

Srovnání vybraných přístupů k realizaci evaluací Comparison of Selected Evaluation Approaches

Jiří Remr*

Abstrakt

Cílem statě je poukázat na rozmanitost, jak mohou evaluátoři přistupovat k řešení evaluačních úloh. Provedená analýza dokládá, jak pestrý je rejstřík používaných přístupů a potažmo také, k jak rozdílnému typu informací mohou evaluace založené na rozdílných přístupech vést. Text obrací pozornost k základům evaluačního myšlení a poukazuje na možnosti volby metodologického a konceptuálního rámce při řešení praktických úloh. Stať tedy nemá ambice poskytovat praktický návod k realizaci evaluací, nýbrž poskytuje zobecněné poznatky abstrahované od detailů rutinních evaluačních projektů.

Evaluace jsou nejprve vymezeny s využitím několika definic akcentujících metodologické priority evaluací a evaluačního výzkumu. Dále je detailní pozornost věnována čtyřem přístupům k provádění evaluací, a to konfirmačnímu (Goal-based), exploračnímu (Goal-free), konstruktivistickému (Constructivist) a k přístupu vedenému teorií (Theory-driven). Každý z přístupů je vždy stručně charakterizován, dále je analyzován kontext jeho vzniku a vztah k ostatním přístupům; v neposlední řadě jsou uvedeny klíčové znaky a odlišnosti od ostatních přístupů. Užitečnou součástí analýzy každého přístupu je ilustrační příklad, z něhož jsou patrné odlišnosti v pojetí evaluací realizované v duchu jednotlivých přístupů.

Abstract

Main goal of the paper is to show the variety of how evaluators could solve the evaluation tasks. Results of the performed analysis present how uneven the outcomes of evaluation could be when different approach is taken. The article turns attention toward the basics of evaluative thinking by opening the range of alternative methodological and conceptual frameworks that could be used in the course of intervention program evaluation.

Evaluation is defined with use of several definitions that put emphasis on methodological aspects of evaluative inquiry. In detail, there are four different approaches compared: goal-based, goal-free, constructivists and theory-driven. Each of the approaches is briefly introduced, context of its dissemination is explained and it is also described how the approach relates to other approaches. Useful component of the analysis is an example that highlights the main distinctions among the four approaches and results that each of the approaches can offer.

Klíčová slova

Konfirmační přístup, explorační přístup, konstruktivistický přístup, evaluace vedená teorií, evaluační přístupy

Keywords

Goal-based evaluation, goal-free evaluation, constructivist evaluation, theory-driven evaluation, evaluation approach

Úvod

Jedny z nejčastěji citovaných definic evaluace (kromě definic dále uváděných v textu také např. Rossi, Lipsey a Freeman 2004, Scriven 1991, Weiss 1972, House 1980) spojuje důraz na proces získávání podkladů (důkazů) a postupy uplatňované při jejich vyhodnocení. Podobně také Guba a Lincolnová (1989) považují evaluaci za činnost, při které se uplatňují vědecké postupy s cílem získat a analyzovat obsah, strukturu a výstupy intervencí. Konstruktivisticky zakotvená definice evaluací odkazuje k metodologickému základu, když vyjmenovává konkrétní postupy, jež se v oblasti evaluací využívají. Navíc vymezuje obsah evaluačních šetření, která v jejich pojetí zahrnují nejen výsledky, ale také obsah, strukturu a výstupy. Stake (2004) ve své definici evaluací a charakterizaci evaluační činnosti explicitně poukazuje, že každý jednotlivec je sice „rodilým evaluátorem“, neboť hodnocení je jako takové přirozenou součástí každodenního života, nikoliv však nutně „dobrým evaluátorem“. Snaha formalizovat jednotlivé postupy používané při evaluaci se projevuje zapojením některých principů vědeckého výzkumu, resp. metod a technik používaných ve společenskovědním výzkumu.

Výše uvedené definice představují evaluaci jako formalizovaný postup, který se vyznačuje mj. systematickostí, kontrolovatelností (přezkoumatelností) a vyloučením či alespoň podstatnou redukcí subjektivismu. Jiné definice kladou důraz na cíl a účel prováděné evaluace. V kontextu hodnocení intervenčních programů tak např. definují Hallová a Hall (2004) evaluace jako specifický typ šetření, jenž ověřuje, zdali, a v jaké míře, byly naplněny cíle a plány stanovené v rámci konkrétní intervence. Takovýmto důrazem na porovnání plánované a skutečně dosažené hodnoty výstupních hodnot I. Hallová a D. Hall redukuje aplikační oblast evaluace na problematiku související s posuzováním výsledků (outcome evaluation). Nicméně je třeba uvést, že výčet aplikačních oblastí je v případě evaluací podstatně širší a zahrnuje mj. také evaluaci dopadů či analýzu potřeb (Rossi, Lipsey a Freeman 2004; Stufflebeam a Sinkfield 2007, Mark, Henry a Julnes 2000, Shadish, Cook, a Leviton 1991).

Ucelenou definici kombinující účel a způsob provedení evaluace navrhuje Patton (2008), podle něhož je evaluace činností, která zahrnuje systematický sběr údajů charakterizujících intervenční programy, způsoby jejich fungování a výsledky, k nimž dospívají. M. Q. Patton dále poukazuje,

že cílem evaluací je zvýšit kvalitu těchto intervenčních programů, posílit jejich efektivitu, popsat mechanismy jejich fungování a přinést konkrétním zainteresovaným stranám podklady, na jejichž základě je možné provádět rozhodování o jejich dalším směřování. Na jiném místě rovněž uvádí, že evaluace mohou (a mají) přinášet objektivní a prakticky využitelné údaje (Patton 1982).

Do praktické roviny se uvedené vymezování evaluací z hlediska způsobu a účelu provedení promítá v podobě široké škály přístupů, které mohou být při provádění evaluací využity. S ohledem na konkrétní cíl tak může být zvolen přístup, který se zaměří buď na co nejpřesnější kvantifikaci dosaženého výsledku nebo na vnímání přínosů dané intervence konkrétními zainteresovanými stranami či na vysvětlení mechanismu, jak daná intervence vyvolává změnu v chování účastníků posuzovaného programu. Následující analýza je věnována vybraným způsobům, jak lze přistoupit k provádění evaluací či jinými slovy jak může být nastaven design konkrétního evaluačního systému. V tomto ohledu poukazuje na obecně-metodologické předpoklady využití jednotlivých forem v závislosti na konkrétních podmínkách neopouštěje přitom teoretickou rovinu.

