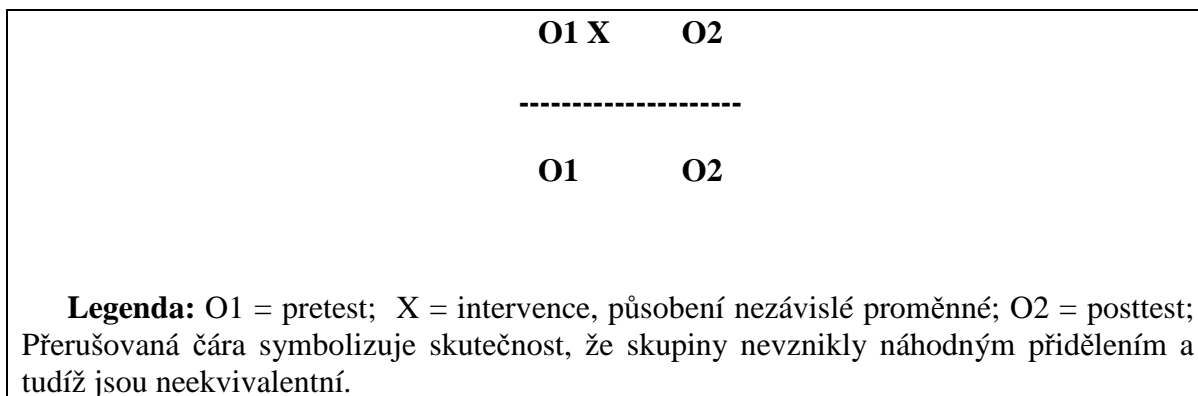
Zdroj: Žižlavský 2003:90

Populace možných subjektů, která je definována a je předmětem výzkumu, tedy může být rozdělena na experimentální a kontrolní skupinu. V případě, že máme subjekty přiděleny do experimentální a kontrolní skupiny, musíme znát (obzvláště v případě programové evaluace) i délku experimentu. Průběh – délka experimentu nesmí být podle Earla Babbie (1998) v žádném případě stanovena rigidně. Jednou z hodnotících charakteristik programu, který je předmětem výzkumu, je i doba potřebná na to, aby se stal efektivním, aby začal fungovat – přinesl výsledky.

6.3.2 Kvazi-experimentální design

V evaluačním výzkumu, jak již bylo řečeno, nelze plně dosáhnout podmínek nutných pro řádný experiment. „Nelze účastníky výzkumu z etických důvodů dlouhodobě držet uzavřené v laboratoři. Navíc mohou účastníci měnit, přizpůsobovat své chování umělému laboratornímu prostředí, takže se chovají jinak než ve svém přirozeném prostředí. Etické důvody také brání náhodnému přidělování do skupin, neboť členům kontrolní skupiny se nedostane potřebné péče nebo intervence, což může zhoršit jejich už obtížnou životní situaci. Těmto okolnostem musíme v praxi přizpůsobit experimentální uspořádání“ (Žižlavský 2003:94).

V praxi to tedy podle zmíněného autora (ibid.) znamená, „...že musíme slevit z některých požadavků, třeba z požadavku na náhodné přidělení do experimentální a kontrolní skupiny. Takové uspořádání nenazýváme experiment, ale *kvaziexperiment*.“



Obrázek 3 - Schéma kvazi-experimentu

Zdroj: Žižlavský 2003:94

Z podstaty věci pak plyne, že „...kvaziexperiment ve srovnání s experimentem má nižší interní validitu a také reliabilitu“ (Žižlavský 2003:94). V kvazi-experimentálním uspořádání tedy bude jako kontrolní skupina sloužit „podobná“ existující skupina. „Můžeme ji vytvořit tak, že klienty v souboru uspořádáme podle doby, po kterou čekají na službu nebo intervenci. Potom soubor rozdělíme na poloviny. První polovinu tvoří klienti čekající déle než klienti v druhé skupině. První skupina je experimentální skupinou, jejím členům se dostane intervence (dříve) a druhá skupina představuje kontrolní skupinu. Členové této skupiny ovšem nemají stejné charakteristiky jako členové experimentální skupiny a interní validita je ohrožena zejména nenáhodným způsobem výběru...“ (ibid.).

7 Kvalitativní výzkumné nástroje v evaluaci

Cílem této kapitoly je základní seznámení s možnostmi kvalitativního aplikovaného výzkumu, a seznámení s pojmy jako je kvalitativní výzkumný design a základní techniky sběru kvalitativních údajů. Opět však, stejně jako u předchozí kapitoly platí, že nejde o vyčerpávající přehled kvalitativního vědecko-výzkumného paradigmatu. V tomto ohledu je potřeba odkázat na metodologické učebnice sociálních věd. Spíše jde o základní seznámení s nástroji a podstatou kvalitativních postupů evaluace. Kapitola se dále věnuje možnosti smíšeného výzkumného designu. Klíčovým pojmem kvalitativního evaluačního výzkumu se pak stává role interpretace.

7.1 Kvalitativní strategie

Kvalitativní metody jsou celkem široce využívány při evaluacích sociálních programů, jelikož dokáží podle Pattona (2002:9) vyprávět „příběh“ sociálního programu zachycením dílčích „příběhů“ jednotlivých participantů. Evaluační případové studie mají všechny náležitosti dobrého příběhu. Vyprávějí, co se stalo, kdy, komu a s jakými důsledky.

Avšak stejně tak jako není jediná forma praxe v kvalitativní evaluaci, není jediná filosofická logika nebo ospravedlnění, která by se dala univerzálně použít pro všechny evaluační výzkumníky, kteří pracují v rámci kvalitativní metodologické strategie. Nicméně existuje soubor filosofických zásad a postojů, které se stávají zdrojem pro kvalitativní evaluační praxi. V tomto filosofickém výzkumném rámci, který se podle Greenové (1998) může nazývat různě: kvalitativní, etnografický, naturalistický – je klíčové slovo „interpretativistický“. Tento pojem představuje jednu z hlavních premis, tedy, že „ve světě lidské zkušenosti, existuje pouze interpretace“ (Denzin 1989:8). Dominance interpretivistické logiky ospravedlnění evaluačního výzkumu může být vystopována podle Greenové (1998:383) k vlivné práci Yvonney Lincolnové a Egonu Guby. Tito autoři byli vůdčími zastánci tohoto alternativního paradigmatu v rámci evaluace.

Většina evaluačních výzkumníků, kteří pracují v rámci kvalitativní strategie, podle Jennifer Greenové (1998:387):

a) používá případových studií, a proto zdůrazňují význam kontextu, ve kterém operují, a ne zobecnitelnost zjištění,

b) silně spoléhá (ale ne výhradně) na kvalitativní metody,

c) bere na vědomí (pokud to přímo nevyzdvihují), že jejich samotná přítomnost v evaluačním výzkumu má na program vliv a

d) primárně se ve své práci zaměřují na zvýšení porozumění praktickému sociálnímu programu.

Kvalitativní přístupy k programové evaluaci preferují kvalitativní techniky sběru informací jako například narativní rozhovory, polostrukturované rozhovory, zúčastněné pozorování či studium dokumentů (k těmto a dalším technikám v samostatné podkapitole). Zmíněné techniky jsou pro mnoho teoretiků a praktiků evaluace nejlépe v souladu s interpretativní perspektivou. Kvalitativní metody se spoléhají na interakční, adaptivní a úsudkové schopnosti samotného výzkumníka.

7.1.1 Holistická perspektiva v kvalitativní evaluaci

Sociální vědci a hodnotitelé, kteří analyzují kvalitativní údaje, se pokoušejí porozumět fenoménu či sociálnímu programu jako celku. To znamená, že je důležitý popis a interpretace sociálního prostředí jednotlivých osob, či organizační kontext pro celkové porozumění tomu, co bylo v terénní fázi výzkumu pozorováno a co bylo řečeno v rozhovorech. Holistický přístup znamená, že celek je chápán jako komplexní systém, který je více než jen suma částí (Patton 2002:59). S touto premisou pracuje systémový přístup, který hovoří o systému jako pojmu vnitřně obsažnějším než je „pouhá“ suma částí. Při využití holistického přístupu v kontextu kvalitativního sociálně-vědního výzkumu jde především o snahu zachytit co nejdříve různé okolnosti a souvislosti. V tomto duchu bylo výše použito i slovo komplexní, které by však při bližší pozornosti nemělo představovat jen synonymum k běžnému hovorovému „složitý“. Komplexitu poměrně hezky ilustruje přirovnání systému (např. živého organismu) např. k hromadě hlíny. Podle Flegra (2006:89) se v přírodě můžeme „...často setkat se systémy, jejichž komplexita je podstatně větší než komplexita kteréhokoli živého organismu. Může se jednat například o hromadu hlíny, Niagarské vodopády nebo atmosférický jev El Niño. Přesto nám komplexita těchto systémů obvykle nepřipadá nijak omračující, tedy alespoň dokud nejsme nuceni modelovat procesy v daném systému, například pro dlouhodobé předpovědi počasí.“ Tato určitá nezajímavost hromady hlíny a naopak fascinace živým složitým organismem je dána podle Flegra (2006:89) jednoduchým důvodem: „To, co ve

skutečnosti obdivujeme u organismů, není jejich komplexita, ale vlastnost úplně jiná, ve své podstatě přesně protichůdná. Je to organizovanost (ve smyslu uspořádanosti). Zatímco pro popis vysoce komplexního systému potřebujeme velké množství informace, pro popis vysoce organizovaného systému potřebujeme této informace naopak málo.“

Strategie holistického porozumění v kvalitativním výzkumu je podle Pattona (2002:59) v protikladu k logice a procedurám evaluačních studií, které jsou vedeny v analytické tradici „vezměte část a sledujte jak funguje.“ Hlavní kritika takovéto logiky je v pohledu kvalitativních výzkumníků podle Pattona (ibid., vl. překl.) ta, že „...takový přístup (1) zjednodušuje komplexitu skutečného světa a zkušenosti jednotlivých participantů, (2) nevěnuje pozornost hlavním faktorům, jež nejsou jednoduše kvantifikovatelné a (3) selhává ve vykreslení smyslu sociálního programu a jeho dopadů jako celku.“

7.1.2 Historie kvalitativních přístupů

Denzin a Lincolnová (2000 dle Patton 2002:79) vystopovali šest fází historie kvalitativního výzkumu, což pomáhá vysvětlit rozdíly se koncepcí utváření kvalitativního výzkumu.

1. Během „tradiční fáze“ koloniálního výzkumu (do 2. světové války) usilovali ve své terénní práci etnografové, ovlivnění pozitivismem, o objektivitu.

2. „Modernistická fáze“ (do 70. let 20. století) byla obdobím, kdy kvalitativně orientovaní sociální vědci zdůrazňovali metodologickou přesnost a procedurální formalismus a usilovali o respekt v rámci sociálních věd a reagovali na nástup postpozitivistického interpretativismu.

3. Během fáze „nejasných žánrů“ (1970-1986) se objevilo množství alternativních přístupů, které způsobily vzájemné půtky a zmatky. Odkazem tohoto období je nejasný žargon různých kvalitativních perspektiv: strukturalismus, symbolický interakcionismus, fenomenologie, etnometodologie, kritická teorie, semiotika, neopositivismus, mikro-makro deskriptivismus, neomarxismus, poststrukturalismus, naturalismus, konstruktivismus a dekonstruktivismus.

4. Pak přišla „krize reprezentace“, která se zaměřila na reflexivitu, moc, privilegia, rasu, gender a socioekonomické třídy – jež všechny podkopaly tradiční důraz na validitu a neutralitu.

5. Fáze pátá popisuje nedávnou historii a je charakterizována jako „trojí krize: krize reprezentace, legitimacy a praxe“ (Denzin, Lincoln 2000:17 dle Patton 2002:79), ve které je podrobena zkoumání podstata kvalitativního psaní studií, včetně perspektivy kvalitativního autora a přicházejí otázky jak hodnotit kvalitu provedeného kvalitativního výzkumu či evaluace.

6. V šesté fázi, kterou Denzin s Lincolnovou nazývají „postexperimentální“, jsou hranice kvalitativního zkoumání rozšířeny tak, že zahrnují i kreativní literaturu faktu, autobiografické etnografie, poezii či multimédia.

