
METODIKA EVALUACE PROGRAMŮ ENVIRONMENTÁLNÍ VÝCHOVY

Jan Činčera, Technická univerzita v Liberci a Centrum pro otázky životního prostředí Univerzity
Karlovy, 2010



Centrum pro otázky životního prostředí Univerzity Karlovy v Praze, 2010

OBSAH

Obsah	1
Cíle metodiky	2
Základní pojmy	3
Program.....	3
Základní vymezení.....	3
Teorie programu	4
Cíle a výstupy programu.....	7
Evaluace	12
Základní vymezení.....	12
Formativní a sumativní evaluace	12
Roviny a oblasti evaluace.....	13
Specifické druhy evaluace	14
Specifika evaluačního výzkumu	15
Praktické aspekty evaluace	16
Evaluační proces	17
Fáze 1 Kontrakt a informace o programu.....	18
Fáze 2: Formulace evaluačních otázek.....	22
Fáze 3 Plán evaluace.....	23
Metodologie výzkumu	23
Typ výzkumu	25
Volba vzorku.....	28
Plán sběru dat	29
Nástroje pro sběr dat	33
Fáze 4 – sběr a Analýza dat	40
Fáze 5 Evaluační zpráva.....	44
Příklad výstupů evaluačních výzkumů	46
Evaluace programu Zelená škola	46
Doporučení	48
Literatura.....	49
Abstrakt.....	53
Abstract	53

CÍLE METODIKY

Každý rok vydávají orgány veřejné správy a samosprávy finanční prostředky na podporu formální i neformální environmentální výchovy. Protože zájem o dotaci je vždy vyšší, než dostupné zdroje, musí odpovědní pracovníci vybírat projekty, které podpoří. Jak ale ukázal výzkum mezi krajskými úřady a magistráty v roce 2009 (Činčera, Kulich & Gallová, 2009), používané metody výběru nezaručují, že budou skutečně podpořeny nejefektivnější programy. Odpovědní pracovníci zpravidla nevědí, jakým způsobem lze vyhodnocovat efektivitu programů či jak z projektové žádosti poznat, zda je zde předpoklad, že program naplní své výchovně vzdělávací cíle. Výsledkem je prostředí, které žadatele o grant nemotivuje ke zkvalitňování svých programů a ve kterém část vynakládaných prostředků mizí bez předpokládaného efektu.

V řadě zemí je přitom běžné, že donoři vyžadují po příjemci finanční podpory doložení efektivity podpořeného programu. Základním prostředkem k tomu jsou evaluační výzkumy. Evaluační výzkumy si zadávají realizátoři či sponzoři programů zpravidla u externích organizací či subjektů. Na jejich základě pak lze získat informace o efektivitě použitých strategií. Ty pak může využít realizátor pro zkvalitňování svého programu i donor pro rozhodování o dalším rozdělování finančních prostředků. Dlouhodobě pak na evaluačních výzkumech získá celá komunita: realizátoři budou připravovat kvalitnější programy a donoři získají za své peníze vyšší efekt.

Předložená metodika je určena především pro pracovníky veřejné správy, kteří rozhodují o distribuci finančních prostředků na podporu formální i neformální environmentální výchovy.

Cílem prezentované metodiky je

- Vymežit základní terminologii pro evaluaci programů environmentální výchovy;
- Popsat základní fáze evaluačního cyklu;
- Charakterizovat základní alternativní postupy při evaluaci programů;
- Identifikovat základní požadavky na evaluační výzkum.

Zpracování metodiky vychází z předpokladu, že se ve střednědobém horizontu stanou evaluační výzkumy běžnou součástí environmentální výchovy. Pro dosažení tohoto cíle je zapotřebí, aby se odpovědní pracovníci orientovali v tématice a získali základní porozumění pro možnosti i nároky evaluačních výzkumů.

ZÁKLADNÍ POJMY

Klíčové pojmy: program, logický model, cíle a výstupy, sumativní a formativní evaluace, oblasti a roviny evaluace.

PROGRAM

ZÁKLADNÍ VYMEZENÍ

Program = „skupina plánovaných aktivit, které naplňují společný strategický cíl. Program bývá obvykle rozdělen na podprogramy nebo projekty.“ (Department of Environment and Conservation, 2004)¹ Program je tedy označení pro výchovně vzdělávací celek, který je realizátory chápán jako jeden produkt a je smysluplnou jednotkou pro evaluaci. Program může být krátký (v řádu vyučovacích hodin), středně dlouhý (několikadenní pobytová akce) i dlouhý (roční i delší integrovaný tematický celek).²

Ekologické výukové programy (EVP) představují tradiční nabídku středisek ekologické výchovy (SEV). Jsou vymezeny jako „výchovně vzdělávací lekce, jejichž smyslem je obohatit vzdělávání na všech stupních škol o ekologický a environmentální rozměr“ o rozsahu nejméně dvou vyučovacích hodin (Máchal, Činčera, Daňková & Broukal, 2008). Typicky tedy pojem EVP označuje krátké programy, v některých materiálech jsou ale pod stejný pojem zařazovány i pobytové akce. V tomto textu budou jako EVP chápány krátké programy v rozsahu několika vyučovacích hodin nepřesahující jeden den.

Pobytové akce jsou typy programů přesahujících délkou jeden den odehrávající se mimo prostory školy (typicky v pobytovém středisku ekologické výchovy). Kromě programů SEV sem mohou patřit i akce táborového a kurzovního typu a víkendové akce, pořádané zejména organizacemi působícími v oblasti neformálního vzdělávání (Hnutí Brontosaurus, Junák).

Jako **projekt** je často označováno několik různých typů programů. Např. metodický manuál Department of Environment and Conservation (2004) jako projekty chápe skupinu menších plánovaných aktivit, které naplňují cíle většího programu. Jindy jsou projekty zaměřovány s integrovanými tematickými celky. Rozlišit je také třeba projekty ve smyslu pedagogickém a

¹ Srovnej: „= Integrovaná sekvence plánovaných výchovně vzdělávacích aktivit a materiálů, jejichž cílem je dosažení určité sady výstupů“ (NAAEE, 1994).

² Pojem program má ale širší význam a nemusí zahrnovat pouze programy, jejichž těžiště je výchově a vzdělávání. V roce 2009 například Ministerstvo životního prostředí České republiky vyhlásilo program Zelená úsporám. Cílem bylo snížit emise skleníkových plynů obyvatel České republiky. Cílovou skupinou byly domácnosti, u kterých se očekávalo, že díky systému státních dotací (aktivity) a související informační kampani zateplí domácnosti a sníží tak spotřebu elektrické energie.

manažerském. V pedagogických projektech se žáci zpravidla podílí na definování cílů programu a samostatně spolupracují na jejich naplnění. Výstupem projektu je určitý produkt – ať už prezentace výsledků či reálná změna prostředí. Projekty jsou typicky mezioborové.

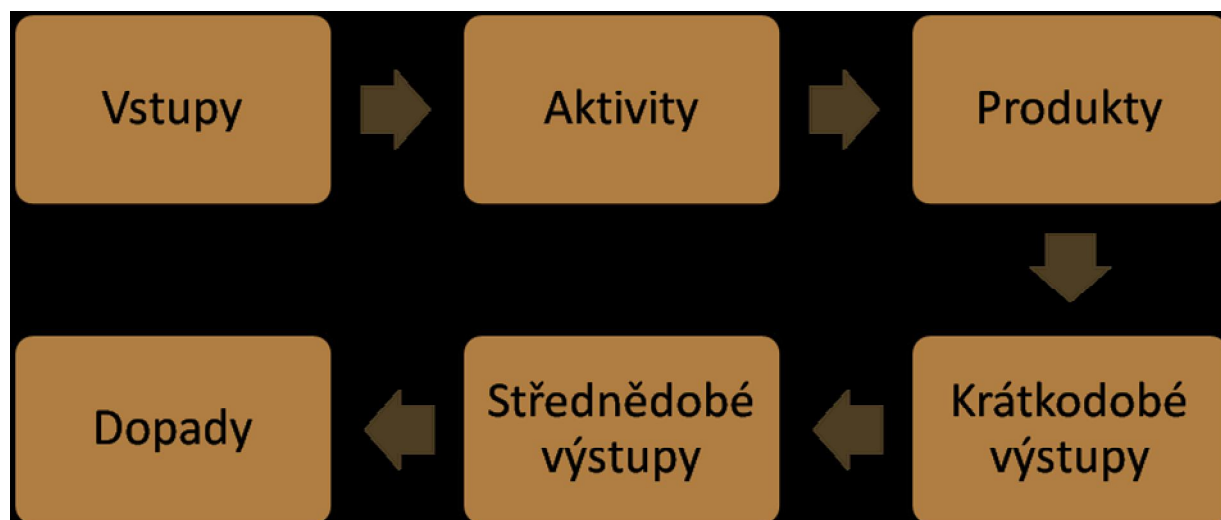
Teorii **integrovaných tématických celků** propracovala zejména S. Kovaliková (1994). Integrované tématické celky (ITC) jsou vymezeny obecným tématem, které je dále rozděleno do dílčích podtémat. Přestože ITC umožňují dětem spolupracovat na vymezení dílčích cílů a vyžadují po nich aktivitu, jsou ITC ve srovnání s projekty řízeny učitelem, který provádí děti sérií úkolů a aktivit. ITC mohou zahrnovat i menší projekty, mohou mít podobu pobytové akce či sledu pobytových akcí, případně sledu ekologických výukových programů nebo provázaných aktivit realizovaných v rámci volnočasové organizace.

Specifickým typem programu pak je **školní program EVVO** skládající se ze strategického a ročního plánu. Školní program může zahrnovat všechny předchozí typy programu.

TEORIE PROGRAMU

Základním východiskem pro tvorbu i analýzu programu je **teorie programu** (program theory). Ta je chápána jako soubor předpokladů o tom, jakým způsobem chce program dosáhnout svých předpokládaných cílů. Teorie programu se skládá z **teorie účinku** (impact theory), vyjadřující předpokládané kauzální vztahy mezi jednotlivými komponenty programu (vstupy, aktivitami, produkty a výstupy na krátkodobé, střednědobé a dlouhodobé úrovni) a **teorie procesu**, vyjadřující organizační plán (například časovou sekvenci jednotlivých aktivit programu (NAAEE, 1994; Rossi, Lipsey & Freeman, 2004). Teorie účinku bývá zpravidla vizualizovaná pomocí **logického modelu**, teorie procesu pomocí různých organizačních diagramů (McLaughlin & Jordan, 2004; NAAEE, 1994; Marcinkowski, 2004; W.K. Kellogg Foundation, 2004; Barch, Duvall, Higgs, Wolske & Zint, 2007).

Podstatou logického modelu je vizualizace předpokládaných kauzálních souvislostí vedoucích až k plánovanému dopadu programu. Nemá proto závaznou grafickou podobu. Jeho základní struktura ale zpravidla odpovídá následujícímu nákresu:



OBRÁZEK 1 LOGICKÝ MODEL PROGRAMU

Jako **vstupy** se zde chápou lidské, finanční i organizační zdroje nutné k realizaci programu. **Aktivita** programu jsou procesy, nástroje, události, technologie či akce, které jsou částí realizace programu. Zdroje a aktivity představují plánovanou práci na programu.

Produkty, výstupy a dopady znázorňují zamýšlené výsledky programu. **Produkty** (outputs) označují přímé produkty programu zahrnující typy, úrovně a cíle služeb, které program realizuje. **Výstupy** jsou specifické změny ve znalostech, dovednostech, chování či statutu účastníků programu. Mohou být krátkodobé (od jednoho do tří let) či střednědobé (čtyři až šest let). **Dopady** jsou dlouhodobé změny, které se objeví v organizaci, komunitě či systémech v horizontu sedmi až deseti let (W.K. Kellog Foundation, 2004)³.

V kontextu programů environmentální výchovy se pak může použít například následující schéma:

³ Mareš a Gavora (1999) překládají „outputs“ a „outcomes“ shodně jako výstupy, výsledky, případně produkty. V terminologii logického modelu mají ale oba pojmy odlišné významy. „Outcomes“ jsou ve standardech NAAEE (2009) chápány jako účinky programu, zahrnující jak obecné cíle (goals), tak očekávané výstupy (objectives). Pro zjednodušení zde používáme pojem „výstupy“. „Outputs“ je pak chápán spíše jako „produkty“, tj. např. počet realizovaných programů, počet absolventů, atd.



OBRÁZEK 2 LOGICKÝ MODEL PROGRAMU ENVIRONMENTÁLNÍ VÝCHOVY

Učením se zde myslí předpokládaný vliv programu na kognitivní (znalosti, porozumění, uvědomění⁴), afektivní (postoje, senzitivitu, hodnoty, odpovědnost, ohnisko kontroly⁵) a konativní (kognitivní dovednosti, emoční dovednosti) proměnné (Marcinkowski, 1997).⁶ Chováním se rozumí dlouhodobý vliv programu na chování absolventa. Dopady pak výsledek změny chování cílových skupin na životní prostředí a komunitu. Při tvorbě programu se doporučuje postupovat zprava doleva, tj. od analýzy potřeb a stanovení žádoucích dopadů, přes identifikaci cílových skupin a jejich žádoucího chování, ke stanovení kognitivních, afektivních a konativních proměnných, které je třeba ovlivnit pro dosažení žádoucího chování, dále k výběru vhodných aktivit potřebných k dosažení efektu na úrovni učení a nakonec k identifikaci zdrojů a předpokladů programu (Braus & Wood, 1993; NAAEE, 2004; W.K. Kellogg Foundation, 2004).

Důležité je zachovat logiku modelu, kdy každá další fáze má vyplývat jako logický důsledek fáze předchozí. Model tedy můžeme číst zhruba takto:

- *Jestliže investujeme sto hodin práce zaměstnanců střediska na přípravu programu a dvacet na jeho realizaci, částku 20.000,- Kč na nákup vybavení a využijeme v době programu prostory organizace (zdroje);*
- *budeme moci připravit, propagovat a administrativně zpracovat nový program (aktivity);*
- *který se bude skládat z jedné tří denní pobytové akce na středisku pro dvacet žáků sedmé třídy (produkty);*
- *po ukončení programu se 70 procent žáků přihlásí do dobrovolného projektu na ekologizaci provozu školy, 80 procent jich správně pozná na přiložených kartičkách platné ekoznačky a v odstavu dvou měsíců po akci 50 procent dětí potvrdí pokles nákupu nápojů v plechovce a v PET lahvích (výstupy);*

⁴ awareness

⁵ Locus of control

⁶ Výčet proměnných je závislý na paradigmatickém modelu environmentální výchovy. Viz např. koncept akčních kompetencí, zahrnujících kognitivní, afektivní i konativní oblast (Jensen & Schnack, 1997).

→ v horizontu jednoho roku od akce dojde k měřitelnému snížení spotřeby vody a energie ve škole, zavedou se kontejnery na separování odpadu a v místním supermarketu se zvýší nabídka biopotravin o nejméně 10 procent (dopady).

CÍLE A VÝSTUPY PROGRAMU

Důležitou součástí formulace teorie programu je přesné vyjádření jeho **cílů a výstupů**⁷ (Rossi, Lipsey & Freeman, 2004).

Cíle (goals) jsou zde chápány žádoucí efekty programu. Cíle jsou vyjádřeny obecně a mohou být formulovány z pohledu programu. V kontextu environmentální výchovy se jimi nejčastěji rozumí obecné stanovisko vyjadřující očekávání, jaké změny program přinese. Vychází ze společenských potřeb pocíťovaných v oblasti environmentální výchovy na regionální, národní či mezinárodní úrovni. Jsou obtížně měřitelné a jejich evaluace je proto náročnější. Cíle se mohou odvozovat od konceptu environmentální gramotnosti nebo proenvironmentálního chování (Bennett, 1989; Sellers, 2007; Schueller, Yaffee, Higgs, Mogelgaard, & DeMattia, 2006).

