

HODNOTÍCÍ (EVALUAČNÍ) VÝZKUM

Ladislav Rabušic

Co je hodnotící výzkum

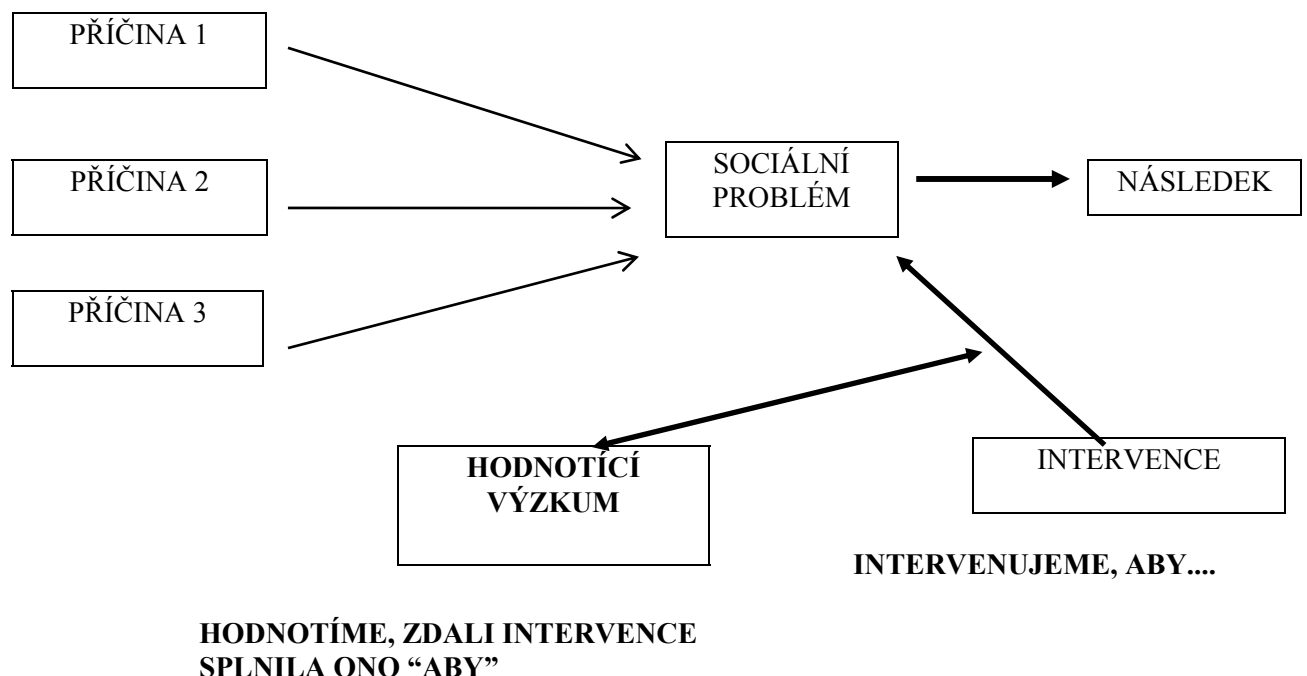
Velké finanční prostředky jsou každoročně investovány do programů, které mají zlepšit podmínky života jedinců i celých sociálních skupin. Jak ale mají být tyto programy hodnoceny? Jak odlišíme účinný program od neúčinného? Jaké informace potřebujeme k tomu, abychom zlepšili úspěšnost našeho rozhodování a zvýšili efektivitu programu? Tyto otázky tvoří jádro výzkumných aktivit, jimž se říká hodnotící výzkum.

Hodnotící výzkum je specifická aktivita, která nasazuje nejrůznější postupy sociálně vědního výzkumu, aby zjistila, do jaké míry určitá činnost (program) naplňuje předem stanovené cíle.

