

EVALUAČNÍ VÝZKUM

EVALUATION RESEARCH

Bc. Jakub KOVÁŘ

Katedra antropologických a historických věd, FF ZČU, Plzeň

Abstract: *The article takes to task the well-established claim that academic anthropology should be free from traps of evaluation. Although, on the other hand, in applied anthropology we could see the strong accent on the evaluational research, as a new option for anthropologist. In this article are discussed the trends in evaluation research by means of their strategies, compared with academic research and with certain emphasis on theory-driven perspective introduced by Huey-tysh Chen. In addition to theoretical reflections on evaluation research, provides the article particular examples based on research conducted in the domain of social programs. The interdependence between methodology and theory of evaluation research is also mentioned.*

Keywords: *evaluation research, history of-, typology of-, evaluation strategies, anthropology*

*„Evaluation has become a mantra of modernity.“
(Pawson a Tilley 1997)*

Při prvním setkání s pojmem evaluace může mnoho antropologů pociťovat odpor. Tento postoj většinou pramení z proklamované snahy nesoudit hodnotové, institucionální, organizační a jiné aspekty zkoumané reality - snahy, jež má základ v tradici kulturního relativismu. V této tradici se každé hodnocení stává implicitně nástrojem kulturní hegemonie a v zásadě otevírá cestu pro nahlížení „těch druhých“ prizmatem našich měřítek. Paradoxem tohoto přístupu je, že sám *hodnotí* tyto snahy jako nesprávné a neuznává jedinou možnost, kterou by evaluace mohla (byť nepřímou) přispět k řešení problémů, které jsou součástí rozhodnutí, jež činíme jako sociální aktéři (Pomerantz 1984).

Problematika evaluace se tak většinou stává součástí sociální práce¹, kde nachází nejširší uplatnění v analýze fungování sociálních programů a politik. Použití evaluace v rámci sociální práce ji ovšem vystavuje zjevné vulgarizaci, při které často dochází pouze k jednosměrnému zhodnocení úspěšnosti implementace konkrétního sociálního programu na konkrétní sociální problém (Rubin a Babbie 1989). Při této redukci významu evaluace se zastírá potenciál tohoto nástroje, a to hlavně ve dvou oblastech. (1) evaluace se může stát prostředkem tvorby *veřejných programů*, které nejsou jen objektivním systémem, vždy ale také kulturním výtvořem lidí (Douglas a Isherwood 1979) a (2) evaluace není „teorie-prostá“, jen se často explicitně nepřiznávají její teoretická a metodologická východiska (Kelly 2004:526). Tím se tedy sama evaluace stává objektem sociálně vědních analýz a úvah. Její vztah k sociálně-vědnímu (základnímu) výzkumu nám, doufejme, její současnou pozici

¹ Kurzy „Evaluálního výzkumu“ jsou vyučovány na katedrách sociální práce a politiky. Okrajově na katedrách sociologie (viz www.fss.muni.cz, www.fsv.cuni.cz).

přiblíží.

Pojem evaluace je v praxi často spjat s konceptem implementace. Charles Lindblom se v 70. letech snažil (pomocí svého inkrementálního modelu tvorby politických rozhodnutí) poukázat na poměrně banální pravdu a sice, že proces implementace odráží kolektivní racionalitu, která „svazuje individuální rozhodování se společensko-politickým kontextem jedince a s praktickou situací, v níž se rozhoduje.“ (Winkler 2002:57) Považování implementace za sféru, v níž se setkává aktér a socio-politická struktura, se stalo rámcem evaluačního výzkumu až v posledních desetiletích. Pro pohled na evaluační výzkum jako autonomní doménu vědeckého bádání bych proto zaměřil na různé evaluační strategie, jejich vývoj, vzájemný vztah i jejich vztah k jiným doménám sociálních věd. Objeví se definice i elementární typologie evaluačního výzkumu.

Návodů na vedení evaluace existuje nepřehledné množství, ale možnosti jejich aplikace je třeba brát s rezervou. Huey-tyh Chen (2004) zpochybňuje univerzální použitelnost těchto manuálů z důvodu odlišných koncepcí jednotlivých sociálních programů. Přesto představím několik užitečných analytických nástrojů, které uvedl do evaluačního výzkumu právě zmiňovaný velikán svého oboru. Jelikož podle mne není lepšího způsobu vysvětlení výzkumných nástrojů, než jejich aplikace, budu se snažit dokazovat a ilustrovat svá tvrzení na konkrétní, již provedené evaluaci. Tímto bych rád uvedl, že překládané příklady jsou součástí projektu *Evaluace preventivních aktivit na taneční scéně*, který proběhl v létě roku 2008 v Terénním programu CPPT o.p.s. Plzeň (dále jen TP).

Evaluační výzkum/základní výzkum (sociálně-vědní výzkum)

Naše každodenní neformální interakce je protnutá mnoha soudy a hodnoceními, které buď účelově, nebo nevědomě pomáhají třídit naši sociální realitu. Tato kategorizace na škále dobrého a špatného, prospěšného a škodlivého atd. v sobě zahrnuje princip hierarchie, jež implicitně vychází z našeho přesvědčení. Evaluační výzkum se liší od našich každodenních soudů tím, že toto přesvědčení (nejčastěji v pozici veřejné politiky, vědeckého diskursu, politické ideologie atd.) reflektuje.

Hlavní odlišnosti evaluačního a základního výzkumu nehledejme v jejich designu, technikách sběru dat, ani metodách jejich analýzy. Hlavním rozdílem je jejich účel. „Evaluace je orientována na aktivitu („*action oriented*“).² Je vedena se záměrem ovlivnit hodnotu („*value*“) nebo dopad („*impact*“) politiky, programu, praxe, intervence nebo služby, je vedena z pohledu tvorby doporučení nebo změny“ (Clarke 1999:vi). Z pozice sociální vědců je často vedena ostrá hranice mezi základním a evaluačním výzkumem s ohledem na to, že základní výzkum se nesnaží být (dokonce se tomu často explicitně vyhýbá) politicky i jinak angažovaným. Další význačný rozdíl spočívá v tom, že sociálně-vědní výzkum má často explorativní charakter. Celý výzkum se pak „snaží sledovat pohybující se cíl“ (Peers a Johnson 1994:23), kdy se priority výzkumu mohou v jeho průběhu měnit a to, co bylo původně centrálním, se stane okrajovým a naopak. V souvislosti s účelem obou druhů výzkumů se liší i způsob publikace jejich výsledků.