Denzin a Lincolnová (2000 dle Patton 2002) očekávají další rozvoj kvalitativního zkoumání novými směry a budoucnost nazývají „sedmou fází“, což možná bude fáze klidu, kdy ustanou debaty o odlišnostech a nastane spíše oslava rozmanitosti výstupů z kvalitativního zkoumání.

Stejně jako je nezbytné zabývat se u kvantitativního výzkumu kvalitou provedeného výzkumu (v pojmech validity a reliability), je potřeba vhodná kritéria hledat také pro kvalitativní evaluační studie. Úsilí o nalezení alternativních kritérií posouzení kvality kvalitativní studie Denzin a Lincolnová (1998) zasazují do páté fáze výše načrtnutých historických proměn kvalitativního výzkumu (blíže k alternativním kritériím pro posouzení kvality kvalitativní studie níže v této kapitole).

Možnosti poznání sociální reality jsou tedy po dlouhou dobu - a do značné míry i v současnosti - v zajetí nekonečných diskusí o možnostech objektivního poznání reality na jedné straně a subjektivitě lidského vnímání světa na straně druhé. Tyto pohledy na svět daly v historii vzniknout mnoha uceleným myšlenkovým směrům, přičemž novodobě-historicky starší (osvícenství) je právě důraz na objektivitu poznání a aplikaci výzkumných metod z přírodních věd do oblasti zájmu věd sociálních. Historicky mladší (romantismus) kvalitativní přístupy si pak musí i v současnosti pracně dobývat své místo na slunci. Je jim vytýkána právě nízká reliabilita a objektivita.

7.1.3 Generalizace vs. extrapolace či transferabilita v kvalitativním výzkumu

Pragmatické kritérium využitelnosti vede k otázce, co můžeme se zjištěními z kvalitativní studie dělat. Zjištění rozhodně osvětlují určitou dílčí situaci či malé množství

případů. Patton (2002:581) se ale ptá: Jak toho využít kromě tohoto malého počtu případů či situací? Mohou být kvalitativní zjištění generalizována?

Shadish (1985 dle Patton 2002:581) argumentuje, že základní principy generalizace jsou aplikovatelné jak na experimenty tak na etnografie (či obecně kvalitativní metody). Jak experimenty, tak případové studie totiž sdílejí problém vysoké míry „lokalizace“. Co je žádoucí ve vědě obecně (generalizace napříč časem či prostorem) vede k úvaze o tom, co je vůbec v sociálních vědách možné. Cronbach (1975:22 dle Patton 2002:582) rovněž porovnává generalizace v přírodních vědách s tím, co by bylo možné ve vědách o člověku a v sociálních vědách. Jeho závěr je, že „...generalizace jsou na ústupu. V dané chvíli závěry popisují existující situaci dobře, později se však závěry musejí vypořádávat s odlišnostmi v realitě a nakonec mohou být platné pouze jako historie.“

V kvalitativním přístupu tak spíše nachází uplatnění koncept extrapolace, tedy přenesení zjištění „vně“ aktuálního rámce (časového i prostorového) předmětu kvalitativní studie, což lze vyjádřit i pojmem „transferabilita“. Konstrukt přenositelnosti (transferability) byl zmíněn jako jedno z alternativních kritérií pro posouzení kvalitativního výzkumu výše a v zásadě komparativně odpovídá principu zobecnitelnosti v kvantitativním výzkumu. Přenositelností se tedy myslí již výše diskutovaná možnost využít závěry z daného případu pro jiný případ, který se mu podobá.

7.1.4 Kreativita v kvalitativním výzkumu

Kvalitativní zkoumání staví jak na kritickém (podstata vědeckého přístupu) tak na kreativním myšlení. Jinými slovy podle Pattona (2002:513) staví jak na vědě tak na „umění“ analýzy. Nicméně technické, procedurální a vědecké aspekty analýzy je jednodušší prezentovat i vyučovat. Kreativitu, tak často používanou v našem každodenním jazyce, je však těžší vyučovat a pravděpodobně ještě těžší se jí naučit (Patton 2002:513).

Obecně se podle Nachtmannové a Veselého (2007:113) uznává, že „...tvořivé myšlení je klíčovým faktorem schopnosti inovovat. Nepřekvapuje tedy, že v posledních deseti či dvaceti letech vzniklo poměrně velké množství prací, jež se zabývají možnými způsoby, kterými lze stimulovat tvořivé myšlení.“ Vstupním předpokladem je, že „...každý člověk může být tvořivý, pokud se dokáže vymanit ze svých myšlenkových stereotypů a pokud je ochoten své tvořivé schopnosti neustále procvičovat a zdokonalovat“ (ibid.).

7.1.5 Vědění versus tvořivost

Naučené vědění a kreativita mohou být v jistém rozporu. Jak uvádí Nachtmannová a Veselý (2007:113): „Nové myšlenky se nenacházejí v oblasti známého. Řešení problémů dosud nevyřešitelných leží mimo ‚galaxii‘ současného myšlení. Jak řekl Einstein: ‚Nemůžeme řešit problém použitím toho, co vás k problému přivedlo‘. Zůstat v oblasti známého je přirozenou tendencí naší mysli, brání nám to však v prozkoumávání nových možností. Jedině když stojíme v neznámém a nevíme si rady, můžeme být tvořiví.“ V této otázce lze využít i poznatků z vývojové psychologie o tvořivosti dětí - děti jsou přirozeně kreativní, jelikož postrádají zásobu vědění dospělého. „Pak se naučí nové věci a svou tvořivost ztrácejí. Vědění je tak v jistém rozporu s tvořivostí: Čím více víme, tím méně si musíme představovat. Čím méně si musíme představovat, tím rutinnější se naše myšlení stává“ (Nachtmannová, Veselý 2007:113). Zde citovaní autoři pak referují o způsobu, jak riziku rutinního myšlení co nejvíce předejít: „Předtím než se pustíte do práce a shánění informací o určité oblasti (problému), zkuste věnovat nějaký čas vlastnímu myšlení. Kladení otázek, hypotéz a pod., které jsou ještě nezátížené vašim věděním, je velmi důležité“.

V tomto ohledu nelze nepřipomenout Scrivenův návrh tzv. „bezcílné evaluace“, který je právě jistou snahou zbavit se „balastu“ vědění, které může bránit kreativitě, ba dokonce bránit vidět souvislosti.

7.1.6 Zásady kreativity v přípravě i realizaci kvalitativního výzkumu

Patton (2002:513) prezentuje celou sadu bodů (zásad), který vyvodil z výzkumů a z výuky kreativního myšlení na základě mnoha publikací (Kelley a Littman 2001, DeBono 1999, Von Oech 1998, Patton 1987):

1. Buďte otevření. Kreativita začíná u otevřenosti množství možností.
2. Vytvářejte různé možnosti. Obvykle existuje více než jeden způsob jak o věci přemýšlet.
3. Divergujte-konvergujte-integrujte. Začněte objevováním množství různých směrů myšlení a možností ještě před tím než se zaměříte na detaily.

4. Užívejte více různých stimulů. Trénink kreativity často zahrnuje mnoho odlišných forem vyjádření: kresba, hudba, hraní rolí...

5. Myslete „cik-cak“ či „kruhově“. Kreativita nebývá výsledkem jasně lineárního postupu logické indukce či dedukce. Kreativní osoba se vrací zpět a zase dopředu, „chodí okolo“ tématu...

6. Změňte své zažité vzorce myšlení. Zvyky, standardní postupy a z toho vyplývající myšlení uzavřené v zažitých vzorcích představuje bariéry v kreativě. Buďte k tomu pozorní a snažte se tyto vzorce myšlení změnit.

7. Propojte jevy. Mnoho tréninkových aktivit kreativního myšlení zahrnuje cvičení jak propojit zdánlivě nepropojitelné.

8. Důvěřujte si. Sebedůvěra v otázce kreativity je cestou ke kreativě. Když si budete říkat „nejsem kreativní“ tak také nebudete.

9. Pracujte na sobě. Kreativita není jen zábava, je to docela těžká práce spočívající ve studiu předchozích výzkumů a odborné přípravě.

10. Hrajte si s kreativitou. Kreativita ale také není jen práce. Může být a měla by být i hrou.

Mezi způsoby „tréninku“ kreativního myšlení lze zařadit např. techniku využití podvědomí při nalézání odpovědí, resp. námětů, které je dále možné vědomě rozpracovat. Podle Nachtmannové a Veselého (2007:114) „...nemusíte na odpovědi pracovat vědomě, vaše podvědomí to udělá za vás. Myšlenky se neberou z ničeho nic, ale vynořují se z podvědomí. Nemůžete však najít řešení problému, o kterém ještě nevíte a který jste si neformulovali.“ Další ze způsobů jak si poradit s problémem je např. „myslet rychle“. „Zkuste přijít s co největším počtem myšlenek během krátkého času... Je mnohem jednodušší přijít s padesáti myšlenkami než s jednou geniální myšlenkou. A je také velmi těžké přijít s padesáti špatnými myšlenkami. Čiré množství takřka zajišťuje, že nějaká myšlenka musí být dobrá“ (Nachtmannová, Veselý 2007:114). Citovaní autoři dále uvádějí postřeh, že „...je jednodušší přijít s padesáti myšlenkami rychle a během krátké doby než během dlouhého času. Je tomu tak proto, že v delším časovém období máte čas na hodnocení kvality těchto myšlenek a hodnocení je jedním z největších nepřátel tvořivosti.“ Vymanit se z konvenčních rámců myšlení může napomoci např. technika tzv. „stoosmdesáti stupňového myšlení“. To znamená, že „...myšlení zaměříme úplně na druhou stranu, než kam nás vede konvenční myšlení. Když konvenční myšlení říká

„měkký“, uvažujte „tvrdý“... Důvod, proč se vydat zcela opačným směrem, a nikoli jen trochu odlišným, je prostý: Je jednodušší stanovit přesný opak něčeho, než stanovit odchylku o deset či dvacet stupňů. Tato technika je tedy spíše začátkem – „spouštěcím mechanismem“ - tvořivého procesu, nikoli výsledkem“ (Nachtmannová, Veselý 2007:115).

Tuto podkapitolu o kvalitativní výzkumné strategii uzavřeme spolu s Pattonem (2002:514) tím, že jak věda, tak umění kvalitativního zkoumání jsou vystaveny časovým tlakům a omezením. Lidé jsou často nuceni pracovat v časovém „presu“ a jejich kreativita je tak zásadně podkopána. Způsob, jakým výzkumník kombinuje kritické a kreativní myšlení je jen z části otázkou jeho osobního „stylu“. Z části je to také otázka situace a množství času, který zbude na kreativní vytváření alternativ.

7.2 Kombinace strategií – metodologický koktejl

Cook a Reichardt (1979:27 dle Clarke 1999:86, vl. překl.) komentují dichotomii mezi kvantitativními a kvalitativními strategiemi tak, že „...je třeba ukončit budování zdí mezi metodami a je třeba začít stavět mosty“. Od té doby je podle Clarke (ibid.) zřetelný v tomto směru určitý posun a to posun k většímu mixování metod. Smíšená metoda výzkumného designu v některých případech integruje kvantitativní a kvalitativní data.

Zastánci metodologického mixu se rekrutují v zásadě z autorů, kteří se přestali zabývat paradigmatickými rozdíly mezi kvalitativními a kvantitativními metodami. Jde především o zastánce utilizace, tedy aby výsledky evaluace byly co nejvíce přínosné pro uživatele. „Válka paradigmat již pro ně nehraje žádnou roli, neexistuje žádný ‚zlatý standard‘ metod. Přijímá se teze, že všechny metody jsou vhodné, pokud přispějí k cílům evaluace“ (Hendl 2005:298).

Většina hodnotitelů se podepisuje pod názor, že není jediná nejlepší metoda, která by pokryla všechny evaluační potřeby. Multimetodové evaluační designy, zahrnující jak kvantitativní tak i kvalitativní přístupy, jsou doporučeny v situacích, kdy nastanou následující situace:

- 1) kombinované metody poskytují možnost rozšířit validitu výzkumu;
- 2) je vyžadován úplný popis programu;
- 3) výzkum je zaměřen jak na výsledky, tak na proces;

4) je potřeba využít smíšené strategie k pokrytí potřeb různých skupin „stakeholderů“ (Smith 1986 dle Clarke 1999:89).