Definice: V hodnotícím výzkumu aplikujeme nejrůznější výzkumné postupy s cílem zhodnotit účinky programů sociálních služeb a sociální intervence.¹

Jiná definice říká: Hodnotící výzkum je aplikovaný výzkum, který se provádí proto, aby mohla být přijata rozhodnutí o jednom nebo několika programech sociálních služeb.

MODEL KONTEXTU HODNOTÍCÍHO VÝZKUMU



¹ Sociální intervenci je možné definovat jako aktivitu prováděnou za účelem dosáhnout zamýšlených výsledků v oblasti sociálních problémů.

Hodnotící výzkum (**HV**) je obvykle aplikován na programy, jejichž cílem je pomoc, kteří se ocitli v nouzi, mají nějaký problém, jsou „utlačováni“.

Témata hodnotícího výzkumu jsou velmi různorodá. Klademe si např. otázky typu:

1. Pomohly jednotlivé programy politiky zaměstnanosti vytvořit v regionu nová pracovní místa a snížit tak nezaměstnanost?
2. Umožňují programy služeb poskytované v domácnostech seniorů oddálit či úplně eliminovat umístování seniorů do domovů důchodců?
3. Redukoval vzdělávací program o sexuálním zneužívání dětí tento typ sociální delinkvence?
4. Přispělo zavedení školného na univerzitách k rychlejšímu rozvoji vysokých škol a ke zlepšení kvality studia?

Hodnotící výzkum, jak napovídá název, je založen na hodnocení. V obecném významu hodnocení znamená srovnávat pozorovanou věc, aktivitu, naměřenou hodnotu s nějakým standardem. Tento standard je definován různými způsoby, v oblasti **HV** často prostřednictvím očekávání různých lidí kteří nějakým způsobem na program participují. Ti jsou samozřejmě různého druhu, v podstatě se ale dělí do tří skupin:

- (1) sponzoři programu, tedy ti, kdo program financují (vládní instituce, nevládní instituce – domácí i zahraniční, privátní sféra);
- (2) lidé, kteří program naplňují („zaměstnanci“ instituce provádějící program);
- (3) příjemci (recipienti) programu (ti, jimž je program určen, pro něj je organizován).

Tyto tři skupiny mohou mít (a obvykle mají) mají vůči programu odlišná očekávání. Hodnotící výzkum často stojí v nelehké úloze rozhodnout, kterou sadu očekávání zvolit jako standard pro srovnávání.

Hodnotící výzkum je aplikován všude tam, kde si přejeme znát výsledky našeho úsilí o sociální reformu. Rozvinul se výrazně v 50. letech s rozvojem sociálního státu a jeho sociálních programů, jejichž cílem bylo odstraňovat sociální problémy. Jeho rozmach nastal v průběhu 70. a 80. let, kdy se začal klást nesmírný důraz na tzv. *accountability* (tedy povinnost skládat veřejné účty ze své činnosti) - tento proces se promítl i do fungování sociálních programů: veřejnost a sponzoři si začali přát vědět, jakých výsledků za jejich peníze tyto programy dosahují a zdali vůbec dosahují naplňují cíle, kvůli nimž byly implementovány. Proto bylo v rámci rozpočtů programů také pamatováno na to, aby 1-3 % rozpočtu byla rezervována pro hodnotící výzkum.

Hodnocení programů jakožto specifický prvek profesionální výzkumné aktivity se zrodil ze dvou lekcí, které toto období velkého investování do sociálních programů přineslo. Tou první bylo uvědomění si faktu, že zde nikdy nebude dost peněz na to, aby se udělalo vše to, co by bylo potřeba udělat; tou druhou pak poznatek, že i kdyby zde byl dostatek peněz, peníze nejsou všechno – pro řešení složitých sociálních problémů samotné peníze nestačí. Hodnotící výzkum tak slouží jako báze pro rozhodování, které věci jsou prvořadé.

