

Sonja Kump

PARADIGME, KONCEPTI IN METODE V EVALVACIJI IZOBRAŽEVANJA

Vir: Zbornik Evalvacija (ur. Darko Štrajn), Pedagoški inštitut, 2000

1. Opredelitev in namen evalvacij

Definicij evalvacije je veliko, mnoge med njimi opredeljujejo evalvacijo kot sistematično uporabo metod družboslovnega raziskovanja za ocenjevanje načrta, implementacije in rezultatov ter učinkovitosti programov, politik oziroma opazovane enote (Rossi, Freeman, 1993). Nekateri avtorji k tej opredelitvi dodajajo, da je evalvacija sistematična, kritična analiza, katere namen je lahko odločanje in strateško načrtovanje (Patton, 1987; Toulemonde, 1995) ali izpopolnjevanje (Kristoffersen, Surssock, Westerheijden, 1998).

Evalvacija lahko poteka na različnih ravneh. V izobraževanju so najbolj pogoste evalvacije na ravni izobraževalnega programa, v zadnjem času tudi na ravni izobraževalne institucije.

Številni avtorji evalvacijo uvrščajo med uporabne družboslovne raziskave. Kot oblika uporabne raziskave se evalvacija od temeljnih raziskav, ki so usmerjene v razvoj teorij, razlikuje po tem, da je usmerjena v spreminjanje obstoječega (Richardson in drugi v Macur, 1996). Predmet evalvacije, kot uporabne raziskave, je unikaten. Znanje, ki ga generiramo, je časovno in prostorsko omejeno, v nasprotju s temeljno raziskavo, kjer je osrednjega pomena pojav in ne situacija. Predmet evalvacije se med raziskovanjem nenehno spreminja, zato evalvacije ni priporočljivo togo strukturirati že na začetku.

Evalvacije izobraževanja so bile v zgodnejših fazah razvoja namenjene predvsem učiteljem, ki naj bi s pomočjo evalvacije izboljšali prakso poučevanja. Od osemdesetih let dalje se moč in kontrola znotraj izobraževanja hitro spreminjata, spreminjajo pa se tudi vloge in pričakovanja. Evalvacije so še vedno namenjene učiteljem, toda poleg njih tudi učencem, vodstvu šol, staršem, delodajalcem, državnim administraciji in politikom izobraževanja. Prevladujoči okvir evalvacij so reforme in inovacije v

izobraževalnih sistemih. Danes politična retorika razglša odgovornost šol in okrepljeno moč managementa v izobraževanju in v tej zvezi postaja evalvacija ena od metod reguliranja izobraževalne dejavnosti.

2. Zgodovinski razvoj evalvacij v izobraževanju

Evalvacije so se, poleg zdravstva, najprej pojavile na področju izobraževanja. V zgodovinskih pregledih evalvacij na področju izobraževanja so po enem viru (Guba, Lincoln, 1989) prve evalvacije začeli izvajati pred sto leti, po drugem viru (Madaus, Stufflebeam, Scriven, 1993) pa celo pred približno dvesto leti, z obdobjem industrijske revolucije in velikih družbenih reform.

Po opisu Madausa, Stufflebeama in Scrivena (1993) je *obdobju reforme* (1800-1900), ko je bila v Ameriki izvedena prva formalna evalvacija izobraževalnih programov, sledilo *obdobje učinkovitosti in testiranja* (1900-1930). To obdobje je bilo povezano z idejo znanstvenega managementa, ki je poudarjala sistematizacijo, standardizacijo in učinkovitost. *Tretje obdobje je povezano z R. Tylerjem* (1930-1945), pionirjem evalvacije v izobraževanju, ki je prekinil tradicijo dragih primerjav eksperimentalne skupine s kontrolno in v evalvacijo uvedel primerjanje nameravanih učinkov z dejanskimi. *Za četrto, t.i. obdobje nedolžnosti* (1946-1957) je značilno ignoriranje evalvacij. Temu sledi *obdobje ekspanzije evalvacij* (1958-1972) in končno *obdobje profesionalizacije* od leta 1973 dalje, ko se je evalvacija razvila v posebno dejavnost, ločeno od znanstvenih raziskav in testiranja.

Guba in Lincoln (1989) sta v svojem zgodovinskem pregledu evalvacije razdelila v štiri generacije. *Prva generacija evalvacij* izobraževanja je bila povezana z družbenim načrtovanjem in nadzorom. Pod vplivom Fredericka Taylorja, ki je leta 1911 objavil knjigo *Načela znanstvenega managementa*, so bili progresivni reformatorji izobraževanja prepričani, da bo do družbenega napredka prišlo le, če bo politično upravljanje izobraževanja in drugih družbenih storitev nadomeščeno z znanstvenim managementom, skratka, da bodo nadzor nad izobraževanjem, namesto politikov, prevzeli profesionalci. Tehnologija merjenja izobraževanja naj bi zagotovila načrte za novi družbeni red in mehanizme povratnih informacij za učinkovito vzdrževanje družbe. Do tridesetih let je bila evalvacija sinonim za testiranje. Za prvo generacijo evalvacij je bilo torej značilno merjenje. Tipičen instrument merjenja so bili šolski testi, s katerimi so ocenjevali uspešnost učencev. Evalvator je imel predvsem tehnično vlogo.

Na evalvacije izobraževanja prve generacije je imela velik vpliv psihometrija. Z merjenjem individualnih razlik naj bi psihometrija zagotavljala tehnologijo za ocenjevanje učinkovitosti in uspešnosti različnih učiteljev, šol, programov in drugih oblik izobraževalne ponudbe. Čeprav je bila uporaba podatkov, pridobljenih s standardiziranimi testi pogosto kritizirana, je ponekod še danes prisotno prepričanje, da so takšni podatki prepričljiv dokaz učinkovitosti šole ali učitelja. Prvo generacijo evalvacij uporabljajo tisti, ki evalvacijo razumejo v tehničnem smislu, torej le kot merjenje.