Bez ohledu na to, zdali vedeme sumativní nebo formativní evaluaci, je často žádoucí dosáhnout vybalancování kvantitativních a kvalitativních přístupů. Nejúčinnější evaluační výzkum je i podle Earla Babbie (1998) takový, který kombinuje kvantitativní a kvalitativní složky. Když hodnotitel provádí vhodné statistické procedury, získává tím cenné informace o zkoumané problematice, ale mnohdy teprve zapojením kvalitativních metod dosáhne hlubokého porozumění procesům, které studuje.

Přístupy k evaluaci a metody, které byly v rámci programové evaluace vyvinuty, jsou tedy hojné. Důležitým bodem pro všechny praktiky, kteří vykonávají evaluaci, je pak otázka, jak zvolit přístup nebo metodu právě pro tu „jejich“ evaluaci. Podle Chena (2005) máme dva obecné pohledy na takovéto rozhodnutí: universalistický a kontingenční pohled.

Univerzalistický pohled na praxi evaluace

Univerzalistický pohled se opírá o univerzální nadřazenost určitých evaluačních přístupů a metod nad jinými. Ve světle tohoto pravidla se musí hodnotitelé snažit aplikovat schválené „nejlepší“ evaluační přístupy a metody při vykonávání jakékoli evaluace. Univerzalistický pohled argumentuje, že vždy existuje „nejlepší“ výzkumná metoda. Například někteří hodnotitelé obhajují kvantitativní postupy (jako je zmiňovaný experiment s náhodným výběrem), zatímco jiní prosazují, že nejlepší jsou kvalitativní nebo smíšené metody.

V tomto případě hodnotitel (bez ohledu na to, k jaké volbě nakonec dojde), který obhajuje jednu nejlepší metodu, uplatňuje univerzalistický pohled na výzkumné metody v programové evaluaci. Formativní versus sumativní evaluace poskytuje podle Chena (2005) podobný příklad. Hodnotitelé, kteří argumentují, že sumativní evaluace je lepší než formativní evaluace – stejně tak jako ti, kteří tvrdí opak – demonstrují univerzalistické principy pro provádění evaluace. Jedním z důvodů pro popularitu univerzalistického pohledu (hlavně během raných fází programové evaluace) je, že univerzalistické principy se jeví jasné a pro praktiky jednoduše proveditelné a sledovatelné. Dalším důvodem je výhodnost, kterou univerzalistický pohled poskytuje evaluačním teoretikům, kteří chtějí zarátovat své argumenty a vyzdvihnout výhody použitých přístupů.

Kontingenční pohled jako perspektiva programové teorie

Alternativním pohledem jak konceptualizovat, rozvíjet a aplikovat principy, přístupy a metody programové evaluace, je kontingenční pohled. Tato perspektiva podle Chena (ibid.) postuluje, že neexistuje jedna nejlepší cesta jak vést programovou evaluaci. Podle této perspektivy by spíše mělo jít o „situační“ volbu mezi různými přístupy k evaluaci programu. Základním principem kontingenčního pohledu je, že podstata každého programu, unikátnost evaluačních záměrů a kontextuální okolnosti, vyžadují celou škálu evaluačních přístupů a metod.

Z pohledu kontingenčního by mohlo být vedení evaluace méně přímé než z pohledu univerzalistického. Na druhou stranu to však vytváří výhodu v možnosti odstranit zjednodušující domněnky. Obecně se podle Chena (ibid.) kontingenční pohled lépe hodí k popisu reality, ve které dochází k programové evaluaci.

7.3 Techniky sběru dat

Data, která hrála klíčovou roli v definici kvantitativního a kvalitativního přístupu, mohou být odlišena nejen podle jejich vnitřních charakteristik, jak jsme učinili v předchozích podkapitolách, ale rovněž podle toho, jak jsou získávána. Konvenčně je činěn rozdíl například mezi primárními a sekundárními daty. Sběr primárních dat zahrnuje použití výzkumných nástrojů, například dotazníků a rozhovorů, které bývají sestrojeny za účelem konkrétní studie.

Sekundární data jsou ta data, která získali jiní výzkumníci, organizace nebo úřady pro své vlastní účely. I tato data mohou hodnotitelům poskytovat užitečné zdroje informací. Jak primární, tak sekundární data mohou nabývat podoby kvalitativní i kvantitativní. Je rovněž možné odlišit mezi strukturovanými a nestrukturovanými daty.

Jen stěží v realitě najdeme evaluační studii založenou pouze na jedné technice sběru dat. Obvykle formuje výzkumnou strategii množství takových technik. Nejčastěji je i ve spojení s technikami sběru dat používán pojem „metody“ sběru dat. Nicméně v této práci se budeme držet pojmu „techniky“ sběru dat. K některým z nich se teď vyjádříme. Vzhledem k tomu, že všechny tyto techniky jsou dostatečně známy i z jiných sociálně-vědních výzkumů a je jim věnována dostatečná pozornost i v česky publikované literatuře, zmíníme se zde o nich jen v krátkosti a „pro úplnost“.

Dotazník

Dotazník je v evaluačním výzkumu jednou z nejfrekventovanějších technik sběru dat. Může být hlavním zdrojem dat, nebo může být užitečným zdrojem dodatečných dat. Dotazník je podle Clarke (1999) hlavním nástrojem ke sběru primárních dat.

Dotazníky jsou schopny získat velké množství vysoce strukturovaných, standardizovaných dat. Kvalita informace, které dotazník poskytuje, je do velké míry závislá na designu dotazníku. Proto by mělo být samozřejmostí, že hodnotitel zvažuje nejen samotný obsah dotazníku, ale rovněž otázky spojené s jeho konstrukcí a administrací. Jak podotýká de Vaus (1996:105 dle Clarke 1999:69, vl. překl.): „...dobré dotazníky se prostě jen tak neobjeví, vyžadují pečlivé rozmyšlení, množství průběžných verzí během evaluace a rozsáhlé testování validity“. Technika dotazníku je populární zvláště v kvantitativně orientovaných výzkumech.

Rozhovor

Rozhovory jsou mezi hodnotiteli další široce používanou technikou. Jsou obzvláště populární mezi kvalitativně orientovanými výzkumníky. Během raných fází evaluace může hodnotitel získat díky rozhovorům s mnoha různými lidmi významné informace. Převážně pomocí rozhovorů s programovými rozhodovateli, administrátory a poskytovateli služeb pak dosahuje hodnotitel plnějšiho porozumění podstatě programu, porozumění hlavním cílům programu a teorii, která stojí za jeho designem a implementací.

Provedení rozhovorů může být důležitým krokem při formulaci samotného evaluačního výzkumného designu. Zdali pak bude rozhovor použit jako primární technika sběru dat, závisí pochopitelně na podstatě evaluačních otázek a kontextu, ve kterém je studie prováděna. Rozhovory poskytují primárně kvalitativní údaje a mohou se vztahovat jak k evaluaci procesu, tak k evaluaci dopadu.

Jak podotýká například Patton (1987:115 dle Clarke 1999:73, vl. překl.): „...účelem kvalitativních rozhovorů v evaluaci je porozumět tomu, jak lidé, kteří jsou zahrnuti do plánování či realizace programu, tento program vidí, naučit se jejich terminologii a zachytit komplexitu jejich individuálních zkušeností“. Tento typ rozhovorů je určitě užitečný v počátcích evaluační studie, jelikož umožňuje obeznámit se s projektem, s programovým personálem a participanty.

Pozorování

V momentě, kdy hovoříme o evaluační praxi, je možné odlišit různá použití pojmu pozorování. V obecném smyslu hodnotitelé používají pozorovací dovednosti bez ohledu na techniky sběru dat, které primárně používají. Hodnotitelé by měli být citliví k různým způsobům, jakými se lidé chovají a jak spolu navzájem interagují. Tato pozorování nejsou systematická, ale vyskytují se spontánně v průběhu vedení celé evaluace. Dále je pochopitelně možné hovořit o použití pozorování jako systematického nástroje, jako přímé techniky sběru primárních dat. Pojem systematické pozorování používáme, když pozorujeme a přitom o tom vedeme záznam podle předem jasně daných procedur (máme stanoveny indikátory toho, co chceme pozorovat atd.). Tento způsob pozorování ve svém důsledku umožňuje opakování i jinými hodnotiteli v jiných geografických oblastech a jiném čase. Výsledné údaje by pak mohly být vzájemně porovnatelné.

Dokumentární zdroje

Hodnotitelům je dostupné velké množství dokumentárních materiálů. Při klasifikaci dokumentárních zdrojů načrtává Burgess (1984 dle Clarke 1999) odlišnost mezi primárními a sekundárními dokumentárními zdroji, veřejnými a soukromými dokumenty, vyžádanými a nevyžádanými dokumenty. Primární zdroje referují o dokumentech sestavených jednotlivci, kteří mají s popisovanými událostmi zkušenost tak zvaně „z první ruky“. Příkladem mohou být zápisy z meetingů, osobní diáře atd.

Sekundární zdroje sestávají z dokumentů, které jsou vyprodukovány někým, kdo nemá osobní zkušenost s danou situací. Příkladem může být něčí stručný výtah z dlouhé oficiální zprávy, který je vyprodukován například za účelem rozšíření informací klíčovými osobám v rámci organizace. Dokumentární zdroje jsou populární především u kvalitativně orientovaných hodnotitelů, ale své uplatnění naleznou i při vedení evaluace kvantitativní výzkumnou strategií.

Jak ukázala tato kapitola zaměřená na metodologické otazníky evaluace sociálních programů, volba metod v evaluačních studiích je ovlivněna mnoha faktory. Mezi tyto faktory patří podle Clarke (1999) obzvláště:

- 1) podstata evaluačního problému,
- 2) potřeba poskytovat informace pro poskytovatele programu,
- 3) kontext evaluace,

- 4) charakteristiky všech programových participantů,
- 5) metodologická orientace výzkumníka,
- 6) podstata programové teorie, která stojí za výzkumem.

Hodnotitel má celou šíři výzkumných nástrojů, ze kterých si může vybírat. Závěrečná část této kapitoly ve stručnosti popsala výběr z hlavních silných stránek a slabých stránek některých z nejběžněji používaných kvantitativních a kvalitativních technik sběru dat. Efektivní evaluace odvisí od připravenosti hodnotitele být schopen vybrat z mnoha metod a technik, a když to bude vhodné, být schopen úspěšně mísit metody v rámci jednoho výzkumného designu.

8 Role hodnotitele v evaluaci

Tato kapitola prezentuje témata jako je role výzkumníka v interpretaci, sociální a politický kontext provedení evaluace, „externí“ vs. „interní“ hodnotitel, role tzv. „stakeholders“ v evaluaci sociálních programů, posouzení vědecké důvěryhodnosti evaluace a posouzení praktické účelnosti evaluace a potřeba jejich vzájemné vyváženosti.

8.1 Role hodnotitele v designu a plánování nových programů

Evaluační výzkumy by měly být integrovanou součástí plánování a implementace nového programu. Žádný program však není vytvořen a implementován podle nějakého souboru „pravidel“. Programy spíše skoro vždy vznikají podle Rossiho a Freemana (1993) ze složitého mixu politických úvah, vlivu klíčových držitelů moci, ekonomických tlaků a dostupnosti kvalifikovaného personálu nebo vhodných technologií. V mnoha případech jsou programy plánovány a následně implementovány bez pomoci evaluačního výzkumu, často ke škodě programu.

Aby maximalizovali hodnotitelé jejich vliv na design a implementaci programů, musí porozumět formálním a neformálním organizačním podmínkám prostředí, ve kterém pracují. Aby mohl být evaluační výzkum mnohdy vůbec proveden, je podle Earla Babbie (1998) třeba, aby získal náležitou logistickou (tj. personální a organizační) podporu v organizaci či instituci, v níž se bude provádět. Při přípravě evaluačního výzkumu je třeba, aby jeho autoři:

1) Získali informace o těch, pro koho mohou mít výsledky evaluace negativní důsledky, tzn. jejichž příjem, status nebo kariéra mohou být těmito výsledky ovlivněny. Tyto lidi je dobré zahrnout do projektu tak, aby na něm smysluplně participovali.

2) Aby si hodnotitelé uvědomili, kdo si evaluaci přeje, proč si ji přeje, ale také, kdo si ji nepřeje a jaké pro to může mít důvody.