Přístupy k realizaci evaluací

Pro jednotlivé přístupy k provádění evaluací je charakteristické, že (1) vypovídají o pojetí externí validity, (2) odrážejí vztah šetření (údajů) a reality (konkrétní hodnocené situace), (3) nastiňují základní metodologickou orientaci a v neposlední řadě také (4) definují ústřední zájem či hodnotu, která určuje celkové nasměrování prováděné evaluace (Alkin, Christie 2005).

Na základě takto stanovených kritérií lze identifikovat celou řadu jednotlivých přístupů, které se v průběhu vývoje metodologie evaluací objevily. Z jejich dlouhého seznamu uvádíme celkem čtyři přístupy, které reprezentují hlavní směry uvažování o způsobech provádění evaluací. Jsme si vědomi toho, že prezentovaný výčet není úplný; chybí zde mj. zplnomocňující přístup (empowerment evaluation), participativní a kolaborativní přístupy a mnohé další. Nicméně primárně sledovaný účel tento výčet beze zbytku naplňuje. Ukazuje totiž na pluralitu forem provádění evaluací,

příčemž čtyři uvedené přístupy současně odrážejí hlavní metodologické akcenty při provádění evaluací.

1. Konfirmační evaluace (Goal-based evaluation; Ralph Tyler)

Tylerův přístup byl v daném historickém kontextu nevídaný. V tomto období byl rozvoj evaluačních šetření svázán výlučně s oblastí školství a vzdělávání; evaluačních šetření v ostatních oblastech, jako např. evaluace intervenčních programů, se objevily až v 60. letech. Podle historického exkursu Guby a Lincolnové (1989) převládaly v meziválečném období evaluace využívající sumarizované výsledky standardizovaných znalostních, kognitivních či výkonnostních testů, jakými jsou např. SRTA (Stone Reasoning Test in Arithmetic, jenž je používán od roku 1908) či SAB (Stanford Achievement Battery, vytvořený v roce 1922). Tyto a mnohé podobné testy byly tehdy hojně používány, neboť bezvadně plnily svůj primární účel: individuální diagnostiku žáků. Kromě toho však výsledky uvedených testů nacházely v sumarizované (agregované) podobě sekundární využití. Na jejich základě byly porovnávány jednotlivé třídy v rámci školy, či dokonce celé školy v rámci dané oblasti či státu. R. Tyler však tento způsob využívání standardizovaných testů odmítl, a namísto toho vypracoval vlastní přístup k realizaci evaluací. Jeho záměrem (úkolem) totiž bylo vyhodnotit nově zaváděné osmileté osnovy a obával se, že diagnostické testy, primárně vytvořené pro měření schopností jednotlivých žáků, nepřinesou potřebná data o účinnosti osnov. R. Tyler v této souvislosti poukazoval především na rozdílný předmět analýzy, tedy schopnosti jednotlivých žáků oproti konkrétně specifikovaným výsledkům používaných (či zaváděných) osnov. Při tomto srovnání by způsobilost používaných diagnostických testů vypovídat o výsledcích (efektivitě) nových osnov nebyla podle R. Tylera obhajitelná.

Tylerův způsob řešení podstatně přesahuje rámec jednoho samotného přístupu k realizaci evaluací. Jeho hlavním odkazem je obecný evaluační postup, který je založen na přiřazování skutečně dosahovaných a empiricky ověřitelných výsledků k předem definovaným cílům (tj. očekávaným výstupům), přičemž obě hodnoty (cíle a výsledky) jsou následně porovnány. Z tohoto důvodu je Tylerův přístup některými nazýván jako objectives-oriented (Alkin 2004), jiní o tomto přístupu píší jako o goal-based evaluation (Stufflebeam, Shinkfield 2007). Obě pojmenování

však stejně dobře vystihují základní podstatu Tylerova přístupu a lze je používat jako synonyma.

Předem definované cíle jsou tedy v rámci tohoto přístupu základním a jediným zdrojem kritérií. Aby bylo možné s cíli v rámci konfirmačního přístupu pracovat, je třeba, aby měly měřitelnou podobu – R. Tyler v tomto ohledu prosazoval takové cíle, které mají odraz v chování jednotlivců¹. Základní otázka konfirmačního přístupu tedy zní: „**Vyskytují se ve skupině účastníků intervenčního programu očekávané prvky chování či nikoliv?**“

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

- 1) definování cílů intervence
- 2) kvantifikace cílů intervence a jejich klasifikace (uspořádání cílů do skupinek dle věcné podobnosti)
- 3) nalezení specifických behaviorálních vzorců pro každou skupinku cílů intervence; jinými slovy definovat znaky (ukazatele) unikátnosti každé skupinky cílů
- 4) identifikace situací, ve kterých je možné výskyt požadovaných behaviorálních vzorců zachytit (pozorovat); nalezení okolností, na nichž lze míru dosažení stanovených cílů dané intervence jasně dokumentovat
- 5) výběr adekvátních metod získání potřebných údajů pro každou skupinku cílů resp. behaviorálních vzorců; výběr či vývoj metodiky měření stupně dosažení cílů intervence
- 6) volba optimální metody sběru dat, provedení předvýzkumu a vlastní sběr dat
- 7) srovnání získaných údajů s předem definovanými cíli; stanovení způsobu využití získaných výsledků a nástin možností a omezení při jejich interpretaci

Tylerův přístup byl novátorský ještě v jednom ohledu: do té doby používané testování neposkytovalo prakticky žádnou zpětnou vazbu, která by

¹ V pedagogickém výzkumu mají takový charakter např. achievement testy.

mohla pomoci dosažené výsledky zlepšovat. Absence akceptačních kritérií (vodítek, která by pomohla určit, jaké výsledky jsou „dobré“, a jaké nikoliv) neumožňovala naplnit bazální princip evaluace, tedy stanovit věčný základ a přínos². Kdybychom i přes tyto výhrady připustili, že prováděné testování bylo evaluací, označili bychom ji zpětně z dnešního pohledu jako evaluaci sumativní. Posuzované hodnoty totiž vytvářely pořadí jednotlivých tříd či škol a ukazovaly na úspěch a neúspěch jednotlivých subjektů. Naproti tomu evaluace prováděná R. Tylerem přinášela zcela jiný typ údajů, neboť zprostředkovávala data poskytující detailně strukturované výsledky o rozdílech mezi skutečnými a očekávanými či požadovanými cíli. Tím provedené evaluace *de facto* poskytly návod, které z výsledků (tedy přesněji, které z behaviorálních vzorců dokumentujících určitou třídu cílů) je vhodné v příštím období zlepšovat. R. Tyler svým pojetím tedy umožnil, aby osmileté osnovy byly nejen porovnány oproti dosavadnímu kurikulu z hlediska rozvoje schopností a dovedností podrobnějších se žáků (škol), ale současně aby osmileté osnovy s intenzivní pomocí jím prováděných evaluací mohly být dále rozvíjeny.