Příkladem cílů může být např.:

- *Cílem programu je vychovat z žáků občany, kteří rozumí environmentálním problémům svého regionu, znají jeho specifika a jsou ochotni nést spoluodpovědnost za jeho další vývoj.*
- *Cílem programu je, aby studenti snížili množství odpadu, který produkují a větší množství odpadu třídili a znovu používali. (Bennett, 1989)*
- *Cílem programu je podporovat odpovědné využívání přírodních zdrojů, rozvinout schopnost kritického myšlení, kooperativního řešení problémů a dovedností potřebných pro rozhodování tak, abychom dosáhli a udrželi rovnováhu mezi kvalitou života a kvalitou životního prostředí. (Simmons et. al., 2004)*

Oproti tomu (očekávané) **výstupy** (objectives) vyjadřují konkrétní a měřitelné (nebo pozorovatelné) výsledky programu. Jsou to tedy určité změny v porozumění, znalostech, postojích, dovednostech či chování, které vznikly v důsledku programu (Department of Environment and Conservation, 2004; Schueller, Yaffee, Higgs, Mogelgaard, & DeMattia, 2006). Formulují se vždy z pohledu cílové skupiny (NAAEE, 2004). Většina autorů dále doporučuje, aby výstupy odpovídaly tzv. **SMART** zásadám, tj. aby byly konkrétní, měřitelné, vyjádřené od cílové

⁷ Mareš a Gavora (1999) překládají „goal“ jako obecný cíl, zatímco „objective“ jako konkrétně vymezený cíl. V textu budeme pojem „goal“ překládat jako „cíl“ a „objectives“ jako „výstupy“.

skupiny, relevantní ve vztahu k oboru a potřebám a časově specifikované. Výstupy mají být dále zpracovány pro každou cílovou skupinu zvlášť a mohou tvořit hierarchické systémy (Bennett, 1989; Braus & Wood, 1993; Department of Environment and Conservation, 2004 Hungerford (2005b); Schueller, Yaffee, Higgs, Mogelgaard, & DeMattia, 2006; Jacobson, McDuff & Monroe, 2006; Sellers, 2007).

Např. ve vztahu ke druhému a třetímu z uvedených příkladů obecných cílů mohou být očekávané výstupy vyjádřeny takto:

- *Studenti uvedou nejméně tři možnosti, jakými libovolné zboží může být znovu využito, recyklováno či může být snížena jeho spotřeba.* (Sellers, 2007)
- *Studenti na základě různých zdrojů identifikují jednotlivá alternativní řešení problému a hodnoty, ze kterých tato řešení vychází. Každé řešení vyhodnotí z hlediska jeho ekologických a kulturních dopadů.* (Simmons et. al., 2004)

Pro formulaci specifických očekávaných výstupů existuje několika sad pravidel, které se v řadě ohledů setkávají. Asi nejznámější je doporučení, aby výstupy byly tzv. **SMART** (Henton, 1996):

- Specifické (Specific)
- Měřitelné (Measurable)
- Dosažitelné (Achievable)
- Relevantní (Relevant)
- Zpětně sledovatelné (Trackable).

Někteří autoři formulují zásady SMART s dílčími odlišnostmi. Simmons (2008) vymezuje A a T jako

- Účastníci (Audience) – definovat výstupy z pohledu účastníka;
- Časové ohraničení (Time-Bounded) – jasné vymezení časového rámce

Stokking et. al. (Stokking, van Aert, Meijberg, & Kaskens, 1999) doporučují držet se následujících zásad:

- Nedefinuj tématem, ale aktivitou studenta;
- Nepopisuj aktivitu učitele, ale studenta;
- Nepopisuj to, co se po žácích chce v průběhu lekce;

- Neodkazuj se na učebnice, ale přesně vymež, co student splní;
- Výstupy formuluj přesně a jednoznačně, aby nebyly předmětem interpretací;
- Používej krátké a jednoduché věty;
- Rozděľuj složité celky na jednoduché;
- Užívej taková slovesa, které referují k pozorovatelným úspěchům;
- Upřesni podmínky, za kterých lze úkol považovat za splněný.

Na míru konkrétnosti výstupů se názory liší. Zajímavé pojetí volí Bennett (1989), který očekávané výstupy rozlišuje na *obecné* a *specifické*. Obecné očekávané výstupy nejsou podle něj definovány pomocí výkonu studenta a měří se subjektivněji než cíle specifické (také zde označované jako *behaviorální*⁸), které lze měřit objektivně a které jsou definovány popisem toho, co žáci po ukončení programu dokážou vykonat. Toto rozlišení může korespondovat s tím, zda očekávané výstupy jsou *předem dané* nebo zda jsou *spoluvytvářeny* v průběhu programu. Zatímco výstupově orientované přístupy v environmentální výchově jednoznačně tíhnou k formulování předem daných, tedy specifických očekávaných výstupů, procesně orientované směry kladou důraz na spoluutváření cílů žáky a zdůrazňují otevřenost výukového procesu.

Oproti tomu Hungerford (2005) obecné (nebehaviorální) výstupy nedoporučuje. Podle něj obecné výstupy zpravidla popisují téma („kurz se bude zabývat...“), aktivity studentů („studenti shlédnou několik filmů...“) nebo obsah toho, co se studenti naučí, znají či porozumí („žáci se naučí ..., žáci získají porozumění..., žáci budou znát...“). Přestože takto formulované výstupy podle něj poskytují jisté informace, není na jejich základě možné určit kritéria pro měření úspěšnosti žáků.

V současné době se proto doporučuje formulovat očekávané výstupy specificky podle SMART zásad. Specifické očekávané výstupy jsou proto vždy vyjadřovány z pozice žáka nebo skupiny žáků a s pomocí akčních sloves. Skládají se ze tří základních částí: *podmínek*, za kterých žák provede příslušný výkon, popisu příslušného *chování žáků* vyjádřeného slovesem a *minimálního stupně úspěšnosti* (Bennett, 1989; Hungerford, 2005b). Mohou být vyjádřeny z pohledu jediného studenta nebo skupiny studentů.

I zde ale probíhají diskuse o vhodné míře konkrétnosti specifických výstupů.

⁸ Stokking et al. (1999) či Henton (1996) ale jako „behavioral objectives“ označují očekávané výstupy v rovině chování. Celá terminologie je značně nejednotná.

Bennett (1989) uvádí příklady obsahující kvantifikovaně vyjádřený minimální stupeň úspěšnosti:

- *Pokud student obdrží set kartiček s obrázky environmentálních problémů a kartiček s obrázky lidí, kteří se podílejí na jejich řešení, bude student schopen přiřadit správně k sobě 75 procent z nich.*
- *Pokud skupina obdrží set kartiček s obrázky environmentálních problémů a kartiček s obrázky lidí, kteří se podílejí na jejich řešení, přiřadí 70 procent studentů správně k sobě 75 procent kartiček.* (Bennett, 1989)

Je třeba podotknout, že s takto kvantifikovanými očekávanými výstupy se zpravidla v české praxi nesetkáváme a obvyklé nejsou ani v zahraniční literatuře. V novějších metodikách se podobné formulace používají spíše pro **indikátory**, tedy pro přímé ukazatele míry úspěšnosti programu.

V související diskusi doporučuje Hungerford (2005) vyhýbat se formulacím jak příliš vágním, tak příliš konkrétním. Podle jeho doporučení by výstupy měly být psány tak, aby mohly být ověřovány více než jednou otázkou v testu, resp. více než jedním indikátorem. Jako příklady vhodných pojetí uvádí:

- *Po ukončení programu student analyzuje environmentální problém a vyjmenuje způsoby, jakými problém ovlivňuje lidstvo i životní prostředí.*
- *Po ukončení programu student analyzuje environmentální konflikt a identifikuje jeho hráče, jejich postoje a hodnoty, ze kterých vycházejí.*
- *Po ukončení programu student připraví seznam pravidel, která by měly být respektovány při přípravě průzkumu či dotazníku.* (Hungerford, 2005)

V některých materiálech (Department of Environment and Conservation, 2004) jsou očekávané výstupy *hierarchizovány*. Tento přístup se používá zejména v souvislosti s aplikací logického modelu programu nebo v komunitních vzdělávacích projektech:

- **Okamžité výstupy** (*Deset domácností změní současný způsob nakládání s vodou, odpady a energií.*)
- **Střednědobé výstupy** (*Dvacet procent domácností v Brumově- Bylnici uplatňuje udržitelný způsob využívání vody, odpady a energie.*)
- **Dlouhodobé výstupy**, které se označují také jako **dopady** (impacts). (*Snížení produkce odpadu, spotřeby vody a energie v Brumově-Bylnici.*)

Jiným kritériem je rozdělení cílů podle zainteresovaných stran. Hentonová (1996) rozděluje tři základní kategorie cílů⁹:

- **Kurikulární**, které formuluje učitel a které pramení z výchovně vzdělávacího kurikula organizace;
- **Skupinové**, které formuluje celá skupina studentů nebo menší podskupiny a která jsou vyjádřením jejich zájmu o určitou oblast nebo které se vztahují k životu a dynamice skupiny;
- **Individuální**, které vyjadřují preference jednotlivých žáků, ať už ve vzdělávací či výchovné (osobnostní rozvoj) oblasti.

Všechny skupiny se vzájemně prolínají, takže určitý očekávaný výstup může ležet současně v kurikulární, skupinové i individuální rovině. Zejména některé výstupy ve skupinové rovině může být snazší vyjadřovat spíše obecně, než specificky, přestože i zde můžeme pracovat s měřitelnými ukazateli:

- *Ve skupině se zlepší vzájemné vztahy a posílí schopnost spolupracovat.*
- *Skupina se posune z fáze bouření do fáze normování.*
- *Skupina si v aktivitě zformuluje vlastní kritéria úspěšnosti a v časovém limitu je dosáhne.*
- *Skupina vyvine a uplatní takový mechanismus diskuse, který umožní každému členovi říct svůj názor, aniž by byl překřikován ostatními.*

Obdobně v případě individuálních výstupů:

- *František prezentuje před ostatními výsledky svého výzkumu formou, která splní existující standardy prezentací.*
- *Anička si holou rukou sáhne na hada a nedá najevo strach ani odpor.*

Celkově lze shrnout, že odborná literatura rozlišuje více způsobů, jak cíle a výstupy programu formulovat. Bez ohledu na přiklonění se k některé z uvedených pojetí je pro úspěšnou evaluaci programu důležité zachování posloupnosti ve zpřesňování a konkretizaci jeho výchovně vzdělávacích cílů.

⁹ Hentonová používá pojem „goals“, ve smyslu prezentované terminologie se ale může jednat i o „objectives“, tedy očekávané výstupy.

EVALUACE

ZÁKLADNÍ VYMEZENÍ

Evaluace = „proces kritického prověřování programu. Zahrnuje sběr a analýzu informací o aktivitách programu, jeho charakteristikách a výstupech. Jejím cílem je formulovat stanoviska o programu ke zvýšení jeho efektivity a / nebo k poskytnutí informací pro rozhodování o programu.“ (Patton, podle Barch, Duvall, Higgs, Wolske & Zint, 2007) Evaluace v oblasti environmentální výchovy slouží k posuzování hodnoty jednotlivých programů – na úrovni produktů, výstupů, procesů, prostředků či vztahu k využití jeho zdrojů (Bennett, 1989).

Evaluace úzce souvisí s pojmy **efektivita** (*effectiveness*) a **úspornost** (*efficiency*). **Efektivitou** se rozumí hodnocení, do jaké míry program dosáhl svých plánovaných výstupů. Program tedy může být posouzen jako efektivní, pokud v jeho důsledku významná část účastníků dosáhla vyššího stupně porozumění dané problematice, posunu v postojích či hodnotové orientaci, rozvinula své dovednosti, změnila chování, atd.

Úspornost posuzuje nákladnost programu. Program může být efektivní, ale neúsporný, pokud bude vyžadovat vysoké finanční, personální, časové či materiální nároky na realizaci (Bennett, 1989; Department of Environment and Conservation, 2004).

Evaluace tedy přináší výpovědi o efektivitě a úspornosti programu. Evaluace úzce souvisí s výzkumem a často je chápána jako jedna ze základních výzkumných oblastí environmentální výchovy (Hart & Nolan, 1999). Podle Bennetta (1989) je případný rozdíl mezi evaluací a výzkumem dán otázkou využití naměřených výsledků: je-li cílem formulovat doporučení a stanoviska k programu, jde o čistou evaluaci, pokud je cílem rozšířit poznání v oblasti teorie a metodiky environmentální výzkumy, jedná se o výzkum. Jiní autoři (např. Patton, 2002) ale upozorňují, že ač evaluační výzkumy používají podobné metody, jako aplikovaný výzkum ve společenských vědách, mají jiné cíle a měly by být poměřovány podle vlastních kritérií.

FORMATIVNÍ A SUMATIVNÍ EVALUACE

Evaluace se nejčastěji dělí na **formativní** a **sumativní** (Bennett, 1989; Stokking, van Aert, Meijberg, & Kaskens, 1999; Department of Environment and Conservation, 2004; Barch, Duvall, Higgs, Wolske & Zint, 2007). Cílem formativní evaluace je zkvalitnit probíhající program. Zadavatelem formativní evaluace proto bývá velmi často realizátor programu (např. středisko ekologické výchovy). Formativní evaluace se pak může zaměřit na kterékoliv oblasti logického modelu programu. Zadavatelem sumativní evaluace bývá často sponzor programu. Jejím cílem je vyhodnotit, do jaké míry ukončený program naplnil své cíle.

Formativní evaluace se provádí v průběhu programu a jejím cílem je modifikovat program v průběhu tak, aby co nejlépe naplnil své cíle. Formativní evaluace často zahrnuje analýzu potřeb a analýzu implementace programu: (Duvall, Higgs a Wolske, 2007)

Sumativní evaluace probíhá po ukončení programu a jejím cílem je vyhodnotit, do jaké míry program naplnil své cíle a dosáhl očekávaných výstupů. Sumativní evaluace se zpravidla zaměřuje na **evaluaci výstupů** a **evaluaci dopadů**.

ROVINY A OBLASTI EVALUACE

Evaluace může být rozdělena do několika typů a úrovní. Kirkpatrick (podle Hoganová, 2003; Simmons, 2008) rozlišuje čtyři úrovně evaluace:



OBRÁZEK Č. 3 EVALUAČNÍ ROVINY PODLE KIRKPATRICKA.

1. **Reaktivní a krátkodobá rovina.** Probíhá bezprostředně po ukončení programu nebo jeho části. Zjišťuje spokojenost účastníků, resp. míru, s jakou akce naplnila jejich očekávání. Součástí jsou dotazy na návrhy změn pro příští program.
2. **Rovina učení.** Měří, jak byly naplněny očekávané výstupy programu v rovině znalostí, dovedností a postojů.
3. **Rovina chování.** Měří míru transferu programu do vlastního života, tedy především změny v jednání účastníků.
4. **Rovina dopadů.** Poměřuje náklady a zisky programu a zaměřuje se na dlouhodobé dopady programu – např. zvýšení kvality práce.

Kirkpatrickovo dělení je možné promítnout i do vymezení **oblastí evaluace** podle logického modelu programu. Nejčastější typy evaluace podle zkoumané oblasti pak jsou následující:

- **Analýza potřeb** - cílem je vyhodnotit, jaké cílové skupiny program potřebují, v jaké míře a jakou formou. Z analýzy také můžeme určit, jaké potenciální cílové skupiny nejsou programem oslovovány a jak by se program měl vzhledem k nim přizpůsobit. Analýzu potřeb je možné realizovat ještě před zahájením programu.
- **Evaluace procesu** nebo **implementace**. Součástí této části evaluace je zpravidla průběžné vyhodnocování, zda program funguje tak, jak bylo plánováno. Výsledky jsou použity pro změny probíhajícího programu. Součástí může být také analýza počtu a typu účastníků a míry, v jaké jsou s programem spokojeni.
- **Evaluace výstupů** měří, do jaké míry program naplnil své očekávané výstupy. Výstupy se myslí krátkodobé a střednědobé změny ve znalostech, postojích, dovednostech či chování účastníků.
- **Evaluace dopadů** měří širší a dlouhodobější dopady programu. Těmi může být např. vznik sítě spolupracujících škol, navázání spolupráce s komunitou, zlepšení vztahů ve třídě, snížení výskytů šikany, zlepšení životního prostředí v regionu, atd.