² V textu budu, častěji než je obvyklé, používat anglické ekvivalenty pro pojmy spjaté s evaluačním výzkumem. Nejde ani tak o reprezentaci jazykových dovedností, ale o pragmatický ústupek anglicismům v případech, pro které nemá čeština přesný překlad, nebo již je habitualizováno jejich používání v anglické verzi.

Zadavateli evaluačních výzkumů jsou nejčastěji lidé nebo organizace (*stakeholders*),³ kteří potřebují konkrétní druh informací o konkrétním problému a nemají potřebu výsledky evaluačního výzkumu kdekoli uveřejňovat.

Výsledky těchto výzkumů tedy nemohou být vystaveny (riziku) popperovské falzifikace a tím vytvářejí specifickou doménu vědění, která často existuje sama pro sebe. Těmto nařčením se samozřejmě evaluační výzkumníci a *stakeholders* brání; nejčastěji s poukazem na to, že evaluační výzkum je vytvářen pro konkrétní (sociální) program a ten se nachází v oblasti hájemství sociální práce, pro kterou se však stalo reformní paradigma⁴ vzhledem ke svému marxistickému nádechu poněkud problematickým (Řezníček 1994:16-19).

PRAXE 1: Evaluace programu DANCE-8 se stala podkladem k inovaci interní metodiky práce, avšak při snaze o publikování jejích výsledků v odborném časopise *Adiktologie* se nestrhla vlna zájmu ze strany *stakeholders*. Zarážející je to zvláště u programů, jejichž existence je záležitostí poměrně mladou (necelý rok) a raritní (v ČR fungují 3 programy podobného zaměření) a tudíž vliv výsledků evaluace na vývoj koncepce takových služeb (politiku jejich vytváření) může být značný.

Klasikové evaluačního výzkumu Peter Rossi a Harrine Freeman shrnují rozdíly mezi základním a evaluačním výzkumem takto:

TABULKA 1: rozdíl mezi základním a evaluačním výzkumem

Zdroj: Rossi a Freeman (1993:405)⁵

Jiří Winkler si v souvislosti s povahou evaluačního výzkumu (a jeho vztahem k „sociální vědě“) klade otázku, zda by měl evaluační výzkum hledat svou vědní etablovanost v managementu, jako „aplikované vědě o ekonomickém řízení“, nebo v „policy science“⁶, jako specifické sociální vědě zabývající se politickým procesem“ (Winkler nedat.:1). Již samotná otázka překlenuje pochybnost o vědeckosti evaluačního výzkumu tím, že ho staví jako pokračovatele starších tradic, z nichž příznačně žádná není dominantní, a obě nakonec spočinou v dialektickém vztahu. Podle Winklera lze právě na evaluaci (potažmo implementaci) v bourdieovském duchu vhodně demonstrovat „významovou distinkci pojetí racionality společenského života, která se stala ústředním faktem nesouladu mezi sociologií a ekonomii...“ (Winkler 2002:22). Diskuze o orientovanosti evaluačního výzkumu směrem k „manažerismu“, nebo „policy science“, ovšem nestírá rozdíl mezi výzkumem

³ *Stakeholders*: lidé, organizace, instituce, kteří/keré jsou nějakým způsobem zaangażováni/y ve fungování (evaluované/m) služby, programu, organizace. Nejčastěji se jedná o politiky, programové administrátory, zaměstnance, donátory.

⁴ „V tomto paradigmatu se představa sociálního fungování pojí s vizí společenské rovnosti...podporou spolupráce a solidarity v rámci určité společenské skupiny...Sociální práce se proto zaměřuje na *zmocňování* (empowerment) jednotlivců i skupin...“ (Matoušek 2001: 187-188)

⁵ Velmi podobnou tabulku uvádí ve své publikaci *Kvalitativní výzkum: základní metody a aplikace* Jan Hendl, ten ji ovšem vydává za svou vlastní (viz Hendl 2005:291).

⁶ Ve slovníku aplikované antropologie by „policy science“ nejspíše figurovala jako „policy analysis“ nebo „analýza politiky“ (viz Veselý 2006).

<i>základní výzkum</i>	<i>evaluační výzkum</i>
vykonává se v oblasti zájmu výzkumníka a přispívá k rozšíření vědomostí o problému	jeho účelem je řešit konkrétní problém
Standarty práce jsou dány vědeckou komunitou	Standarty jsou pragmatické. Evaluace musí být 'dostatečně dobrá' aby odpověděla na otázky týkající se našeho zájmu
výzkumník většinou trénován v jedné disciplíně, které je oddaný	výzkumníci se pohybují v různých programových oblastech
užívá se specifitější sada metodologických procedur	různé metody
využitelnost je hodnocena v rámci vědecké komunity	využitelnost je hodnocena kolegy (evaluátory), sponzory a také v rámci výzkumu rozpoznání politiky, praktického vývoje a řešení sociálních problémů

evaluačním a základním, neboť ani „*policy science*“ nefunguje jako standardní sociální věda – nepracuje s paradigmaty, diskursy...nestaví hypotézy, ani nekonstruuje teorie.

Evaluační výzkum jako samostatná doména

Definice evaluačního výzkumu

V přístupech k určení podstaty evaluačního výzkumu uvedu jen několik jeho elementárních definic a dále se budu věnovat spíše různým druhů evaluačních výzkumů. Takovým postupem se nechci vyhnout exaktní definici tohoto druhu výzkumu, pouze jsem přesvědčen o nutnosti postoupit od ‚fetišizace pojmů‘⁷ k metodologickému pohledu.