Je zřejmé, že **HV** se pohybuje ve výrazném politickém kontextu – je to jeho výrazná charakteristika, která jej odlišuje od jiných výzkumných aktivit. Hodnocení programů je velmi úzce propojeno s politickým rozhodováním týkajícím se sociálních priorit, moci a alokace (přidělování) zdrojů. Výsledky hodnotícího výzkumu se mohou stát municí, s níž oponenti programu mohou program „sestřelit“ a docílit tak jeho zrušení. Výsledky **HV** tak mohou ovlivnit existenci lidí - jejich zaměstnání v programech, výši jejich mezd (financování programu), mohou ovlivnit osudy recipientů programu.

Hodnotící výzkum se tak pohybuje v silném politickém a ideologickém rámci a byť jsou jeho výsledky dosahovány prostřednictvím vědeckých (tedy objektivních) postupů, jejich přijímání je většinou silně subjektivně zbarveno. Např. Richard Nixon jako president USA ustavil národní komisi pro studium následků existence pornografie. Po mnohostranném a zevrubném zkoumání dospěla komise k závěru, že podle všech jejich výsledků pornografie nemá žádné zjevné negativní sociální konsekvence, které jsou jí připisovány. U konzumentů pornografie nebyla např. zjištěna vyšší pravděpodobnost spáchání sexuálního zločinu. Na základě těchto výsledků by se dalo očekávat, že dojde k jisté liberalizaci zákonů v oblasti šíření pornografie. President Nixon nicméně prohlásil, že závěry komise jsou mylné.

Tyto případy, kdy výsledky hodnotícího výzkumu jsou ignorovány, jsou údajně časté, jak napovídají učebnice. Důvody jsou trojí:

1. Výsledky výzkumu nejsou vždy prezentovány v takové podobě, aby jim rozuměli i lidé, kteří nevědí, jak se výzkum dělá.
2. Výsledky jsou někdy v ostrém protikladu s hluboce zakořeněnými představami (což byl jistě případ výsledků zkoumání efektů pornografie).
3. Využití výsledků výzkumu brání tzv. osobní zájmy (*vested interests*). Představme si skupinu nadšenců, kteří poté, co byli vyškoleni v zahraničí, založí poradnu pro mladá manželství, v níž chtějí aplikovat nejmodernější matrimoniologické postupy. Pro svou ideu získají sponzory, kteří do takového podniku financují peníze (např. sociální odbor magistrátu) a získají také souhlas psychologické obce, že takovou praxi smějí provozovat. Jsou přesvědčeni, že jejich systém bude efektivní a že pomůže lidem řešit jejich složité manželské problémy, čímž snad napomohou ke snižování rozvodovosti. Jakkup se asi budou cítit, když objektivní hodnotící výzkum dojde k závěru, že jejich program nefunguje? „Stáhnou roletu“, omluví se klientům, že jim radili špatně a zavřou krám? Jistě ne. V tomto projektu mají své *vested interests*, své osobní zájmy, a proto je možné očekávat, že prohlásí, že výzkum nebyl metodologicky správně proveden a že je tudíž bezcenný. Pravděpodobně také začnou intenzivní lobbying u vlivných osob, aby docílili toho, že program bude i nadále fungovat.

Aby mohl být hodnotící výzkum mnohdy vůbec proveden, je třeba, aby získal náležitou logistickou (tj. personální a organizační) podporu v agentuře či instituci, v níž se bude hodnotící výzkum provádět. Při přípravě hodnotícího výzkumu je třeba, aby jeho autoři:

1. získali informace o těch, pro něž mohou z evaluace vyplynout důsledky – jejichž příjem, status nebo kariéra mohou být ovlivněny. Tyto lidi je dobré zahrnout do projektu tak, aby na něm smysluplně participovali;
2. si uvědomili, kdo si evaluaci přeje, proč si ji přeje, ale také kdo si ji nepřeje a jaké pro to může mít důvody;
3. předali těmto lidem projekt k připomínkám;

Typy evaluačního výzkumu

Až dosud jsme hovořili především o politice hodnotícího výzkumu. Podívejme se nyní, jaké typy výzkumu rozeznáváme. Hodnotící výzkum je často prováděn na základě modelu CIPP (*Context–Input–Process–Product*):