Glavna poteza *druge generacije evalvacij* je bilo opisovanje. Prvič se je pojavil pojem programske evalvacije, v kateri je imel evalvator vlogo opisovalca vzorcev prednosti in slabosti glede na določene cilje programa. Merjenje je v evalvaciji ostalo, toda samo kot eno od orodij evalvacije. Najbolj tipičen model druge generacije evalvacij je ciljni model, ki ga je razvil Ralph Tyler. V ZDA so leta 1932 začeli izvajati osemletno evalvacijsko študijo, ki je bila usmerjena v ugotavljanje uspešnosti različnih tipov šolanja. Tyler, vodja projekta, je evalvacijo preusmeril od osredotočenosti na individualne sposobnosti k oblikovanju kurikuluma. Evalvacijo ni razumel kot tehnologijo razločevanja med posamezniki, ampak bolj kot sredstvo za ocenjevanje stopnje, do katere so bili cilji in nameni kurikuluma uresničeni v praksi. Testiranja ni ovrgel, toda najprej je opredelil cilje in željene spremembe v vedenjskih vzorcih učencev. Učitelj je potem lahko s pomočjo evalvacije določil stopnjo, do katere so se željene spremembe v vedenju učencev resnično zgodile. Testiranje je postalo le eden od načinov za zagotavljanje evidence o vedenju učencev.

Čeprav je bil Tyler pod vplivom progresivnih idej znanstvenega managementa, so bila njegova zgodnja dela tesno povezana z razredom in s tem, kaj lahko za izboljšanje programa naredi posamezni učitelj. Toda sčasoma je prišlo do premika od evalvacije kurikuluma za razrednega učitelja k evalvaciji kot instrumentu politike in javne odgovornosti. Ker se je moč odločanja o programu premaknila od razreda k politiki in javnosti, se je spremenilo tudi pojmovanje utilitarnosti evalvacij.

Kljub temu, da je Tylerjev pristop k oblikovanju programa in evalvacije pritegnil številne posnemovalce in ostaja vpliven še danes, so mnogi avtorji njegov model kritizirali. Tako je npr. Elliot Eisner menil, da so določena področja vednosti, zlasti umetnost, neprimerna za vnaprejšnje specifikacije rezultatov, kajti

ustvarjalnost je po naravi nepredvidljiva (Eisner v Madaus, Stufflebeam, Scriven, 1993). Lee Cronbach je v svojem delu *Izboljšanje programa s pomočjo evalvacije* opozoril, da postaja evalvacija preveč odvisna od rutin in ritualov testiranja (Cronbach v Madaus, Stufflebeam, Scriven, 1993). Po njegovem mnenju naj bi evalvacija igrala pomembno vlogo v procesu izpopolnjevanja programa, ne pa, da je njena vloga končna ocena. Prepričan je bil, da so rezultati poučevanja vedno multidimenzionalni in zato jih je neustrezno opisati le s študijami, ki združujejo različne vrste izvedbe poučevanja v posamezne testne točke.

Najbolj oster kritik ciljnega modela evalvacije je bil Lawrence Stenhouse. Prepričan je bil, da izpopolnjevanje kurikulumu ne more potekati brez izpopolnjevanja učitelja (Stenhouse v Norris, 1993). Ciljni model je s povečano jasnostjo rezultatov izobraževanja poiskoval izboljšati izobraževalno prakso, toda izkušnje so pokazale, da se s tem pristopom praksa ni izboljšala. Stenhouse je menil, da je najbolj učinkovit način izboljšanja prakse možen le s pomočjo učiteljeve kritike prakse. Predlagal je strategijo, ki temelji na učitelju kot raziskovalcu, ne pa na zunanji evalvaciji kot instrumentu prenove kurikulumu in kontrole kakovosti.

Po mnenju Norrisa je ciljni model evalvacije oblika empiricizma v metodologiji evalvacije (Norris, 1993). Predpostavlja študijski program, ki naj bi izražal opazne in izmerljive napovedi študentovega vedenja. Ciljni model s perspektive razvijalcev programov pomeni racionalni pristop k načrtovanju kurikulumu. Toda ciljni model je več kot le tehnika evalvacije, saj je predpisujoč glede narave znanja in njegovega pridobivanja.

Presojanje je značilnost *tretje generacije evalvacij*, ki se je pojavila po izstrelitvi Sputnika leta 1957. V ZDA je narastla zaskrbljenost glede tehnološkega zaostanka. Američani so namreč sklepali, da so Rusi v prednosti pri odkrivanju vesolja ravno zaradi uspešnega izobraževalnega sistema. Prepričani so bili, da ciljno usmerjena opisna evalvacija ni razkrila pomanjkljivosti ameriškega izobraževanja in zato so v evalvacijo vključili presojo. Evalvator je poleg tehnične in deskriptivne vloge dobil še vlogo razsodnika. Ameriška vlada je izdatno povečala finančno podporo za izpopolnitev izobraževalnih programov in za evalvacijske študije, s pomočjo katerih naj bi spremljali in vodili razvoj programov v izobraževanju. Drugi pomembni mejnik evalvacijskih študij v ZDA je bil sprejem Zakona o osnovnošolskem in srednješolskem izobraževanju leta 1965, ki je vključeval elemente evalvacije. Temu so sledile obsežne evalvacije

izobraževanja na zvezni in lokalnih ravneh. Mnoge evalvacije so se izkazale za nekoristne, toda apetiti zveznih agencij za evalvacijo se niso zmanjšali, ampak ravno obratno: evalvacija je postala velik posel.