Pro úplnost tedy předkládám tři definice evaluačního výzkumu:

1. sociálně-vědní aktivita zaměřená na sběr, analýzu, interpretaci a poskytování informací o fungování a efektivitě sociálních programů.

(Rossi et al. 2004:2)

2. používám pojem evaluace dosti široce tak, aby zahrnoval jakékoli úsilí ke zvýšení lidské efektivity prostřednictvím systematického empirického zkoumání...pokud někdo zkoumá a soudí výsledky a efektivitu, tak se zabývá evaluací. Pokud je toto zkoumání vedeno systematicky a empiricky prostřednictvím pečlivého sběru dat a pozorné analýzy, tak lze mluvit o výzkumu.

(Patton 1990:11)

3. profesionální evaluace je definována jako systematické ovlivnění kvality nebo hodnoty něčeho.

(Scriven 1991)

Každá z jednotlivých definic referuje o specifickém způsobu hodnocení, ale žádná z nich nepodává příliš konkrétní a aplikovatelný obraz evaluace.⁸ Přínosnějším tedy

⁷ S tímto slovním spojením operuje Charles W. Mills (viz Mills 2008).

⁸ Názvy prací většiny autorů zabývajících se evaluací přitom paradoxně aplikovatelností a pragmatičností jen hemží (viz seznam literatury).

bude zaměřit se na parciální charakteristiky evaluace.

Evaluační strategie

Jednotlivé evaluační strategie se vždy formovaly kolem význačných osobností daného oboru. Vždy také platí, že jednotlivé strategie se ve své aplikaci málokdy vyskytují ve své „čisté“ formě a jejich užití se v praxi řídí spíše charakterem problému, než snahou o jejich rigidní následování. Z pohledu některých autorů (viz Green 1998) se toto *ex post* rozdělení nedá považovat za přínosné a to zvláště z toho důvodu, že všechny evaluační strategie v podstatě sdílejí podobnou metodologii a příliš nereflektují otázku programové ideologie. Je tomu však skutečně tak?

- **A vědecko-experimentální modely (postpositivismus)**

Jedná se o pravděpodobně nejstarší přístup k evaluaci, založený na metodách (nejčastěji) sociálních věd. Upřednostňují nestrannost, exaktnost, objektivitu a validitu analyzovaných a předkládaných dat. Mezi nejčastěji používané metody tradičně patří experimenty a quasi-experimenty, ekonometricky orientované perspektivy (analýzy nákladů x efektivity, výsledků). Mezi proponenty tohoto přístupu řadíme Michaela S. Scrivera a Donalda T. Campbella.

- **B modely orientované na management systém (pragmatismus)**

Mezi nejčastější přístupy v rámci těchto modelů patří *Program Evaluation and Review Technique* (PERT) a *Critical Path Method* (CPM). Oba se hojně používají v business sféře a u vládních programů. Tento přístup je částečně reakcí na zklamání z nadužívání vědeckých metod, což je patrné v odklonu od teorie a v preferenci (povětšinou) ekonomických kritérií hodnocení programů. Mezi hlavní reprezentanty tohoto proudu patří v první řadě Michael Patton.

- **C kvalitativní/antropologické modely (interpretativismus)**

Nevýhodou předchozích přístupů byla neschopnost zahrnout do hodnocení programů stanoviska aktérů. Nastává tedy obrat k aktérovi a jeho hodnotám. Mezi upřednostňované metody výzkumu patří pozorování, ale i potřeba znovu zahrnout fenomenologický rámec evaluačního kontextu. Autoři patřící k tomuto směru evaluace se nazývají „Čtvrtou generací“ (byť jsou ve všech přehledech řazeni na třetí místo) a jejich pluralitní přístup se nejjasněji značí v dílech Egona Guby a Yvonne Lincoln (Smutek 2005).

- **D na účastníka orientované modely (integrativní přístup)**

V této poslední a nejrecentnější kategorii se evaluátoři soustředí v první řadě na stakeholders, ale také na samotné klienty programů a technologii (včetně technologie moci, ideologie). Tyto proudy se někdy prohlašují za otevřeně ideologické, se snahou o emancipaci konkrétních sociálních skupin. Ve svých pracích se autoři jako Peter Rossi nebo Lee Cronbach snaží integrovat výhody předešlých přístupů ve snaze vytvářet komplexní evaluační výzkum.

U výše zmíněných evaluačních strategií se zdá na první pohled nápadná analogie s chronologickým a logickým vývojem paradigmat v sociálních vědách jako takových. Dobrodružství evaluace dalo by se považovat za bezchybný příklad toho,

co Kaplan (1964) nazval „zákonem kladiva“⁹. Zástupy programů, služeb, organizací a politik, které by měly být evaluovány se rozrůstají a současné době již nepochází tato snaha pouze od centrální vlády, která je stále dominantním tvůrcem veřejných politik, ale stále častěji prorůstá i do komerční sféry (Kelly 2004). Velký rozkvět evaluací nastal v 60. letech minulého století jako reakce na sociální plán ‚velké společnosti‘ (v

PRAXE 2: Požadavek na maximálně participativní model evaluace se stává v některých kontextech neadekvátním. U evaluace programu DANCE-8, který funguje v oblasti (sekundární) protidrogové prevence bychom tak vedle klientů programu a *stakeholders* měli v ideálním případě zahrnout do evaluace i dealery omamně psychotropních látek (OPL), stejně jako policisty z Národní protidrogové centrály (NPC).