CIPP jako model hodnotícího výzkumu:

Prvek	Kladené otázky
-------	----------------

Kontext (Context)	Jaké jsou cíle programu. Odpovídají potřebám příjemců?
Vstupy (Input)	Jaké prostředky jsou požadovány pro naplnění cílů programu (počet zaměstnanců, investice, rozpočet)?
Průběh (Process)	Jak byli účastníci informováni o průběhu aktivit? Jak byly přidělovány zdroje? Jak byly připravovány materiály?
Produkt (Product)	Jaká existuje evidence o výsledcích programu. Měl by být program ukončen, měl by pokračovat, měl by být zásadně revidován? Měly by být zdroje sníženy či navýšeny? Měl by být program sloučen s jiným programem?

Kroky při provádění evaluačního výzkumu

Jelikož má **HV** charakter výzkumu, jsou jeho postupy do jisté míry podobné. Odlišné jsou však zde charakteristiky prostředí – je to svým způsobem výzkum instituce, jejích zaměstnanců a aktivit s možností velmi přímého dopadu na osudy konkrétních osob

A) **Příprava** – během této fáze by měl hodnotitel (výzkumník nebo spíše výzkumný tým) jasně definovat, co má být hodnoceno, měl by si ujasnit účel hodnocení, měl by si definovat ty, kdo mají na hodnocení zájem a měl by si ujasnit, v jakých podmínkách a omezeních bude hodnotící výzkum probíhat. (Výzkumný tým je obvykle kontaktován ředitelem programu, představitelem financující skupiny–instituce, politikem).

B) **Popis evaluanda** (toho, co se má hodnotit) – na základě nejrůznějších materiálů je třeba získat detailní informace o programu a instituci (chartu, zakládající listiny, cíle programu, výroční zprávy, předchozí evaluační materiály apod.). Z těchto dokumentů je třeba popsat program a o tomto popisu je třeba debatovat se zadavatelem. Je třeba jej upozornit, že věci na papíře jsou jedna věc a skutečné aktivity věc druhá. Z toho důvodu je třeba získat svolení pozorovat program v akci, a tím získat detailnější představu o průběhu programu. Zde se často ukáže, že různí lidé s různým vztahem k programu budou mít na program odlišné názory. Na to je třeba klienta předem upozornit a připravit. Tato deskripce různorodých percepčí programu je již sama o sobě důležitá.

C) **Účel hodnocení** – je třeba zjistit, co byla hlavní příčinou pro objednávku hodnotícího výzkumu a jakým způsobem mají být výsledky hodnocení využity. Dále kdo tyto výsledky využije a k jakému účelu?

D) **Identifikace hlavních hráčů** – ti mohou být: sponzoři, správní rada, státní nebo regionální úřady, politikové, administrativa na různé úrovni, zaměstnanci programu, recipienti (někdy i rodiče recipientů), zástupci místní veřejnosti

E) **Podmínky a limity evaluace** – každé hodnocení se děje v konkrétních podmínkách, které vytváří určité limity: těmi bývají peníze („na váš výzkum máme tolik a tolik peněz“) čas („hodnocení musí být hotovo do tří měsíců“), osoby („naši zaměstnanci vám mohou být k dispozici tehdy a tehdy“), existující data („jsou uložena tam a tam“), politika („pokud bude program redukován, bude možné využít finance pro založení nového programu“). Politický aspekt výzkumu je zde výrazně přítomen – proto je třeba během výzkumu si uvědomit, kdo z hodnocení získá a kdo může ztratit, kdo hodnocení podporuje a kdo ho naopak odmítá (a snad i lehce sabotuje) apod.

Je dobré na základě těchto pěti druhů zjištění sestavit celkovou syntézu, jak vy chápete program a jeho fungování. Tu je dobré předložit jednotlivým skupinám „hráčů“ a žádat je o zpětnou vazbu.