3) Aby předali těmto lidem celý projekt evaluace k připomínkám.

8.2 Externí a interní hodnotitel

Často je uváděn rozdíl mezi externím a interním hodnotitelem. Externí hodnotitel je v podstatě nezávislý konzultant, který je pověřen vést evaluaci pro poskytující organizaci, nebo pro organizaci, která financuje program. Jak vyzpozoval Patton (1986 dle Clarke 1999), definující charakteristikou externích hodnotitelů je, že nemají dlouhodobou průběžnou pozici v rámci zkoumaného programu nebo v rámci organizace, která je hodnocena.

V protikladu k tomu jsou interní hodnotitelé zaměstnanci na plný úvazek, kteří vedou evaluace v rámci svých vlastních organizací podle nařízení svého vlastního managementu. Ačkoli interní i externí hodnotitelé používají stejných výzkumných metod a střetávají se s podobnými metodologickými problémy, jejich vztah k organizaci nebo agentuře, ve které vedou evaluaci, je rozdílný. Výhody a nevýhody, které se pojí s dichotomií externí a interní hodnotitel, jsou uvedeny v následující tabulce.

Tabulka 3 - Výhody a nevýhody zapojení interních a externích hodnotitelů

Výhody
<p><i>Interní hodnotitelé budou:</i></p> <ul style="list-style-type: none">- seznámení s historií, backgroundem, hodnotami a kulturou organizace,- budou odpovědnější k implementaci evaluačních doporučení, budou totiž odpovědní i za vyprodukování těchto doporučení,- budou se spíše zaměřovat na celkové zájmy, tak jak je vnímá management.
<p><i>Externí hodnotitelé budou:</i></p> <ul style="list-style-type: none">- mít nezávislý postoj a budou poskytovat „čerstvý“ pohled na program,- mít objektivní, kritický přístup,- mít přehled o větším množství podobných organizací a mohou poskytnout hodnocenému programu srovnání s jinými,- mít znalosti a zkušenosti o širokém spektru evaluačních technik,- budou resistantní k „zastařování“ managementem, který by mohl (explicitně nebo implicitně) požadovat co nejpříznivější hodnocení programu.
Nevýhody

Interní hodnotitelé by:

- mohli mít skryté zájmy na určitých výsledcích,
- mohli být negativně ovlivněni historií a znalostí organizačních hodnot,
- mohli být ovlivněni vlastním pohledem na management,
- nemuseli mít zkušenost s velkou šíří evaluačních metod a technik,
- nemuseli mít velký zájem na provedení evaluace,
- mohli inklinovat k favorizování programů, které byly vyvinuty v rámci jejich vlastní organizační jednotky nebo sekce v rámci celé organizace,
- by mohli mít obtíže se zahrnutím „stakeholderů“ z jejich vlastní organizace do aktivní participace na evaluačním procesu.

Externí hodnotitelé by:

- mohli ignorovat některé vnitřní faktory, takže jejich soudy nemusí reflektovat komplexní realitu dané situace,
- nemuseli rozpoznat, kdo je skutečným „klíčovým“ hráčem a mohli by tak být snadněji uvedeni v omyl určitými zájmovými skupinami,
- se mohli více zajímat o závěrečnou zprávu samotnou a její prezentaci než o implementaci zjištění ve zprávě obsažených,
- mohli být ovlivněni potřebou zajistit si hodnotící kontrakt do budoucna,
- mohli být nesensitivní k organizačním normám a vnitřním vztahům,
- mohli být primárně odpovědní nějaké jiné externí organizaci.

Zdroj: Feek (1988) a Love (1991) dle Clarke 1999:23, vl. překl.

Jak již bylo naznačeno i v předchozích kapitolách, hodnotitel může být integrovanou součástí běhu programu, může více či méně kooperovat se „stakeholdery“, musí se snažit o co nejvyšší míru využití zjištění z evaluace, kterou provádí. Jako takový je tedy významným aktérem, který ovlivňuje chod programu a jeho budoucí nasměrování.

V této podkapitole jsme jen stručně shrnuli, jakých dvou základních podob (interní, externí) může hodnotitel nabývat. Shrnuli jsme slabé i silné stránky obou možností. Zde nezbyvá, než opět upozornit na souvislost s předchozími kapitolami, kde byly

zdůrazňovány i přístupy k evaluaci, založené na tzv. „organizačním učení“, kdy se v rámci programu postupně vychovávají osoby, které budou následně vykonávat hodnocení pod svou vlastní taktovkou. Na tomto případě je pak nejlépe vidět vztah interního a externího hodnotitele, kdy se jejich role mohou doplňovat (kombinovat) a nemusí se nutně navzájem vylučovat.

8.3 Role „stakeholderů“

Role tzv. „stakeholderů“ v programové evaluaci je významná hlavně nepřímo. Nepřímo ovlivňuje, jaký typ evaluace vlastně bude prováděn, jak budou prezentovány výsledky atp. Jinými slovy, to jak hodnotitel vnímá svou roli v evaluaci programu, nemusí být ovlivněno pouze volbou výzkumných metod, ale bude to mít i souvislost s podstatou vztahu hodnotitele s různými skupinami „stakeholderů“.

Od hodnotitele se mnohdy očekává, že bude mít přímý kontakt s praktiky programu a dalšími participanty. To vše s cílem lépe identifikovat hlavní praktické problémy v rámci určitého evaluačního kontextu. Hodnotitel je zde od toho, aby poskytoval odborné služby a pokryl tak potřeby svých klientů, tedy realizátorů, financovatelů, celkově vzato „stakeholderů“. Odpovědný hodnotitel nepřistupuje k evaluaci bez jasných představ o tom, jaké by měly být evaluační otázky, nebo jak by mohl být měřen úspěch programu. Pomocí detailní a systematické pozorovací terénní práce a hloubkových interview s programovými participanty, postupně objeví klíčové evaluační otázky. Tímto způsobem se pak prezentují různé skupiny „stakeholderů“ v celém evaluačním procesu.

Pro praxi programové evaluace jsou podle Chena (2005) důležité tři kroky, z nichž první se týká právě „stakeholderů“:

1) *Systematicky identifikovat potřeby „stakeholderů“.* Většinou je evaluace sociálního programu iniciována s cílem vyhovět specifickým potřebám „stakeholderů“. Pokud má být programová evaluace užitečná právě pro tyto „stakeholdery“, jsou to právě jejich očekávání, která musí mít hodnotitelé na paměti, když vytvářejí design evaluace.

Potřeby „stakeholderů“ ohledně evaluace se v různých fázích vývoje programu mění. Například potřeby ohledně evaluace během fáze plánování a rozvoje programu se výrazně liší od evaluačních potřeb ve fázi pokročilé implementace a zaběhaného programu. Zásadní věcí je tedy precizní komunikace se „stakeholdery“ a porozumění jejich očekáváním.

2) *Volba nejvhodnějších evaluačních strategií.* Jde o výběr metod a postupů, které nejlépe sedí určitým potřebám, a o zvážení jejich výhod a nevýhod (časová náročnost, finanční náročnost, vědecká přesnost...). Následně pak k výběru jejich nejoptimálnější kombinace, kdy určitá rizika zvolených metod jsou co nejlépe vyvážena jinými pozitivy (trade off efekt). Jakmile jsou tedy identifikovány potřeby „stakeholderů“, musí hodnotitelé vybrat strategii k pokrytí každé této potřeby. Dostupných je mnoho možností, zmíněné efekty trade-off tedy občas musí nastat.

3) *Uvedení vybrané aktivity v život.* Praktická realizace zvolené strategie (metod a technik).

Hodnotitel by se měl podle Chena (2005:6) snažit zajistit kvalitní zpětnou vazbu programu tím, že se ujistí, že programová evaluace je orientována na budoucí akci, bere v úvahu jak vědeckou důvěryhodnost, tak důvěryhodnost evaluace pro „stakeholdery“ a používá holistický přístup.

8.4 Vědecká důvěryhodnost a důvěryhodnost pro „stakeholdery“

Chen (2005) se snaží odlišit mezi důvěryhodností (kredibilitou) vědeckou a důvěryhodností evaluace pro „stakeholdery“. Vědecká důvěryhodnost programové evaluace podle něj (2005:8) odráží rozsah, ve kterém se evaluace řídí vědeckými principy. Čím více je výzkum veden za využití vědeckých principů, tím vyšší je jeho vědecká důvěryhodnost. Nicméně vědecká důvěryhodnost je pouze první druh důvěryhodnosti, který musí programová evaluace splňovat. Jakožto aplikovaná věda, musí programová evaluace rovněž pracovat s důvěryhodností zainteresovaných aktérů („stakeholderů“). Tento druh důvěryhodnosti k programové evaluaci odráží rozsah, v jakém „stakeholderi“ věří, že evaluační design bude skutečně odrážet jejich záměry a potřeby. Obecně se tedy dá říci, že čím vyšší je shoda evaluace s pohledem na věc, zájmy a potřebami „stakeholderů“, tím vyšší bude jejich důvěryhodnost k prováděné evaluaci.

Podle Chena (2005) pak ideální evaluace dosahuje jak vysoké vědecké důvěryhodnosti, tak vysoké důvěryhodnosti pro „stakeholdery“. Ovšem tyto dvě komponenty nejdou ruku v ruce automaticky. Evaluace může dosahovat vysoké vědecké důvěryhodnosti, ale nízké důvěryhodnosti pro „stakeholdery“. Nastává to v případě, že hodnotitelé sledují pokud možno všechny vědecké principy, ale zaměřují svou pozornost na kritéria evaluace jednoduše podle toho, která považují sami za správná. Takováto evaluace bývá často

zainteresovanými aktéry odmítána, navzdory její nesporné vědecké důvěryhodnosti. Je tomu tak jednoduše proto, že nereflexuje jejich zájmy ohledně hodnocení programu.

Takže cílem (a výzvou pro design) evaluace je dosáhnout vybalancovaného stavu obou těchto druhů důvěryhodnosti. Možnou strategií jak toho dosáhnout, je snažit se dosáhnout důvěryhodnosti pro „stakeholdery“ v raných fázích designu evaluace a posléze evaluaci legitimizovat vědeckými principy. Nejprve by tedy měli hodnotitelé dostatečně komunikovat s programovými „stakeholdery“ a porozumět jejich pohledu na věc, jejich záměrům a potřebám. Hodnotitelé by posléze měli začlenit toto porozumění jejich potřebám do výzkumných otázek a celkově do evaluačního designu, to vše podle potřebných vědeckých principů.

Jak bylo v předchozím textu zmíněno, zájmy „stakeholderů“ jdou nesmírně důležitou charakteristikou. Podobně důležité pak je, v jakém stádiu se program zrovna nachází. Od něj se pak budou odvíjet nejhodnější evaluační strategie.

9 Prezentace evaluačních zjištění a jejich praktické využití

V této kapitole jsou probírána témata, jako jsou typy zjištění evaluační studie, faktory ovlivňující využití zjištění evaluačního výzkumu, nástroje maximalizace využití evaluačních zjištění, závěrečná zpráva evaluační studie, náležitosti závěrečné zprávy a různé podoby a rozsah závěrečné zprávy podle okruhu příjemců zjištění, možnosti rozšíření (disseminace) zjištění mezi potenciální příjemce zjištění (stakeholdery).

V předchozích kapitolách jsme se koncentrovali na definování podstaty evaluačního výzkumu a na identifikaci některých praktických obtíží. Diskutovali jsme mnoho různých přístupů a teorií. Nicméně lze říci, že panuje obecný konsensus v tom, že jedna z hlavních definujících charakteristik evaluačního výzkumu je v tom, že se předpokládá jeho užitečnost, resp. přenesení výsledků do praxe.

Evaluační teoretici mohou poskytnout různé definice evaluace (jak bylo uvedeno v první kapitole), ale v zásadě téměř veškerá literatura o evaluaci o ní hovoří jako o tom typu aktivit, které mají posloužit rozhodovatelům a všem „stakeholderům“ ohledně dalšího fungování sociálního programu. Primárním účelem evaluace tedy je hodnotit záležitosti spojené se sociálním programem tak, aby poskytovala informace a pomohla těm lidem, kdo jsou odpovědní za rozhodování o budoucnosti programu. Jak proklamují Rossi a Freeman (1993:443, vl. překl.): „Nakonec budou všechny evaluace posuzovány podle jejich použitelnosti“. Zaměříme se teď tedy na obecné otázky „použití“ zjištění z evaluačního výzkumu.