USA) (Pawson a Tilley 1997:3). V té době se Campbell v USA proslavil svou knihou „Reforms and Experiments“ s vizí „experimentálního přístupu k reformám společnosti“ (Pawson a Tilley 1997:6), čímž se dopustil heroického omylu zanedbáním role teorie a analýzy veřejných politik. Tehdy proponované metodologické přístupy k evaluaci (strategie A) stavěly na tradicích ‚molární‘ kauzality a to v dobách, kdy antropologie již za sebou měla okouzlení (strukturálním) funkcionalismem a do módy přicházel strukturalismus. David L. Martinson ve své studii z roku 1974 analyzoval data ze 1400 evaluačních manuskriptů a poukázal na další negativum takto vedených evaluačních výzkumů. Většina z evaluačních výzkumů té doby totiž podle něj ukazuje, že vlastně ‚nic nefunguje‘ tak, jak by mělo¹⁰. V důsledku toho dochází k transformaci či rozpadu původních programových cílů v průběhu jejich realizace (Winkler 2002). Tato skepse vedla ke vystřízlivění z čistě kvantitativních (experimentální) přístupů k evaluaci, mezi jejichž typické reprezentanty můžeme řadit metodu pre-test → post-test.

V 70. letech tedy vstupuje do evaluačního výzkumu politika („*politics*“) a její nahlížení prizmatem manažerského paradigmatu (strategie B), pro který se stává centrálním způsob stanovení cílů politického programu. Evaluace se začíná zaměřovat na procesualnost politiky, na mechanismy a neměnné elementární struktury politického systému, které determinují vytváření sociálních programů. Pro evaluaci se v tomto směru stali východiskem manažerské postupy, stejně jako pro strukturální antropologii perspektivy lingvistické (viz Lévi-Strauss 2006). Důraz na ekonomickou a utilitární představu evaluace se projevuje v eklekticismu metod s důrazem na srozumitelnost pro tvůrce veřejných politik, kteří výzkumné zprávy zkrátka nečtou a je proto nutné „infiltrovat jejich chápání světa nekanonizovanými výzkumnými zjištěními“ (Weiss a Buculavas 1980:90).

Přibližně po dalších deseti letech se evaluační výzkumy staly dějištěm dalšího vystřízlivění, když se do struktury začal stále častěji promítat vliv aktéra. Objekt rozhodování byl postaven na roveň rozhodovacímu subjektu (strategie C)¹¹. Netřeba připomínat, že jedním z prvních, kdo rehabilitoval jednajícího aktéra v sociálních vědách je Pierre Bourdieu.

Konstruktivistický zájem se přesunul od zjištění ‚zda něco funguje‘ k procesu, ve kterém se implicitní ideje aktérů (na pozicích pracovníků programů) kříží

⁹ Jakmile dáte dítěti do ruky kladivo, tak brzy pozná tu univerzální pravdu, že vše co lze, musí být přítlučeno (Kaplan 1964).

¹⁰ V kontextu (současného) evaluačního výzkumu toto není chápáno jako neúspěch. Jinými slovy, nelze zde použít rétorickou figuru o konstrukci úspěchu skrze tvrzení, že ‚vše funguje‘. Je však nezbytné do analýzy příčin ‚nefungování‘ zahrnout i vliv samotné programové teorie, potažmo veřejných politik.

¹¹ Na úrovni rozhodovacího subjektu můžeme také uvažovat na úrovni (i) individuálních aktérů veřejné politiky (konkrétní ministr) a (ii) nad-individuálními subjekty veřejné politiky (vláda) (Ochraňa 2005).

s oficiálními schématy a politikami. Sociální realita spočívá na vzájemném vyjednávání („*negotiation*“) významů. Pokud bychom označili postpozitivistický proud evaluace v intencích analýzy symbolické komunikace za *analýzu obsahu sdělení* a pragmatický proud za *analýzu povahy komunikátora*, tak interpretativní proud může sloužit jako analogie *analýzy publika*.

Podstatným rysem interpretativního modelu je rehabilitace kvalitativních metod. Jako v každém jiném výzkumu se kritéria pro aplikaci konkrétní metody odvíjejí od povahy výzkumného problému. Ať již se jedná o sběr dat – prostřednictvím rozhovorů, *focus groups*, pozorování, video nebo audio nahrávek; stejně jako užití analytických rámců – zakotvená teorie, kritická diskursivní analýza, konverzační analýza, etnografie nebo narativní analýza, musíme mít vždy na zřeteli naše možnosti i účel evaluačního výzkumu. Eleanor Chelimsky v této souvislosti poněkud cynicky poznamenává, že „je sotva prozíravé vstupovat do žhavé politické debaty vyzbrojena pouze případovou studií“ (Chelimsky 1987:27). Použití mikro-sociologických a antropologických metod na druhé straně nabízí mnoho výhod. David Silverman (1994) argumentuje, že při užití kvantitativních metod je v evaluačním výzkumu: (1) kladen důraz na to, jak se konkrétní fenomén konstituuje v každodenních situacích, (2) jsme schopni lépe odpovědět na otázku „proč?“, neboť z dat získaných z každodenních situací máme znalost toho „jak“ a (3) pro nás snadnější vést dialog s pracovníky programu, protože víme, jak se v daných situacích chovají.

Na přelomu 80. a 90. let se prosazuje komplementární hledisko, které se snaží syntetizovat všechny předešlé proudy (strategie D). Skutečným titánem sociologické metodologie tohoto období byl bezpochyby Lee J. Cronbach¹², jenž se k závěru své akademické kariéry dopracoval k funkcionálně diachronnímu přístupu k sociálnímu zkoumání. Cronbach poskytuje evaluátorům soubor symbolů k definování domén evaluace. Domény tvoří jednotky („*units*“), řešení („*treatment*“), pozorování („*observations*“), výstupy („*outcomes*“) a nastavení („*settings*“) – s běžně používanou zkratkou pro celý soubor UTOS. Zajímavějším se však zdá Cronbachovo přirovnání evaluátora ke čtenáři. Sociální program, stejně jako veřejné politiky, se pro něj stávají textem (Shadish et al. 1991). V tomto kontextu se inspirativními stávají jednak studie francouzských dekonstruktivistů, tak i antropologicky vedené pokusy o rozumění kultuře, skrze to, jak jsou tvořeny její texty a texty o ní (viz Clifford, Marcus 1986). Jean-Francois Lyotard vidí v metanaracích moderní společnosti legitimizaci mocenských zájmů. Jednou z těchto metanarací se stává i vyprávění o Rovnosti a Pokroku, na kterém stojí sociální programy a jejich evaluace. Cronbachův přístup by spíše než v derridovském řešení odkazoval k Habermasovo modelu plurality veřejných prostorů. Z tohoto diskursivního pojetí veřejného prostoru, do kterého by měla zasahovat i evaluace, se etablovaly například feministické kritiky veřejných politik (Benhabib a Cornell 1987).