Sredi sedemdesetih let je bila v Cambridgu organizirana konferenca o alternativnih evalvacijskih pristopih. Sklepi udeležencev konference so bili objavljeni v publikaciji *Beyond the Numbers Game* (Hamilton, et al., 1977). Prispevki različnih avtorjev so predstavljali paradigmatičen premik od evalvacijske metodologije, ki vrednoti numerično, k opisnemu vrednotenju (Norris, 1993). Poudarek je bil na bolj humanističnih raziskovalnih tradicijah. V tem času se pojavijo *evalvacije četrte generacije*, zlasti responzivna evalvacija, ki predstavlja pragmatičen pristop kot nasprotje metodološkemu dogmatizmu. Različni teoretiki evalvacije so poudarjali različne vidike responzivnosti. Robert Stake, avtor pojma "responzivna evalvacija", je poudaril pomen personalizacije in humanizacije evalvacijskega procesa (Stake, 1976). Evalvacijo je opredelil kot interaktiven, pogajalski proces, ki vključuje vse udeležence (stakeholders). Responzivnost zahteva neposredni stik z ljudmi, partnerski odnos med vsemi udeleženi v programu. Po mnenju Stakea ta pristop žrtvuje nekaj natančnosti pri merjenju, v upanju, da bo povečal uporabnost izsledkov za osebe, ki so povezane s programom.

Evalvacije v 20. stoletju so bile pogosto utemeljene z reformami izobraževalnih sistemov, ponekod, kot npr. v ZDA, so spodbude za evalvacije izhajale iz zakonodaje. V šestdesetih in sedemdesetih letih so bile evalvacije povezane s problemi pravičnosti in enakosti, t.j. s prepričanjem, da lahko z redistribucijo izobraževalnih možnosti vplivamo na spreminjanje družbene strukture, s tem pa zagotavljamo človeške vire za trajno ekonomsko rast. Številne evalvacije, zlasti v ZDA in Veliki Britaniji, so imele majhen neposredni vpliv na izobraževalno politiko. Po mnenju Norrisa pa so tam, kjer so imele vpliv, bolj komplicirale kot racionalizirale politične procese (Norris, 1993). Obsežne evalvacije izobraževanja v šestdesetih in sedemdesetih letih so bile na splošno neuspešne pri zagotavljanju jasne evidence o učinkovitosti programov. S spreminjanjem politične klime se je spreminjala tudi klima za evalvacijo. Od osemdesetih let dalje narašča politična pomembnost odgovornosti in porast managerializma, kar predstavlja resne ovire za formativne in demokratične namene evalvacije.

3. Metode, ki izhajajo iz dveh osnovnih paradigem

Večna tema razprav v evalvacijskem raziskovanju je dilema, ali uporabljati kvantitativne ali kvalitativne metode. Sodobne evalvacije so v glavnem kombinacija obeh metod, sam problem pa je običajno globlji, saj zadeva osnovni paradigmi družboslovnega raziskovanja. Pri tem je potrebno opozoriti, da meje med paradigmo, pristopom in metodo pri različnih avtorjih niso vedno jasne (Macur, 1996). V literaturi o evalvaciji lahko zasledimo dve prevladujoči paradigmi, t.j. pozitivistično in konstruktivistično, ki predstavljata dve različni izhodišči evalvacije in vodita do različnih pričakovanj o njenih rezultatih.

Prvo vrsto paradigme, t.j. *pozitivistično paradigmo*, različni avtorji označujejo s številnimi izrazi, kot so npr.: klasično znanstvena, tradicionalna (Parlett in Hamilton, 1977), racionalistična (Guba, Lincoln, 1989), racionalno-tehnična (Everitt, 1996), pozitivistična (Scriven, 1993) ali eksperimentalna paradigma (Radaelli in Dente, 1996).

Pozitivistična paradigma temelji na predpostavkah in metodologiji pozitivističnega koncepta znanosti, namen evalvacije pa je merjenje in napovedovanje rezultatov. Iz zgodovinskega razvoja je poznano, da so družboslovne discipline zaradi prevlade dominantne znanstvene kulture začele posnemati znanstvene, naravoslovne metode, ne glede na popolnoma drugačno naravo vednosti. V tistem času so se pojavile prve evalvacije, ki so se ukvarjale predvsem z merjenjem. Najbolj tipična metoda je bil (kvazi) eksperiment z naključnim vzorčenjem. Pri eksperimentu se naključno vnaprej izberejo posamezniki iz ciljne skupine, ki so potem razvrščeni v eksperimentalno in kontrolno skupino. Pri kvaziekperimentu pa je kontrolna skupina oblikovana tako, da je čimbolj podobna eksperimentalni skupini. Predpostavka teh metod je, da so rezultati merljivi in preverljivi, kar pomeni, da lahko tudi drugi znanstveniki z izvedbo določenega eksperimenta pridejo do istih rezultatov, zanemarjene pa so subjektivne definicije udeležencev programa (izvajalcev, uporabnikov, naročnikov itd.).

Bistvo evalvacijske raziskave, ki izhaja iz pozitivistične paradigme, je odkriti vzročne povezave ali testirati hipoteze o njih. Za pozitivistično paradigmo je značilen dualizem med evalvatorjem in objektom evalvacije. Evalvator zavzame objektivno, vrednostno nevtralno stališče o predmetu evalvacije. Objektivna, zunanja drža evalvatorja je pogoj objektivnosti spoznanja, vrednostna nevtralnost pa je možna, ker so izločene vrednote evalvatorja ali kateregakoli drugega udeleženca (izvajalca, udeleženca programa itd.).

Tudi druga vrsta paradigme, t.j. *konstruktivistična paradigma*, ima več izrazov: alternativna, naturalistična (Guba, Lincoln, 1989), konstruktivistična (Radaelli in Dente, 1996), respozivna (Stake, 1993), relativistična (Scriven, 1993), interpretativna (Everitt, 1996) ali socialno antropološka oziroma iluminativna paradigma (Parlett in Hamilton, 1977).