9.1 Vyjasnit si účel studie: Typologie účelu výzkumu

Rozhodnutí o designu, analýze a způsobu prezentování výsledků vždy závisí na typu studie, kterou provádíme. Takže prvním krokem, který bychom měli učinit, je ujasnit si účel výzkumu. Alternativní účely můžeme podle Pattona (2002:213) rozdělit podél kontinua od teorie k akci:

1. Základní výzkum: Má přispět k základnímu poznání reality a teorii.
2. Aplikovaný výzkum: Má osvětlit společenský zájem (veřejný zájem).

3. Sumativní evaluace: Má stanovit efektivitu programu.

4. Formativní evaluace: Má zlepšit fungování programu.

5. Akční výzkum: Má vyřešit specifický problém.

Sociální vědci v základním a aplikovaném výzkumu publikují v odborných časopisech, jejichž čtenáři jsou opět výzkumníci, kteří pak posuzují jejich přínos a používají k tomu kritéria jako je vědecká přesnost, validita a jasné a srozumitelné využití teorie. V protikladu k tomu podle Pattona (2002:213) evaluační výzkumníci a akční výzkumníci publikují zprávy pro specifické „stakeholdery“, kteří pak na základě jejich zpráv činí rozhodnutí, zlepšují fungování programů a řeší problémy.

U jednotlivých účelů výzkumu jsou odlišná jak očekávání od výzkumu, tak i příjemci zjištění. Liší se i způsob podávání zpráv z výzkumu a způsob šíření těchto informací. Právě díky těmto odlišnostem si musí být výzkumník hned na počátku vědom, jaké jsou priority, jaký je účel výzkumu. Žádná jednotlivá studie podle Pattona nepokryje všechny tyto odlišné účely a okruh příjemců zjištění. Patton (2002:224) uvádí přehlednou typologií účelů výzkumu (viz tabulka níže).

Tabulka 4 - Typologie účelu výzkumu

Typ výzkumu	Účel	Zaměření výzkumu	Požadované výsledky	Způsob generalizace	Klíčové předpoklady	Způsoby publikace	Standardy posuzování
Základní výzkum	Poznání jako takové; objevit pravdu	Otázky jsou důležité pro samotnou disciplínu nebo osobní intelektuální zájem výzkumníka	Příspěvek k rozvoji teorie	Zobecnění v čase i prostoru (v ideálním případě)	Realita obsahuje „vzorce“, tyto „vzorce“ jsou pochopitelné a vysvětlitelné	Odborné časopisy v jednotlivých oborech, odborné knihy	Vědecká přesnost postupu výzkumu, univerzalita a ověřitelnost teorie
Aplikovaný výzkum	Porozumět podstatě a příčinám lidských a sociálních problémů	Tyto otázky jsou důležité pro společnost jako celek	Příspěvek k teorii, což pomůže formulovat programy zaměřené na řešení problémů a intervence	Tak obecné v čase a prostoru jak je to jen možné, ale v jasně ohraničeném kontextu aplikace zjištění	Lidské a sociální problémy lze pochopit a lze je řešit pomocí znalostí o problematice	Specializované akademické časopisy, časopisy referující o aplikovaném výzkumu, interdisciplinární časopisy zaměřené na řešení problémů	Vědecky přesný a hluboký teoretický vhled do problému
Sumativní evaluace	Posoudit efektivitu našich intervencí a aktivit (programů, personálu, produktů)	Cíle intervence	Závěrečné soudy a generalizace týkající se efektivních typů intervence a posouzení podmínek, za jakých jsou tyto intervence	Všechny intervence s podobnými cíli	Co funguje zde za specifikovaných podmínek, může fungovat i jinde	Evaluační zprávy pro financovatele programu a tvůrce politik, specializované časopisy	Zobecnitelnost do budoucna a na jiné programy za specifikovaných podmínek

			efektivní				
Formativní evaluace	Zlepšit zásah do společnosti: program, intervenci či produkt	Silné a slabé stránky specifického programu, či personálu, který je předmětem studie	Doporučení pro zlepšení fungování programu	Zobecnitelnost je omezená na specifické podmínky programu, které jsou studovány	Lidé jsou schopni využít informace ke zlepšení toho co v každodenní realitě dělají	Ústní předání informací, vnitřní zprávy pro konkrétní organizaci, jiní hodnotitelé	Využitelnost a skutečné využití zainteresovanými aktéry v konkrétním studovaném programu
Akční výzkum	Řešit problémy v programu, organizaci či komunitě	Organizační problémy a problémy v komunitě	Bezprostřední akce, řešení problémů tak rychle jak to jen jde	Jen teď a tady	Lidé v konkrétních studovaných podmínkách jsou schopni řešit problémy studiem sebe sama	Interpersonální interakce mezi participandy na výzkumu, neformální nepublikované výstupy	Vnitřní pocity spokojenosti mezi participandy výzkumu, proveditelnost navržených řešení

Zdroj: Patton 2002:224

Podle Šimíkové, Navrátila a Winklera (2004:5) je „...významným cílem většiny evaluačních pracovníků dosáhnout stavu, kdy jejich výzkumné výsledky nejsou partnery z praxe bezvýhradně přijímány. Jednoznačně pozitivní ohlasy reprezentantů cílových skupin nebo politiků a programových pracovníků jsou ukazatelem malé praktické užitečnosti evaluací. Podobně jako jejich souhrnné odmítnutí.“

Problém použití evaluace se stal předmětem mnoha debat. Hlavním rysem Pattonova přístupu orientovaného na využití (Utilization-focused evaluation) je to, že klade na první místo použití. Celý proces evaluace se stává interaktivním, v souvislosti s tím jak hodnotitelé pracují v těsnějším kontaktu se „stakeholdery“. Patton uvádí v život přístup k evaluaci, který doporučuje spolupráci mezi výzkumníky a sponzory nebo uživateli výzkumu. V obecném smyslu dává zaměření na použití evaluačních zjištění vyniknout třem hlavním otázkám: Jakými způsoby mohou být evaluační zjištění užitečná? Jaké jsou hlavní faktory, které ovlivňují použití? Co je možné udělat ke zvýšení vlivu evaluačních zjištění?

Nejběžněji vnímaný výstup z evaluační studie je závěrečná zpráva. Tu je možno chápat jako tzv. „veřejně politický dokument“.

9.2 Veřejně politické dokumenty

Evaluační studie a závěrečné zprávy tedy patří mezi tzv. „veřejně politické dokumenty“ (dále jen VPD). Podle funkce a způsobu vzniku je možno podle Veselého, Drhové a Nachtmannové (2007:50) členit VPD do čtyř velkých typů: (1) legislativní (ústavní

zákon, zákon, nařízení, vyhláška, mezinárodní dohody), (2) strategické (strategie, koncepce, dlouhodobé záměry, plány, akční plány, plány rozvoje, strategické plány, národní politiky), (3) výzkumné (veřejněpolitické analýzy, výzkumné zprávy) a (4) deklarativní a advokační (programové prohlášení vlády, dokumenty politických stran).

Evaluační studie tedy patří mezi „výzkumné“ veřejně politické dokumenty. Hlavním smyslem výzkumných VPD je „...co nejobektivnější analýza situace v určité oblasti. Jde o dokumenty, které mají velmi silný analytický základ, tj. opírají se o dostupné teorie a empirická data. Velmi často je jejich součástí sada indikátorů, které popisují stav v dané problematice. Hlavním kritériem je objektivita, úplnost a přesnost. To je také jejich hlavní přednost“ (Veselý, Drhová, Nachtmannová 2007:54).

Na druhou stranu za slabiny výzkumných VPD citovaní autoři (ibid.) považují jejich „...často malou bezprostřední převoditelnost do politické praxe. Často mají podobu spíše jakéhosi ‚obecného vodítka‘ či podkladu a nikoli bezprostředního návodu. Ačkoli obsahují návrhy opatření (jinak by nešlo o VPD), ta jsou často na poměrně vysoké míře obecnosti.“ Výzkumné VPD týkající se „velkorozsahových“ veřejných politik tak ze své podstaty vznikají nejčastěji na půdě univerzit, výzkumných ústavů či tzv. think-tanků. Studie týkající se malorozsahových sociálních programů (lokální projekty sociálních služeb atd.) však také lze považovat za VPD, jelikož ovlivňují lokální politiku (např. lokální politiku sociálních služeb). Takovéto studie pak samozřejmě vznikají v jakékoliv specializované lokální instituci např. neziskového typu, nebo v samotných organizacích, které sociální program realizují.

9.3 Typy použití evaluačních zjištění

Jak tvrdí Clarke (1999), v raných dobách evaluačního výzkumu existoval obecný předpoklad, že největšími příjemci informací poskytovaných hodnotiteli budou tvůrci programů a programoví manažeři. A ti budou připraveni činit svá rozhodnutí na základě zjištění plynoucích z evaluačních studií. Podle tohoto pohledu na věc by úspěšné programy měly být na základě evaluačních zjištění podporovány a programy, které by byly shledány neefektivními nebo by měly nepředvídané negativní vedlejší efekty, by byly buď modifikovány nebo zrušeny. V praxi se však tento předpoklad podle Clarke (ibid.) nepotvrdil. Ve skutečnosti programový management považuje spíše za vysoce nevhodné

činit rozhodnutí o budoucnosti programů pouze na základě výsledků jedné evaluační studie. Nicméně to pochopitelně nemusí nutně znamenat, že evaluace nemají svůj vliv.

Carol Weissová (1988 dle Clarke 1999), jedna z nejznámějších autorek, která píše o evaluaci, popisuje čtyři způsoby, kterými jsou evaluační informace používány pro rozhodovací proces (d-m proces):

1) Informace poskytovaná hodnotiteli může působit jako varování, že se něco nedaří. Když tedy takováto výzkumná zjištění naznačují, že například některé skupiny v rámci cílové populace nedostávají slibované služby, pak může dojít na korektivní akce, a to ještě dříve, než se situace dále zhorší.

2) Evaluační zjištění mohou poskytovat návody ke zlepšení programu. V případech, kde je například k dispozici větší množství různých způsobů poskytování služeb cílové populaci, může komparativní analýza pomoci identifikovat, který přístup funguje nejlépe.

3) Třetí příspěvek evaluace je v poskytování nového pohledu na známý problém. Weissová o tom referuje jako o „rekonceptualizaci“.

4) Evaluace může být také použita k mobilizaci podpory pro program.

Když hovoří o problematice použití zjištění z programové evaluace Rossi a Freeman (1993:174), vydělují tři typy použití:

1) přímé neboli instrumentální použití,

2) konceptuální a

3) přesvědčující (persvazivní) použití.

1) Přímé neboli instrumentální využití je na místě, když jsou evaluační zjištění používána přímo programovými plánovači nebo manažery k tomu, aby učinili změny ve způsobu fungování programu. Jak bylo poznamenáno už v úvodu této kapitoly, ačkoli programy se jen zřídka zruší na základě samotných evaluačních zjištění, evaluace mohou mít dopad na rozpočty, nebo vést ke změnám v interní struktuře programu.

2) Konceptuální využití, nebo „rekonceptualizace“ (abychom využili terminologii Weissové), se objevuje, když evaluace ovlivňuje smýšlení lidí o programu. Nejsou to jen zjištění z evaluace, která mohou mít konceptuální dopad. Je jím rovněž samotný fakt, že evaluace může způsobit, že osoby zahrnuté do realizace programu najednou vidí v jiném

světla, proč jsou úkoly a aktivity organizovány právě daným způsobem, což může zmírnit napětí mezi managementem a dalším programovým personálem.

3) Přesvědčující (persuasive) využití poukazuje na užití evaluační informace k podpoře nebo potlačení nálad pro nebo proti programu.

Podle Weissové (1986:31-36 dle Clarke 1999:175-176) potřebuje koncept „využití výzkumu“ vyjasnit. Weissová identifikuje dokonce sedm perspektiv ve využití výzkumu. Ačkoli referuje o sociálně-vědním výzkumu obecně, různé modely, které takto identifikuje, mají svou relevanci i v kontextu evaluačního výzkumu. Sedm různých významů využití výzkumu je popisováno následovně:

1) Model vedený znalostí (knowledge-driven model): Tento model se odvozuje od fyzikálních věd, kde základní výzkum poskytuje zdroj pro aplikovaný výzkum, který eventuálně vede k praktické aktivitě.