Ve snahách o integrativní uchopení evaluačního výzkumu se často kritizuje jeho proveditelnost. Proto se, v některých studiích, setkáváme s poměrně bizarním požadavkem na evaluaci evaluace (analýzu její proveditelnosti). Na tyto kritiky reagoval Peter H. Rossi, když uvedl pojem „dost dobrá“ evaluace, který pragmaticky řešil otázku výzkumného designu. Společně s Huyem Chenem postavili základ evaluace vedenou teorií („*theory-driven*“) (Chen a Rossi 1983). Evaluační výzkum se

¹² Ve statistice se dodnes používá jeho koeficient reliability ‚Cronbachovo alfa‘ a teorie zobecnitelnosti. V oblasti evaluace se zabýval hlavně evaluací vzdělávacích programů. V závěru své akademické dráhy se stává pesimisticky k možnostem sociálního výzkumu, když tvrdí, že „výzkumný design je spíše uměním, než vědou“ (Cronbach 1982 in Shadish et al. 1991:349).

dostává do role tvůrce programové teorie, kterou následně testuje a tím na jedné straně odhaluje významy, které mohou být aktérům skryté, a na druhé straně prověřuje empiricky jejich platnost (Chen a Rossi 1983). Pro „dost dobrou“ *theory-driven* evaluaci je nutné provést podrobnou analýzu programu a zpracovávat tak co nejlépe definované soubory informací.

Integrativní přístup se projevil i na metodologické rovině, kdy dochází ke kombinaci kvantitativní a kvalitativní metodologie („*mixed methods*“) (Rallis a Rossman 2003). Newman a Benz (1998:9) nazývají bifurkaci kvantitativních a kvalitativních metod jako „falešnou dichotomii“ vytvořenou předchozími paradigmaty. V případě *theory-driven* evaluace se zdá použití *mixed-methods* nezbytné (vytváření i testování teorie).

PRAXE 3: Evaluační výzkum byl často kritizován za svou neteoretičnost a „bezhodnotovost“ (Pawson a Tilley 1997:13). V oblasti protidrogové prevence se vyskytují dva komplementární modely (teorie) závislosti: (i) medicínský model – závislost je nemoc (ii) bio-psycho-socio-spirituální model – určení závislosti kontextuálně. (Kudrle 2003) V oblasti a terciální (i primární) prevence, na které pracuje program DANCE-8, se dlouhodobě uplatňuje teorie *harm reduction* – snižování rizik spojených s užíváním OPL, vycházející částečně ze strategie ochrany veřejného zdraví (*Public Health*) (Hrdina a Korčíšová 2003). Všechna tato teoretická ukořtení je třeba reflektovat při vztahování evaluace k programové teorii.

Další typologie evaluačního výzkumu

Předchozí kapitola prezentovala vývoj evaluačního výzkumu prostřednictvím jeho strategií. Nyní bych se chtěl zaměřit na statictější typologie, které jsou přínosné nejen v tom ohledu, že je s nimi běžně zacházeno v praxi evaluačního výzkumu a jejich definice slouží k jejich snadnějšímu čtení i vytváření.

Formativní a sumativní evaluace

Tato dichotomie představuje nejelementárnější rozdělení evaluačního výzkumu na (i) orientování na proces („*action-oriented*“), tedy formativní a (ii) orientování na výsledky („*conclusion-oriented*“), tedy sumativní. Michael Scriven (1991) vnímá oba typy jako oddělené větve, které se v evaluačním výzkumu neprotínají. Nejprve totiž získáváme data a poté k nim přiřazujeme hodnoty, významy (které mají pro stakeholders, politiky, nebo klienty). Základní rozdíly nabízí následující tabulka:

TABULKA 2: rozdělení sumativní a formativní evaluace

	<i>Formativní</i>	<i>Sumativní</i>
Cíloví příjemci	Programoví manažeři/ pracovníci programů	Tvůrci politik, veřejnost, financovatelé
Role sběru dat	Vyjasnění cílů, podstata implementace, identifikace výsledků	Implementace, měření výsledků
Role evaluátora	Interaktivní	Nezávislá
Metodologie	Kvantitativní i kvalitativní	Kvantitativní
Četnost sběru dat	Kontinuální monitoring (longitudinální)	Omezená (průřezové studie)
Podávání zpráv	Neformální, diskusní skupiny	Formální zprávy
Četnost podávání zpráv	Kontinuálně	Na závěr evaluace