Druga vrsta paradigme zagovarja pristop, ki temelji na spoznanjih iz humanističnih ali antropoloških ved. Temelji na opisovanju in interpretaciji, ne pa na merjenju in napovedovanju. Evalvator uporablja predvsem kvalitativne metode, njegova naloga pa je, da opiše in interpretira različne perspektive različnih udeležencev (izvajalcev, uporabnikov, naročnikov evalvacije). Priljubljene metode so kvalitativne, kot na primer nestrukturirani intervju.

Konstruktivistična paradigma temelji na naturalizmu (zato tudi naturalistična paradigma), ki poudarja, da so družbeni pojavi drugačni od fizičnih, zato se jih ne da pojasnjevati z uporabo znanstvenih metod naravoslovnih disciplin. Pri raziskavi je potrebno prikazati konstrukcije (stališča) različnih strani, ki so vpletene v evalvacijo. Konstruktivistična paradigma predpostavlja obstoj večplastne, družbeno konstruirane realnosti, ki je ne vodijo naravni zakoni (Guba, Lincoln, 1989). Zavrača obstoj ene, objektivne realnosti ali resnice, ki bi jo bilo moč odkriti s pomočjo znanstvenega raziskovanja, kajti posamezniki in skupine konstruirajo svojo lastno, subjektivno realnost. Zaradi tega konstruktivistična paradigma izhaja iz prepričanja, da obstajajo različne subjektivne realnosti. Od znanstvene paradigme se razlikuje tudi glede pojmovanja vrednot; vrednote ne moremo preprosto izločiti iz raziskovalnega procesa. Po mnenju Gube in Lincolna raziskovanje upošteva vrednote na več načinov: skozi vrednote raziskovalca, ki se kažejo v izbiri problema in fokusu raziskovanja; z izbiro paradigme, ki vodi raziskovanje do problema; z izbiro teorije, ki jo uporabimo pri zbiranju, analizi in interpretaciji rezultatov in z vrednotami, ki so neločljivi del konteksta (Guba, Lincoln, 1993).

Room je opozoril, da ima izbira ene ali druge paradigme posledice na celoten proces evalvacije (Room v Macur 1996):

* Pozitivisti obravnavajo predmet evalvacije kot neodvisno spremenljivko, medtem ko predstavniki konstruktivistične paradigme predpostavljajo, da predmet evalvacije raste in se razvija v določenem okolju, ki ima na predmet odločujoči vpliv. Ta vpliv je prav tako pomemben kot sam predmet evalvacije, merimo

pa ga lahko šele po implementaciji programa. Naloga evalvatorja je osvetlitev procesov delovanja konfliktov, ki se odvijajo med okoljem in predmetom evalvacije.

* Pozitivisti proučujejo proces, da bi določili, do katere mere je bil program uresničen. Pri tem ne upoštevajo, da se cilji skozi življenjsko obdobje programa lahko spreminjajo. Za konstruktivistično paradigmo pa predstavlja reorientacija (zamenjava ciljev ali odklon od prvotno zastavljenih ciljev) ključni element obravnave, pri čemer je še posebej pomembna utemeljitev, zakaj je do spremembe ciljev prišlo.

* Paradigmi se razlikujeta tudi po presojanju veljavnosti zaključkov evalvacije. Veljavnost je v prvem primeru zagotovljena le z doslednostjo pridobivanja kvantitativnih podatkov. Konstruktivistična paradigma je do tega pristopa kritična zato, ker sklepi evalvacije posplošujejo rezultate in ne upoštevajo posebnosti programa. Konstruktivistična paradigma pripisuje večji poudarek orisom posameznih parametrov.

Evalvacijo kakovosti izobraževanja torej lahko izvedemo s kvantitativnimi in kvalitativnimi metodami ter pri tem ugotovljamo kvantiteto in kvaliteto spremembe. Kvantiteta spremembe lahko vključuje npr. število prebranih knjig, točke na standardiziranem testu dosežkov, število pravilno izgovorjenih besed, število odnosov z drugimi učenci ali učitelji. Vsak od teh rezultatov ustreza dimenziji kakovosti, ki zahteva opis namesto merjenja. To pomeni npr. opis, kako so prebrane knjige vplivale na učenca - osebno in intelektualno. Namesto števila stikov z učitelji lahko opišemo, kaj ti stiki učencem pomenijo. Odgovori na takšna vprašanja zahtevajo opis učenčeve perspektive in okoliščin.

Če evalvator kombinira uporabo kvantitativnih in kvalitativnih metod, se izogne nevarnosti, da bo vse svoje prizadevanje osredotočil le na zbiranje števil, ki naj bi po mnenju nekaterih predstavljale natančnost in točnost. Resnično pomembna vprašanja v evalvaciji so: Kaj je vredno vedeti o programu? Kateri podatki bodo najbolj uporabni? Kateri evalvacijski model bo najbolj ustrezal evalvacijski situaciji in razvojni stopnji programa?