2) Model řešení problémů (problem-solving model): Podle tohoto modelu empirická výzkumná zjištění poskytují druh informace, které potřebují rozhodovatelé. Tam, kde je výzkum veden s cílem vysvětlit specifický problém, mohou mít jeho zjištění přímý dopad.

3) Interaktivní model: Nejsou to jen sociální výzkumníci, kdo poskytují znalost a informace pro rozhodovatele. V tomto modelu jimi jsou sociální vědci, kteří se angažují v interaktivním hledání informací spolu s politiky, administrátory, klienty, praktiky, žurnalisty, nátlakovými skupinami a podobně.

4) Politický model: Za určitých okolností jsou managementem programu přejímány politické a ideologické pozice, což může způsobit, že budou méně vnímaví k výzkumným zjištěním. Proto se výzkum může stát silnou zbraní v ruce té strany, která je odpůrcem stávajícího managementu a vidí, že by jim evaluační závěry mohly pomoci.

5) Taktický model: Tento model použití výzkumu nemá co do činění s obsahem nebo zjištěními studie, ale referuje o tom, že samotný fakt existence výzkumu může být rozhodovateli použit k taktickému odklonění kritiky od programu.

6) Informační, resp. osvětový model (enlightenment model): Tento model reprezentuje mechanismus, kterým si sociálně-vědní výzkum hledá cestičku do cyklu tvorby politiky. Představou je, že zobecnění činěná sociálním výzkumem se profiltrují skrze informovanou veřejnost a formují tak způsob, jakým lidé smýšlejí o sociálních otázkách. Nejsou to tedy jednotlivé studie, které mají v rámci tohoto modelu výrazný vliv, více jde o případy

konceptů a teorií, které jsou vytvářeny v rámci sociálních věd obecně a informují posléze rozhodovatele.

7) Model intelektuální iniciativy (intellectual enterprise model): V této interpretaci využití výzkumu jsou sociální vědy viděny jako jeden element v rámci všech intelektuálních aktivit společnosti. Jak výzkumníci, tak tvůrci politiky, ovlivňují a jsou sami ovlivněni trendy ve veřejném mínění.

Weissová připouští, že žádná z těchto perspektiv plně nevysvětlí využití výzkumu. Každý zmíněný model však vytváří určitý příspěvek našemu porozumění použití, které může sociálně-vědní výzkum mít a potažmo i programová evaluace.

9.4 Jak maximalizovat využití?

Chen (2005:59-61) při svých tipech, jak maximalizovat využití výsledků programové evaluace, odkazuje na již zmíněné kroky nakládání se svou praktickou taxonomií. Podle něj by mohlo vést ke zdárnému výsledku při využití evaluačních zjištění přidržení se následujících kroků:

1. Identifikovat stádium programu, který je předmětem zájmu (iniciační stádium, stádium počáteční implementace, stádium pokročilé implementace, stádium již zaběhaného programu).
2. Vybrat evaluační strategii, která koresponduje s interními/externími záměry „stakeholderů“ (potřeba zlepšit aktuální chod programu, vyjádření odpovědnosti za výsledky běhu programu).
3. Vybrat evaluační přístup a výzkumné metody, které poskytují akceptovatelné kompromisy (trade-offs), např. mezi časovou náročností a vědeckou přesností, mezi důkladností a výdaji.
4. Komunikovat se „stakeholdery“ fakta o zvolené evaluační strategii/přístupu a výzkumných metodách („stakeholderi“ by měli být co nejlépe srozuměni s druhy informací, které budou finálním produktem evaluace, komunikace pomáhá zabránit nedorozuměním mezi „stakeholdery“ a evaluátory).

Alkin (1990 dle Clarke 1999) dále poskytuje některé obecné návody pro využití evaluačních zjištění, které obsahují následující prvky:

- 1) identifikovat klíčové rozhodovatele,
- 2) ujistit se, že evaluace se zaměřuje na otázky, které vznesli sami „stakeholdeři“,
- 3) omezit evaluační studii na proměnné, nad kterými mají rozhodovatelé nějakou kontrolu,
- 4) výsledky podávat brzo a ujistit se, že zpráva je krátká, stručná a jasná.

Podobně na základě textu Solomona a Shortella (1981), uvádějí i Rossi a Freeman (1993:449) seznam pěti pravidel, která by měli hodnotitelé sledovat s cílem maximalizovat využití evaluace:

1) Hodnotitelé musí porozumět tomu, jak pracují a co chtějí sami rozhodovatelé. Informace by měly být prezentovány ve formě nejpřijatelnější pro cílovou populaci. Například klíčoví rozhodovatelé nemají čas probírat se skrze dlouhé a detailní výzkumné zprávy z evaluace, stručné shrnutí je pro ně efektivnější.

2) Evaluační výsledky musí být včasné a dostupné v době, když jsou nejvíce potřeba. Hodnotitel může být požádán, aby data byla dostupná už v době, kdy ještě není výzkumná zpráva kompletně hotová. Zatímco konzervativněji založení vědci to mohou vidět jako porušení vědecké přesnosti, pragmatičtější orientovaní hodnotitelé pravděpodobně uznají, že poskytnout některé informace je lepší, než neposkytnout žádné.

3) Evaluace musí respektovat závazky stakeholderů k programu. Jednou z cest, jak zvýšit využití evaluace, je podnítit hlavní skupiny stakeholderů k participaci na evaluačním procesu už od raných fází evaluace. To zajistí, že evaluace je citlivá k různým pohledům a zájmům různých skupin stakeholderů.

4) Hodnotitelé by měli diskutovat s potenciálními uživateli o tom, jak by mohly být výsledky interpretovány.

5) Evaluace by měly zahrnovat hodnocení jejich vlastního využití (doporučuje se, aby se hodnotitelé a potenciální uživatelé předem dohodli, na jakých základech budou posuzovat užitečnost evaluace).

9.5 Závěrečná zpráva z evaluace

Výzkumník hodnotící sociální program by měl dobře plánovat adekvátní způsoby publikování a rozšiřování (disseminace) evaluačních výsledků a doporučení. Jedním z dilemat může v tomto ohledu být i kruciólní otázka: Napsat zprávu či nenapsat zprávu?

Patton (2002) sám tvrdí, že ve své dosavadní práci získal tu zkušenost, že závěrečné zprávy mají často menší dopad než přímá interakce tváří v tvář s příjemci evaluace. Uspořádat ústní „brífingy“ se stává stále důležitější kompetencí evaluátora (Hendricks 1982 dle Patton 2002:510). Závěrečné zprávy často představují důležitou funkci disseminace směrem k příjemcům zjištění, kteří nejsou přímými tvůrci rozhodnutí v sociálních programech, ale stojí tak zvaně „opodál“. Závěrečné zprávy však nejsou automaticky a nezbytně primárním zdrojem informací pro ty, u kterých se předpokládá bezprostřední využití evaluačních zjištění.

Podle Rossiho a Freemana (1993:452) je v zásadě možné vytvořit hned několik oddělených závěrečných zpráv a stručných shrnutí evaluace, a to v různých stylech a formátech pro různé cílové skupiny. Proto je možné odlišit mezi primárním a sekundárním rozšířením materiálů mezi příjemce (primární a sekundární disseminace).

Primární disseminace je určena k publikaci plných verzí zprávy, technicky detailních verzí. Druhý způsob informuje příjemce o evaluačních zjištěných způsoby, které jsou přizpůsobené potřebám specifických „stakeholderů“. Sekundární disseminace nezahrnuje podle Rossiho a Freemana (1993) jen psanou zprávu, ale může nabývat mnoha rozdílných forem, včetně slovních zpráv, obrazové prezentace výsledků nebo dokonce filmů a videa.

Když tedy píšeme pro sponzory a obecně stakeholdery programu, měly by být podle Clarke (1999) minimalizovány technické termíny a tam, kde je to potřeba, by měl být uveden i glosář hlavních pojmů. Je rovněž vhodné zahrnout do závěrečné výzkumné zprávy stručné shrnutí, jelikož se dá předpokládat, že mnoho politiků, rozhodovatelů a programového personálu nebude číst celou zprávu.

Miles a Huberman (1994 dle Hendl 2005:315) „...považují za minimální požadavek, že zpráva o kvalitativní studii má:

1. Vysvětlit, čím se výzkum zabýval;
2. Informovat jasným způsobem o sociálním a historickém kontextu prostředí, kde se výzkum prováděl;

3. Sdělit přirozenou historii výzkumu, aby se vědělo, co se dělalo, kdo to dělal a jak (popisujeme, jak se našly klíčové koncepty, které proměnné se objevily nebo se postupně ztratily, jaké kódy vedly k důležitému poznání);

4. Obsahovat základní data (ve formě krátkých vyprávění, organizovaných úryvků z interview, grafů, fotografií, atd.), aby si čtenář mohl sám vyvodit závěry paralelně s autorem;

5. Formulovat závěry a popsat jejich širší význam pro vědu nebo v souvislosti s efekty, které mohou mít.“

Odborný či vědecký článek může být vhodnou formou referování o proběhlém výzkumu. V takovém případě mívá často kompoziční strukturu zvanou IMRAD, což je zkratka anglických slov Introduction, Methods, Results a Discussion, tedy úvod, metody, výsledky a diskuse. V tomto čtyřčlenném modelu odborné publikace dominují rysy logičnosti a objektivity, proto lze tvrdit, že se tento model uplatňuje zejména v časopisech přírodovědného zaměření. Nicméně lze podle Hendla (2005) přijmout názor, že „...tato standardní forma vědeckého článku je kompromisem, jenž dovoluje v mnoha případech i kvalitativnímu výzkumníkovi sdělit to podstatné o provedeném výzkumu.“

V evaluační zprávě je mimořádně důležité zpracovat odstavec týkající se doporučení. Doporučení mají být podle Hendla (2005) odvozena z dat a jasně formulována. Ohledně struktury závěrečné zprávy pak Hendl (2005:327) uvádí, že „...začátek zprávy o evaluaci obsahuje akční souhrn. Na 2 až 5 stránkách výzkumník uvádí v komprimované podobě záměr evaluace, použité metody, popisuje nejdůležitější výsledky, závěry a doporučení. Přitom se snaží odlišovat výsledky (nalezená fakta), interpretace, posouzení a závěry. Zpráva o evaluaci má také obsahovat část, která osvětluje, co se má zlepšit ve fungování programu.“

Klíčové části evaluační zprávy pak lze shrnout např. takto (Hendl 2005:328):

1. „Akční souhrn. Přehled celého hodnocení v rozsahu 2 až 5 stránek. Je určen pro zadavatele hodnocení. Vyjadřuje se netechnickým způsobem.

2. Popis programu. Části a prvky programu, aktivity, materiální předpoklady. Cíle (pokud možno zobrazená logika programu). Cílová skupina, místo. Organizace programu, personální obsazení.

3. Cíle hodnocení. Zdůvodnění a volba přístupu k evaluaci. Popis metod. Popis metod sběru dat, vzorkování. Popis metod analýzy.

4. Výsledky. Popis charakteristik účastníků a míst. Hodnocení účasti. Popis výsledků a jejich komparace. Popis témat, která se vyskytla, jejich ilustrace citacemi, úryvky nebo vyprávěním. Zobrazení výsledků pomocí tabulek a grafů.

5. Závěry a doporučení. Diskuse všech důležitých výsledků s ohledem na původní zadání projektu evaluace. Validita výsledků a jejich omezení. Je nutné uvést zdůvodnění pro každé doporučení.“

Závěrečná zpráva je podle Clarke (1999) konečným produktem evaluace (o „konečném“ produktu evaluace je možné hovořit maximálně jen na straně hodnotitele, který dále nebude participovat na rozvoji programu, pro program by naopak měla být závěrečná zpráva vstupem pro další změny - pozn. aut.) a je důležité, aby byla jasně prezentována a dobře napsána. Forma zprávy může být předem předepsána (managementem programu), nebo může být v rukou hodnotitele. Níže jsou v tabulce uvedeny obecné rady ohledně obsahu a prezentace závěrečné zprávy a ohledně další z možností jak strukturovat finální výzkumné evaluační zprávy.