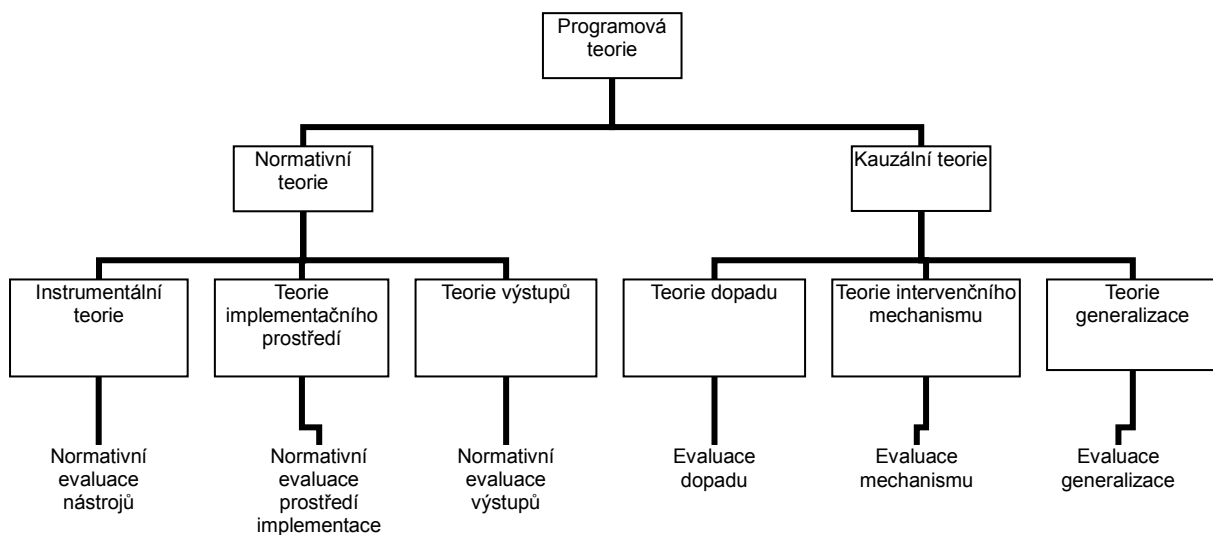
Zdroj: Clarke 1999:8, převzato a modifikováno ze Smutek 2005:29

Byť se někteří autoři (Nachmias 1980) snaží, vedle rozdělení na sumativní a formativní evaluaci, propagovat další dva odlišné typy (i) evaluaci procesu a (ii) evaluaci dopadu, tak je nutno poznamenat, že tyto typy se vztahují spíše na problematiku „implementačního deficitu“, než k problematice hodnocení sociálního programu jako takového. Důležitým se však stává požadavek na provázanost obou přístupů. Jinými slovy, sumativní evaluace se neobejde bez evaluace formativní, protože v momentě, kdy nám sumativní evaluace sociálního programu odhalí, že program nepracuje, jak by měl (což by vhodně vedená evaluace měla), tak bez zhodnocení formy fungování nejsme schopni posoudit, zda se tak stalo v důsledku špatného nastavení programu (veřejné politiky), nebo (jen) program nefunguje, jak by měl (Rossi et al. 1999:40).

Chenova typologie evaluace

Komplexní typologii evaluace teorií vedenou („theory-driven“) představil již zmiňovaný Huey Chen. Programová teorie podle Chena (2006:16-20) obsahuje implicitně dva druhy teorií (1) deskriptivní [normativní teorie], jež operuje s tím, jak by program měl fungovat, s jeho cíli a hodnotami a (2) deskriptivní [kauzální teorie], která se orientuje na kauzalitu fungování programu a mechanismus naplňování cílů sociálního programu¹³. Pod křídly jednotlivých teorií se formují specifické domény, na které navazují konkrétní metody hodnocení programů.

DIAGRAM 1: Chenova hierarchie v evaluaci



¹³ Analogii k Chenovo konceptu *preskriptivní a deskriptivní* teorie by mohla představovat sémiotická teorie kultury Clifforda Geertze. Geertz tvrdí, že na jedné straně existují symboly, které fungují jako „modely něčeho“ (výklad a interpretace reality), a na straně druhé mohou symboly fungovat jako „modely pro něco“ (návod a informace k organizaci reality) (Geertz 2000).

Zdroj: převzato a modifikováno z Chen 1990: 53

Hranice mezi jednotlivými přístupy jsou pevné v tom smyslu, že nedovolují například vztahovat výsledky z evaluace intervenčního mechanismu k normativní teorii. Ale naopak v rámci jedné evaluační studie bychom měli používat kombinaci přístupů z obou větví programové teorie. Hogwood a Peters (1985) také upozorňují na hraniční charakter evaluačního výzkumu, když tvrdí, že se nachází mezi analýzou politického procesu a analýzou pro politický proces.

Předtím, než se pustíme do samotné evaluace, bychom se podle Chena (2006) měli zastavit u identifikace stádia, ve kterém se evaluovaný program nachází. Mezi konkrétní evaluační cíle toto zjištění zařazuje i Winkler (2006:2) a uvádí též přehlednou typologii.

TABULKA 3: fáze programu a jeho evaluace

<i>Fáze programu</i>	<i>Otázky (funkce)</i>	<i>Evaluační cíle</i>
1 Posouzení sociálních problémů a potřeb	V jakém rozsahu jsou komunity potřebné a jaký standard uspokojení dosahují?	Diagnóza problému - deskripce potřeb
2 Určení cílů programu	Co se musí stát, aby se sladily potřeby a standardy uspokojení?	Analýza potřeb pro potřeby programu
3 Tvorba programových variant	Jaké služby/aktivity by měly být použity k dosažení žádoucí změny?	Posouzení plánu/ programové teorie
4 Výběr programové alternativy	Jaký možný programový přístup je nejlepší?	Studie proveditelnosti programové varianty
5 Implementace programu	Jak by měl být program realizován? Jaké jsou implementační deficity?	Procesní evaluace
6 Řízení programových aktivit (operations, outputs)	Jak probíhá výkon jednotlivých činností vzhledem k plánu?	Monitoring programu
7 Programové výsledky	Dosahuje program žádoucích efektů? Plní svůj účel? Jakých efektů dosahuje?	Evaluace efektů
8 Programová účinnost a efektivnost	Jsou programové efekty dosahovány s rozumnými náklady?	Cost-benefit analýza, Cost- effectiveness a.