SPECIFICKÉ DRUHY EVALUACE

Akční výzkum = v kontextu evaluační teorie takový typ evaluace, ve kterém hlavní roli hrají samotní realizátoři programu, případně ve spolupráci s externím evaluátorem. Evaluace pak se soustředí na vybraný dílčí aspekt fungování programu. Realizátor navrhne změnu, dokumentuje průběh implementace a vyhodnotí výsledky. Na jejich základě pak podle potřeby navrhuje další změny.

Expertní analýza = analýza teorie programu provedená na základě hodnocení vztahů mezi jednotlivými kategoriemi logického modelu, použitých strategií a dalších podkladů získaných bez samotného výzkumu programu. Může vycházet i ze specifických oborových standardů pro program.

Evaluace orientovaná na uživatele, participační evaluace (utilization-focused evaluation) = evaluace, do jejíhož procesu se zapojuje co nejvíce se zainteresovaných stran, tj. například realizátoři, sponzoři i účastníci programu diskutují spolu s externím evaluátorem design evaluace.

Evaluace úspornosti programu (cost-efficiency analysis) = evaluace, které vyhodnocují poměr mezi náklady a efekty programu.

Metaevaluace = evaluační výzkumy založené na shrnutí výsledků dříve provedených evaluačních výzkumů

SPECIFIKA EVALUAČNÍHO VÝZKUMU

Oproti běžným aplikovaným výzkumům v sociálních vědách mají evaluační výzkumy určitá specifika:

- **Pragmatická orientace** – evaluace vychází z konkrétní potřeby zadavatele a jeho možností. Důsledkem může být hledání kompromisu mezi rigorózním a pragmatickým polem práce. Evaluační výzkum je limitován požadavky zainteresovaných stran, finančním rámcem i praktickým kontextem. Často například není možné provést randomizaci účastníků, eliminovat intervenující proměnné, sebrat dostatečné množství dat nebo provést triangulaci. Některé programy mají příliš málo účastníků nebo se jednotlivá provedení programu příliš liší v závislosti na skupině či jiných faktorech. Další programy připouštějí individuální míru zapojení škol či účastníků a může být obtížné definovat, co znamená jej absolvovat. Jindy nemají některé zainteresované strany (např. učitelé) na evaluaci zájem spolupracovat. Hladinu přijatelného kompromisu je proto třeba stanovovat pro každou situaci zvlášť. V zásadě ale platí, že se evaluátor musí snažit najít v daném kontextu nejlepší možné řešení.
- **Důraz na analýzu a interpretaci** sebraných dat. V těchto fázích evaluačního procesu by evaluátor neměl dělat kompromisy, současně tyto oblasti (spolu s expertním posudkem) kladou na evaluátora největší odborné nároky.
- **Důraz na doporučení k modifikaci programu.** Klíčová součást evaluační zprávy, ve které evaluátor klade výsledky evaluace do kontextu analýzy programu a snaží se navrhnout změny v teorii či implementaci programu.
- **Vymezení role evaluátora.** Namísto původního oddělení výzkumníka a zkoumaného subjektu doporučuje Patton (2002) evaluátorům co nejvíce přímých zkušeností s programem a ve vztahu k lektorům „empatickou neutralitu“. Wiltzová (2006) pro evaluace v prostředí neformálních organizací navrhuje nahradit tradiční roli „tajemného evaluátora“ rolí „kritického přítele“. Robottom (1985) doporučuje pro formativní evaluaci programů environmentální výchovy využívat akční výzkum, ve kterém se evaluátory stávají samotní lektori a učitelé. Při spolupráci se zadavatelem je doporučován participativní přístup (utilization-focused evaluation), ve kterém evaluátor po celou dobu spolupracuje s realizátorem (Rossi, Lipsey & Freeman, 2004). Velmi doporučován je tento přístup i v kontextu evaluace programů environmentální výchovy.

Evaluátor se může stát nejenom hodnotitelem, ale i konzultantem realizátora a vést ho k samostatnému vedení evaluačních výzkumů (empowerment evaluation) (Rossi, Lipsey & Freeman, 2004; Wiltz, 2006). Evaluátor, který je zpravidla expertem na teorii programu a zná znalostní bázi oboru, ve kterém se pohybuje, navíc může se střediskem spolupracovat na modifikacích samotného programu.

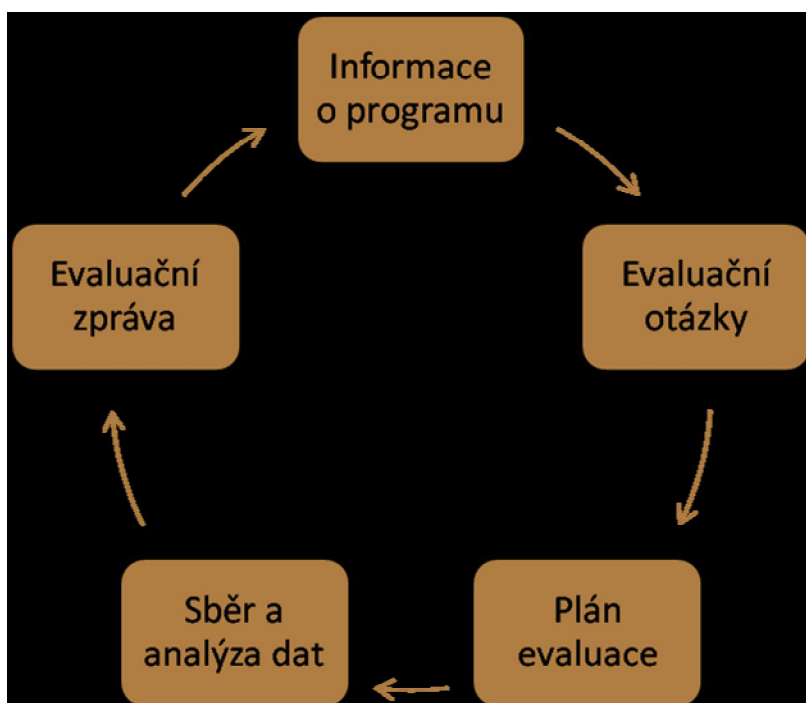
PRAKTICKÉ ASPEKTY EVALUACE

Evaluační výzkumy jsou odborně, časově i finančně náročné. Přestože v důsledku mohou znamenat značnou časovou i finanční úsporu, je zapotřebí pro ně počítat s vysokou vstupní investicí:

- **Evaluátor** by měl ovládat nejenom principy evaluačního výzkumu, ale i tvorby a vedení programu. Současně se musí dobře orientovat v teorii environmentální výchovy a znát příklady domácích i zahraničních evaluačních výzkumů. V České republice je kritický nedostatek evaluátorů. **V současné době proto není reálné evaluační výzkumy požadovat jako závaznou součást grantové žádosti.** Efektivnější strategií může být zvýhodnění programů, které evaluačním výzkumem procházejí či prošly a zejména podpora zvyšování počtu kvalifikovaných evaluátorů.
- Zpracování evaluačního výzkumu trvá přibližně **šest až dvanáct měsíců**, může být ale i delší. Kvalitní evaluace jsou cyklické, tj. výsledky jednoho výzkumu vedou ke změnám programu, které jsou následně ověřovány další evaluací. Evaluace by proto neměly být považovány za jednorázovou záležitost.
- Zpracování evaluace vyžaduje **finanční prostředky** související se sběrem a zpracováním dat i s odbornou prací evaluátora. V zahraničí se cena evaluace odhaduje na 10-15 procent ceny programu. V České republice dosud není ustálený trh. Reálná cena se ale pohybuje v rozmezí 20-70.000, Kč za evaluaci v závislosti na jejím rozsahu a kvalitě.

EVALUAČNÍ PROCES

Proces evaluace můžeme znázornit jako cyklus čtyř základních kroků: (Ecosystem, 2004)



OBRÁZEK Č. 4 PROCES EVALUACE PROGRAMU.

- V první fázi formulujeme cíle evaluace. Součástí této fáze je také shrnutí všeho, co víme o programu včetně případného zpracování jeho **logického modelu**.
- Ve druhé fázi cyklu formulujeme **evaluační otázky** a definujeme základní **indikátory** evaluace.
- Ve třetí fázi určujeme **plán (design) evaluace**, **metody** a nástroje pro sběr dat a postupy pro jejich **vyhodnocení**.
- Ve čtvrté fázi zpracováváme evaluační zprávu a plánujeme další postup. Ten může znamenat změny v realizovaném programu, které bude třeba dále evaluovat, nebo další pokračování evaluace.

Vybranými body procesu se budeme zabývat v rámci následujících kapitol.

FÁZE 1 KONTRAKT A INFORMACE O PROGRAMU

Na počátku evaluačního procesu bývá zpravidla uzavření **kontraktu** mezi evaluátorem a zadavatelem evaluace. V jejím průběhu dochází k dohodě o cílech a podmínkách evaluace. Základní **cíle evaluace** bývají následující:

- Zkvalitnění programu
- Úspory v organizaci
- Prezentace
- Komunikace se sponzory

Vzhledem k malému rozšíření evaluačních výzkumů v České republice nelze předpokládat, že hodnocené programy budou fungovat zcela podle očekávání realizátorů. Evaluátor by proto měl odmítnout kontrakt v případě, že realizátor očekává, že mu výzkum doloží efektivitu hodnoceného programu.

Při vyjednávání o **podmínkách** evaluace se zpravidla diskutuje:

- Dostupné zdroje pro evaluaci
- Časový rámec evaluace
- Role zainteresovaných stran při evaluaci
- Formální výstup evaluace

V současné době se v zahraničí jednoznačně doporučuje zapojovat do evaluace v co největší míře všechny zainteresované strany (evaluace zaměřená na uživatele).

Evaluátor se navíc snaží získat co nejvíce **informací o samotném programu**. V průběhu hledá odpovědi na následující otázky:

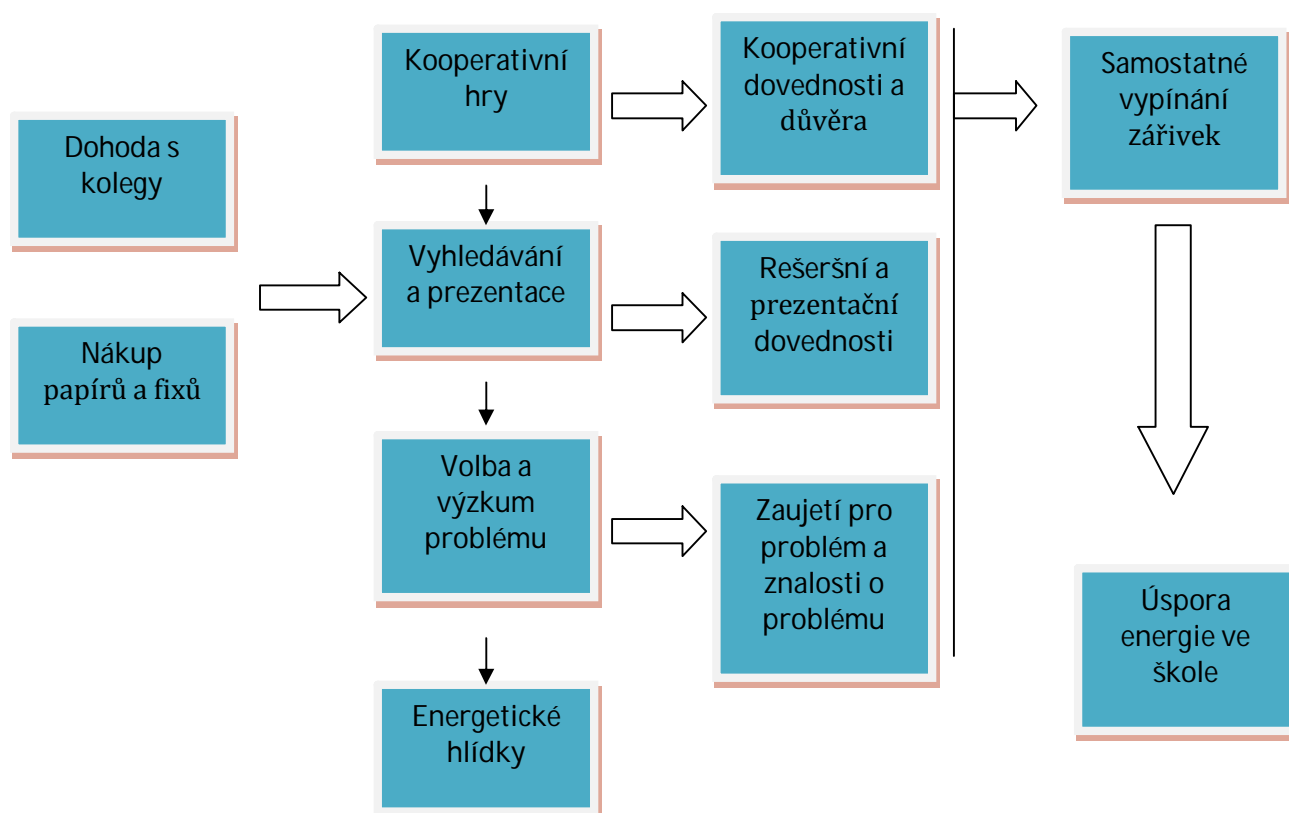
- Jaká je teorie programu?
- Je pravděpodobné, že bude fungovat?
- Má program nějaká slabá místa, která by bylo vhodné změnit před zahájením evaluace nebo na které by bylo dobré se zaměřit při evaluaci?
- Je evaluace proveditelná?

To často předpokládá provedení expertní analýzy programu zahrnující následující body:

- **Paradigma programu** – některé přístupy environmentální výchovy kladou důraz na to, aby program dosahoval předem stanovených **výstupů** a cílů. Existují ale i alternativní proudy, které kladou větší důraz na **proces** výchovy a předpokládají, že hlavní hodnotou jsou

přínosy programu pro jednotlivé účastníky, které se objevují v průběhu programu. Volba dalších kroků evaluačního cyklu by měla tento paradigmatický rámec programu respektovat.

- **Logický model programu** – pro úspěšnou evaluaci je nezbytné porozumět teorii programu. Protože logické modely nejsou dosud v české praxi rozšířeny, musí evaluátor zpravidla logický model ve spolupráci s realizátorem vytvořit.



OBRÁZEK Č. 5 LOGICKÝ MODEL FIKTIVNÍHO PROGRAMU HLÍDAČI SVĚTEL

- **Organizační model (sekvence) programu** – zde si evaluátor ujasňuje sekvenci aktivit v programu.
- **Předchozí evaluace** a zkušenosti s programem.
- Na základě předchozích bodů evaluátor může provést expertní analýzu a vyhodnotit **reálnost teorie účinku programu**. Zde evaluátor čerpá z existující odborné literatury

zabývající se efektivními strategiemi environmentální výchovy či evaluačními výzkumy podobných programů. Na jejím základě může dojít k přesvědčení, že není pravděpodobné, aby program v současné podobě naplňoval svoji teorii. Protože by evaluace byla v takovém případě zbytečná, může evaluátor navrhnout zadavateli změny v programu a přistoupit k evaluaci až po jejich provedení. Typické důvody pro doporučení programu k přepracování jsou:

- Program má příliš mnoho výstupů a je nerealistické, aby jich dosáhl;
 - Výstupy na rovině učení nekorespondují s výstupy na rovině chování;
 - Aktivity programu nekorespondují s výstupy na rovině učení;
 - Program používá strategie, které jsou v odborné literatuře považovány za neefektivní či vyvrácené. Příkladem je například
 - Aplikace K-A-B modelu (program rozvíjí znalosti s předpokladem, že tak ovlivní postoje dětí a v důsledku jejich chování).
 - Orientace programu na obecné uvědomění (které většina cílových skupin má z médií) a moralizující postoje, současně absence konkrétních dovedností, které cílová skupina potřebuje pro dané proenvironmentální chování.
 - Nároky programu se mívají s možnostmi cílové skupiny (např. z hlediska věku, skupinové dynamiky).
 - Program uplatňuje neetické postupy (např. manipuluje dětmi a nutí je do určitého chování, straší je, prezentuje jednostranné informace).
- **Analýza evaluovatelnosti** – zde by měl evaluátor vyhodnotit, zda je možné provést evaluační výzkum programu. Typickými důvody pro neevaluovatelnost může být:
 - Program existuje ve velkém množství různých variant;
 - Program je natolik procesně orientován, že probíhá s každou skupinou jinak;
 - Program nemá ujasněnou teorii a formulovaný logický model;
 - Programem zatím prošlo příliš málo účastníků;
 - Na evaluaci nejsou dostupné potřebné zdroje.