Najbolj pogosta kritika kvalitativnih metod se nanaša na subjektivnost evalvatorja. Znanost pripisuje veliko vrednost objektivnosti in resnici. Prizadevanje za resnico predvideva en sam pravilni odgovor. Toda kvalitativne metode dopuščajo raznovrstne perspektive in raznovrstne "resnice" glede na različna mnenja. Kako lahko v takšnih primerih en odgovor ustreza objektivnosti in resnici? Danes filozofi znanosti dvomijo

o možnosti, da bi bila lahko katerakoli metoda resnično "objektivna". Filozof in evalvator Scriven z veliko mero prepričljivosti trdi, da kvantitativne metode niso več sinonim za objektivnost, prav tako pa tudi kvalitativne metode niso sinonim za subjektivnost (Scriven v Patton, 1987). Sestavljanje testov in vprašalnikov ni nič manj odprto možnosti vdora evalvatorjeve pristranosti, kot pri postavljanju vprašanj v intervjuju ali opazovanju pojava. Številke ne pomenijo zaščite pred pristranostjo; včasih jo celo prikrivajo. Vsi statistični podatki temeljijo na opredelitvi nekoga o tem, kaj se bo merilo in kako se bo merilo. "Objektivna" statistika je vedno oblikovana na subjektivnih odločitvah, katere postavke bodo vključene. Guba (1978) in drugi (npr. House, 1980) so pojma subjektivnost in objektivnost nadomestili s pojmom "nevtralnost" evalvatorja. Nevtralni evalvator vstopa v evalvacijo brez skritega namena, brez teorije, ki naj bi jo dokazoval, in brez vnaprej pričakovanih rezultatov, ki naj bi jih podpiral. Tradicionalno prizadevanje za resnico so nadomestili s prizadevanjem za uporabnost in koristnost, objektivnost pa so nadomestili s poštenostjo in natančnostjo pri upoštevanju mnogovrstnih perspektiv, interesov in možnosti.

4. Vrste, modeli evalvacije

V literaturi srečamo veliko različnih modelov in vrst evalvacije. Evalvacije so se namreč avtonomno razvijale na mnogih področjih (Macur, 1996). Strokovnjaki različnih ved so zato pri evalviranju uveljavili svoje metode: sociologi študijo primera, ekonomisti cost-benefit analizo, psihologi eksperiment itd., vse pa so postale "evalvacijske" metode. Toda večina modelov izhaja iz izobraževanja, saj lahko na tem področju zasledimo največ primerov evalvacij. Različne taksonomije praviloma razvrščajo evalvacijske modele od pozitivističnih, začenši z eksperimentalno evalvacijo, do konstruktivističnih, kot je npr. responzivna evalvacija. Vsaka evalvacija ima posebne značilnosti. Čeprav meja med njimi ni ostro začrtana (včasih lahko uporabimo kombinacijo modelov) in kljub temu da se tudi posamezni model v različnih okoliščinah lahko razlikuje, so med evalvacijami vendarle razlike.

Znotraj pozitivistične paradigme zasledimo eksperimentalni model ali tudi funkcionalistični model.

Eksperimentalni model, ki se je razvil v prvi generaciji evalvacij, je opisan kot merjenje. Uporablja načela eksperimentalne znanosti na področju evalvacije socialnih programov. Cilj evalvacije je priti do splošnih zaključkov o učinkih določenega programa s kontroliranjem zunanjih spremenljivk in

izoliranjem vpliva programa. Evalvator uporablja kvantitativne metode in rezultate predstavlja s pomočjo kvantitativnih podatkov. Najpomembnejši element tega modela je tehnična veljavnost njenih sklepov in možnost posploševanja. Evalvacija temelji na konsenzu glede ciljev programa. Da bi ugotovili vzročno-posledični odnos med programom in njegovimi učinki, je potreben dober eksperimentalni načrt.

Znotraj konstruktivistične paradigme zasledimo naslednje vrste evalvacije: študija primera, responzivna evalvacija, demokratična evalvacija, pluralistična evalvacija, iluminativna evalvacija in naturalistična evalvacija. Skupna poteza naštetih evalvacij je upoštevanje različnih perspektiv vseh ljudi, ki so vpleteni v program.

Študija primera je usmerjena v sam proces programa in v različne poglede na program. Ima lastnost naturalistične ali konstruktivistične paradigme. Študija primera je običajno podroben opis in analiza programa ali institucije v kontekstu njenega okolja. House meni, da je osnovno vprašanje študija primera "Kakšen se program zdi različnim ljudem, ki ga poznajo?" (House, 1980). Standardna metodologija so intervjuji z različnimi udeleženci, opazovanje kraja izvajanja in predstavitev rezultatov študije primera. Metode in predstavitve so torej skoraj izključno kvalitativne. Namen tega pristopa je izboljšati razumevanje bralca ali občinstva evalvacije predvsem s tem, ko jim prikažemo, kako drugi razumejo program, ki se evalvira. Evalvator se običajno pogaja z različnimi udeleženci glede tega, kaj je potrebno storiti za izboljšanje kakovosti.

Patton je *responzivno evalvacijo* označil kot premik od hipotetično-deduktivne paradigme k paradigmi izbire, s poudarkom na mnogovrstnosti metod in alternativnih pristopov (Patton, 1981). Poudaril je, da mora izbrana evalvacijska metoda ustrezati specifični evalvacijski situaciji. Patton je "paradigmo izbire" oz. alternativno paradigmo opisal kot situacijsko responzivnost, s poudarkom na kreativnih vidikih v evalvaciji, v nasprotju z evalvacijskim modelom, ki je bil voden po pravilih.

Guba in Lincoln sta naturalistični pristop in responzivno evalvacijo povezala v celoten okvir za izboljšanje koristnosti evalvacijskih rezultatov (Guba, Lincoln, 1982). Odprtost naturalističnega pristopa omogoča evalvatorju, da je še posebej občutljiv za različne perspektive raznih udeležencev. Ta senzitivnost evalvatorju omogoča, da zbira podatke in poroča o ugotovitvah, pri tem pa ima jasno v mislih različne

perspektive udeležencev. Dolžnost evalvatorja je identificirati vse javnosti, ki imajo svoj delež ali interes v določeni dejavnosti, in jim omogočiti, da opredelijo svoje probleme oziroma vprašanja, ki jih zanimajo. Po prepričanju Gube in Lincolna je pošteno in pravično, da ima vsak soudeleženi svoj glas v evalvaciji. Menila sta, da je evalvacija storitev za udeležence programa. Vsakemu udeležencu mora biti jasno, *kaj* bo evalvacija obravnavala, *komu* je namenjena in *kakšne metode* bodo uporabljene.