Zdroj: Winkler 2006:7

Na tomto místě bych zmínil, že evaluace (zvláště v podání Chena) se považuje za holistickou a interdisciplinární. Interdisciplinarita se projevuje v integrovanosti mnoha metodologických přístupů a snahou o využití perspektivy mnoha různých oborů (viz kap. 3.2). Holismus evaluace spočívá v nutnosti zahrnovat do hodnocení co nejvíce zainteresovaných subjektů. Peter Rossi a Harrine Freeman (1993) předkládají 9 různých kategorií, které by se v zásadě daly syntetizovat do tří skupin subjektů determinující (evaluovaný) sociální program. Já k těmto devíti přidávám kategorii desátou a to „osobu evaluátora“, jenž vystupuje v mnohdy rozdílných úlohách. Například může být interním zaměstnancem evaluovaného sociálního programu. Tuto možnost navrhuje například Aaron Wildavsky, když tvrdí, že taková osoba by měla

být „oddána určitým modům řešení problémů“ (Wildavsky 1980 in Smutek 2005:114) a do budoucna by se měla stát fundamentem pro model sebeevaluuující organizace (Wildavsky 1980 in Smutek 2005:114). V této pozici se však evaluátor může velice snadno dostat do konfliktu rolí (evaluátor x zaměstnanec). V souvislosti se zainteresovaností evaluátora na chodu programu lze rozlišovat i evaluaci participativní („*participatory*“ nebo „*collaborative*“) a nezávislou („*independent*“) (Rossi et al. 2004:57).

1. Stakeholders

- Tvůrci politických rozhodnutí
- Tvůrci programu
- Sponzoři programů
- Sponzoři evaluačního výzkumu
- Manažeři programu
- Zaměstnanci programu
- Konkurenti programu
-

2. Klienti

- jednotlivci, domácnosti, komunity, organizace

3. Výzkumný pracovník (evaluátor)

- Evaluační a výzkumné komunity
- Osoba evaluátora

V průběhu evaluace dochází k ne/přímé konfrontaci mezi jednotlivými skupinami subjektů (jejich zástupců) i mezi subjekty uvnitř těchto skupin. Kontext evaluace často proměňuje charakter těchto konfrontací.

PRAXE 4: Program DANCE-8 se v době své evaluace nacházel ve stádiu počáteční implementace (fáze programu 2 a 3). V této fázi byla tedy součinnost *stakeholders* a evaluátora velice intenzivní. Jako nejkomplicovanější se v průběhu evaluace ukázala míra zapojení klientů (přímých i nepřímých v pozicích majitelů klubů a pořadatelů tanečních akcí). U sběru dotazníků (prováděných s přímými klienty) bylo nezbytné využít služeb dobrovolníků (aby nedocházelo ke konfrontaci mezi klienty a pracovníky programu) a v případě *focus groups* s majiteli klubů a pořadatelů tanečních akcí se v jejich průběhu ukázala nevhodnou moderace diskuze evaluátorem, který byl zároveň interním zaměstnancem programu. Většina otázek pak směřovala na jeho osobu a nedocházelo tak k interakci mezi účastníky *focus group*.

Evaluátor vystupuje v několika rolích a to zvláště jako poskytovatel informací,

vyhledávač potencionálních obtíží („*troubleshooter*“), facilitátor, konzultant a školitel („*illuminator*“) (Chen 2005). Podobnost s rolemi antropologa jak je vymezil John van Willigen (2002) se vyloženě nabízí - aplikovaný antropolog by měl fungovat ve veřejné i soukromé sféře jako instruktor, výzkumník nebo konzultant.

Závěr

Klasik sociologické teorie Robert K. Merton se, v návaznosti na mnoho svých předchůdců, věnoval problematice nepředvídaných důsledků záměrného sociálního jednání. Merton viděl dvě metodologické nástrahy samotné analýzy tohoto problému. První se týkala „problému zjištění kauzální souvislosti, problému zjištění míry, v níž lze ‚důsledky‘ oprávněně připsat jistému jednání“ (Merton 2007:120). Druhý problém spočívá ve „zjištění skutečných záměrů daného jednání“ (Merton 2007:120). Merton (2007) přisuzuje v konečném důsledku nejvyšší kritérium weberovskému „smyslu“, který by měl být zjevný ze srovnání naší znalosti aktéra, kontextu a samotného jednání s deklarovaným nebo dedukovaným záměrem tohoto aktéra.

V evaluačním výzkumu se tyto dvě metodologické nástrahy stávají jeho obsahem. Evaluace formativní (procesu) a sumativní (výsledku) tvoří základní rámec, ze kterého poté vychází konkrétní analýza, hodnocení daného sociálního programu. Ve svém vývoji si evaluační výzkum prošel několika stádii a současná snaha o jeho hypertrofii do (bezmála) všech oblastí veřejného i soukromého života se setkává s deziluzí, často i ze strany samotných teoretiků evaluace. „Jsme často zklamaní. Po celé té *Bouři a Vzduchu*, který se s evaluací zvedne a poté, co analyzujeme její výsledky, nakonec zjistíme, že nám k ničemu nebyla. Věci jdou stejně dál vlastní cestou, aniž by je evaluace nějak ovlivnila.“ (Weiss a Buculavas 1980:171). Stala-li se evaluace mantrou modernity, tak předně v oblasti analýzy úspěšnosti implementace veřejných sociálních politik „shora-dolů“. Postupný důraz na aktéra a jeho vliv na (organizační) strukturu, jak ho chápal například Michel Crozier (1964) obrací pozornost na tvorbu rozhodnutí „zdola-nahoru“. Analýza obou těchto směrů, stejně jako důraz na integrativní výzkumné strategie je mantrou dnešního evaluačního výzkumu.

Aplikovaná antropologie staví deklarativně evaluační výzkum a roli antropologa jako evaluátora do pole své působnosti (viz Ervin 2005, van Willigen 2002). Evaluátor-antropolog může (oproti evaluátorovi vzešlému z oblasti sociální politiky a práce) využít své obeznámenosti s teoriemi a koncepty, kterými se řídí socio-kulturní realita, a tím poskytnout cennou zpětnou vazbu pro administrátory sociálních programů o možných dopadech jejich rozhodnutí. U evaluačních přístupů vedených teorií („*theory-driven*“) se tato perspektiva jeví nezbytnou. Zároveň může antropolog na základě terénního výzkumu odhalit, jak vnímají provedené intervence jejich objekty, klienti („*insiders-perspective*“). Evaluační výzkum má samozřejmě své limity, na některé jsem se také snažil poukázat, ale deziluze o evaluaci by neměla být vedená snahou „hlavně nehodnotit“.