Otázky kladené při hodnotícím výzkumu:

Otázky deskriptivní: Jsou to otázky, jimiž zjišťujeme, co program vlastně je a co naplňuje.

Např.: Jaké aktivity program podporuje? Za jakým účelem? Kdo tyto aktivity provádí (naplňuje)? Jak jsou tyto aktivity rozsáhlé, kolik stojí a koho (jaké recipienty) obsáhnou? Jsou podmínky a aktivity a klienti programu relativně stále prvky nebo zde existuje podstatná variabilita?

Otázky na provádění programu (implementační): jaké aktivity byly prováděny, zdali to jsou ty, které byly původně zamýšleny a zda se dotýkají původně zamýšlených problémů a populace. Např.: Jsou deklarované aktivity skutečně prováděny? Jsou tyto aktivity v souladu se zákonem a administrativními regulacemi? Jsou tyto aktivity v souladu s profesionálními standardy? Jsou zdroje pro program efektivně využívány?

Otázky na to, jaký má program vliv: jsou to otázky, které osvětlí skutečné účinky programu. Např.: Dosahuje program zamýšlených účelů a výsledků? Jaké jsou celkové výsledky programu? Liší se výsledky v závislosti na druhu příjemců? Má program nějaké vedlejší (a nezamýšlené – ať pozitivní či negativní – efekty)? Ovlivnil nějak příbuzný program s podobným cílem? Dosahuje tento program podobných účinků jako program alternativní? Jsou účinky programu hodny vynaloženým finančním a jiným nákladům?

Sběr dat

Na základě stanovených otázek se musíme rozhodnout, která data a z jakých zdrojů budeme sbírat, kdo je bude sbírat. Způsoby sběru dat jsou nejrůznější, uplatní se zde nejrůznější postupy sociálně vědního výzkumu. Záleží na invenci výzkumného týmu.

Z jiného pohledu je možné **HV** rozdělit do následujících tří typů podle toho, zda výzkum hodnotí:

1. konečné účinky programu
2. problémy spojené s implementací programu
3. informace potřebné pro zavedení nového programu a pro jeho rozvoj (*needs assesment*)

Ad 1. Hodnocení konečných účinků programu

- zda program plní vytčené cíle
- zda nemá nějaké vedlejší (nezamýšlené) škodlivé
- zda je program úspěšně realizován s rozumnými finančními náklady
- zda jsou náklady a výnosy tohoto programu srovnatelné a náklady a výnosy programu jiného, avšak s podobnými cíli.

V tomto režimu práce budou formální cíle programu závisle proměnnou, která musí být definována prostřednictvím nejrůznějších měřitelných indikátorů. Naší snahou bude zjistit, zda účinky programu byly vskutku způsobeny tímto programem.

Použité metody:

- experiment – experimentální skupinou budou příjemci programu, kontrolní skupinou ti, kdo jsou v podobné životní situaci, ale na programu neparticipují.
- quasi-experiment – výzkum s designem, kdy $N=1$.

- analýza nákladů a výnosů (*cost-benefit analysis*) – slouží ke zjištění efektivity programu: někdy je ovšem obtížné objektivně určit, jak velké jsou např. výnosy, neboť např. zlepšené zdraví nebo eliminace komplexu méněcennosti je obtížné vyčíslit v penězích. I do této čistě ekonomické metody vstupují subjektivní hodnocení.

S jakými problémy se můžeme setkat?

Někdy jsou cíle programu popsány v tak grandiózních proklamacích, že je málokdo bere vážně - byly tak zformulovány pro politické cíle nebo z toho důvodu, aby přesvědčily zákonodárce nebo jiné vlivné osoby, že pro program by se měly najít peníze. Jindy jsou cíle programu stanoveny velmi vágně, že hodnotitelé se neshodnou na pozorovatelných indikátorech, jejichž prostřednictvím by měřili jeho úspěch či neúspěch.