Tylerův odkaz je dodnes živým způsobem provádění evaluací. Jednak tím, že inspiroval mnohé pokračovatele k dalšímu rozvíjení konfirmačního přístupu, mezi všemi jmenujme např. Blooma et al. (1956) či Magera (1962), a dále tím, že umožnil jiným, aby vůči cílenému přístupu k evaluacím vymezili své vlastní přístupy (viz např. M. Scriven či E. Guba, o kterých je pojednáno ve větším detailu dále). Koneckonců současná podoba evaluací tak, jak je požadována Federální vládou USA (a do určité míry i Evropskou komisí) vychází do značné míry právě z Tylerova konfirmačního přístupu³. V obecnějších souvislostech lze hovořit o tzv. *objectives-oriented tradition*, která se objevuje i v jiných oborech: přestože nebyl doposud prokázán přímý vliv, nelze si nepovšimnout shody základních principů např. s dnes rozšířeným MBO (Management by Objectives) v manažerských studiích.

² Věčný základ (merit) charakterizuje samotnou podstatu předmětu evaluace a odkazuje k jeho funkčnosti, zatímco přínos (worth) odkazuje k tomu, jak je věčný základ vnímán konkrétními zainteresovanými stranami.

³ R. Tyler se sám stal v 60. letech vedoucím NAEP (National Assessment of Educational Progress), což byl federální program pro sběr dat o výkonech žáků amerických škol.

Nespornou výhodou Tylerova konfirmačního přístupu je jeho relativní jednoduchost a možnost snadného provedení, třeba i jen formou interně prováděných evaluací, bez pomoci speciálně školených evaluátorů. Za jeho silnou stránku lze považovat rovněž fakt, že tento přístup vede zúčastněné zainteresované strany (zejména pak manažery posuzovaných intervenčních programů) k zamyšlení nad jejich původními záměry a skutečně dosaženými cíli a podněcuje je k jejich zpřesnění a sdílení s ostatními zainteresovanými stranami.

Někteří ze současných evaluátorů (Rodriguez-Campos 2005, Schwandt 1984) při hodnocení Tylerova odkazu poukazují na skutečnost, že jeho přístup vlastně není žádnou specifickou inspirací pro provádění evaluací, neboť jde o pouhou deskripci postrádající klíčové atributy samotné evaluace. Nejen R. Tylerovi, ale obecně celému konfirmačnímu přístupu je dále vytýkáno, že zcela odhlíží od kontextu, v němž jsou evaluace realizovány, a že nebere v potaz nezamýšlené dopady, ke kterým mohou posuzované intervenční programy vést (Scriven 1993). Konfirmačnímu přístupu je rovněž vytýkán relativně vysoký sklon k managerialismu⁴ (House 1993; Guba, Lincoln 1989) a tendence k příliš obecnému posuzování sledovaných intervencí. Stake (1975) konfirmačnímu přístupu vyčítá, že jako formativní způsob evaluace vůbec nepočítá s individuálními hodnotami a hodnotovými systémy samotných klientů (v Tylerově případě žáků) a nebere v rámci posuzované situace ohled na jejich svobodné rozhodování a jednání (viz např. odlišné aspirace žáků škol v městských aglomeracích oproti aspiracím žáků na venkově).

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

Uvedený ilustrační příklad je uveden ve snaze o poskytnutí konkrétního obrazu shora uvedených charakteristik. Uvedený příklad je založen

⁴ Tedy k přílišnému zaměření na finální výstupy odhlízejícího od mechanismů a způsobů jejich dosažení.

na předpokládaném provedení evaluace zaměřené na posouzení efektů rekvalifikačního kursu určeného nezaměstnaným osobám starším 50 let, který je věnován práci s výpočetní technikou. Příklad odpovídá situaci, kdy takovýto rekvalifikační kurs pilotně proběhl v jedné lokalitě a na základě provedené evaluace bude rozhodnuto buď o jeho zastavení, nebo o rozšíření kursu do dalších okresů.

Ilustrační příklad konfirmačního přístupu⁵:

V rámci konfirmačního přístupu je vypracován test počítačové gramotnosti, který se po skončení kursu předloží každému z účastníků. Po dohodě s iniciátorem a s jeho souhlasem se test zaměřuje na používání textového editoru a tabulkového procesoru a jsou definovány konkrétní úlohy, které indikují zvládnutí obsahu kursu. V neposlední řadě je stanoveno, že pokud 80 procent účastníků kursu tento test zvládne (tj. dosáhne alespoň 75procentní úspěšnosti při řešení úloh zadaných v testu), bude možné považovat kurs za vhodný k rozšíření do dalších lokalit.

Výsledky účastníků kursu v závěrečném testu jsou tak v rámci konfirmačního přístupu interpretovány jako důkaz účinnosti výukových metod (kurs skutečně vzdělává) a jejich přínosu (osvojení znalostí a dovedností účastníky kursu). Na základě toho se může realizátor kursu (resp. sponzor) rozhodnout o rozšíření kursu do jiných okresů (bude-li kurs posuzován jako efektivní), popř. o změně obsahu kursu či o změně jeho výukových metod (bude-li kurs shledán neefektivním).

2. Explorační přístup (Goal-free evaluation; Michael Scriven)

Jak již bylo uvedeno, od dob R. Tylera byla významná část evaluací založena na porovnávání předem vytyčených cílů příslušnými zainteresovanými stranami se skutečně dosaženými výsledky. Tento způsob provádění evaluací, který Guba a Lincolnová (1989) nazývají „popisem“ a považují ho za tzv. druhou generaci evaluační metodologie, je využíván v řadě situací z rozličných důvodů dodnes a rozhodně ho nelze považovat ani za překonaný, ani za aplikačně omezený či nevhodný. Evaluátoři se však pochopitelně snažili navrhovat nové přístupy překonávající své ascendenty v charakteristikách, jež byly předmětem největších kontroverzí a jejichž

⁵ Příklad je průběžný a opakovaně se k němu vracíme v každém z dále prezentovaných přístupů. Postupně jsou tedy uvedeny čtyři variace na stejné téma. Díky tomu si čtenář může vytvořit představu o odlišnostech, které mezi jednotlivými přístupy panují.

výsledky zůstávaly v mnohých ohledech za původním očekáváním. V tomto ohledu je tedy explorační přístup třeba chápat jako Scrivenovu odpověď na konfirmační přístup, jako odpověď, jež přinesla nový pohled na způsob provádění evaluací.