Responzivni evalvator ne verjame, da obstaja en sam odgovor na vprašanje v zvezi s programom. Njegov namen je razumeti program z različnih vidikov. Pri tem uporablja kvalitativne metode raziskovanja: neposredno in posredno opazovanje dogodkov in deskriptivno opisovanje podatkov. Responzivni evalvator interpretira program z vidika uporabnika, izvajalca, odločevalca, sponzorja in drugih udeležencev programa. Prednost responzivne evalvacije je njena občutljivost za različna stališča in možnost njihovega medsebojnega soočanja. Guba in Lincoln se od drugih avtorjev razlikujeta po tem, da poudarjata nujnost soočanja različnih udeležencev, pri čemer pa naj bi bil rezultat evalvacije konsenz, ki ga med seboj dosežejo. Slabost responzivne evalvacije je v pomanjkanju prioritete pri opazovanju in posredovanju rezultatov.

Demokratično evalvacijo je MacDonald, ki je avtor izraza, opredelil kot informiranje celotne skupnosti o značilnostih izobraževalnih programov. Financer evalvacije ne postavlja nikakršnih zahtev (MacDonald, 1976). Demokratični evalvator priznava vrednostni pluralizem in si prizadeva predstaviti različne interese. Osnovna vrednota evalvacije so informirani državljani. Evalvator deluje kot borzni posrednik v izmenjavi informacij med različnimi skupinami. V vlogi "brokerja" izmenjuje informacije med različnimi skupinami, ki bi rade vedele več ena o drugi. Tehnike zbiranja in predstavitve podatkov morajo biti razumljive in dostopne nespecializirani javnosti. Ključni pojmi demokratične evalvacije so "zaupnost", "pogajanje", "dostopnost", "pravica vedeti".

Helen Simons je demokratično evalvacijo opisala kot sredstvo za izenačevanje odnosov moči, znotraj katerih potekajo programi (Simons, 1987). Po njenem mnenju predstavlja demokratična evalvacija napad na avtoriteto znanosti in pravico raziskovalcev, da postavljajo dokončne sodbe o družbenem svetu. Metode in poročila evalvacije naj bi bile dostopne nespecializirani javnosti. S tem se demokratizira dostop do procesa in produktov evalvacije.

Pluralistična evalvacija prepozna in upošteva interese različnih skupin, vpletenih v program. Uspeh programa je pluralističen pojem, saj ga ocenjujemo po različnih kriterijih glede na različne pomene, ki so pripisani tem kriterijem.

Iluminativna evalvacija je bila reakcija na psihometrično tradicijo in predstavlja vnos humanističnih vrednot namesto tehnoloških potez evalvacije. Parlett in Hamilton sta iluminativno evalvacijo opisala kot "paradigmo za ljudi in ne za rastline" (Parlett, Hamilton, 1977). Pri iluminativni paradigmi so pomembni tudi kvantitativni podatki, a sami niso dovolj. Vedno so postavljeni v določen kontekst, ki ga je potrebno natančno opisati.

Pri *naturalistični evalvaciji* je poudarek na informiranju javnosti, ne pa le na izbranih skupinah. Osnovna vrednota naturalistične evalvacije je participatorna demokracija, druga vrednota je odločanje, ki temelji na informacijah in vodi v izboljšave, tretja vrednota je vrednostni pluralizem.

Poleh naštetih vrst evalvacije lahko v strokovni literaturi zasledimo še druge modele in vrste evalvacije.

Model ciljno usmerjene evalvacije jemlje cilje programa kot določene in skuša izmeriti, v kolikšni meri so bili cilji uresničeni. Tako cilji (tisti, ki jih postavijo oblikovalci programov) postanejo kriterij uspeha. "Jasni programski cilji so ekvivalent hipotezi v raziskavi" (House, 1980). Uspeh programa torej merimo z (ne)skladnostjo učinkov z danimi cilji. Cilji pa se ne nanašajo le na končne učinke, ampak poznamo tudi vmesne - procesne cilje. Ta model se zdi podoben eksperimentalnemu modelu, a med njima obstaja ključna razlika. Pri eksperimentalni evalvaciji izobraževalnih programov lahko izoblikujemo teste, s katerimi merimo uspešnost različnih programov. Ciljno usmerjena evalvacija pa upošteva definirane cilje posameznega programa in ocenjuje, ali so bili posamezni cilji doseženi.

Model evalvacije, usmerjene na odločanje, strukturira evalvacijo glede na konkretno odločitev, ki jo je potrebno sprejeti, kar običajno pomeni odločitve politikov izobraževanja ali vodstva šol. Predpostavka tega modela je, da bo takšna evalvacija povečala izkoriščenost njenih rezultatov. Med metodami se uporabljajo

vprašalniki ali intervjuji, evalvator daje prednost proučevanju naravne variacije v programskem okolju tako, kot da bi organiziral eksperiment.

Model uporabniško usmerjene evalvacije (ali model evalvacije brez ciljev in tudi ocena potreb) se je razvil kot odgovor na ciljno usmerjeno evalvacijo. Scriven je prepričan, da evalvator, ne le da ne bi smel evalvacije utemeljiti na programskih ciljih, ampak bi moral biti namenoma neinformiran o ciljih programa. S tem bi se izognil pristranosti. Na ta način lahko evalvator meri vse učinke programa - nameravane in stranske, ki so lahko tako pozitivni kot negativni (Scriven v House, 1980). Patton navaja štiri vzroke za uporabo tega pristopa: izognemo se tveganju, da bi proučevali le ozko določene programske cilje, pri tem pa zgrešili pomembne nepredvidene učinke; odstranimo negativno konotacijo, ki je pripisana odkritju nepredvidenih učinkov; izključimo pristranost v dojetanju, ki jo v evalvacijo prinese poznavanje ciljev; skozi pogoje osvobojenosti od ciljev ohranimo objektivnost in neodvisnost evalvatorja (Patton, 1990).