Literatura

- BENHABIB, S. a D. CORNELL. Eds. 1987. *Feminism as critique: on the politics of gender*. Minneapolis: University of Minnesota press.
- BENZ, C. R. a I. NEWMAN. 1998. *Qualitative-Quantitative Research Methodology: Exploring the Interactive Kontinuum*. Carbondale, IL: Southern Illinois University Press.
- CLARKE, A. 1999. *Evaluation Research*. London: Sage.
- CLIFFORD, J. a G. E. MARCUS. Eds. 1986. *Writing culture: the poetics and politics of ethnography*. Berkeley: University of California Press.
- CROZIER, M. 1964. *The Bureaucratic Phenomenon*. Chicago: University of Chicago Press.
- DOUGLAS, M. a B. ISHERWOOD. 1979. *The World of Goods: Towards an Anthropology of Consumption*. New York: Basic.
- ERVIN, A. M. 2005. *Applied Anthropology: Tools and Perspectives for Contemporary Practice*. Pearson/Allyn and Bacon.
- FREEMAN, HARRINE a ROSSI, PETER. 1993. *Evaluation: A systematic approach* (5th.ed.). Neubury Park, CA: Sage.
- HENDL, J. 2005. *Kvalitativní výzkum: základní metody a aplikace*. Praha: Portál.
- HOGWOOD, B. W. a G. B. PETERS. 1985. *The Pathology of Public Policy*. Clarendon Press.
- HRDINA, P. a B. KORČIŠOVÁ. 2003. „Terénní programy,“ in *Drogy a drogové závislosti – mezioborový přístup*, Eds. J. Radimecký a K. Kalina. Praha: Úřad vlády ČR.
- CHELIMSKY, E. 1987. The politics of program evaluation. *Society* 25(1): 24-32.
- GEERTZ, C. 2000. *Interpretace kultur*. Praha: SLON.
- CHEN, H-T. 1990. *Theory-driven evaluation*. London: Sage.
- CHEN, H-T. 2006. *Practical program evaluation*. London: Sage.
- CHEN, H-T a P. H. ROSSI. 1983. Evaluating with sense: The theory-driven approach. *Evaluation revue* 7:283-302.
- KAPLAN, A. 1964. *The conduct of inquiry*. San Francisco: Chandler.
- KELLY, M. 2004. „Qualitative evaluation research,“ in *Qualitative research practice*, Eds. D. Silvermanem, C. Seale, G. Gobo a J.F. Gubriem. London: Sage.
- KUDRLE, S. 2003. „Úvod do bio-psycho-socio-spirituálního modelu závislosti,“ in *Drogy a drogové závislosti – mezioborový přístup*, Ed. J. Radimecký a K. Kalina. Praha: Úřad vlády ČR.
- LÉVI-STRAUSS, C. 2006. *Strukturální antropologie*. Praha: Argo.
- MERTON, R. K. 2007. *Studie ze sociologické teorie*. Praha: SLON.
- MILLS, CH. W. 2008. *Sociologická imaginace*. Praha: SLON.
- OCHRANA, F. 2005. *Evaluace ve veřejné politice*. Praha: CESES.
- PATTON, M. Q. 1990. *Qualitative Evaluation and Research Methods: A Systematic Approach*. Kalifornia: Sage.
- PAWSON, R. a N. TILLEY. 1997. *Realistic evaluation*. London: Sage.
- PEERS, S. a M. JOHNSTON. 1994. „Theory, utility and stakeholders: methodological issues in evaluating a community project on HIV/AIDS,“ in *Challenge and innovation: methodological advances in social research on HIV/AIDS*, Ed. M. Boulton. London: Tylor & Francis.
- POMERANTZ, A. 1984. “Agreeing and disagreeing with assessment: Some features of preferred/dispreferred turn shapes”. in *Structure of Social Action: Studies in Conversation Analysis*. Eds. M. Atkinson a J. Heritage. Cambridge: Cambridge University Press.

- RALLIS, S. S. a G. B. GROSSMAN. 2003. „Mixed methods in evaluation contexts: a pragmatic framework,“ in *Handbook of Mixed methods in social & behavioral context*, Eds. A. Tashakori a Ch. Teddlie. London: Sage.
- RUBIN, A. a E. BABBIE. 1989. *Research methods for social work*. California: Wadsworth Publishing Company.
- ŘEZNÍČEK, I. 1994. *Metody sociální práce*. Praha: SLON.
- NACHMIAS, D. 1980. *The Practice of policy evaluation*. New York: St. Martin's Press.
- SCRIVEN, M. 1991. *Evaluation thesaurus*. Newbury park: Sage.
- SHADISH, W.R., T. D. COOK. a L. D. LEVITON. 1991. *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage Publications.
- SILVERMAN, D. 1994. „Anylysing naturally-occurring data on AIDS counselling; some methodological and practical issues,“ in *Challenge and innovation: methodological advances in social research on HIV/AIDS*, Ed. M. Boulton. London: Tylor & Francis.
- SMUTEK, M. 2005. *Evaluace sociálních programů*. Hradec Králové: Gaudeamus.
- ROSSI, P. H., H. E. FREEMAN a M. W. LIPSEY. 2004. *Evaluation: a systematic approach*. Thousand Oaks: Sage.
- VESELÝ, A. 2006. Veřejná politika a policy analysis. *Pražské sociálně vědní studie. Veřejná politika a prognostika* 1(15):1-26.
- WEISS, C. H. a M. BUCULAVAS. 1980. *Social Science Research and Decision Making*. New York: Columbia University Press.
- WILLIGEN, J. van. 2002. *Applied antropology: an introduction*. Greenwood Publishing group.
- WINKLER, JIŘÍ. 2006. *Učební text předmětu SPP 405*. Dostupné na: <<http://is.muni.cz>>, stáhnuto 18.12. 2009.
- WINKLER, J. 2002. *Implementace veřejných programů a přístupy k jejich hodnocení*. Brno: ESF MU.