Podstatou rozdílu mezi konfirmačním a exploračním přístupem je stupeň pozornosti věnovaný skutečným výstupům, výsledkům a dopadům posuzovaných programů. M. Scriven reagoval na typický nedostatek konfirmačních evaluací, které sice dokázaly přesně odpovědět na otázku, zdali bylo dosaženo vytyčených cílů a v jaké míře, nicméně vůbec nebraly v potaz vyskytnuvší se nezamýšlené důsledky působení sledovaných programů či intervencí⁶. Explorační přístup k provádění evaluací tedy staví do pozice hodnocených prvků cíle samotné, přičemž odmítá jejich apriorní danost. Scriven (1973) zdůrazňoval, že evaluátor by dokonce cíle posuzovaného intervenčního programu neměl znát, aby jimi nebyl ovlivněn: „Goals bias the evaluator“⁷. Toto je v kontextu Scrivenovy evaluační teorie tzv. kvalitativní význam objektivity; za objektivní je v tomto ohledu považován takový výsledek, u kterého lze prokázat, že není vychýlený⁸. Evaluátor by tedy v souladu se základním principem tohoto přístupu měl věnovat pozornost **všem** efektům, k nimž daná intervence směřuje, a to jak k efektům zamýšleným, tak také nezamýšleným. Díky tomu mají evaluace prováděné v souladu s principy tohoto přístupu mnohem lepší předpoklady zachytit i vedlejší a neočekávané důsledky posuzovaných intervencí. Konkrétní evaluační praxe ukazuje, že explorační přístup dosahuje velmi dobrých výsledků zejména při identifikaci výstupů a výsledků, nicméně pro hodnocení dopadů se výhody jejich aplikace výrazně snižují. Základní otázka exploračního přístupu tedy zní: „**Jaké jsou veškeré efekty dané intervence?**“

M. Scriven v souvislosti s exploračním přístupem poukazuje na potřebu vymanit se z rámce, který soustřeďuje pozornost realizátora evaluací na jednotlivé dílčí oblasti, a doporučuje provádět evaluace takovým způs-

⁶ Podoba evaluačních šetření byla v USA v 60. letech ovlivněna zákonnými pravidly, která přesně specifikovala jejich podobu. Specifikace evaluací vycházela, s ohledem k účelu využití (např. rozhodnutí o pokračování či zastavení finanční podpory konkrétních škol z veřejných zdrojů), z Tylerovy metodiky.

⁷ M. Scriven v této souvislosti někdy píše o tzv. bias-free evaluation.

⁸ Kromě této kvalitativní objektivity je možné rozlišit také objektivitu kvantitativní, která spoléhá na shodu intersubjektivit (House 1980). Kvantitativní pojetí objektivit však M. Scriven zásadně odmítá.

sobem, při kterém by samotné hodnocení nebylo nijak svázáno s původními programovými cíli. Taková představa má navíc oporu v samotné normativní rovině evaluací; apriorní definice cílů, jejichž dosažení má dané evaluace potvrdit či vyvrátit, může totiž vést k jednostranné akcentaci zájmů konkrétních zainteresovaných stran (např. iniciátorů intervence nebo manažerů intervenčních programů), což by bylo v rozporu nejen se Standardy definovanými Joint Committee on Standards for Educational Evaluation (cit. dle Sanders 1994), ale i s obecnými zásadami nestrannosti. M. Scriven sám zdůrazňoval, že explorační přístup je díky eliminaci rizika vychýleného hodnocení spravedlivějším přístupem (Scriven 1993).

Z předchozího je patrné, že pokud je evaluace realizována v souladu s principy exploračního přístupu, měla by být minimalizována interakce evaluátora s ostatními zainteresovanými stranami (tedy zejména s manažery daného intervenčního programu a s jeho pracovníky, resp. se zadavateli daného evaluačního šetření). Jen tak je totiž možné zajistit (a prokázat), že způsob provedení evaluace je skutečně nezávislý, tj. založený výlučně na odbornosti evaluátora neovlivněného preferencemi a očekáváním některé ze zainteresovaných stran.

Explorační přístup nevyklučuje při své aplikaci simultánní využití přístupu konfirmačního; z řady konkrétních případových studií je skutečně patrná snadná možnost oba přístupy kombinovat. Výhodou takovéto kombinace je zvýšení užítku z prováděné evaluace.

Explorační přístup čelí kritice poukazující na nedostatečně transparentní postup při stanovení indikátorů. Jak je nalézt? Jak provést operacionalizaci, když nejsou zadane (resp. evaluátorovi nejsou známe) cíle posuzované intervence?

Uvedení exploračního přístupu není ve vývoji evaluační metodologie ničím nahodilým; rozvinutí tohoto přístupu jasně vychází z diskursu 60. a 70. let, jenž byl veden na téma účelu a poslání evaluací jako takových. Explorační přístup je koncipován jako imanentní součást intervenčních programů, jako průvodce zainteresovaných stran napomáhající v jejich rozhodovací činnosti ve všech fázích rozvoje a v celé šíři implementace posuzovaných intervencí. M. Scriven svým přístupem za celý evaluační obor jasně deklaroval ambici zapojit evaluace nejen do procesu implementace intervenčních programů. Tím se pokusil marginalizovat tehdejší úzus

přípravy a provedení evaluačních šetření až v pokročilých fázích implementace intervencí.

Ilustrační příklad exploračního přístupu:

Evaluátor, který použije explorační přístup, nezná záměry iniciátora hodnocených kursů, a tak věnuje pozornost všem zaznamenanatelným efektům, ke kterým kurs vede. Evaluátor dojde k závěru, že kurs sice nenaučí jednotlivé účastníky formátovat odstavce a vytvářet tabulky, nicméně vede ke zdokonalení obecných uživatelských návyků (pravidelné zálohování, rychlejší psaní na klávesnici apod.). Současně je díky exploračnímu přístupu odhalen jeden z nezamýšlených efektů kursu, kdy si účastníci mezi sebou vyměňují zkušenosti z průběhu pohovorů u potenciálních zaměstnavatelů a sdílí mezi sebou aktuální nabídky volných pracovních pozic, což jim napomáhá k rychlejšímu získání nových pracovních míst.