SUPEREKO Program – příklad nerealistické teorie programu

Věk: 3.-9. třída¹⁰

Délka: 4 hodiny¹¹

Cíle:¹²

- Seznámíme děti s hlavními problémy životního prostředí.
- Děti se dozvědí, jak správně třídit odpad.
- Děti získají vztah k přírodě a budou si jí více vážit.¹³
- Děti si zkusí vyrobit ptačí budku.¹⁴

Aktivity:

V průběhu programu si děti

- zasoutěží ve znalostním kvízu¹⁵,
- budou skákat na louce v pytlích¹⁶
- a každé si namaluje jedno zvířátko.

¹⁰ Věková skupina je příliš široká.

¹¹ Délka nerealisticky krátká vzhledem k výstupům.

¹² Ani jeden z uvedených výstupů není SMART.

¹³ V daném časovém rámci není realistické takto ovlivnit postoje cílové skupiny.

¹⁴ Chybí odpovídající aktivita.

¹⁵ Není jasné zaměření testu.

¹⁶ Není jasný vztah aktivity k environmentální výchově.

FÁZE 2: FORMULACE EVALUAČNÍCH OTÁZEK

Ve fázi evaluačních otázek evaluátor ve spolupráci se zadavatelem formuluje otázky, na které bude při evaluaci hledat odpovědi. Evaluační otázky jsou tedy určitým ekvivalentem výzkumných otázek či hypotéz známých z metodiky kvantitativního, respektive kvalitativního výzkumu.

Evaluační otázky se mohou vztahovat k výstupům programu, ale mohou se zaměřovat i na jiné aspekty fungování programu. V zásadě by evaluační otázky měly odpovídat některé z oblastí logického modelu programu.

Přestože ke každému programu je možné najít velké množství evaluačních otázek, doporučuje se, aby evaluátor pomohl jejich počet redukovat na rozumnou míru. Ta se může pohybovat kolem pěti otázek na výzkum.

Příklady evaluačních otázek:

1. Do jaké míry program **rozvíjí znalosti** dětí o globálních problémech?
2. **Chovají** se absolventi programu k životnímu prostředí jinak než děti, které ho neabsolvovaly?
3. Co si děti v průběhu programu **uvědomily nového**?
4. Jakým způsobem interpretují účastníci **přínosy** studia pro svoji profesní dráhu?
5. Které konkrétní aktivity děti nejvíce **ovlivnily**?
6. Jakým způsobem ovlivňuje **skupina** vnímání programu účastníky?
7. Jak program ovlivnil **názory** účastníků na efektivní environmentální výchovu?
8. Jakým způsobem ovlivňují efektivitu studia **prekoncepty** účastníků?

Při formulaci výzkumných otázek identifikuje evaluátor **proměnné** (kategorie), které bude zkoumat. Jejich správná identifikace je velmi důležitá. Pro evaluátora je velmi důležité, aby na základě svých znalostí oboru pomohl proměnné správně formulovat.

Například evaluační otázka č. 2 odkazuje na výstupy programu v rovině chování. Je ale možné, že si realizátor neuvědomuje, že výzkum chování absolventů programu přesahuje jeho finanční a časové možnosti. V takovém případě by bylo vhodnější otázku přeformulovat a na místo „chování“ zkoumat jinou proměnnou, tj. např. „odhodlanost k jednání“. Pro zvýšení vypovídací hodnoty výsledků pak evaluátor může doporučit výzkum souvisejících kognitivních, afektivních či konativních proměnných, tj. např. akčních znalostí a dovedností, přesvědčení o zvládnutí určitých dovedností, ohnisko kontroly atd.

Evaluátor by také s přihlédnutím k zaměření programu mohl navrhnout specifikovat sledované chování pouze na určitou kategorii – např. spotřebitelské chování, ekomanagement atd.

Současně probíhá rozvaha o míře jejich **obecnosti**. Obecněji formulované proměnné (např. odhodlanost k proenvironmentálnímu jednání) umožňují použít při evaluaci už hotové nástroje a porovnat výsledky s jinými výzkumy. Specifičtější formulované proměnné (např. odhodlanost k zateplení domu) zase přesněji odpovídají teorii programu.

Formulace evaluačních otázek představuje základní krok k rozvaze o plánu evaluace. Porovnáme-li například otázku č. 2 a 4 vidíme, že přestože obě otázky zkoumají výstupy programu v rovině chování, jejich formulace předpokládá jinou strategii vedení výzkumu.

FÁZE 3 PLÁN EVALUACE

Při stanovování plánu (designu) evaluace evaluátor vymezuje postup pro sběr a zpracování dat výzkumu. Řeší tedy tyto problémy:

- metodologie výzkumu;
- typ výzkumu;
- volba vzorku;
- plán sběru dat;
- evaluační nástroj;
- plán analýzy dat.

METODOLOGIE VÝZKUMU

Při volbě metodologie výzkumu volíme mezi výzkumem kvantitativním, kvalitativním a smíšeným.

Kvantitativní přístup je charakteristický využíváním numerických dat a statistických metod. Standardní postup kvantitativního pedagogického výzkumu postupuje od stanovení problému k formulaci hypotéz a jejich testování k vyvození závěrů (v případě evaluačních výzkumů se místo hypotéz odpovídá na evaluační otázky). Nejpoužívanější formou jsou dotazníky či testy absolventů. Kvantitativně orientovaná evaluace nám odpoví na otázky typu

- *Do jaké míry bylo dosaženo očekávaných výstupů programu?*

Je vhodná zejména tehdy, pokud máme relativně standardizovaný program, který se opakuje ve více skupinách a umožňuje nám proto sebrat větší množství dat. Programy, pro které se používá kvantitativní evaluace, jsou zpravidla orientovány na výstupy, nikoliv na procesy.

Příklady evaluačních otázek vhodných pro kvantitativní evaluaci:

1. Do jaké míry program **rozvívá znalosti** dětí o globálních problémech?
2. **Chovají** se absolventi programu k životnímu prostředí jinak než děti, které ho neabsolvovaly?

Kvalitativní přístup užívá v evaluaci narativní formy, jako jsou vyjádření pocitů či myšlenek účastníků programu. Kvalitativní výzkumy zpravidla nejsou chápány jako ověřování hypotéz, ale začínají sběrem dat ze zkoumaného prostředí a postupují formulování nových hypotéz či teorií (Patton, 2002; Švaříček & Šed'ová, 2007; Miovský, 2006; Hendl, 2008; Corbin & Strauss, 2008). Data jsou často sbírána pomocí pozorování, rozhovorů či videonahrávek.

Kvalitativně orientovaná evaluace nám např. odpoví na otázky:

- Proč účastníci reagovali při programu tak, jak reagovali?
- Jak účastníci interpretují to, co se dozvěděli?

Kvalitativní přístup je vhodný zejména tehdy, pokud chceme analyzovat program do hloubky, nebo když pracujeme s relativně malými skupinami a program nelze dostatečně standardizovat, aby byl opakovatelný. Kvalitativní přístup je vhodný také tehdy, pokud program používá obecné očekávané výstupy či pracuje s výstupy, které jsou průběžně definovány ve spolupráci s žáky. Programy, které je vhodné vyhodnocovat kvalitativně, se ve své metodice zpravidla soustředí na proces, nikoliv na výstupy. Kvalitativní evaluace nám umožňují pochopit, jaký smysl přisuzují účastníci programu, který absolvují či absolvovali.

Příklady evaluačních otázek:

1. Co si děti v průběhu programu **uvědomily nového**?
2. Jakým způsobem interpretují účastníci **přínosy** studia pro svoji profesní dráhu?
3. Které konkrétní aktivity děti nejvíce **ovlivnily**?
4. Jakým způsobem ovlivňuje **skupina** vnímání programu účastníky?
5. Jak program ovlivnil **názory** účastníků na efektivní environmentální výchovu?
6. Jakým způsobem ovlivňují efektivitu studia **prekoncepty** účastníků?

Je třeba zmínit, že přestože v teorii environmentální výchovy je od devadesátých let patrný trend rostoucího využívání kvalitativních výzkumů (Hart & Nolan, 1999), zaměřuje se většina teoretických i metodických prací o evaluaci především na její kvantitativní formu. Důvodem je možná nedostačující transparentnost výsledků kvalitativního výzkumu pro účely evaluace. Na druhou stranu je patrné, že pro některé metodické přístupy je kvantitativní výzkum obtížně použitelný. Východiskem může být **smíšený design**, který uplatňuje prvky kvantitativní i kvalitativní. Smíšený design je sice náročnější pro evaluátora, ale zpravidla bývá doporučován, protože přináší nejvíce informací a vyvažuje jednostranné nevýhody obou přístupů (Bennett, 1989).

TYP VÝZKUMU

Kvantitativní i kvalitativní metodologie nabízí určité možnosti výběru vhodného typu výzkumu. Jeho volba se odvíjí od evaluačních otázek, na které má přinést odpovědi.

V **kvantitativním** výzkumu se nejčastěji používají následující typy:

- Průzkumy;
- Monitoring výstupů;
- Experimenty a kvazi-experimenty;
- Korelační analýzy.

Průzkumy předpokládají jednorázový sběr dat provedený na definovaném vzorku populace. Jejich využití leží zejména v **analýze potřeb**, například ve výzkumu environmentální gramotnosti cílové skupiny. Důležitým předpokladem je **reprezentativnost** vzorku, který by měl odrážet klíčové vlastnosti základní množiny.

Monitoring výstupů představuje jednoduchý typ výzkumu, ve kterém se měří hodnoty sledované proměnné a porovnávají s určitým **indikátorem** vyjadřujícím žádoucí míru úspěšnosti. **Indikátory** jsou přesná a měřitelná **kritéria** pro vyhodnocení míry dosažení daného výstupu. Definují určitou měřitelnou vlastnost programu a současně podmínku pro hodnocení jejího naplnění, tj. **hladinu úspěšnosti**.

Příklad indikátoru

Očekávaný výstup programu: Po ukončení programu žáci identifikují činnosti, kterými ovlivňují životní prostředí.

Indikátor: 70% žáků na myšlenkové mapě uvede správně alespoň tři různé činnosti, kterými ovlivňuje životní prostředí s uvedením alespoň jednoho environmentálního dopadu každého z nich.

Monitoring výstupů je vhodný zejména pro **jednoduché** evaluace krátkých programů, kdy předpokládáme vstupní úroveň účastníků a není pravděpodobný vliv jiných faktorů. Současně je zřejmé, že monitoring výstupů není vhodný tehdy, pokud

- Neznáme vstupní úroveň cílové skupiny;
- Program je dlouhodobější a nelze vyloučit vliv jiných faktorů na sledované proměnné.

Korelační analýzy

Korelační analýzy se používají pro zjištění, zda existuje vztah mezi dvěma či více soubory dat. Příkladem může být vztah mezi kvantitativně vyjádřenou mírou zapojení účastníka do programu a jeho výsledky v testu.

Korelace může být pozitivní, negativní či žádná. Při pozitivní korelaci ovlivňuje nárůst hodnot na jedné proměnné růst na druhé proměnné, při negativní způsobuje nárůst jedné proměnné pokles hodnot druhé proměnné atd. Síla korelace může být hodnocena jako slabá, střední či žádná a je vyjádřena korelačním koeficientem.

Korelační analýzy mohou být velmi výhodné například tehdy, pokud program připouští různou míru zapojení, program nemá jasné zakončení nebo není možné jasně rozlišit absolventy programu od jiných skupin.

Experimenty a kvazi-experimenty

Experimenty se považují za nejdůležitější typ kvantitativního výzkumu. Experimenty předpokládají **porovnání** dvou či více skupin **náhodně** (randomizovaný výběr) vybraných absolventů programu (experimentální skupina) a respondentů, kteří programem neprošli (kontrolní skupina). Porovnání zpravidla probíhá **opakovaně**, standardně před programem a po něm.

Kvazi-experimenty jsou experimenty, ve kterých jsou zkoumány skupiny, které nebyly vybrány náhodně, ale jinak. Kvazi-experimenty se zpravidla považují za méně spolehlivé, ale akceptovatelné alternativy experimentů.

Kvalitativní evaluační výzkumy mohou uplatňovat širokou paletu přístupů, které se navíc mohou navzájem doplňovat či kombinovat. Mezi nejznámější typy patří zejména:

- Etnografický výzkum;
- Biografický výzkum;
- Případová studie;
- Zakotvená teorie;
- Akční výzkum.

Etnografické výzkumy zkoumají kulturu určité populace. V kontextu evaluace může jít například o fungování a vnitřní pravidla ekologického kroužku, ekotýmu, atd.

Biografické výzkumy se zaměřují na životní dráhu jednotlivce či jednotlivců. V rámci evaluace se mohou zaměřovat například na analýzu dlouhodobého vlivu absolvovaného programu na další život a chování vybraných absolventů.

Případové studie se zaměřují na hloubkovou analýzu určité jednotky či jednotek. Těmi může být například jedno konkrétní provedení programu.

Zakotvená teorie je někdy chápána jako typ výzkumu, ale jde spíše o samostatnou metodologii výzkumu (existující ve více verzích), která může mít kvalitativní i kvantitativní charakter. Jejím cílem je budovat na základě existujících dat teorii, vysvětlující fungování určitého fenoménu. Základem teorie zpravidla je identifikace jedné centrální kategorie a několika dalších klíčových kategorií, které umožňují srozumitelně popsat fenomén. V kontextu evaluace může být využita například pro analýzu toho, jak a proč interpretují účastníci určitý program.

Akční výzkum předpokládá, že evaluaci z významné části vede sám realizátor programu, případně ve spolupráci s externím evaluátorem. Evaluace pak probíhá v cyklu identifikace problému – plán řešení – implementace plánu – vyhodnocení změny.

Smíšené výzkumy kombinují kvantitativní i kvalitativní přístup. Obvyklé je, že v evaluaci převažuje jeden z přístupů a druhý je doplňující. Například evaluace může sbírat kvantitativní data pomocí testu a dotazníku, ke kterým může být připojena otevřená otázka. Ta pak může být vyhodnocena buď v souladu s kvalitativní metodikou, nebo kvantitativně.

Například odpovědi na otázku „Jaký nejsilnější zážitek se Vám s programem spojuje?“ mohou být nejprve **kategorizovány**, tj. rozděleny například na „vztahy ve skupině“, „náhodné incidenty“, „úspěchy“. V každé kategorii pak může být spočítána četnost pojmů v kategorii a případně porovnána s relativní četností pojmů ve stejných kategoriích v kontrolní skupině.

VOLBA VZORKU

Specifickým rysem evaluace programů ve srovnání s jinými výzkumnými oblastmi je, že pracujeme zpravidla s relativně malou základní množinou. Ani u kvantitativně orientovaných evaluací nejsou proto výjimkou vzorky s méně než sto respondenty, evaluace pracující se vzorky v řádech stovek respondentů jsou poměrně vzácné. U programů, které vychází z metodologických paradigmat zdůrazňujících proces a samostatnost žáků, může být každé provedení programu jedinečné a tudíž nesrovnatelné.