Ad 2) Monitorování implementace programu

Někdy je dobré před tím, než se pokusíme hodnotit celkové výsledky programu, získat data o tom, jak je program uskutečňován a prováděn. Místo toho tedy, abychom řekli zda je program dobrý či špatný, ukážeme, jaká nastala chyba při jeho zavádění.

Otázky, které si zde klademe:

- Jaká strategie na získávání peněz vede k největšímu balíku sponzorských financí?
- Jaký podíl populace, jíž je program určen, jej skutečně dostává?
- Jaký druh jedinců na programu neparticipuje?
- Proč je jich tak mnoho?
- Jsou klienti spokojeni se službou?
- Proč tak mnoho klientů opouští program před jeho dokončením?

Metody: všechny možné - survey, zúčastněné pozorování, analýza dokumentů,

Ad 3) Hodnocení potřeb (*needs assesment*)

Hodnotící výzkum se týká nejen programů, které již běží, ale také situace, kdy se uvažuje, že bude třeba zavést nový program, který by minimalizoval nějaký sociální problém. Před tím, než je nový sociální program zaveden, je nutné diagnostikovat situaci: zjistit, zdali je v daném prostředí vůbec potřeba program zavádět.

Příklad: chceme vytvořit regionální program pomoci pro bezdomovce. K tomu budeme muset zjistit některá fakta. Např.:

Kolik bezdomovců se vlastně v regionu vyskytuje?

Jaké příčiny vedou k bezdomovectví?

Jak velký podíl z bezdomovců se pro tento způsob života rozhodl dobrovolně?

Jak velký podíl z bezdomovců jsou ti, kteří jsou mentálně choří nebo jsou závislí na alkoholu či drogách?

Jak velký podíl z bezdomovců jsou ti, kteří ztratili práci, rozpadla se jim rodina a oni se propadli do chudoby?

Jak dlouho jsou lidé bezdomovci?

Jak velký podíl je mezi nimi rodin s dětmi?

Jaké účinky má bezdomovectví na děti, na jejich školní výkon, zdraví, výživu, sebevědomí?

Atd., atd.

Jelikož se v tomto typu hodnotícího výzkumu pohybujeme v oblasti sociálních potřeb, je třeba si uvědomit, že v tomto kontextu mohou být potřeby definovány z hlediska **normativního** a z hlediska **poptávkového**.

Normativní definice potřeb znamená, že při hodnocení potřeb cílové populace (ta, pro niž program chceme zavést) budeme srovnávat její životní podmínky s takovou úrovní, která je považována experty na daný problém za lidsky akceptovatelná a důstojná. V případě bezdomovectví se tento program může začít zavádět i tehdy, když bezdomovci sami nijak nevyjadřují potřebu změnit svůj způsob života.

Při **definici potřeb z hlediska poptávky** či požadavku vycházíme pouze ze situace postižených, kteří nějakým způsobem sdělují, že se jim něčeho nedostává, že jsou tedy potřební. Opět v našem případě bezdomovectví tedy ti, kdo dávají přednost životu na ulici, se mezi potřebné nebudou počítat. Definice potřeb z hlediska vyjadřované poptávky může být ovšem někdy ošidná. Jedinci či skupiny nemusejí vyjadřovat svůj požadavek na nějakou potřebu jednoduše proto, že ani netuší, že by jim někdo mohl pomoci. Nebo: jelikož se často jedná o skupiny marginalizované, které jsou sociálně vyloučeny, jejich hlas není vůbec slyšet.

Na způsobu definice potřeby budou záviset metody, s jejichž prostřednictvím získáme relevantní data pro náš cíl.

Pro hodnocení potřeb existuje pět druhů výzkumných technik:

- (1) technika klíčového informátora,
- (2) technika komunitního fóra,
- (3) technika zkoumání případů, které již sociální službu někde využívají,
- (4) technika sociálních indikátorů,
- (5) technika průzkumu cílové populace.