Explorační přístup tak spíše než k efektivitě kursu obrací pozornost k jeho celkové užitečnosti. Zadavatel evaluace (resp. iniciátor kursů) může na základě výsledků takto provedeného šetření odhadnout míru efektivnosti kursu (i když významně méně přesně než při použití konfirmačního přístupu) a také může identifikovat konkrétní formy, jak kurs upravit pro ostatní regiony. Explorační přístup pomůže odhalit a zvýraznit takové parametry kursu, které by i přes svůj vysoký význam zůstaly v implicitní rovině.

3. Konstruktivistický přístup (Constructivist evaluation; Egon Guba, Yvonna Lincolnová)

Konstruktivistický přístup (samotnými autory označovaný také jako „evaluace čtvrté generace“ či původně publikovaný pod názvem „naturalistický přístup“) je reakcí na obecně převládající pozitivistické chápání reality jako objektivní danosti a možnosti jejího zkoumání a hodnocení. Čelní představitelé konstruktivistického přístupu, E. Guba a Y. Lincolnová, se přibližně od počátku 80. let snažili rozproudít vývoj evaluačního oboru diskuzí o hodnotovém pluralismu, kritikou přílišné koncentrace evaluátorů na výzkumnou metodologii a poukazováním na nevhodný design evaluací, který častokrát zůžil posuzované situace natolik, že dosažené záměry ztratily vnější validitu.

V rámci historického kontextu je třeba poukázat na skutečnost, že 80. léta v USA znamenala oproti předchozím dvěma desetiletím zásadní redukci

počtu externě prováděných evaluací požadovaných federální vládou. Snížení poptávky po evaluacích bylo způsobeno jednak rozpočtovými škrty Reaganovy administrativy (bylo méně intervenčních programů, jež by bylo třeba evaluovat), a dále také posílením role interních evaluačních týmů působících v rámci jednotlivých ministerstev. Rovněž je třeba upozornit na rostoucí význam evaluací⁹. Evaluátoři se tak více než dříve obraceli k principiálním otázkám evaluací na rozdíl od praktických aspektů jejich provádění, jež dominovaly v předchozím období.

E. Guba a Y. Lincolnová při formulaci svého přístupu vycházeli jednak ze Stakeova responzivního přístupu¹⁰, ze kterého převzali tezi, že evaluačního šetření by se měl účastnit nejen evaluátor a zadavatelé, ale také všechny další zainteresované strany, kterých se prováděná evaluace dotýká. Konstruktivisté v tomto ohledu dokonce přiznávají všem zainteresovaným stranám, které jsou zapojeny do prováděné evaluace, plně angažovanou účast zahrnující jak spolurozhodování v otázkách cílů, tak také právo ovlivňovat jednotlivé činnosti směřující k závěrům daného evaluačního šetření. Druhým pilířem, o nějž se konstruktivistický přístup opírá, je metodologie založená na hermeneutické dialektice. Evaluace založené na konstruktivistickém přístupu tak nejprve sumarizují a analyzují mentální konstrukce jednotlivých zainteresovaných stran. V druhém kroku jsou pak tyto mentální konstrukce podrobeny vzájemnému srovnávání, kdy jsou identifikovány oblasti shody mezi zainteresovanými stranami a témata, v nichž se zainteresované strany liší. Následují diskuse zainteresovaných stran facilitované evaluátorem, jejichž cílem je sblížit mentální konstrukce zainteresovaných stran a dosáhnout shody.

E. Guba s Y. Lincolnovou tak zpochybňují epistemologické možnosti nezávislého a nezáčastněného poznávání (objektivní) reality a kladou je do protikladu s transakčním subjektivismem, o nějž se opírá jimi zdůrazňovaný intersubjektivní (sdílený) charakter individuálních konstrukcí.

⁹ E. House (1993) komentuje: „Evaluace začaly být až příliš důležitou činností na to, aby byly ponechány jen v rukách akademiků.“

¹⁰ Responzivní přístup je dalším ze způsobů, jak mohou být evaluace koncipovány. Toto pojetí je typické tím, že nemá pevný design, ten se utváří postupně krok za krokem až v průběhu realizace dané evaluace vždy na základě dílčích závěrů z předchozích etap.

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

Konstruktivistický přístup klade do středu své pozornosti zainteresované strany a věnuje značnou pozornost analýze jejich role v posuzovaném intervenčním programu a sledování konkrétních zájmů, jimiž jsou tyto zainteresované strany vedeny. Naproti tomu explorační (M. Scriven) a konfirmační (R. Tyler) přístup obrací svou pozornost k cílům sledovaných intervenčních programů a k metodám, jak co nejlépe ověřit jejich dosažení. Zainteresované strany vystupují v takto pojatých evaluacích jen jako jeden z činitelů, v němž se zrcadlí přínos (worth) posuzovaných intervenčních programů, nikoliv však jejich věcný základ (merit).

Mnoho evaluátorů poukazuje na značnou obtížnost provedení evaluací způsobem, který by využíval konstruktivistický přístup, přesto se však lze s konkrétními příklady jejich praktického využití setkat. Významnou bariéru aplikačního rozvoje konstruktivistického přístupu představuje principiální cíl takto prováděných evaluací, jímž je konvergence protichůdných zájmů (subjektivních mentálních konstrukcí) a dosažení konsenzu. Pawson a Tilley (1997) proto zařazují konstruktivistický přístup založený na negociaci mezi nositeli jednotlivých mentálních konstrukcí do „útluného předměstského světa, v němž se odehrávají setkání mezi řediteli škol, učiteli a rodiči žáků, při kterých marní čas diskuzemi o charakteru domácích úkolů pro pátou třídu“. Rovněž další z předpokladů, hermeneutický dialektický proces, činí z konstruktivisticky prováděných evaluací činnost, která je v rámci obvykle požadovaných časových rámců (řádově měsíce) prakticky neuchopitelná a neproveditelná. Konstruktivistický přístup tak sice příliš nezměnil evaluační praxi a vžitě způsoby provádění evaluací, nicméně významným způsobem ovlivnil evaluační teorii. Je téměř nemožné najít monografii či učebnici evaluací vydanou v posledních pětadvaceti letech, která by na konstruktivistický přístup neodkazovala. Je třeba na tomto místě poukázat na skutečnost, že samotný přístup se stal

inspirací pro mnohé další (např. participativní přístup) a má rovněž své přímé pokračovatele – např. Bhola (1998), Schwandt (1984) či Huebner, Betts (1999).