Tabulka 5 - Struktura typické evaluační zprávy

Oddíl	Obsah
9.6 Titulní strana	<ul style="list-style-type: none"> - název programu, jeho umístění a trvání - jméno organizace, která zaštiťuje studii - jméno(a) hodnotitele(ů) - datum podání zprávy
Obsah	<ul style="list-style-type: none"> - soupis hlavních oddílů a pododdílů - seznam všech grafů, obrázků a tabulek
Stručné shrnutí (Summary)	<ul style="list-style-type: none"> - stručný přehled hlavních zjištění a doporučení - zmínka o těch rysech výzkumného designu nebo metodologie, které mohou hrát roli při interpretaci zjištění
Úvod	<ul style="list-style-type: none"> - okolnosti výzkumu <i>O programu:</i> - popis původu programu, jeho hlavní elementy a zamýšlené cíle

	<ul style="list-style-type: none"> - vysvětlení odůvodnění programu - náčrt okolností, za kterých program běží <p><i>O evaluaci:</i></p> <ul style="list-style-type: none"> - podstata, rozsah a typ použité evaluace - shrnutí hlavních evaluačních otázek - informace o jiných evaluacích podobných programů
Metodologie	<ul style="list-style-type: none"> - diskuse o výběru výzkumného designu - popis typu použitého měření výsledků, podstaty použitých nástrojů ke sběru dat a informace o otázce reliability a validity výzkumu (či alternativních kritérií pro posouzení kvality výzkumu – pozn. aut.) - náčrt časového plánu sběru dat - metody použité k analýze dat
Zjištění	<ul style="list-style-type: none"> - výsledky - diskuse o zjištěních - doporučení budoucích aktivit
Přílohy	<ul style="list-style-type: none"> - příklady použitých výzkumných nástrojů (dotazníky, hodnotící škály atd.) - grafy, obrázky a tabulky, které nebyly zahrnuty do hlavního „těla“ výzkumné zprávy - technické informace vztahující se k otázkám výzkumného designu nebo analýzy dat - odkazy (zdroje informací, jiné výzkumy)

Zdroj: Clarke 1999:182, vl. překl.

Bez ohledu na to, jaké metody disseminace jsou nakonec vybrány, vždy bude potřeba vyprodukovat nějakou formální, psanou zprávu. Uživatelé evaluace, stejně tak jako samotní hodnotitelé, pak mají zájem na co nejvyšší kvalitě závěrečné zprávy. Proto se přistupuje i k hodnocení kvality evaluačních zpráv. V podstatě se tak evaluuje samotná evaluace.

V tabulce uvedené níže Monnier (1997 dle Clarke 1999:184) popisuje osm kritérií, které je dobré brát v úvahu při hodnocení evaluační zprávy. Stručně řečeno, pokud chtějí

hodnotitelé maximalizovat dopad své evaluace, musí identifikovat správné problémy, vymyslet správné evaluační otázky, přijmout a použít přísné a systematické metodologické procedury a dát si pozor při prezentaci informací, aby k ní došlo způsobem, který bude srozumitelný různým skupinám „stakeholderů“.

Tabulka 6 - Hodnocení kvality evaluačních zpráv

Úkol	Indikace splnění úkolu
Potřeby pokrytí témat	Adresuje evaluace informace adekvátně potřebám zadavatele?
Relevantní rozsah	Jsou otázky ohledně programu, jeho design a výstupy, výsledky a dopady plně prozkoumány? Zahrnují jak zamýšlené tak nepředpokládané politické interakce a konsekvence?
Obhájitelný design	Je evaluační design vhodný a přiměřený? Umožňuje v rámci metodologických omezení odpovědět na hlavní evaluační otázky?
Reliabilní data	V jakém rozsahu jsou primární a sekundární data sesbírána? Poskytují adekvátní stupeň validity a reliability pro zamýšlené účely?
Jasná analýza	Jsou kvantitativní a kvalitativní informace přiměřeně a systematicky analyzovány podle profesionálních standardů, tak aby evaluační otázky byly zodpovězeny validním způsobem?
Důvěryhodná zjištění	Vyplývají zjištění a soudy logicky z analýzy dat a je interpretace založena na popsanych předpokladech a racionálním základě?
Nezaujaté závěry	Jsou doporučení jasná a nezkreslená personálem nebo „stakeholdery“? Jsou doporučení detailizována a aplikovatelná do praxe?
Jasná závěrečná zpráva	Popisuje zpráva jasně program, který byl hodnocen, včetně jeho kontextu a účelu, spolu s procedurami a zjištěními z evaluace, takže poskytované informace mohou být snadno pochopitelné?

Zdroj: Monnier 1997 in Clarke 1999:184, vl. překl.

Obzvláště v případě kvalitativních studií je obtížné dostat ideální modelové kvalitě závěrečné studie. Vhodným způsobem, jak se podle Hendla (2005:333) „...naučit psát dobrou zprávu o kvalitativním výzkumu, je studium publikovaných příkladů. Podotýkáme, že ve zprávě o kvalitativním výzkumu se častěji používá Ich-forma. Tento způsob psaní pomáhá situovat kvalitativního výzkumníka do procesu výzkumu a nutí ho zodpovědně přiznat jeho aktivní roli při výzkumu. Dále je nutné používat pseudonymy při označování

jedinců, institucí atd., aby byly zohledněny etické aspekty při zveřejňování privátních informací.“

V této kapitole bylo uvedeno vícero praktických rad k vedení evaluace. Nejen k otázce obecného zvýšení (maximalizace) použití výzkumných zjištění, ale i praktických informací k sepsání závěrečné zprávy, kde má být shrnuto vše podstatné z právě provedené evaluace. Obzvláště poslední tabulka, která shrnuje kritéria pro posuzování provedené programové evaluace, je dobrým vodítkem pro konstrukci dobrých evaluací. Ale neměli bychom zapomínat, že často zde zmiňovaná slova jako je validita a reliabilita, by mohly mít své alternativy v podobě důvěryhodnosti, přenositelnosti či spolehlivosti.

10 Závěr

Lze předpokládat, že v některých programech bude hrát evaluace výraznou roli (jakožto institucionalizovaná aktivita), v jiných případech ovšem bude evaluace přítomna pouze jaksi formálně, nebo bude využívána velmi málo či vůbec. V takových případech lze předpokládat, že nebude mít evaluace tak výrazný vliv na změny v sociálních programech. Podle Matouška (in Smutek 2005:10) „...je stav u nás takový, že ti kdož rozhodují, rozhodují podle svých vazeb na poskytovatele, v lepším případě podle důkazů, jež mají charakter anekdot. Realizátoři konstruují program také podle subjektivních preferencí a nemusejí se obávat kritické prověrky založené na nezávislém posouzení toho, jak program zlepšil kvalitu života uživatele, resp. jak prospěl obci. Uživatelé nemají k dispozici ani kriteria, ani data ukazující kvalitu programů.“ K těmto problémovým bodům se snaží přispívat např. proces plánování rozvoje sociálních služeb (komunitního plánování), proces standardizace a vyhodnocování kvality sociálních služeb a další způsoby vlivu na kvalitu sociálních programů na mikro či mezzo rovině. Jde obecně tedy o přístupy, které vycházejí ze zásad Nového veřejného managementu (NPM). Na makro rovině zpětné vazby (národní a mezinárodní rovina) pro „velkorozsahové“ sociální programy se dá říci, že službu zhodnocení formování i možných dopadů reforem vykonávají evaluační studie přicházející z univerzit, think-tanků či výzkumných ústavů (např. Výzkumný ústav práce a sociálních věcí, dále jen VÚPSV). Zde lze doporučit nahlédnout do řady studií, provedených na zakázku VÚPSV a které mají charakter sumativních i formativních studií (viz www.vupsv.cz). Jako zdroj nástrojů pro provedení evaluace v kontextu veřejné politiky by bylo možné doporučit např. „Banku evaluačních nástrojů“, která je uveřejněna v tematické oblasti protidrogové politiky (www.drogy-info.cz) nebo podobně v oblasti vzdělávací politiky na stránkách Národního ústavu odborného vzdělávání (www.nuov.cz).

Takováto systematická zpětná vazba by měla přispět k tomu, že sociální programy budou vnitřně konstruovány tak, aby byly co nejfunkčnější či dokonce trvale udržitelné. Dobrý program musí podle Matouška, Koláčkové a Kodymové (2005:278) „...pokrývat relevantní okruh klientových potřeb a zahrnovat okruh pro klienta nejvýznamnějších osob. To znamená, že by organizátor programu měl umět předběžně tyto potřeby a tyto osoby definovat. Programy reagující na jedinou potřebu, opírající se o jednodimenzionální model klientova problému, nebývají účinné.“

Při konstrukci funkčního a trvale udržitelného sociálního programu by neměla být podceňena role teoretických vysvětlení příčin a souvislostí sociálního problému. Formulace relevantního okruhu potřeb a osob by podle Matouška, Koláčkové a Kodymové (2005:278) „...měla vždy vycházet z konzistentní a komplexní teorie vysvětlující příčiny... chování i možnosti jeho ovlivnění.“ Funkční a udržitelný program musí mít dále „...dostatečnou intenzitu a délku. Musí mít jasnou strukturu, jasné zacílení. Cíl musí být zřejmý klientovi a spolupracujícím osobám z klientova prostředí, klient by s ním měl být srozuměn. Pro provádění programu musejí být pracovníci dostatečně vyškoleni a musí být průběžně monitorováno, zda program běží konzistentně s plánem a zda je dosahováno plánovaných cílů“ (ibid.).

Všechna tato doporučení a mnohá další vyplývají z rozsáhlého rozboru problémových otázek a širších souvislostí evaluace sociálních programů, kterým se věnovaly všechny předchozí kapitoly. Nicméně lze vysledovat i kritické odkazy na racionální manažerské přístupy k realizaci sociálních programů. Jde o obecnou kritiku Nového veřejného managementu, který zdůrazňuje výstupy, jejich měření, tedy snahy o objektivní podchycení výstupů, jejich kvantifikaci, tvorbu časových řad atd. Tím se celý veřejný sektor a jeho řízení dostává do sfér pozitivní vědy a kvantitativních metod, což je podnětem ke zřetelným kritikám, hlavně ze strany konstruktivistických a interpretativních proudů sociálních věd. Finlayson (2003 dle Fitzpatrick 2005:163) uvádí rozsáhlou kritiku nových forem manažerismu ve veřejném sektoru:

1) Důraz ve veřejném sektoru se posunul ze vstupů (výdaje a sociální struktury) směrem k výstupům (výkonnost a efektivita). S limity sociálních výdajů, které byly brány v úvahu, bylo dosaženo tvůrci politiky růstu produktivity za stávajících zdrojů. Výstupy byly v rostoucí míře předmětem kvantifikace. Klíčovému měření výstupů je nastavení specifických cílů, které mohou někdy odrážet skutečné aktivity institucí pokřiveným způsobem. Ty aktivity organizací, které nejsou jednoduše měřitelné, byly vyčleněny z kategorie výstupů měřitelných, přitom ale velká část právě těchto aktivit je nakonec skutečnou, reálnou službou, např. neměřitelná psychická podpora klienta či rodiny. Jeden z důsledků tohoto důrazu na kvantifikaci může být, že tyto „nevýstupové“ aktivity vyzní do ztracena, jelikož k nim nejsou přiřazeny žádné „body“ a jelikož „požívají“ čas a zdroje, které by mohly být využity jinde, a proto mohou dokonce způsobovat sankční reakce inspekčních (hodnotících orgánů).

2) Ethos veřejných služeb orientovaný na sociální potřeby, byl částečně nahrazen souborem formálních, kontraktuálních vztahů. Tento přístup je bližší konkrétně vyjádřeným preferencím uživatelů, „zákazníků“ služeb. Zákazníci služeb jsou považováni za potenciálně schopné posoudit kvalitu veřejných služeb stejně jako jsou schopni posoudit kvalitu např. zubní pasty.

3) Soupeření v rámci organizací a mezi organizacemi veřejného sektoru se stalo stále běžnějším jevem. Od organizací veřejného sektoru se očekává, že budou více podnikat a budou se méně bát rizika, tedy že budou kopírovat příklady ze soukromého sektoru. Budou spolupracovat se soukromým sektorem různými podobami partnerství a budou umožňovat působení soukromým firmám, když budou samy selhávat.