Formativna in sumativna evalvacija. Formativna ali tudi procesna evalvacija evalvira sam proces izvajanja programa ali določenega političnega ukrepa, v nasprotju s sumativno evalvacijo, ki evalvira končne rezultate programa ali ukrepa. Namen formativne evalvacije je izpopolnjevanje programov, namen sumativne evalvacije pa je ugotavljanje učinkovitosti programov, ki vodi v odločitve, ali naj se programi ukinejo ali nadaljujejo.

Formativna evalvacija uporablja kvalitativne metode. Zagotavlja podrobnosti o prednostih in pomanjkljivostih programa: Kaj deluje in kaj ne deluje dobro? Kako program doživljajo udeleženci in kako učitelji? Formativne evalvacije so zelo koristne v zgodnjih fazah programa, ko se pričakuje, da se bo program spreminjal in razvijal. Učitelji si prizadevajo uporabljati nove pristope in se prilagajati potrebam udeležencev programa. Formativna evalvacija lahko zagotavlja povratne informacije o procesu in učinkih, ki jih ima program na udeležence. V tem primeru je zelo primeren naturalistični evalvacijski pristop, ki evalvatorju omogoča spremljanje razvoja programa.

Notranja in zunanja evalvacija. Ta delitev se nanaša na izvajalce evalvacije. O zunanji evalvaciji govorimo takrat, ko jo izvajajo vladne komisije ali ustanove, različni neodvisni inštituti ali agencije, notranjo evalvacijo (samoevalvacijo) pa izvajajo sami izvajalci programov s pomočjo zunanjih izvedencev. Zunanjo

evalvacijo torej izvaja evalvator, ki ni iz institucije, kjer se izvaja ali oblikuje program - običajno so to izvedenci iz univerzitetnih, vladnih ali nevladnih inštitutov. Notranjo evalvacijo izvajajo izvajalci programov, skratka ljudje iz institucij, kjer se program izvaja. Prednost notranje evalvacije je temeljita seznanjenost evalvatorja s programom, ki ga bo evalviral. Prednost zunanje evalvacije pa je morda večja usposobljenost izvedencev za evalvacijo, čeprav se mora tudi zunanji evalvator dobro seznaniti s predmetom evalvacije in udeleženci programa.

5. Načrtovanje evalvacije

Formule, kako izvesti evalvacijo, ni, ker preprosto ne obstaja en sam ali najboljši načrt. Konkretni evalvacijski načrt je odvisen ne le od izbranega modela, ampak tudi od razpoložljivih virov, časa in sredstev, ki so na voljo, pa tudi od kreativnosti evalvatorja. Patton zato govori o umetnosti evalvacije (Patton, 1987). Priporoča, da se v procesu načrtovanja evalvacije najprej obravnavajo konceptualni in tehnični problemi evalvacije. Konceptualni problemi so osredotočeni na naslednja vprašanja, ki zadevajo razmišljanje udeležencev o evalvaciji:

- * Katere so glavne skupine, ki so zainteresirane za evalvacijo?
- * Kakšen je namen evalvacije?
- * Kateri pristop, model ali okvir evalvacije bo uporabljen?
- * Katera so osnovna evalvacijska vprašanja?
- * Katere politične premisleke naj bi upoštevali?
- * Kakšni viri so na razpolago za evalvacijo?

Na našeta vprašanja ne morejo odgovoriti evalvatorji in metodologi. Odgovori na konceptualna vprašanja se lahko oblikujejo le v okviru pogajanj med vsemi zainteresiranimi udeleženci v evalvaciji. Evalvator lahko takšna pogajanja spodbuja, lahko je tudi eden od udeležencev v pogajalskem procesu, toda odgovorov na konceptualna vprašanja naj ne bi nikoli enostransko narekoval. V nasprotnem primeru lahko pride do "kratkega stika" v procesu oblikovanja političnega in konceptualnega temelja za evalvacijo, ki je odločilen za kredibilnost in uporabnost evalvacije.

Tehnični načrt evalvacije izvira iz konceptualne usmerjenosti evalvacije. Tehnične in metodološke probleme evalvacije je najbolje obravnavati v kontekstu konceptualnih, političnih in utilitarističnih premislekov. Obravnava tehničnih problemov evalvacije vključuje načrt za zbiranje podatkov in njihovo analizo. Tehnični načrt predstavlja odgovore na naslednja vprašanja:

- * Katera bo metoda (oziroma metode) proučevanja?
- * Katera bo osnovna enota (oziroma enote) analize ?
- * Kakšna bo strategija vzorčenja?
- * Kakšne primerjave (če sploh) bodo izvedene?
- * Katere vrste podatkov bodo zbrane? Kdo jih bo zbiral? Kdaj? S katerimi instrumenti?
- * Kako bo zagotovljena kakovost in natančnost podatkov?
- * Kako bo zagotovljena veljavnost in zanesljivost podatkov?
- * Katere vrste analiz bodo izvedene?
- * Katere vrste ugotovitev bodo sledile analizam?

6. Kako izbrati ustrezno metodološko orodje?

V literaturi o evalvacijskih metodah lahko zasledimo burne razprave med zagovorniki kvantitativne metodologije in zagovorniki kvalitativne metodologije. V zadnjih letih so se te razprave umirile (Patton, 1987). Postopoma prihaja do soglasja, da je veliko bolj pomembno izbrati takšne metode, ki bodo ustrezale konkretnim evalvacijskim vprašanjem in problemom, kot zagovarjati univerzalnost kateregakoli posameznega metodološkega pristopa, ki bi ustrezal vsem evalvacijskim okoliščinam.