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

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

Ilustrační příklad konstruktivistického přístupu:

*Při uplatnění konstruktivistického přístupu diskutuje evaluátor během společného brainstormingu s vybranými účastníky daného kursu; nezaměstnanými, kteří nebyli do kursu přijati; lektory; zástupci potenciálních zaměstnavatelů; pří-
mými nadřízenými zaměstnanců, kteří se v minulosti kursu zúčastnili a získali nové zaměstnání; pracovníky daného úřadu práce; zástupci iniciátora daného intervenčního programu a s rodinnými příslušníky účastníků kursu. V průběhu diskuzí, které evaluátor moderuje, se pozornost zaměřuje na očekávání, s nimiž jednotlivé zainteresované strany vstupovaly do dané situace (tím dojde ke konfrontaci subjektivit resp. mentálních konstrukcí). Evaluátor zjišťuje, která z těchto*

očekávání se naplnila a která nikoliv, následně hledá způsoby, jak koncipovat či vylepšit posuzovaný kurs, aby uspokojil potřeby každé ze zainteresovaných stran. Tyto diskuze jsou opakovány do doby, než je nalezeno kompromisní uspořádání kursu, které maximalizuje užitečnost pro každou ze zainteresovaných stran. Výsledkem je tedy návrh změn v uspořádání kursu, které vede k tomu, že iniciátor považuje kurs za užitečný, (potenciální) účastníci v něm spatřují cestu k získání nového zaměstnání, pracovníci úřadu práce jsou přesvědčeni, že absolventy takto koncipovaného kursu dokáží snáze umístit na trhu práce a zaměstnavatelé budou preferovat absolventy kursu oproti jiným uchazečům, neboť jsou přesvědčeni o tom, že tito absolventi skutečně disponují potřebnými schopnostmi a dovednostmi.

Zadavatel evaluace se dozví jen velmi málo o efektivnosti kursu, a tak rozhodnutí o rozšíření do dalších regionů bude muset opřít o jiné poznatky, nicméně díky konstruktivisticky pojaté evaluaci získá podobné informace o tom, jaký je přínos kursu pro jednotlivé zainteresované strany. Takovéto poznatky mohou pomoci jak při zlepšování kursu, tak také při komunikaci s jednotlivými zainteresovanými stranami.

4. Evaluace vedená teorií (Theory-driven evaluation; Huey-Tsyh Chen, Peter Rossi)

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

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

tak také jinak koncipovaná šetření. Významnou změnu oproti ostatním popisovaným přístupům představuje způsob využití takto získaných údajů. Ty jsou totiž vztahovány ke kontextu dané intervence a slouží k vysvětlení způsobu, jak daná intervence funguje. Zájem vysvětlení je v evaluačních ojedinělý. Z mnoha pramenů je patrné, že smysl, cíle a poslání evaluací jsou tradičně spatřovány zejména v popisu (deskripci) a v jisté omezené míře v exploraci, nicméně pokus o vysvětlení (např. způsobu fungování intervenčního programu) se v uchopitelné a systematicky pojaté podobě objevuje až u H. Chena. Na tento přístup a jeho nosnou ideu o evaluačních vysvětlujících mechanismy fungování posuzovaného intervenčního programu navázal na realistický přístup, který přinesl do evaluační teorie klíčové prvky emergentního realismu (zejm. pak generativní pojetí kauzality, roli kontextu či rozlišení mechanismu a pravidelnosti).

Chenův přístup směřuje k eliminaci interpretační nejistoty (zejména) v případě přijetí nulové hypotézy o fungování intervence programu v podmínkách experimentálního designu¹¹. Pokud totiž není shledán statisticky významný rozdíl mezi výstupy v testovací a kontrolní skupině, nabízí se vždy několik zásadních otázek či pochybností: Znamená výsledek testu, že daný intervenční program skutečně nevyvolává očekávané efekty? Co když je neprůkazný výsledek způsoben navrženým evaluačním designem? Co když jsou použity takové nástroje měření, které jen nejsou dostatečně citlivé, aby mohly způsobený efekt zachytit? Mohli bychom dosáhnout jiného výsledku, tj. prokázat funkčnost dané intervence, pokud bychom zvýšili počet případů/velikost testovací a kontrolní skupiny? H. Chen proto v této souvislosti nabízí řešení, které spočívá v doplnění poznatků získaných z evaluací založených na experimentálním (či kvaziexperimentálním) designu o nový typ informací a znalostí – o tzv. „teorii“ působení dané intervence (Donaldson 2007; Rogers 2008).

Ústřední pojem – teorie – tedy v podání autora znamená pravděpodobný a obhajitelný model očekávaného působení daného intervenčního programu (Chen, Rossi 1983). Chenova „teorie“ je představou o mechanismech, kterými by daný intervenční program mohl působit, a které by mohly vyvolávat pozorovatelné výstupy. Kondicionál v poslední větě je

¹¹ Tedy při ověřování účinnosti intervence pomocí porovnávání výsledků dosažených v testovací skupině (jež byla intervencí podrobena) oproti výsledkům skupiny kontrolní (ve které nebyla intervence aplikována).

zvolen záměrně, neboť H. Chen sám nechápal svou definici teorie deskriptivním způsobem, tedy jako konstatování faktického stavu, nýbrž jako vyjádření žádoucího, lepšího či optimálního stavu (Chen 1990). Chenovu teorii je tedy možné si představit jako soubor vzájemně provázaných hypotéz o vztazích mezi jednotlivými událostmi mezi prvky dané intervence. Vlastní proces evaluačního šetření je pak *de facto* verifikací tohoto souboru hypotéz a modifikací výchozí představy na základě empiricky získávaných údajů.

Přístup k provádění evaluací, který je veden teorií, je založen na třech elementárních principech. Tím prvním je zaměření na změnu (*action oriented*); zahrnuje totiž detailní popis jednotlivých mechanismů, pomocí nichž intervenční programy mohou působit (tj. vyvolávat změnu). V tomto ohledu přitom nejde ani o popis výchozí struktury, ani o zachycení jednotlivých prvků samotných intervenčních programů, ale jde o zachycení dynamiky intervence. Druhý princip Chenova přístupu je založen na konceptualizaci samotné intervence a na detailním rozboru použité implementační strategie. V tomto ohledu se teorií vedený přístup zaměřuje na analýzu konkrétně použitých nástrojů dané intervence a na zkoumání intenzity či způsobu jejich použití. Třetím z principů tohoto přístupu je volba kritérií, na jejichž základě jsou výsledky daného intervenčního programu posuzovány. Zde H. Chen navazuje jak na R. Tylera, tak také na M. Scrivena, kteří v rámci svých přístupů věnovali volbě a adekvátnímu stanovení indikátorů značnou pozornost. H. Chen však zvolil odlišný způsob jejich stanovení – navrhuje totiž, aby kritéria pro posouzení účinnosti dané intervence navrhly samotné zainteresované strany.¹²

H. Chen v rámci svého přístupu definoval celkem šest oblastí¹³, jimž je třeba při provádění evaluací věnovat pozornost. Jde (1) o samotnou povahu intervence (*treatment*), jejíž poznání a analýza je nezbytným východiskem pro zachycení změn vyvolaných daným intervenčním programem, (2) o implementační prostředí, které jasně odkazuje ke kontextu, v němž se posuzovaná intervence, resp. intervenční program odehrává, a (3) o výsledky, ke kterým daný intervenční program dospěl. Druhá trojice

¹² Pokud tato výzva zní jako utopie, která nemůže vést v praxi k adekvátním závěrům, odkazujeme na pravidla stanovení indikátorů využívaných pro monitoring a evaluaci operačních programů Evropské unie.