4) Moc se posunula směrem k manažerům, takže se předpokládá, že si zvnitřní technologie (slovník a postupy) managementu jako odborné disciplíny. Profesionální autonomie řadových pracovníků tak byla v zásadě redukována. Režimy auditů a inspekcí se staly nekonečným kolotočem. A jsou to právě tyto inspekce (evaluace), které umožňují dlouhodobou institucionalizaci manažerských režimů. To pak může představovat i to riziko, že transformují své aktivity do podoby informací, které režim vyhodnotí jako užitečné, bez ohledu na potřeby reality.

Kritické poznámky k manažerismu tohoto typu uvádí dále např. Fitzpatrick (2005:163). Přežití různých organizací v rámci soupeřícího prostředí (které svým působením manažerismus podněcuje) se stává stále důležitějším prvkem reality. Různí aktéři tak mohou mít zájem na simulaci informací, které pak poskytují pro externí inspekce (hodnotitele). Tato simulovaná data pak mohou zajistit jejich přežití. Takže spíše než by odpovídala strategie nového veřejného managementu pragmaticky na svět okolo, nemusí podle Fitzpatricka (2005) dělat nic více, než vytvářet mlhavé prostředí, ve kterém statistiky nahrazují realitu a kvantita kvalitu, kde politika vedená znalostí/zkušeností (evidence based) se stává znalostí vedenou politikou (policy-led) a kde dosažené cíle se stávají samy o sobě známkou progresu, bez ohledu na skutečnost.

Nové podoby vládnutí ve společnosti (governance) a manažerismus ve veřejném sektoru jsou často považovány za pragmatické filosofie založené na evidenci (znalostech, zkušenostech), jsou tedy spojeny jen s tím co funguje ve společnosti. Sociologicky informovaná kritika trvá na tom, že diskurzivní techniky governance a manažerismu vytvářejí sociální svět, kterému se pak samy nabízejí jako řešení.

V žádném případě neměly vést předchozí odstavce k nějaké definitivní skepsi ohledně využití systematické evaluace sociálních programů. Celý tento text se snažil poukázat na to, že systematická evaluace je neocenitelná jak pro současné tak pro budoucí snahy o zlepšování společnosti. Měli bychom, společně s Rossim, Lipseyem a Freemanem (2004) poznamenat, že myšlenka, že sociální problémy společnosti, jako je kriminalita, nezaměstnanost či bezdomovectví budou zmírněny či úplně odstraněny působením efektivních a účinných sociálních programů, je velmi lákavá. Stejně tak je lákavé představovat si jistou formu „morální rekonstrukce“ společnosti cestou systematického hodnocení a racionální výstavby sociálních programů. Podle citovaných autorů je ale „...katastroficky naivní myslet si, že efektivní a účinné sociální programy toho dosáhnou“ (Rossi, Lipsey, Freeman 2004:418).

11 Literatura

- ALKIN, M. C. (Ed.). *Evaluation Roots - Tracing Theorists Views and Influences*. Thousand Oaks : Sage Publications, 2004.
- BABBIE, E. *The Practice of Social Research*. 8th Edition. Wadsworth Publishing Company, 1998.
- CLARKE, A. *Evaluation Research*. London: Sage Publications, 1999.
- COLEBATCH, H. K. *Úvod do policy*. Brno: Barrister and Principal, 2005.
- DENZIN, N. K.; LINCOLN, Y. S. (Eds.). *Collecting and Interpreting Qualitative Materials*. Thousand Oaks : Sage Publications, 1998.
- FIALA, P.; SCHUBERT, K. *Moderní analýza politiky. Uvedení do teorií a metod policy analysis*. Brno: Barrister and Principal, 2000.
- FITZPATRICK, T. *New Theories of Welfare*. Houndmills: Palgrave-Macmillan, 2005.
- FLEGR, J. *Zamrzlá evoluce aneb je to jinak, pane Darwin*. Praha: Academia, 2006.
- GREENE, J. C. Qualitative Program Evaluation. In: Denzin, N. K.; Lincoln, Y. S. (Eds.). *Collecting and Interpreting Qualitative Materials*. Thousand Oaks: Sage Publications, 1998.
- HENDL, J. *Kvalitativní výzkum*. Praha: Portál, 2005.
- HORÁK, P.; KULHAVÝ, V. *Analýza cílů veřejných programů současné veřejné a sociální politiky z vnitroorganizační perspektivy*. Fórum sociální politiky č. 2, 2008.
- CHEN, H.-T. *Practical Program Evaluation*. Thousand Oaks : Sage Publications, 2005.
- KELLER, J. *Teorie modernizace*. Praha : Sociologické nakladatelství, 2007.
- MARK, R. *Research Made Simple*. Thousand Oaks: Sage Publications, 1996.
- MATOUŠEK, O. Hodnocení sociálních služeb – výzva pro blízkou budoucnost. In: Smutek, M. (Ed.). *Možnosti sociální práce na počátku 21. století*. Hradec Králové : MaV, 2005.
- MATOUŠEK, O. a kol. *Sociální služby*. Praha : Potrál, 2007.
- MATOUŠEK, O. Vývoj hodnocení sociálních programů. In: Matoušek, O. a kol. *Sociální služby*. Praha: Portál, 2007.
- MATOUŠEK, O.; KOLÁČKOVÁ, J.; KODYMOVÁ, P. (Eds.). *Sociální práce v praxi*. Praha: Portál, 2005.
- NACHMIAS, D. (Ed.). *The Practice of Policy Evaluation*. New York: St. Martin's Press, 1980.
- NACHTMANNOVÁ, M.; VESELÝ, A. Klíčové dovednosti analytika veřejné politiky. In: Veselý, A., Nekola, M. (Eds.). *Analýza a tvorba veřejných politik - přístupy, metody a praxe*. Praha: SLON, 2007.
- PALUMBO, D.; SHARP, E. Evaluation of Process Versus Evaluation of Impact of Community Corrections. In: Nachmias, D. (Ed.). *The Practice of Policy Evaluation*. New York: St. Martin's Press, 1980.

- PATTON, M. Q. *Qualitative Research and Evaluation Methods*. Thousand Oaks : Sage Publications, 2001.
- PATTON, M. Q. *Qualitative Research and Evaluation Methods*. 3rd ed. Thousand Oaks: Sage Publications, 2002.
- PATTON, M. Q. *Utilization Focused Evaluation*. Thousand Oaks: Sage Publications, 1997.
- PETRUSEK, M. *Teorie a metoda v soudobé sociologii*. Praha: Karolinum, 1993.
- POLLITT, C. *The Essential Public Manager*. Maidenhead: Open University Press, 2003.
- ROSSI, P. H.; FREEMAN, H. E. *Evaluation – A Systematic Approach*. London: Sage Publications, 1993.
- ROSSI, P. H.; LIPSEY, M. W.; FREEMAN, H. E. *Evaluation – A Systematic Approach*. London : Sage Publications, 2004.
- ROSSI, P. H.; WILLIAMS, W. (Eds.). *Evaluating Social Programs*. New York: Seminar Press, 1972.
- SHAW, I. F. *Qualitative Evaluation*. London : Sage Publications, 1999.
- SMUTEK, M. (Ed.). *Možnosti sociální práce na počátku 21. století*. Hradec Králové: MaV, 2005.
- ŠIMÍKOVÁ, I.; NAVRÁTIL, P.; WINKLER, J. *Hodnocení programů zaměřených na snižování rizika sociálního vyloučení romské komunity – evaluace programů*. Praha: VÚPSV, výzkumné centrum Brno, 2004.
- VESELÝ, A. Veřejná politika. In: Veselý, A., Nekola, M. (Eds.). *Analýza a tvorba veřejných politik - přístupy, metody a praxe*. Praha: SLON, 2007.
- VESELÝ, A.; DRHOVÁ, Z.; NACHTMANNOVÁ, M. „Vědění o politice“: Teoretické poznatky o tvorbě veřejných politik. In: Veselý, A., Nekola, M. (Eds.). *Analýza a tvorba veřejných politik - přístupy, metody a praxe*. Praha: SLON, 2007.
- VESELÝ, A., NEKOLA, M. (Eds.). *Analýza a tvorba veřejných politik - přístupy, metody a praxe*. Praha: SLON, 2007.
- VESELÝ, A.; PATEROVÁ, H.; NEKOLA, M. „Vědění pro politiku“: praktické přístupy k analýze a tvorbě politik. In: Veselý, A., Nekola, M. (Eds.). *Analýza a tvorba veřejných politik - přístupy, metody a praxe*. Praha: SLON, 2007.
- WEINBACH, R. W. *Evaluating Social Work Services and Programs*. Boston: Pearson, 2005.
- WINKLER, J. *Implementace – Institucionální hledisko analýzy veřejných programů*. Brno: Masarykova univerzita, 2002.
- ŽIŽLAVSKÝ, M. *Metodologie pro Sociální politiku a sociální práci*. Brno: Masarykova univerzita, 2003.

12 Rejstřík

A

Alkin, 11, 37, 38, 39, 40, 75

B

Babbie, 14, 18, 47, 57, 62, 86
Berman, 31
Burgess, 60

C

Campbell, 41
Clarke, 9, 10, 11, 12, 23, 26, 27, 28, 29, 30, 35, 36, 43,
56, 57, 59, 60, 63, 64, 71, 72, 73, 75, 76, 78, 79, 80,
86
Colebatch, 23, 24, 86
Cook, 56

D

de Vaus, 59
DeBono, 54
Denzin, 49, 51, 52, 86
Drhová, 71, 87

F

Feek, 64
Fiala, 8, 15, 86
Finlayson, 83
Fitzpatrick, 25, 83, 84, 86
Flegr, 86
Foster, 24
Freeman, 13, 16, 18, 20, 30, 68, 72, 75, 85, 87

G

Greene, 15, 16, 86
Guba, 38, 39
Gunn, 23

H

Hendl, 22, 56, 76, 77, 86
Hendricks, 76
Herman, 27
Horák, 24, 86
Huberman, 43, 76

Ch

Chen, 11, 16, 36, 66, 74
Christie, 38

J

Jann, 15

K

Kelley, 54
Kodymová, 86
Koláčková, 86
Kulhavý, 24, 86

L

Lincoln, 52, 86
Lipsey, 20, 85, 87
Littman, 54
Love, 64

M

Mark, 86
Matoušek, 13, 20, 21, 86
Miles, 43, 76
Monnier, 79, 80

N

Nachmias, 18, 29, 86
Nachtmannová, 54, 55, 71, 86, 87
Navrátil, 87
Nekola, 10, 26, 86, 87

P

Palumbo, 31, 32, 86
Paterová, 26, 87
Patton, 8, 9, 12, 26, 30, 32, 33, 34, 50, 51, 52, 53, 54,
59, 63, 69, 70, 76, 87
Petrusek, 44, 45, 87
Plowden, 24
Pollitt, 24, 87

R

Reichardt, 56
Rossi, 12, 13, 16, 18, 20, 30, 33, 36, 68, 72, 75, 85, 87

S

Scriven, 26, 29, 30, 33, 34
Shadish, 21, 35, 53
Sharp, 31, 32, 86
Shaughnessy, 46
Schubert, 8, 15, 86
Smith, 57
Smutek, 1, 13, 82, 86, 87, 90
Suchman, 18

Š

Šimková, 87

T

Thatcher, 22
Tyler, 41

Williams, 33, 87
Winkler, 87

V

Veselý, 9, 10, 15, 26, 54, 55, 71, 86, 87
Von Oech, 54

Zechmeister, 46

Z**W**

Weinbach, 19, 87
Weiss, 11, 21, 33, 36

Žižlavský, 46, 47, 48

Ž

Redakční rada Edice texty k sociální práci:

Mgr. Karel Bauer; Mgr. Radka Janebová, Ph.D.; PhDr. Martin Smutek, Ph.D.;

Mgr. Zuzana Truhlářová, Ph.D.



Řada: Výzkumné metody v sociální práci – sv. 4

Název: **Evaluace sociálních programů**

Rok a místo vydání: 2014, Hradec Králové

Vydání: první

Náklad: 200

Vydalo nakladatelství Gaudeamus při Univerzitě Hradec Králové jako svou 1413. publikaci.

ISBN 978-80-7435-473-1