Rozhodnutí o výběru vzorku souvisí především se zvolenou metodologií evaluace. Pokud není možné provést **census** – sběr dat od všech absolventů programu, je třeba vybrat určitý **vzorek**. Pro **kvantitativní** evaluace zaměřené na výstupy zpravidla usilujeme o co nejrozsáhlejší vzorek, který by dostatečně reprezentoval absolventy programu. Výběr by měl být **náhodný**. Z tabulky doporučené velikosti vzorku v poměru k počtu absolventů programu vidíme, že pro evaluace programů založené na analýze jedné skupiny (třídy), bychom měli do vzorku zahrnout de facto všechny jeho absolventy (Barch, Duvall, Higgs; Wolske & Zint, 2007):

Počet absolventů	Velikost vzorku	Požadovaný procentuální podíl
10	10	100
20	19	95
50	44	88
100	80	80
250	152	61
500	217	43
1000	278	28
2500	333	13
5000	370	4

OBRAZEK Č. 6 DOPORUČENÁ VELIKOST VZORKU PODLE VELIKOSTI ZÁKLADNÍ MNOŽINY.

V případě, že ověřujeme dopady programu, nebo když uplatňujeme **kvalitativní** formu evaluace, zpravidla pracujeme s malými vzorky, které vybíráme cíleně (**záměrný výběr**). Nejčastěji se využívají některá z těchto kritérií výběru (Patton, 2002; Miovský, 2006; Barch, Duvall, Higgs; Wolske & Zint, 2007; Švaříček & Šed'ová, 2007):

- Extrémní – sledovaná kritéria se u nich objevují v maximální míře;
- Typické – odpovídající průměru;
- Maximálně variantní – co nejodlišnější případy;
- Vhodné – dobře dostupné případy;
- Homogenní – odpovídající jednomu ze zvolených kritérií.

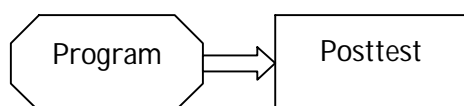
Na rozdíl od kvantitativní evaluace není velikost vzorku předem určena. V zakotvené teorii se doporučuje sbírat data tak dlouho, dokud nedojde k tzv. **teoretickému nasycení** (saturaci) sledovaných kategorií, tj. dokud nová data nepřestanou přinášet nové porozumění sledovaných kategorií.

Součástí každého evaluačního procesu by mělo být nastavení etických kritérií. Jeho základem je především anonymizování výpovědí respondentů v evaluační zprávě, případně dohoda mezi evaluátorem a respondenty o způsobu nakládání s daty. Zejména kvalitativní evaluace při tom vyžaduje vysokou citlivost. Ve Spojených státech je například podmínkou získat od respondentů tzv. **informovaný souhlas** (Barch, Duvall, Higgs; Wolske & Zint, 2007).

PLÁN SBĚRU DAT

Plán sběru dat vychází ze zvolené metodologie výzkumu a z možností (zejména finančních) evaluace. **Kvantitativní** evaluace uplatňují lineární postup, ve kterém se data sbírají v různých kombinacích před, při či po programu. Můžeme rozlišit následující typy plánů sběru dat (které se mohou dále vzájemně kombinovat) (Bennett, 1989; Barch, Duvall, Higgs, Wolske & Zint, 2007):

Program – Posttest



OBRÁZEK Č. 7 EVALUAČNÍ DESIGN PROGRAM-POSTTEST.

Velmi jednoduchý design vhodný pro krátké programy a v situacích, kdy omezené zdroje a časové možnosti neumožňují zvolit jinou možnost. Nevýhodou je zejména to, že nemůžeme ověřit, zda se hodnoty naměřené posttestem liší od předchozích znalostí, postojů, hodnot a

dovedností účastníků, ani zda (v případě delších programů) nemohla být změna vyvolána jiným faktorem.

Design je často realizován tak, že jako posttest je využit úkol či aktivita na konci programu. Využívá se zde poslední fáze procesu učení, kterou může být v závislosti na využitém modelu učení (Činčera, 2007b) aplikace nově naučeného konceptu (např. model IAA), reflektivní přeformulování původního konceptu (model EUR) nebo debriefing. Výhodou je nenásilné zapracování evaluace do toku programu. Slabinou je naopak určité smísení procesu učení s evaluací. Lze předpokládat, že např. při prezentaci nového porozumění problematice (v EUR modelu) dochází ve skupině stále ještě k ujasňování nově konstruovaného významu a monitorované výsledky nemusí proto přesně odrážet skutečný cílový stav.

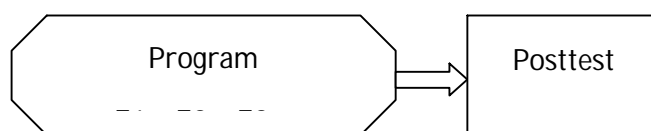
Určitou alternativou k program – posttestu designu je varianta, kdy v posttestu žádáme účastníky, aby sami hodnotili svůj posun oproti stavu před programem. V dotaznících se k tomu někdy používají doplňovací tvrzení typu

- *Naučil jsem se ...*
- *Nové pro mě bylo,*

Pro některé účastníky ale může být takové vyjádření obtížné, často se objevují odpovědi:

- *Asi nic, všechno už jsem znala. Jenom jsem si uvědomila, že...*

Program s formativním testováním a posttestem

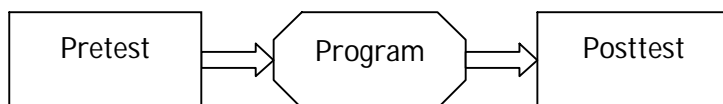


OBRÁZEK Č. 8 EVALUAČNÍ DESIGN S FORMATIVNÍM TESTOVÁNÍM A POSTTESTEM.

Formativní evaluace je výhodná zejména pro delší programy, protože nám umožňuje zaznamenávat postupný pokrok účastníků programu. V navrženém modelu budou pravděpodobně špatně odlišitelné ostatní faktory nesouvisející s programem. Může být vhodné

kombinovat kvalitativní metody měření ve formativní evaluaci s kvantitativními v sumativní, tj. např. v průběhu pracovat s pozorováním, rozhovory, reflexemi či měřením spokojenosti účastníků a po ukončení programu zvolit určitou formu dotazníku či testu.

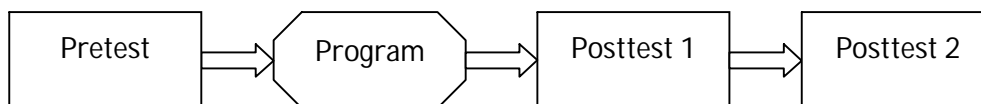
Pretest-program-posttest



OBRÁZEK Č. 9 EVALUAČNÍ DESIGN PRETEST- PROGRAM-POSTTEST.

Varianta umožňující jednoduše vyhodnotit posun účastníků v měřených znalostech, postojích či dovednostech. Stejně jako v předešlém modelu není ale ani zde zřejmé, zda za nalezeným rozdílem nemohou být nějaké jiné příčiny. Pokud uplatňujeme kvalitativní postupy, mohou účastníci v posttestu paradoxně vyjádřit horší míru sebehodnocení. To proto, že se účastníci před programem často z důvodu neznalosti tématiky nadhodnocují a teprve v jeho průběhu si uvědomují zjednodušenost svých prekonceptů. V takovém případě posttestem de facto zjišťujeme, zda došlo k narušení původních prekonceptů a k důležité zpochybnující fázi procesu učení.

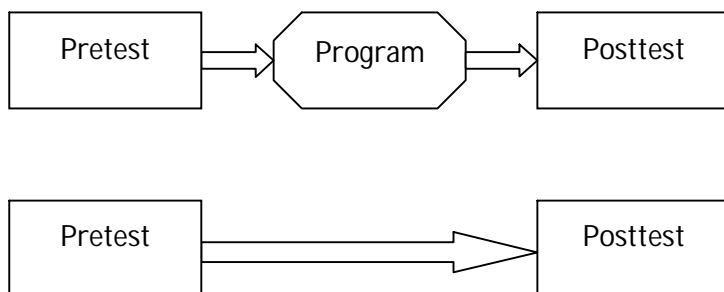
Pretest – program – posttest 1 – posttest 2



OBRÁZEK Č. 10 EVALUAČNÍ DESIGN PRETEST- PROGRAM-POSTTEST –POSTTEST 2.

Vhodná varianta pro měření střednědobých výstupů a dlouhodobých dopadů programu. Důležitá zejména tehdy, pokud klademe důraz na změny v chování účastníků programu nebo na dopady programu.

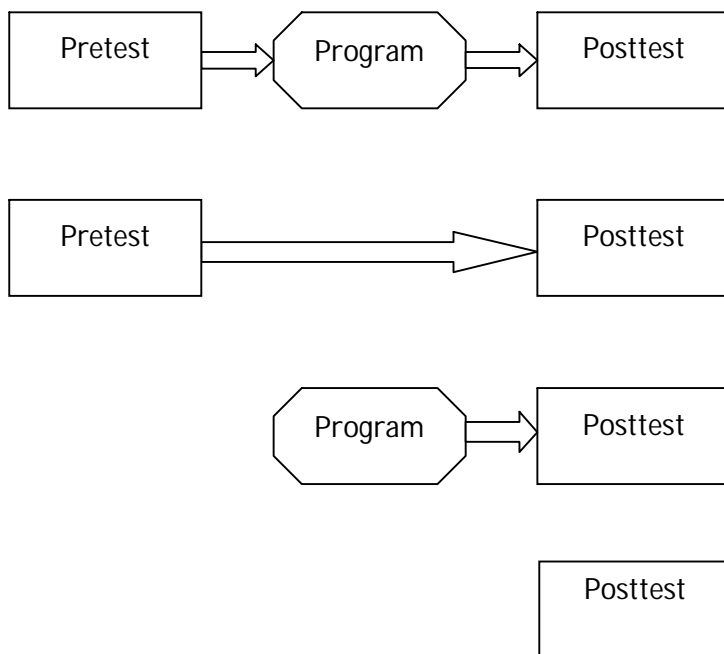
Srovnání s kontrolní skupinou



OBRÁZEK Č. 11 EVALUAČNÍ DESIGN EXPERIMENTÁLNÍ SKUPINA – KONTROLNÍ SKUPINA.

Srovnání účastníků programu a srovnatelné skupiny, která programem neprošla, je metoda umožňující dobře vyhodnotit efektivitu programu a současně eliminovat riziko vedlejších faktorů nezahrnutých do evaluace. Předpokladem je, že máme možnost najít skupinu, která je ve všech důležitých parametrech srovnatelná (např. věk, počet, škola atd.). Tento typ představuje určitý „zlatý evaluační standard“ a většina autorů jej proto doporučuje.

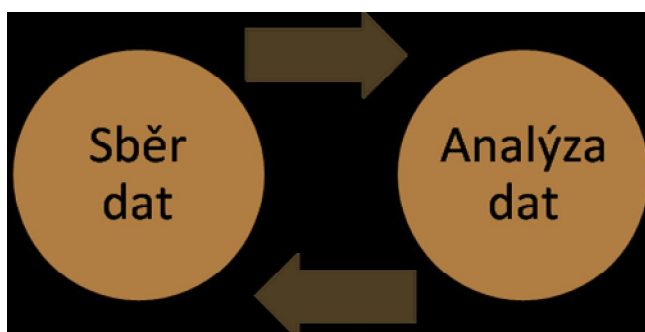
Šalamounův test čtyř skupin



OBRÁZEK Č. 12 EVALUAČNÍ DESIGN ŠALAMOUNŮV TEST ČTYŘ SKUPIN.

Tento typ testu doporučuje pro evaluace environmentální výchovy Smith-Sebasto (2005). Je výhodný zejména tehdy, pokud můžeme očekávat, že výsledky posttestu testovaných skupin budou ovlivněny předchozím pretestem. Program prokáže svoji efektivitu tehdy, pokud obě experimentální skupiny projeví statisticky významné rozdíly v posttestu oproti pretestu, výsledky přitom budou srovnatelné. Současně budou srovnatelné výsledky obou kontrolních skupin.

Kvalitativní evaluace zpravidla uplatňují cyklický sběr dat. Evaluátor data průběžně sbírá a vyhodnocuje. Na základě průběžné analýzy modifikuje nástroj pro sběr dat (např. formuluje nové otázky) a navrhuje další respondenty. Sběr probíhá tak dlouho, dokud nedojde k teoretické saturaci.



OBRÁZEK Č. 13 CYKlickÝ PLÁN SBĚRU DAT V KVALITATIVNÍM VÝZKUMU

NÁSTROJE PRO SBĚR DAT

Výběr vhodných nástrojů pro sběr dat je další fází evaluačního procesu. Zvolené nástroje by měly nejenom odpovídat vybraným indikátorům a evaluačním otázkám, ale měly by splňovat i další kritéria: měly by být finančně i technicky dostupné, adekvátní pro danou kategorii respondentů, **validní** (měřící příslušný očekávaný výstup) a **reliabilní** (spolehlivé a přesné) (Chrástka, 2007; Bennett, 1989).

Přestože existuje několik druhů validity, v kontextu evaluačních výzkumů hraje největší roli tzv. **obsahová validita**, ve které jsou vlastnosti nástroje propojeny se sledovanými proměnnými programu jeho realizátory a tvůrci. U určení **reliability** testu se zpravidla vyhodnocuje vnitřní integrita (tzv. Cronbach alfa test) a test-retest.

V kvalitativní evaluaci se pak zpravidla usiluje o tzv. **důvěryhodnost** dat. Obecně platným doporučením je usilovat o **triangulaci** – sběr dat několika způsoby a analýzu více metodami či hodnotiteli.

V této kapitole se nebudeme zabývat popisem široké škály nástrojů pro sběr dat, které se používají v pedagogickém výzkumu, ale zaměříme se pouze na nástroje nejčastěji používané v evaluaci programů environmentální výchovy.

Písemné dotazníky a studentské zprávy

Používají se nejčastěji pro měření postojů, hodnot, výpovědí respondentů o jejich chování. Přestože se v nich nejčastěji používají uzavřené otázky (např. různé škály) vhodné ke kvantitativnímu zpracování, zpravidla se doplňují i otázkami otevřenými, jejichž zpracování může být částečně kvalitativní. Výhodou dotazníků je snadná administrace, nevýhodou náročnost jejich vytvoření. Další nevýhodou je vyšší časová i finanční náročnost evaluace. Nevýhodou může být také určité zkreslení výpovědí (Bennett, 1989; Stokking, van Aert, Meijberg, & Kaskens, 1999; Simmons, 2008).

Výhoda dotazníků je i v možnosti využít **standardizované** nástroje. Mezi nejznámějšími nástroje patří:

- Nové ekologické paradigma (NEP) určené na měření obecných proenvironmentálních postojů;
- The Middle School Environmental Literacy Instrument (MSELS), uplatňovaný ve Spojených státech pro měření environmentální gramotnosti žáků;
- The Children's Attitudes Toward the Environment Scale (CATES), určený na měření environmentálních postojů malých dětí;
- Odhodlání k proenvironmentálnímu jednání (PBS), určená na měření odhodlání k měření politických a právních akcí a další.

Typem dotazníku je tzv. **studentská zpráva** (learner report). Studentské zprávy se zpravidla používají bezprostředně po ukončení programu nebo jeho části. Jejich cílem je zjistit, jak účastník hodnotí svoji zkušenost s programem a co se domnívá, že se v rámci programu naučil. Jejich využití je tedy zejména na reaktivní úrovni evaluace. Využívají buď uzavřené (výběrové), otevřené (*Co jsi se v programu naučil?*) nebo semi-otevřené otázky (*Naučil jsem se, že..., Tři nové věci, které jsem si uvědomil, jsou...*) (Stokking, van Aert, Meijberg, & Kaskens, 1999; Cachelin, Paisley & Blanchard, 2009).