Klíčový informátor

Získáváme názory expertů, kteří mají specifické vědomosti o cílové populaci, o jejich problémech a potřebách. K tomu používáme metody rozhovorů nebo dotazníkového šetření.

Výhody: rychlá a nenákladná informace, navázání kontaktů s odborníky, kteří mohou být využiti i v dalších fázích realizace programu.

Nevýhoda: Expert nepochází přímo z populace, která je cílem programu, kvalita jeho informace závisí na jeho objektivitě a hloubce znalostí.

Komunitní fórum

Uspořádá se mítink, na němž všichni ti, koho se problém dotýká vyjadřují svá mínění.

Z vědeckého hlediska je tento přístup poněkud riskantní. Ti, totiž, kdo se tohoto mítinku zúčastní, nemusí být dobrými reprezentanty lidí, jichž se problém skutečně dotýká. Nejhlasitěji a nejviditelněji se budou pravděpodobně vyjadřovat ti, kdo mají největší *vested interests*. Je proto lepší, v případě této techniky, pořádat ne jedno otevřené setkání, ale řadu menších, uzavřených mítinků pro různé druhy populace, která bude ve hře.

Zkoumání případů, které již sociální službu někde využívají

Tento přístup se snaží odhadnout potřeby pro program a charakteristiky potenciálních klientů na základě toho, že studuje charakteristiky účastníků podobného programu, který byl již implementován někde jinde, nejlépe v prostředí, které je podobné tomu našemu. Hlavní technikou sběru dat pro tento přístup je sekundární analýza případových studií z této komunity. Jsou zde některá nebezpečí: 1. data, která agentura provádějící program ve studované komunitě, archivuje, nemusejí být reliabilní. Např. proto, že to není vůbec žádnou prioritou organizátorů programu (mají úplně jiné starosti) nebo také proto, že naopak mohou zveličovat počty klientů, aby od sponzora programu získali dodatečné finance. Největší nevýhodou ovšem je to, že naše analýza je založena na klientech, kteří program skutečně využívají. Nevíme však, zdali to jsou skutečně všichni

potřební. Možná by jich ve skutečnosti bylo daleko více, avšak možnosti agentury pokrýt potřebu jsou jen omezené, takže uspokojuje jen část populace *in need*.

Sociální indikátory

To je přístup, který již velmi dobře známe — využívá existujících agregovaných statistik, které mohou indikovat problém. Např. je-li v nějakém regionu vyšší kojenecká úmrtnost než je průměr v dané zemi, je to znamení, že něco zde není v pořádku a region se s tává z tohoto hlediska potenciálním “klientem” intervenujícího programu (lékařského, sociálního, apod.).

Průzkum cílové populace

Nejpřímějším způsobem, jak zjistit charakteristiky o potřebách z hlediska existujících sociálních problémů a jejich percepci populací, je uskutečnit průzkum na reprezentativním vzorku této populace. Pokud se jedná o menší region, připadá v úvahu i census.

Jako všechny výzkumné postupy, i v případě hodnocení potřeb je třeba raději kombinovat více naznačených postupů, abychom dostali jasnější a spolehlivější obrázek o situaci.

Hodnotící výzkum je dnes velmi významnou výzkumnou činností ve společenských vědách. Největšími evaluačními společnostmi jsou Amerika, Kanada a Austrálie. V roce 1995 se konala první mezinárodní evaluační konference.

Zpracováno podle:

- Rubin, A., Babbie, E. 1989. *Research Methods for Social Work*. Wadsworth Publishing Company, Belmont, California, str. 480–507.
- Mark, R. 1996. *Research Made Simple*. A Handbook for Social Workers. Sage Publications, str. 230–240.
- Babbie, E. 1998. *The Practice of Social Research*. 8th Edition. Wadsworth Publishing Company, str. 333–355.
- Mertens, D. M. *Research Methods in Education and Psychology*. SAGE Publications, London 1998, kap. 9: „Evaluation“.