¹³ H. Chen je nazývá domény (*domains*).

oblastí zahrnuje (4) dopady, jejichž detailní analýzu H. Chen považuje za ústřední téma celé evaluace, (5) intervenující mechanismus, což je v podstatě kauzální vztah spojující povahu intervence a její výsledek (*outcome*) a (6) zobecnění, v souvislosti s nímž H. Chen upozorňuje, že získané poznatky mohou být využívány i s časovým odstupem, pro jiné populace či v jiných kontextech.

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

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

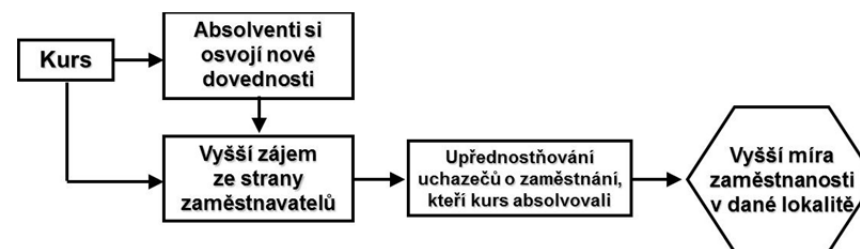
Přístup k provádění evaluací, který je veden teorií, lze tedy využívat zejména v takových případech, kdy není možné testovat vztahy mezi jednotlivými proměnnými přímo; v těchto situacích dokáže teorií vedený přístup zvýraznit výsledky sledovaných programů a zároveň také přesně specifikovat jednotlivé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

Ilustrační příklad teorií vedeného přístupu:

Evaluátor provádějící evaluace v duchu tohoto přístupu po úvodním seznámení s danou situací připraví „teorii“, jak daný kurs může dosahovat očekávaných výsledků, popř. převezme (a upraví) teorii připravenou zadavatelem evaluace.

Následně připraví takový design evaluace, který umožní vytvořenou teorii ověřit. Takový model může mít třeba podobu, jež je patrná z následujícího obrázku:

Obrázek 1: Návrh konceptuálního schématu



Zdroj: autor

V průběhu evaluace je výchozí model (dokumentující teorii změny) postupně upravován podle získávaných poznatků tak dlouho, až vznikne finální podoba „ověřené“ teorie. Evaluátor během provádění takovéto evaluace spolupracuje s jednotlivými zainteresovanými stranami, získává konkrétní podkladová data o jednotlivých prvcích své teorie a klade přitom důraz na přesné zachycení a popis klíčových procesů (vazeb mezi jednotlivými prvky). V závěrečné zprávě pak evaluátor zdůrazní kritické momenty, které se spolupodílejí na úspěšnosti kursu.

Zadavatel tak získá podklad verifikující, zdali předpokládaný mechanismus působení kursu skutečně funguje či nikoliv. Teorii vedený přístup tedy neověřuje dosažené cílové hodnoty (jako tomu je v přístupu konfirmačním), nehledá ani jednotlivé efekty, k nimž kurs vede (viz explorační přístup), ale prověřuje platnost předpokládaných vztahů mezi aktivitami a výstupy, mezi výstupy a výsledky a mezi výsledky a dopady. Zadavatel na základě výsledků takto pojaté evaluace zváží, zdali i v ostatních regionech panují podobné okolnosti jako v pilotní lokalitě (výchozí úroveň počítačové gramotnosti, počet a struktura pracovních sil atd.) a pokud ano, rozhodne o rozšíření kursu.

Souhrn

Z nastíněného přehledu jednotlivých přístupů je patrné, že každý z nich je možné si představit jako reflektor, který osvětluje konkrétní část dané intervence. Ze zkušeností a z dostupných podkladů se ukazuje, že žádný z existujících přístupů neposkytuje recept na všechny situace,

ve kterých evaluace jsou a mohou být prováděny. Existence několika evaluačních přístupů je tak odrazem mnohosti hodnocených situací; pro každou situaci je vhodný jiný evaluační přístup. Současně to znamená, že v dané situaci lze využít několika různých přístupů a zohlednit tak konkrétní parametry dané aplikační situace. Tato variantnost v přístupech k realizaci evaluací na jedné straně představuje jistá rizika, neboť volba konkrétního přístupu může být v dané aplikační situaci provedena laikem, ať už na straně realizátora či na straně zadavatele bez adekvátní rozvahy nad implikacemi takovéto volby. Na druhé straně platí, že zmiňovaná mnohočetnost přístupů, jakkoliv činí oblast evaluací nepřehlednou, je smysluplná a v konečném důsledku užitečná, neboť přibližuje evaluační metodologii konkrétním aplikačním situacím.

V praxi se obvykle zohledňují pragmatická kritéria pro volbu přístupu či pro nastavení designu dané evaluace. V řadě případů je typ (přístup) definován zadavatelem, v jiných situacích nebývá čas či prostředky k provedení evaluace vícero přístupů, v některých případech může chybět povědomí o alternativních způsobech provedení evaluace či znalost konkrétních metodologických nástrojů nutných pro realizaci daného evaluačního přístupu. Způsob provedení evaluace je dále determinován jejím cílem, což některé přístupy v dané aplikační situaci diskvalifikuje. Přesto však je vhodné o alternativních přístupech vždy uvažovat, neboť jejich efektivní realizací může být významně zvýšena užitečnost (a využitelnost) provedené evaluace.