Testy

Testy můžeme rozdělit na *písemné vědomostní a dovednostní* testy. Písemné vědomostní testy se používají k měření znalostí a kognitivních dovedností. Výhody i nevýhody jsou obdobné, jako u písemných dotazníků. Testy a dotazníky mohou být administrovány obdobným způsobem, dotazník někdy může obsahovat testové otázky. Typické příklady otázek jsou otázky s vícečetným výběrem (multiple choice), otázky pravda / nepravda, doplňovací otázky, přiřazování a další. Jiným typem jsou otázky založené na analýze textu nebo problémové úkoly (Bennett, 1989; Stokking, van Aert, Meijberg, & Kaskens, 1999; Simmons, 2008; Marcinkowski, 1997).

Dovednostní testy se používají pro evaluaci praktických dovedností, tj. v kontextu environmentální výchovy např. měření, zpracovávání vzorků, orientace v terénu (Bennett, 1989).

Rozhovory a ohniskové skupiny

Rozhovory i ohniskové skupiny jsou typicky kvalitativní nástroje, jejichž cílem je porozumět pocitům, prožitkům či hloubce porozumění účastníka. Zatímco rozhovory se vedou s jednotlivými účastníky, ohniskové skupiny jsou založeny na řízeném rozhovoru mezi malou skupinou vybraných účastníků programu. V obou případech je nezbytností úplný záznam, který potom bude dále zpracováván (Patton, 2002; Simmons, 2008; Frechtling, 2002).

Rozhovory můžeme rozdělit do několika typů:

- Nestrukturované, ve kterých tazatel nemá předem připravené otázky a vede rozhovor přirozeně;
- Polostrukturované, resp. rozhovor s návodem, kdy má tazatel připravenou strukturu rozhovoru, ale pružně jí upravuje;
- Strukturované, kdy tazatel klade předem připravené otázky v daném pořadí a v přesných formulacích.

V rozhovorech je třeba se zejména vyhnout otázkám, které jsou

- Manipulativní
- Uzavřené
- Složené z více částí.

Pozorování

Cenná technika zejména pro formativní hodnocení nebo pro měření dovedností. Bennett (1989) rozlišuje mezi *systematickým* pozorováním, založeným na práci s připraveným

seznamem kritérií, které pozorovatel hodnotí (tzv. **zcitlivující koncepty**), a *informálním* pozorováním, ve kterém si pozorovatel poznamenává zaznamenané postřehy.¹⁷ Alternativou k písemnému zaznamenávání pozorovaného může být použití videozáznamu nebo zvukové nahrávky. Systematické (standardizované) pozorování je zpravidla založeno na principu *ratingu*, kde pozorovatel hodnotí míru pozorovaných jevů do předem připraveného formuláře přiřazováním určitých kvantitativních jevů (Chráska, 2007).

Pozorování je při evaluaci zpravidla tzv. *zúčastněné* (Švaříček & Šed'ová, 2007) Při plném zúčastněném pozorování je pozorovatel přímo součástí dění a je v plné interakci s účastníky.¹⁸ Častější je ale situace, kdy se pozorovatel neúčastní přímo aktivit a sleduje chování skupiny.

V praxi jsou často jako výstupy z pozorování využívány postřehy pedagoga. To je ale problematická metoda, protože pro vedoucího programu je velmi náročné, aby současně program vedl, monitoroval a dělal zápis z pozorování. Praxe ukazuje, že postřehy vedoucího programu a pozorovatele jsou často rozdílné. Je proto velmi vhodné, aby pozorovatelem byl někdo jiný než pedagog, který program vede.

Nevýhodou pozorování také je, že přítomnost pozorovatele může ovlivnit chování skupiny (Simmons, 2008). To platí zejména tehdy, pokud je z pozorování pořizován filmový záznam.

Analýza artefaktů, průběžné písemné reflexe

V rámci programů environmentální výchovy studenti často vyrábí různé artefakty, které demonstrují určitý stupeň porozumění problematice, zvládnutí kognitivních či manuálních dovedností či osvojení určitých postojů.

Artefakty mohou být například postery, které účastníci prezentují v závěru programu, elektronické prezentace, nástěnky, studentské noviny či články v nich, výtvarná díla, slohové práce, eseje či volný text na dané téma, zkrášlený školní pozemek, vyrobené cedule naučné stezky atd.

Zejména v rámci delších programů jsou účastníci často požádáni o vedení určité formy samostatných písemných reflexí. Ty se zpravidla označují jako **logbook** či **portfolio**.

¹⁷ Chráska (2007) rozlišuje jako pozorování standardizované a nestandardizované, případně Švaříček & Šed'ová (2007) mluví o pozorování strukturovaném a nestrukturovaném.

¹⁸ O nezúčastněném pozorovateli by skupina neměla vědět – může být např. skrytě snímána kamerou.

Logbook¹⁹ je strukturovaný či nestrukturovaný sešit, do kterého si účastník zapisuje, jaké programy absolvoval a reflektuje je (Stokking, van Aert, Meijberg, & Kaskens, 1999).

Portfolia mohou mít více typů, ale zpravidla se chápou jako složka materiálů, vytvářená studentem, dokumentující jeho posun v oblasti, nejlepší práce, proces řešení projektu, atd. (Jančaříková, 2007).

Portfolia i logbook umožňují sledování dlouhodobého vývoje žáka a jejich význam leží především ve využití pro kvalitativní evaluaci.

Pro krátké programy může podobnou úlohu hrát závěrečná *písemná zpětná vazba*, do které účastníci formou volného textu píšou své pocity a objevy z programu. Zpětné vazby by pak měly být analyzovány kvalitativním způsobem na principu kategorizace.

Pro využití v evaluaci je třeba připravit seznam tzv. hodnotící tabulku (rubric), obsahující hodnocených kritérií (např. prezentovaných dovedností) s hodnotící škálou, definovanou jednotlivými indikátory (Bennett, 1989).

Příklad hodnotící tabulky pro program zaměřený na přípravu žádostí o grant

Kategorie / indikátor	Nedostačující	Slabé	Průměrné	Dobře zvládnuté	Výborné
Logická stavba projektového záměru	Žádost o grant obsahuje jen vágní vymezení cílů projektu a prostředků k jejich naplnění	Žádost o grant obsahuje konkrétní vymezení prostředků, ale není jejich jasný jejich vztah k cílům / cíle jsou vágní a nekonkrétní	Žádost o grant má konkrétní cíle i prostředky, nezabývá se ale metodami evaluace ani indikátory	Žádost o grant obsahuje logický rámec projektu neobsahující některé náležitosti – např. metody měření či s drobnými chybami ve formulacích cílů a vazbami mezi nimi.	Žádost o grant obsahuje korektně zpracovaný logický rámec projektu zachycujícího cíle a prostředky na různých úrovních, indikátory i metody ověřování.

Závěrečná diskuse

¹⁹ Mareš a Gavora (1999) překládají jako „deník“.

V rámci programů se často používají aktivity, které umožňují účastníkům přímo se vyjadřovat k programu. U programů ovlivněných metodikou Prázdninové školy Lipnice (Franc, Zounková & Martin, 2007) je obvyklé bodování programu podle spokojenosti účastníků či jiných kritérií, tzv. *hitace*. Technika je tedy určena pro hodnocení na reaktivní rovině.

Jiné metody mohou souviset s uplatněním Kolbova cyklu, resp. cyklu učení prožitkem (Kolb, 1984; Henton, 1996). Podle tohoto modelu učení je účastník veden od konkrétní zkušenosti k jejímu reflektivnímu a zobecňujícímu zpracování až k plánování dalšího postupu. Konkrétní techniky procesu zpracovávání zkušenosti, tzv. *debriefingu*²⁰ je možné využít i pro evaluaci získaných znalostí, dovedností či postojů z celého programu. (Činčera, 2007b; Reitmayerová & Broumovská, 2007). Např. při technice 4F účastníci postupně formulují, jaká fakta (Facts) a emoce (Feelings) se v programu objevily, co pro ně bylo nové (Findings) a jaké změny na základě programu plánují udělat (Future).

Debriefingových technik je velké množství a není možné se zde jimi podrobně zabývat. Je ale třeba si uvědomit, že debriefing je součástí cyklu učení, protože pomáhá účastníkům formulovat nové porozumění a plán příštích kroků. Ve využití pro evaluaci programu je nezbytné pokusit se postřehy účastníků z této fáze zaznamenávat, např. metodou pozorování.

Pro evaluaci pak je dále možné využít diskusní aktivity, ve kterých studenti vyjadřují své postoje. Příkladem jsou u nás velmi známé aktivity krok vpřed či lidské kontinuum, při kterém se účastníci staví na pomyslnou čáru vyjadřující míru jejich souhlasu či nesouhlasu s prezentovaným tvrzením (Thomson & Hoffman, 2008). V modelu EUR (Činčera, 2007), který u nás používá v programech několik středisek ekologické výchovy, může být tato aktivita se stejnými otázkami použita v evokační a v reflektivní fázi. Tím máme šanci měřit případný posun účastníků na čáře. Předpokladem je ale opět možnost prvotní a konečné postavení studenta zaznamenat.

Další alternativní nástroje

Zajímavou evaluační technikou vhodnou například na měření postojů je využití **řízené imaginace** žáků. Metoda byla vyzkoušena například pro evaluaci programu zaměřeného na problematiku mokřadů. Žáci si představovali cestu mokřadem a pak zaznamenávali své pocity. Výzkumníci ale narazili na obtíže v rozlišení mezi prožitkem imaginárního a skutečného mokřadu (Cachelin, Paisley & Blanchard, 2009). V České republice tuto metodu použil například Krajhanzl (2007).

²⁰ Reitmayerová a Broumovská (2007) navrhují pojem „cílená zpětná vazba“. Dále se používá např. *processing* (Henton, 1996), reflexe, rozbor či review. Každý z uvedených pojmů má své obhájce i kritiky.

Pojmové mapy jako evaluační metody doporučuje například Simmonsová (2008). Andrews, Tressler a Mintzes (2008) využili pojmové mapy pro evaluaci pobytového programu zaměřeného na mořské ekosystémy. Při analýze se hodnotily rozdíly v pretestu a posttestu a mezi experimentální a kontrolní skupinou ve strukturální komplexitě (např. počet vazeb, prvků) a obsahové validitě mapy.

Pro analýzu postojů či například environmentální senzitivity mohou být využity například **obrázky** zadané na jednotné téma. I přes vysoké nároky na analýzu mohou být vhodným nástrojem pro sběr dat od malých dětí. Metodu v českém kontextu vyzkoušeli například Činčera a Komárková (2010).

Určitým typem pozorování jsou metody, při kterých je zaznamenáváno chování studentů mimo program, tzv. **neobtěžující metody**. Příkladem může být např. evidence výpůjček knih o environmentální tematice ve školní knihovně, přehled zboží zakoupeného ve školním bufetu, účast na nepovinném kroužku či akci. Některé situace, vyžadující po studentech samostatné rozhodování, mohou být učitelem předem připraveny –např. na trávník, kolem kterého děti půjdou, může učitel umístit odpadky a pak vyhodnocovat, zda je děti samy od sebe začnou sbírat (Bennett, 1989).

FÁZE 4 – SBĚR A ANALÝZA DAT

Sběr dat se v kvantitativní evaluaci označuje jako **administrace**, v kvalitativní se mluví o **tvorbě dat**. Při administraci je důležité zejména zajistit stejné podmínky všem skupinám, kterým je zadáván test, vysvětlit jeho důvody a zabezpečit anonymitu. Při pořizování rozhovorů doporučuje většina autorů (s výjimkou Glasera, 1998) pořizování nahrávek, které jsou potom přepsány (**transkripce**).

Po naměření či získání dat pro evaluaci je třeba je vhodným způsobem zpracovat. Zpracování dat pro **kvantitativní** design vychází ze statistických metod, zatímco v kvalitativním výzkumu zpravidla analyzujeme texty, kategorizujeme a interpretujeme.

V kvantitativních evaluacích se používají dva druhy statistiky: **deskriptivní** a **inferenční** (Benett, 1989; Barch, Duvall, Higgs; Wolske & Zint, 2007).

Deskriptivní statistika se zabývá popisem dat. Ta mohou mít charakter buď

- *nominální* (také *kategorický*), ve kterých čísla nemají kvantitativní význam a jsou používána jako „nálepky“ pro určité charakteristiky;
- *ordinální* (pořadový), ve kterých čísla vyjadřují pořadí určitého kritéria;
- *metrický* či *intervalový*, kdy čísla vyjadřují rozdíly mezi objekty (Chráška, 2007).

V rámci deskriptivní statistiky zpravidla vyhodnocujeme *četnost* a *modus*²¹ (pro nominální data), pro metrická data dále *aritmetický průměr*, *median*²² a *směrodatnou (také standardní) odchylku*²³.

Deskriptivní statistika může být v některých případech pro analýzu dat dostačující (monitoring výstupů). To platí zejména tehdy, pokud jsme specifické očekávané výstupy programu formulovali tak, že v sobě zahrnují procentuálně vyjádřený minimální stupeň úspěšnosti. Takovému kritériu se říká také **hladina úspěšnosti** (Department of Environment and Conservation, 2004). Vrátime-li se k již uvedenému příkladu specifického výstupu, vidíme, že k jeho ověření nám stačí prostý aritmetický průměr:

²¹ Modus udává nejčastěji se vyskytující hodnotu.

²² Median udává tzv. střední bod – hodnotu, která je ve vzestupné řadě naměřených hodnot přesně v polovině.

²³ Směrodatná odchylka (standard deviation, SD) udává míru rozptylu naměřených hodnot. Čím více se hodnoty blíží průměru, tím je směrodatná odchylka menší.

- Pokud skupina obdrží set kartiček s obrázky environmentálních problémů a kartiček s obrázky lidí, kteří se podílejí na jejich řešení, přiřadí 70 procent studentů správně k sobě 75 procent kartiček (Bennett, 1989).

Pokud používáme evaluační design typu pretest – program – posttest nebo srovnání s kontrolní skupinou, budeme potřebovat nástroje inferenční statistiky pro porovnání výsledků před programem a po něm a mezi experimentální a kontrolní skupinou.

Nejčastěji porovnávané, s jakou pravděpodobností mohly naměřené rozdíly mezi průměry či četnostmi obou či více skupin vzniknout náhodně (a jsou-li tedy rozdíly tzv. **statisticky významné**) a zda jsou podstatné či zanedbatelné (tzv. **velikost efektu**, effect size, také **praktická významnost**).

Při vyhodnocování mohou nastat **chyby** dvou typů: může se stát, že neoprávněně odmítneme nulovou hypotézu ve prospěch alternativní (mylně tvrdíme, že mezi oběma skupinami je rozdíl), nebo že naopak neoprávněně odmítneme hypotézu alternativní ve prospěch nulové (mylně tvrdíme, že mezi skupinami není rozdíl).

Protože ve statistických výzkumech nemáme jistoty, ale pouze větší či menší míru pravděpodobnosti, musíme nejprve určit tzv. **hladinu významnosti (α)**, která vyjadřuje maximální přípustnou míru rizika, že neoprávněně odmítneme nulovou hypotézu. V evaluačních výzkumech se zpravidla hladina významnosti stanovuje na $\alpha=0,05$ (pravděpodobnost, že omylem přijmeme alternativní hypotézu je maximálně pět procent), případně $\alpha=0,01$ (Chráska, 2007; Barch, Duvall, Higgs; Wolske & Zint, 2007).

Před samotným výzkumem či zpětně po fázi sběru dat je možné určit ještě tzv. **sílu testu (power analysis, β)**, která určuje pravděpodobnost vzniku chyby druhého typu (mylné odmítnutí alternativní hypotézy ve prospěch nulové) při dané hladině α , velikosti výzkumného vzorku a použité technice zpracování dat. Za vhodnou hodnotu se považuje $\beta>0,8$. Zvýšit sílu testu je možné použitím jiné metody zpracování dat, ale zejména zvětšením vzorku (Barch, Duvall, Higgs; Wolske & Zint, 2007).

Pro vyhodnocení míry statistické významnosti rozdílů mezi skupinami se při evaluaci programů používá několik typů testů v závislosti na počtu testovaných skupin a charakteru měřených dat (Hendl, 2004):

- Pro porovnávání procentuální míry četnosti dvou či více skupin nominálních dat se nejčastěji používá *chi kvadrát*, *kontingenční tabulka* nebo *Fisherův test*.

- Pro práci s pořadovými daty se často používá *Wilcoxonův test* nebo *Mann-Whitneyho test*.
- Pro srovnání průměrů dvou skupin intervalových dat se nejčastěji používá *t-test*.²⁴
- Porovnááme-li více skupin, než dvě, používá se zpravidla *ANOVA* či *ANCOVA*.²⁵
- Pro korelační analýzu se zpravidla používá *Spearmanův* (pro ordinální data) či *Pearsonův* (pro intervalová data) korelační koeficient.

Výstupem testů statistické významnosti je vypočtená hodnota **pravděpodobnosti (p)**, že rozdíly mezi skupinami mohly vzniknout náhodně. Je-li hodnota p rovna či nižší než zvolená hladina statistické významnosti, můžeme naměřený rozdíl považovat za signifikantní.

Testy statistické významnosti ale nehodnotí, jakou váhu statisticky významný rozdíl má. Tu měří tzv. **velikost efektu**. Velikost efektu se počítá jako rozdíl průměrů obou skupin dělený směrodatnou odchylkou kontrolní skupiny. Výsledná hodnota se zpravidla interpretuje podle rozmezí (Barch, Duvall, Higgs; Wolske & Zint, 2007):

- <0.1 – triviální efekt
- 0,1-0,3 – malý efekt
- 0,3-0,5 – střední efekt
- >0.5 – velký efekt.

Pro všechny uvedené funkce je možné na internetu najít řadu volně přístupných kalkulátorů.

Pro zpracování **kvalitativních dat** v evaluaci programů environmentální výchovy existuje více doporučených postupů. Některé jsou založeny na převodu textu do kvantifikovatelné podoby, jiné se snaží získaná data uspořádat a interpretovat.

Stokking et. al. (Stokking, van Aert, Meijberg, & Kaskens, 1999) doporučuje odpovědi na otevřené otázky **kategorizovat**, tj. rozdělit do cca 5-8 kategorií a v nich pak kvantifikovat četnost. Další postup pak může být analogický, jako u zpracování kvantitativních dat.

²⁴ T-test i některé další testy mohou být tzv. jednostranné či oboustranné. Oboustranné testy se používají tehdy, pokud alternativní hypotéza byla formulována tak, že výsledky experimentální skupiny mohou být stejné, vyšší i nižší, než skupiny kontrolní. Jednostranné testy předpokládají, že platit může pouze jedna z možností. Při evaluaci programů je vhodnější uvažovat raději o oboustranných testech. T-test dále může být párový a nepárový. Nepárový test užíváme pro porovnání dvou různých skupin, tj. např. experimentální a kontrolní. Párový test se užívá při opakovaném měření na stejné skupině, tj. např. pretest-posttest. (Chráška, 2007)

²⁵ ANOVA na porovnání rozptylu mezi skupinami s rozptylem uvnitř skupin. ANCOVA se používá při vícenásobném měření.

Pro hodnocení testů kognitivních znalostí či esejů je vhodné používat předem připravenou **hodnotící tabulku (rubric)**, podle kterých jsou pracím přiděleny body, které pak mohou být dále porovnávány.

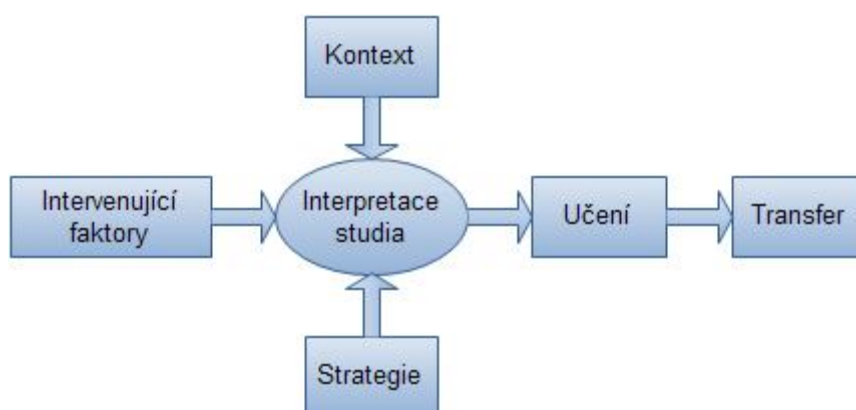
Pro kvalitativní výzkum Barch et. al. (2007) uvádí následující doporučený postup:

1. Redukce dat – výběr významových celků, korespondujících s evaluačními otázkami;
2. Zobrazení dat – v takové organizaci či uspořádání, které umožňuje identifikovat společné motivy, struktury a vazby;
3. Shrnutí závěrů a ověření.

Při obsahové analýze se doporučuje všimnout si především opakujících se motivů, vzorců a toho, jak se vztahují k evaluačním otázkám, odlišnostem, zajímavostem, shodám či rozdílům oproti literatuře, atd.

Společným jmenovatelem metod je opět kategorizace textu. Kategorie mohou být vytvořeny před zpracováváním textu nebo mohou být vytvářeny postupně (emerging categories) (Taylor-Powell & Renner, 2003).

Často se využívá prvků tzv. **zakotvené teorie** (Glasser, 1998; Corbin & Strauss, 2008), ve které se k významově nosným částem textu (**významové jednotky**) nejprve v procesu tzv. **kódování** přiřazují tématicky odpovídající **koncepty**, ty jsou dále shlukovány do obecnějších **kategorií**, mezi kterými jsou nakonec hledány vztahy a souvislosti, vedoucí k formulování nových teorií. Z kategorií je pak vybrána jedna **centrální**, která umožňuje interpretovat vztahy mezi ostatními kategoriemi.



OBRÁZEK Č. 14 PŘÍKLAD VZTAHŮ MEZI KATEGORIEMI PŘI ANALÝZE PROGRAMU POMOCÍ ZAKOTVENÉ TEORIE

FÁZE 5 EVALUAČNÍ ZPRÁVA

Evaluační zpráva vzniká jako výstup práce evaluátora. Zpravidla má následující strukturu:

1. **Abstrakt** – obsahuje krátké shrnutí cílů, metodiky a výsledků evaluace;
2. **Základní cíle, oblast a typ evaluace;**
3. **Popis programu**, včetně popisu teorie programu, logického modelu, případně vstupní expertní analýza programu a výsledky předchozích evaluací programu;
4. **Evaluační otázky;**
5. **Plán evaluace**, zahrnující volbu a odůvodnění evaluační metodologie,
6. **Charakteristika vzorku;**
7. **Metody sběru dat;**
8. **Prezentace výsledků** – zpravidla členěné podle evaluačních otázek;
9. **Interpretace, diskuse** – uvedení výsledků do kontextu jiných evaluačních výzkumů a existující teorie, otázky evokované výsledky evaluačního výzkumu, možná vysvětlení zjištěných jevů, limity evaluace;
10. **Doporučení** – k modifikaci programu i pro další evaluace;
11. **Závěry** – stručné shrnutí hlavních výsledků evaluace;
12. **Přílohy** – mohou obsahovat použité nástroje, podrobná naměřená data, příklady transkripce rozhovorů.

Je zřejmé, že body 9 a 10 kladou velké nároky na znalosti evaluátora v oboru environmentální výchovy a metodiky tvorby programů. Evaluátor zde v zásadě analyzuje příčiny neúspěšnosti programu ve sledovaných kategoriích. Ty mohou být následující:

- Chyba v měření. Evaluátor by měl vyjasnit **limity** evaluačního výzkumu a stanovit, do jaké míry mohly mít vliv na naměřené hodnoty.
- Chyba v **teorii** programu. Evaluátor může na základě výsledků identifikovat možnou chybu v teorii programu a navrhnout změnu.
- Chyba v **implementaci** programu. Program může mít oprávněnou teorii, ale může být implementován (uváděn) neadekvátním způsobem. Tento typ chyb je možné poznat z evaluace implementace programu, tj. v podstatě na základě pozorování průběhu programu.

Interpretace nálezů evaluace vyžaduje i značnou opatrnost v přisuzování příčin selhání programu. Zejména v situaci, kdy evaluace odhalí chybu v implementaci programu, nelze říct, zda program obsahuje chybu také ve své teorii. V takovém případě je třeba doporučit odstranit implementační nedostatky (tj. například znovu vyškolit lektory) a poté zopakovat evaluační výzkum. Je také patrné, že každá evaluace musí reflektovat vlastní limity, které jsou často způsobeny nedostatkem finančních zdrojů, neochotou respondentů či specificky programem.

PŘÍKLAD VÝSTUPŮ EVALUAČNÍCH VÝZKUMŮ

EVALUACE PROGRAMU ZELENÁ ŠKOLA

Středisko ekologické výchovy Živica koordinuje na Slovensku program Zelená škola (Eco-school). Program předpokládá, že žáci a učitelé vytvoří týmy, které budou postupně ekologizovat provoz školy. Výsledkem má být rozvoj akčních znalostí a dovedností žáků, interní ohnisko kontroly a v důsledku proenvironmentální chování v oblasti ekomanažementu. Úspěšné školy (tj. školy, které získají určitý počet bodů) získávají certifikát Zelená škola, který je spojený s jistou prestiží. Do programu je zapojeno více než sto škol. Program byl evaluován v Irsku, Skotsku i České republice. Výsledky prokázaly pozitivní efekty programu na většinu sledovaných proměnných.

Při vstupní expertní analýze zarazil evaluátora systém bodování škol v programu. Na jedné straně měla každá akce směřující k ekologizaci provozu školy přesně vyčíslenou bodovou hodnotu. Na druhé straně nebyly nijak bodovány aktivity žáků ani nijak hodnocena míra jejich zapojení. To se evaluátorovi zdálo v rozporu s doporučeními z teoretické literatury. Podle nich by měla být iniciativa na řešení delegována především na žáky, aby se posílilo jejich přesvědčení, že jsou schopni ovlivňovat dění. Iniciativa je předpokladem pro vnitřní motivaci, ta je jedním z faktorů umožňujících transfer do vlastního chování.

Po dohodě s koordinátorem programu byly stanoveny evaluační otázky. Ty sledovaly vliv programu na akční znalosti, akční dovednosti, ohnisko kontroly a proenvironmentální chování. Kromě toho měla být porovnána česká a slovenská verze programu. V evaluačním plánu jsme dohodli, že jako experimentální skupinu vezmeme žáky ze škol, které aktuálně drží titul Zelená škola. Kontrolní skupinou měli být stejně staří žáci ze škol, které byly v programu méně než půl roku.

Evaluačním nástrojem byl test a dotazník doplněný o otevřenou otázku, sledující zážitky z programu. Výzkum byl tedy převážně kvantitativní s kvalitativním prvkem.

Výsledky ukázaly, že mezi experimentální a kontrolní skupinou nebyly statisticky významné rozdíly ve prospěch experimentální skupiny, s výjimkou položky sledující třídění odpadu. V několika případech dosáhli lepších hodnot žáci z kontrolní skupiny. V otevřené otázce žáci reflektovali program jako pestrou koláž aktivit, zahrnující stavění májky, folklórní tance nebo přenášení žab přes silnici. Většina žáků (včetně žáků z experimentální skupiny) o programu Zelená škola nic nevěděla. Evaluátorovi, který odpovědi na otevřenou otázku

kategorizoval a kvantifikoval, se zdálo významné, že uvedené aktivity neodkazují na vlastní iniciativu dětí, ale vždy na program připravený učitelem.

Podle evaluátorovy interpretace program neplní své cíle. Příčinou je chyba v implementaci programu na školách. Učitelé patrně nesprávně chápou cíle programu jako ekologizaci školy, namísto změny postojů, znalostí, dovedností a chování dětí. Tento závěr by ale bylo třeba ověřit dalším kvalitativním výzkumem mezi učiteli. Chyba v implementaci ale patrně vyplývá z nevhodně nastaveného bodovacího systému, tedy z teorie programu. Bodovací systém bylo doporučeno změnit tak, aby přenášel těžiště aktivity z učitele na žáka.

Koordinátor programu tyto výsledky přijal. Korespondovaly s jeho pocitem, že program není zcela v pořádku. Na jejich základě pak přistoupil ke změnám pravidel pro získání titulu.

DOPORUČENÍ

- Podporovat šíření metodiky evaluačního výzkumu.
- Podporovat vzdělávací programy a kurzy určené pro budoucí evaluátory environmentální výchovy.
- Podporovat spolupráci mezi školami, středisky a univerzitami na evaluačních výzkumech.
- Zvýhodňovat v grantových soutěžích programy, které prošly či procházejí evaluačním výzkumem.
- Poskytovat finanční zdroje realizátorům environmentální výchovy na evaluační výzkumy svých programů.
- Podporovat různost přístupů v evaluačních výzkumech i designu programů, inovativnost a kreativní hledání nových metod a strategií efektivní environmentální výchovy.

LITERATURA

- [1] ALLUM, Liz; LOWE, Barbara; ROBINSON, Louise. *How do we know it's working? A toolkit for measuring attitudinal change in global citizenship from early years to KS5*. RISC, 2008. 115 s. ISBN 978-1-874709-10-6.
- [2] ANDREWS, Kristen E., TRESSLER, Kurt D., MINTZES, Joel J. Assessing environmental understanding: an application of the concept mapping strategy. *Environmental Education Research*. Vol. 14, no. 5, October 2008, p. 519-536.
- [3] AVERCH, H.A. *Using Expert Judgment*. In WHOLEY, J.S.; HATRY H.P.; NEWCOMER, K.E. Handbook of Practical Program Evaluation. San Francisco: Jossey-Bass, 2004. S. 292-309.
- [4] BARCH, Brian; DUVAL, Jason; HIGGS, Amy; WOLSKE, Kim; ZINT, Michaela. Planning and Implementing an EE Evaluation. [online] [Poslední aktualizace 2007-11-06] [Cit. 2008-07-20]. Dostupné na <http://66.135.39.45:7080/meera-dev/knowledge-base/plan-an-ee-evaluation/>
- [5] BENNETT, Dean B. *Evaluating Environmental Education in Schools. A practical guide for teachers*. [online]. UNESCO - UNEP, Division of Science, Technical and Environmental Education, 1989. Dostupné na <http://unesdoc.unesco.org/images/0006/000661/066120eo.pdf>
- [6] BRAUS, J.E., WOOD, D. *Environmental education in the schools – Creating a program that works!* Washington: Peace Corps, 1993.
- [7] CACHELIN, Adrienne; PAISLEY, Karen; BLANCHARD, Angela. Using the Significant Life Experience Framework to Inform Program Evaluation: The Nature Conservancy's Wings & Water Wetlands Education Program. *The Journal of Environmental Education*. Winter 2009, Vol. 40, n. 2., p. 2-15.
- [8] CORBIN, Juliet; STRAUSS, Anselm. Basics of Qualitative Research. 3e. Techniques and Procedures for Developing Grounded Theory. Thousand Oaks: Sage, 2008. 400 s. ISBN 141290644X
- [9] CORDANO, M., WELCOMER, S. A., SCHERER, R. F. An analysis of the predictive validity of the new ecological paradigm scale. *The Journal of Environmental Education*. 2003, vol. 34, is. 3, s. 22–28. ISSN 0095-8964.
- [10] ČINČERA, Jan. Evaluace programu Ekoškola [online]. *Envigogika*. Praha: Centrum pro otázky životního prostředí Univerzity Karlovy. Roč. 3, 2008, č. 2. Dostupné z <<http://www.czp.cuni.cz/envigogika>> ISSN 1802-3061.
- [11] ČINČERA, Jan; MACHÁČKOVÁ, Petra. Evaluace pobytového programu Podblanického ekocentra [online]. *Envigogika*. Praha: Centrum pro otázky životního prostředí Univerzity Karlovy. Roč. 4, 2009, č. 3. Dostupné z <<http://www.czp.cuni.cz/envigogika>> ISSN 1802-3061.
- [12] ČINČERA, Jan. Evaluační strategie středisek ekologické výchovy [online]. *Envigogika*. Praha: Centrum pro otázky životního prostředí Univerzity Karlovy. Roč. 3, 2008, č. 2. Dostupné z <<http://www.czp.cuni.cz/envigogika>> ISSN 1802-3061.
- [13] ČINČERA, Jan. Práce s hrou. Pro profesionály. Praha: Grada, 2007.
- [14] ČINČERA, Jan. Výstupy či procesy. Paradigmatický (kvazi) spor environmentální výchovy [online]. *Envigogika*. Praha: Centrum pro otázky životního prostředí Univerzity Karlovy. Roč. 4, 2009, č. 2. Dostupné z <<http://www.czp.cuni.cz/envigogika>> ISSN 1802-3061.
- [15] ČINČERA, Jan; GILAR, Petr; SOKOLOVIČOVÁ, Júlia. Specializační studium pro koordinátory environmentální výchovy, vzdělávání a osvěty: interpretace a efektivita z pohledu absolventů. *Envigogika*. Praha: Centrum pro otázky životního prostředí Univerzity Karlovy. Roč. 5, 2010, č. 1. Dostupné z <<http://www.czp.cuni.cz/envigogika>> ISSN 1802-3061.
- [16] ČINČERA, Jan; KULICH, Jiří; GOLLOVÁ, Dita. Efektivita, evaluace a podpora programů environmentální výchovy [online]. *Envigogika*. Praha: Centrum pro otázky životního