Porovnání některých charakteristik jednotlivých teoretických přístupů k realizaci evaluací

	Konfirmační (R. Tyler)	Explorační (M. Scriven)	Konstruktivistický (E. Guba, Y. Lincolnová)	Teorii vedený (H. Chen)
Definice	Shoda mezi chováním a cíli	Srovnání výsledku s potřebami (<i>needs</i>)	Evaluační realita není objektivně daná, ale je mentální konstrukcí	Mechanismus intervence je zachytitelný pomocí teorie změny
Předmět evaluace	Osvědčení cílů	Identifikace dosažených efektů	Vnímání efektů zainteresovanými stranami	Vysvětlení mechanismu dosažení efektů
Klíčová informace	Cíle a výstupy, indikátory	Zaměřené a nezaměřené důsledky	Konstrukce reality	Teorie působení intervence
Teoretický základ evaluace	Pozitivismus	Nespecifický	Fenomenologie	Vědecká metoda
Metodologický akcent	Vnitřní validita	Nespecifický	Hledisko validity je odmítáno	Vnější validita
Role evaluátora	Nezúčastněná, objektivní	Nezávislá autorita; minimální interakce se zadavatelem	Je jednou ze zainteresovaných stran	Nezúčastněná
Role kontextu	Bezvýznamná	Nespecifická	Dána interakcí mentálních konstrukcí zainteresovaných stran	Zásadní
Výhody	1. Je systematický 2. Stojí na tradičních vědeckých základech 3. Legitimizuje i jiné než konvenční metody sběru dat	1. Je proveditelný i bez znalosti cílů 2. Upozorňuje na roli vedlejších a neočekávaných efektů 3. Při stanovení indikátorů je flexibilní	1. Klade důraz na zájmy všech zainteresovaných stran	1. Pomáhá pochopit mechanismus působení intervence
Nevýhody	1. Nezahrnuje způsob definice cílů 2. Příliš akcentuje pre-post design 3. Vede ke konvergentním efektům (předčasné závěry, odhlíží od dopadů)	1. Neposkytuje návod na provedení v praxi 2. Vyžaduje kompetentního evaluátora (příliš spoléhá na jeho kvalitu) 3. Definice potřeb může být zkrácena výzkumným procesem	1. V praxi je obtížně proveditelný 2. Zájmy zainteresovaných stran někdy konvergovat nemohou 3. Riziko ztráty nestrannosti evaluátora je velmi vysoké	1. Zaměřuje se na nepozorovatelné procesy 2. Riziko opakování záměrů iniciátora intervence je vysoké

Zdroj: autor

Literatura

- [1] Alkin M. C. (2004) *Evaluation Roots*. London: SAGE Publication.
- [2] Alkin M. C., Christie C. A. (2005) *Theorists' Models in Action*. San Francisco: Jossey-Bass.
- [3] Bloom B. S., Engelhart M. D., Furst E. J., Hill W. H., Krathwohl D. R. (1956) *Taxonomy of Educational Objectives: Handbook I. Cognitive Domain*. New York: David McKay.
- [4] Bholá H. S. (1998) "Program evaluation for program renewal: A study of the national literacy program in Namibia". *Studies in Educational Evaluation* 4
- [5] Chen H-T. (1990) *Theory-Driven Evaluations*. London: SAGE Publishing.
- [6] Chen H-T., Rossi P. H. (1983) "Evaluating with Sense: The Theory-driven Approach". *Evaluation Review* 7
- [7] Donaldson S. I. (2007) *Program Theory-Driven Evaluation Science*. New York: LEA.
- [8] Guba E. G., Lincoln Y. S. (1989) *Fourth Generation Evaluation*. London: SAGE Publishing.
- [9] Hall D., Hall I. (2004) *Evaluation and social research – introducing small-scale practice*. New York: Palgrave MacMillan.
- [10] House E. R. (1980) *Evaluating with Validity*. London: SAGE Publishing
- [11] House E. R. (1993) *Professional Evaluation. Social Impact and Political Consequences*. London: SAGE Publishing.
- [12] Huebner A.J., Betts S.C. (1999) "Examining Fourth Generation Evaluation: Application to Positive Youth Development". *Evaluation* 1999 5: 340-358.
- [13] Lay M., Papadopoulos I. (2007) "An Exploration of Fourth Generation Evaluation in Practice". *Evaluation* 2007 13: 495-504.
- [14] Mager R. F. (1962) *Preparing Instructional Objectives*. Palo Alto: Fearon.
- [15] Mark, M. M., Henry, G. T., Julnes G. (2000) *Evaluation: An Integrated Framework for Understanding, Guiding and Improving Policies and Programs*. San Francisco: Jossey-Bass

- [16] Molas-Gallard J., Davies A. (2005) "Toward Theory-Led Evaluation. The Experience of European Science, Technology, and Innovation Policies". *American Journal of Evaluation* 2: 1-18.
- [17] Patton M. Q. (1982) *Practical Evaluation*. London: SAGE Publishing.
- [18] Patton M. Q. (2008) *Utilization-Focused Evaluation*. London: SAGE Publishing.
- [19] Pawson R., Tilley N. (1997) *Realistic Evaluation*. London: SAGE Publications.
- [20] Rodríguez-Campos L. (2005) *Collaborative Evaluations*. Florida: Llumina Press
- [21] Rogers P. J. (2008) "Using Programme Theory to Evaluate Complicated and Complex Aspects of Intervention". *Evaluation* 2008 14: 29-48.
- [22] Rossi P. H., Lipsey M. W., Freeman H. E. (2004) *Evaluation. A Systematic Approach*. London: SAGE Publications.
- [23] Sanders R. J. (1994) *The Program Evaluation Standards* 2. London: SAGE Publishing.
- [24] Schwandt T. A. (1984) *An Examination of Alternative Models for Sociobehavioral Inquiry*. Indiana: IN.
- [25] Scriven M. (1973) "The Methodology of Evaluation". In Worthen B. R., Sanders J. R. (eds.) *Educational Evaluation: Theory and Practice*. Belmont: Wadsworth.
- [26] Scriven M. (1991) *Evaluation Thesaurus*. London: SAGE Publishing.
- [27] Scriven M. (1993) *Hard-Won Lessons in Program Evaluation*. San Francisco: Jossey-Bass.
- [28] Shadish W. R., Cook T. D., Leviton L. C. (1991) *Foundations of Program Evaluation*. London: SAGE Publishing
- [29] Stake R. E. (1975) *Program Evaluation, Particularly Responsive Evaluation*. Kalamazoo: Western Michigan University Evaluation Center
- [30] Stake R. E. (2004) *Standards-Based & Responsive Evaluation*. London: SAGE Publishing.
- [31] Stufflebeam D. L., Shinkfield A. J. (2007) *Evaluation, Theory, Models & Applications*. San Francisco: Jossey-Bass.
- [32] Weiss C. H. (1972) *Evaluation Research*. New Jersey: Prentice-Hall