- prostředí Univerzity Karlovy. Roč. 4, 2009, č. 2. Dostupné z <<http://www.czp.cuni.cz/envigogika>> ISSN 1802-3061.
- [17] ČINČERA, Jan; KOMÁRKOVÁ, Magdaléna. Využití kresby pro evaluaci programu na rozvíjení environmentální senzitivity mladších dětí. *Envigogika*. Praha: Centrum pro otázky životního prostředí Univerzity Karlovy. Roč. 5, 2010, č. 2. Dostupné z <<http://www.czp.cuni.cz/envigogika>> ISSN 1802-3061.
- [18] *DOES Your Project Make a Difference?* [online] Sydney: Department of Environment and Conservation, 2004. [Cit. 2008-07-20]. Dostupné na <http://www.environment.nsw.gov.au/resources/community/projecteval04110.pdf>
- [19] DUNLAP, R.E, Van LIERE, K.D. The new environmental paradigm: a proposed measuring instrument and preliminary results. *The Journal of Environmental Education*, 1978, iss. 9, p. 10-19.
- [20] DUNLAP, Riley E.; van LIERE, Kent D.; MERTIG, Angela G.; JONES, Robert E. Measuring Endorsement of the New Ecological Paradigm: A Revised NEP Scale – Statistical Data Included. *Journal of Social Issues*. 2000, Vol. 56, n. 3. ISSN 0022-4537.
- [21] EXCELLENCE IN ENVIRONMENTAL EDUCATION - GUIDELINES FOR LEARNING (PRE K-12). North American Association for Environmental Education, 2004. [Cit. 2008-07-21]. Dostupné na http://www.naaee.org/npeee/learner_guidelines.php
- [22] FITZPATRICK, Jody L.; SANDERS, James R.; WORTHEN, Blaine R. *Program Evaluation. Alternative Approaches and Practical Guidelines*. Boston: Pearson Education, 2004. 555 p. ISBN 0-321-07706-7.
- [23] FRANC, Daniel; ZOUNKOVÁ, Daniela; MARTIN, Andy. *Učení zážitkem a hrou. Praktická příručka instruktora*. Brno: Computer Press, 2007. ISBN 978-80-251-1701-9.
- [24] FRECHTLING, Joy et. al. *The 2002 User-Friendly Handbook for Project Evaluation*. The National Science Foundation, 2002. [online] [Cit. 2009-01-05] Dostupné na <http://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf>
- [25] GLASER, Barney B. *Doing Grounded Theory: Issues and Discussions*. Mill Valley: Sociology Press, 1998. ISBN 1884156118.
- [26] *Guidelines for the Preparation and Professional Development of Environmental Educators*. Washington: NAAEE, 2004.
- [27] HART, Paul; NOLAN, Kathleen. A critical analysis of research in environmental education. *Studies in Science Education*, 1999, iss. 34, p. 1-69.
- [28] HENDL, Jan. *Kvalitativní výzkum. Základní teorie, metody a aplikace*. Praha: Portál, 2008. ISBN 978-80-7367-485-4.
- [29] HENDL, Jan. *Přehled statistických metod zpracování dat*. Praha: Portál, 2004.
- [30] HENTON, Mary. *Adventure in the Classroom. Using Adventure to Strengthen Learning and Build a Community of Life-Long Learners*. Dubuque: Kendall/Hunt, 1996. ISBN 0-7872-2459-6.
- [31] HOGAN, Christine. *Practical facilitation. A toolkit of techniques*. London: Kogan Page, 2003. ISBN 0-7494-3827-4.
- [32] HORNOVÁ, Karolína. Evaluace výukového programu environmentální výchovy. *Envigogika*, 2007, č. 3. Dostupné na <http://envigogika.cuni.cz> . ISSN 1802-3061
- [33] HUNGERFORD, Harold R. The General Teaching Model (GTM). In HUNGERFORD, Harold H.; BLUHM, William J.; VOLK, Trudi L.; RAMSEY, John M. *Essential Readings in Environmental Education*. Champaign: Stipes, 2005. ISBN 1-58874-469-8. P. 423-443.
- [34] CHRÁSKA, Miroslav. *Metody pedagogického výzkumu*. Praha: Grada, 2007.
- [35] JANČAŘÍKOVÁ, Kateřina. Žakovské portfolio – vhodná forma hodnocení environmentální výchovy. [online] *Envigogika*, 2007, č. 3. Dostupné na <http://envigogika.cuni.cz> . ISSN 1802-3061
- [36] KOLB, David. *Experiential Learning. Experience as The Source of Learning and Development*. Prentice Hall, 1984. 256 s.

- [37] KOVALIK, Susan J.; OLSEN, Karen D. *Kid's eye view of science. A Teacher's Handbook for Implementing an Integrated Thematic Approach to Teaching Science, K-6*. Kent: Center for the Future of Public Education, 1994.
- [38] KRAJHANZL, Jan. *Děti a příroda: Prožívání a zkušenost* [online]. <Cit. 2010-05-25>. Dostupné na <<http://www.vztahkprirode.cz/view.php?cisloclanku=2007050045>>
- [39] MÁCHAL, Aleš; ČINČERA, Jan; DAŇKOVÁ, Lenka; BROUKAL, Václav. *Ekopedagogova osmero 2007*. [Interní materiál]. Praha: Pavučina, 2008.
- [40] MARCINKOWSKI, Thomas. Assessment in Environmental Education. In WILKE, Richard J. *Environmental Education. Teacher Ressource Handbook. A Practical Guide for K-12 Environmental Education*. Thousand Oaks: Corwin, 1997. P. 143-198.
- [41] MARCINKOWSKI, T. *Using a Logic Model to Review and Analyze an Environmental Education Program*. Washington: North American Association for Environmental Education, 2004. ISBN 1-884008-86-0.
- [42] MAREŠ, Jiří; GAVORA, Peter. *Anglicko-český slovník pedagogický*. Praha: Portál, 1999.
- [43] McLAUGHLIN, J.A.; JORDAN, G.B. Using Logic Models. In WHOLEY, J.S.; HATRY H.P.; NEWCOMER, K.E. *Handbook of Practical Program Evaluation*. San Francisco: Jossey-Bass, 2004. S. 7-32.
- [44] MEERA – *My Environmental Education Evaluation Ressource Assisntent*. University of Michigan, 2008. Dostupné na <http://meera.snre.umich.edu>
- [45] MIOVSKÝ, Michal. *Kvalitativní přístup a metody v psychologickém výzkumu*. Praha: Grada, 2006. ISBN 80-247-1362-4
- [46] MUSSER, Lynn M., MALKUS, Amy J. The Children's Attitudes Toward the Environment Scale. *The Journal of Environmental Education*, 1994, vol. 25, no 3, p. 22-26.
- [47] MUSSER, Lynn M.; DIAMOND, Karen E. The Children's Attitudes Toward the Environment Scale for Preschool Children. *The Journal of Environmental Education*. 1999. Vol. 30, No. 2; p. 23-30. ISSN 0095-8964.
- [48] PATTON, M. Q. *Qualitative Research and Evaluation Methods*. Thousand Oaks: Sage, 2002. 700 s.
- [49] PRŮCHA, J., WALTEROVÁ, E., MAREŠ, J. *Pedagogický slovník*. 6. aktualizované a rozšířené vyd. Praha : Portál, 2009. ISBN 978-80-7367-647-6.
- [50] REICHARDT, C.S.; MARK, M.M. Quasi-Experimentation. In WHOLEY, J.S.; HATRY H.P.; NEWCOMER, K.E. *Handbook of Practical Program Evaluation*. San Francisco: Jossey-Bass, 2004. S. 126-149.
- [51] REITMAYEROVÁ, Eva; BROUMOVÁ, Věra. *Cílená zpětná vazba*. Praha: Portál, 2007.
- [52] ROBOTTOM, Ian. Evaluation in Environmental Education: Time for Change in Perspective? *The Journal of Environmental Education*, Fall 1985, vol. 17 n. 1, p. 31-36.
- [53] ROSSI, P. H., LIPSEY, M.W., FREEMAN, H.E. *Evaluation. A Systemic Approach*. Thousand Oaks: Sage, 2004.
- [54] SELLERS, Jennifer. Goals. In *MEERA Evaluation Glossary*. [online] <Poslední změna 2007-12-18>. Dostupné na <http://meera.snre.umich.edu>
- [55] SCHUELLER, S.K.; YAFFEE, S.L.; HIGGS, S. J. ; MOGELGAARD, K.; DEMATTIA, E.A. *Evaluation Sourcebook: Measures of Progress for Ecosystem- and Community-Based Projects*. Ann Argot: Ecosystem Management Initiative, University of Michigan, 2006.
- [56] SIMMONS, Bora. *Designing Evaluation for Education Projects*. [online] Office of Education and Sustainable Development. [Cit. 2008-06-25] Dostupné na <http://wateroutreach.uwex.edu/use/documents/NOAAEvalmanualFINAL.pdf>
- [57] SMITH-SEBASTO, N.J.; SEMRAU, H.J. Evaluation of the Environmental Education Program at the New Jersey School of Conservation. *The Journal of Environmental Education*, Fall 2004, 36, 1, p. 3-6.
- [58] STOKKING, H.; van Aert, L.; MEIJBERG, W; KASKENS, A. *Evaluating Evironmental Education* [online]. IUCN Commission on Education and Communication CEC, 1999. Dostupné na <http://www.iucn.org/wp2003/books/pdf/cec4.pdf>

- [59] ŠVAŘÍČEK, Roman; ŠEĐOVÁ, Klára. Kvalitativní výzkum v pedagogických vědách. Praha: Portál, 2007.
- [60] TAYLOR-POWELL, Ellen; RENNER, Marcus. *Analysing qualitative data*. [online] Madison: University of Wisconsin, 2003. [Cit. 2008-04-02] Dostupné na <http://learningstore.uwex.edu/pdf/G3658-12.PDF>
- [61] THOMSON, Gareth; HOFMAN, Jenn. *Measuring the Success of Environmental Education Programs*. [online] Canadian Parks and Wilderness Society; Sierra Club of Canada. [Cit. 2008-04-02] Dostupné na http://www.peecworks.org/PEEC/PEEC_Inst/100052276.0/ee-success.pdf
- [62] USAID. *Performance Monitoring and Evaluation Tips. Selecting Performance Indicators*. [online] USAID Center for Development Information and Evaluation, c1996 [Cit. 2008-07-20] Dostupné na http://pdf.dec.org/pdf_docs/pnaby214.pdf
- [63] W.K. KELLOG Foundation. *Logic Model Development Guide. Using Logic Models to Bring Together Planning, Evaluation, and Action*. [online] Battle Creek: W.K. Kellogg Foundation, 2004. [Cit. 2008-07-20] Dostupné na <http://www.wkkf.org>
- [64] WHOLEY, J.S.; HATRY H.P.; NEWCOMER, K.E. *Handbook of Practical Program Evaluation*. San Francisco: Jossey-Bass, 2004.
- [65] WILTZ, K. L. I need a bigger suitcase: The evaluator role in nonformal evaluation. *In New Directions for Evaluation*. 2005, iss. 108, p. 13-28.
- [66] ZAHRADNÍK, Martin; PACHMANOVÁ, Lenka. Příklad využití wiki-prostředí v environmentálním vzdělávání na vysoké škole: hodnocení kurzu Organizace a řízení ochrany životního prostředí. [online] *Envigogika*, 2009, roč. 4, č. 3. Dostupné na <http://envigogika.cuni.cz> . ISSN 1802-3061

ABSTRAKT

Programy environmentální výchovy a vzdělávání musí často doložit svou efektivitu – základním prostředkem k tomu jsou evaluační výzkumy. Evaluační výzkumy si zadávají zpravidla realizátoři či sponzoři programů a na jejich základě pak získávají informace o efektivitě použitých vzdělávacích metod. Výsledky pak může využít realizátor programu pro zkvalitňování svého programu i externí posuzovatel – například donor pro rozhodování o dalším rozdělování finančních prostředků. Dlouhodobě je pak zisk z evaluačních výzkumů oboustranný: realizátoři připravují kvalitnější programy a program lze obhájit jako přínosný. Předložená metodika má sloužit oběma těmto účelům; může být využita i pracovníky veřejné správy, kteří formální i neformální programy environmentální výchovy z různých důvodů posuzují.

Cílem prezentované metodiky je vymežit základní terminologii pro evaluaci programů environmentální výchovy, popsat základní fáze evaluačního cyklu, dále charakterizovat základní alternativní postupy při evaluaci programů, a nakonec identifikovat základní požadavky na evaluační výzkum.

Zpracování metodiky vychází z předpokladu, že se ve střednědobém horizontu stanou evaluační výzkumy běžnou součástí environmentální výchovy. Pro dosažení tohoto cíle je zapotřebí, aby se odpovědní pracovníci orientovali v tématice a získali základní porozumění pro možnosti i nároky evaluačních výzkumů.

ABSTRACT

Environmental education often has to demonstrate its effectiveness - an essential means to this end is evaluation research. Evaluation studies are typically commissioned by the promoters or sponsors of programs who, on this basis, get necessary feedback. The results can then be used by those who implement the program to improve it, or by its external evaluators, such as the donor, to assess its efficiency. The long-term gain of the evaluation research accrues to both sides: for the implementers of such programs it reflects the quality and could assist better preparations of a program; externally it can also be shown to be beneficial. The proposed methodology is intended to serve both these purposes and may be used by public administration officials who evaluate both formal and informal environmental education programs for various reasons (incl. making decisions on the distribution of funds to support formal and informal environmental education programs).

The aim of the presented methodology is to define the basic terminology for evaluating environmental education programs, describe the basic stages of the evaluation cycle, characterize the basic alternative procedures for evaluating programs, and finally, to identify the basic requirements for evaluation research.

Development of the methodology is based on the assumption that the evaluation research in the medium term will become an integral part of environmental education programs. To achieve this objective requires that officials are familiar with the theme and gain a basic understanding of the opportunities and needs of the assessment procedures.