Izbira primerne evalvacijske metode je odvisna od odgovorov na naslednja vprašanja:

- * Kakšen je namen evalvacije?
- * Kdo bo uporabil ugotovitve evalvacije?
- * Kakšne vrste informacij nameravamo zbirati?
- * Koliko sredstev in časa je na razpolago za izvedbo evalvacije?

Odgovori na ta vprašanja bodo opredelili, katera vrsta podatkov bo najbolj ustrezala določeni evalvaciji. Pri izbiri evalvacijskih metod naj torej ne bi veljala toga pravila. Umetnost evalvacije pomeni oblikovanje modela in zbiranje informacij, ki bodo ustrezale posebnim okoliščinam in določenemu kontekstu, znotraj katerega deluje program. Vsak posamezen evalvacijski model predstavlja medsebojno delovanje metodoloških izbir, ustvarjalnosti, virov, praktičnosti in sodb ljudi, ki so udeleženi.

Na vprašanja, kot so npr.: Kateri model evalvacije je najboljši? Katera strategija bo zagotovila najbolj uporabne odgovore? ni enostavnih, neposrednih in univerzalnih odgovorov. Odgovor bo v vsakem primeru posebej odvisen od tega, kaj različni udeleženci (vodstvo šol, financerji, osebje, udeleženci programa, starši itd.) želijo vedeti, od namena evalvacije, razpoložljivih sredstev ter od interesov in sposobnosti vključenih evalvatorjev.

Seveda pa izbira ustreznega metodološkega pristopa ni enostavna naloga. Zaradi različnih okoliščin in različnih vrst programov so zaželjeni različni pristopi in različne tehnike zbiranja podatkov. Usposobljen evalvator bo kombiniral različne metode, s katerimi bo dobil različne podatke za osvetlitev zastavljenih vprašanj. To pa predpostavlja, da evalvator obvladuje vse razpoložljive družboslovne raziskovalne metode, kot so: raziskovalne tehnike vprašalnikov, socialni kazalci, cost-benefit analize, standardizirani testi, eksperimentalni ali kvazi-eksperimentalni modeli, opazovanje z udeležbo, poglobljeni intervjuji itd., ki jih lahko uporabi za raznovrstne probleme. V sodobnih evalvacijah na področju izobraževanja se pogosto uporablja triangulacija, ki naj bi prispevala k metodološki natančnosti.

Triangulacija pomeni uporabo raznovrstnih strategij zbiranja podatkov. S triangulacijskimi evalvacijskimi pristopi povečamo natančnost in veljavnost evalvacije. Denzin je identificiral štiri vrste triangulacije (Denzin, 1978):

* *podatkovna triangulacija* vključuje uporabo različnih virov podatkov o istem vprašanju, kot so npr.: intervjuji z ljudmi na različnih položajih ali z različnimi pogledi na obravnavano vprašanje;

* *evalvatorska triangulacija* pomeni vključenost evalvatorjev različnih strok, da bi se izognili pristranostim v primerih, če evalvacijo izvaja ena oseba;

* *teoretična triangulacija* predstavlja uporabo raznovrstnih perspektiv ali teorij pri interpretaciji ene vrste podatkov;

* *metodološka triangulacija* vključuje uporabo raznovrstnih metod pri proučevanju določenega programa oziroma enote evalvacije, kot so npr.: intervjuji, vprašalniki, dokumentacija, opazovanje.

Idealen evalvator naj bi bil metodološko prožen, izkušen in sposoben uporabiti različne metode pri proučevanju kateregakoli posebnega evalvacijskega vprašanja, pri tem pa naj bi bil nenehno soočen z realnostjo omejenih virov in ponudbo skromnega usposabljanja, ki je na razpolago večini evalvatorjev. Ti zadržki pa pomenijo, da predstava o usposobljenem in izkušenem evalvatorju vključuje tudi evalvatorja kot pogajalca, ki si prizadeva doseči najboljši model oziroma načrt in najbolj uporabne odgovore znotraj resničnega sveta politike, različnih interesnih skupin in metodoloških nepopolnosti.

LITERATURA:

Denzin, N.K. (1978). The logic of naturalistic inquiry. V: Denzin, N.K. (ed.), *Sociological methods: A sourcebook*. New York: McGraw-Hill.

Everitt, A. (1996). Developing Critical Evaluation. *Evaluation*, 2, 173-188.

Guba, E.G. (1978). *Toward a methodology of naturalistic inquiry in educational evaluation*. Los Angeles: Center for the Study of Evaluation.

Guba, E.G., Lincoln, Y.S. (1982). *Effective Evaluation*. San Francisco: Jossey Bass.

Guba, E.G., Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. Newbury Park, London, New Delhi: Sage Publications.

Guba, E.G., Lincoln, Y.S. (1993). Epistemological and Methodological Bases of Naturalistic Inquiry. V: Madaus, G.F., Stufflebeam, D., Scriven, M.S. (eds). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Boston, The Hague, Dordrecht, Lancaster: Kluwer-Nijhoff Publishing.

House, E. (1980). *Evaluating with validity*. Newbury Park, CA: Sage.

Kristoffersen, D., Surssock, A., Westerheijden, D. (1998). *Manual of Quality Assurance: Procedures and Practices*. London: Quality Support Centre, EC/Phare/ETF.

MacDonald, B. (1976). Evaluation and the control of education. V: Tawney, D. (ed.) *Curriculum Evaluation Today: Trends and Implications*. London: Macmillan Educational.

Macur, M. (1996). *Evalvacija kulturne politike na področju gledališke dejavnosti*. Magistrska naloga. Ljubljana: Univerza v Ljubljani, Fakulteta za družbene vede.

Madaus, G.F., Stufflebeam, D., Scriven, M.S. (eds)(1993). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Boston, The Hague, Dordrecht, Lancaster: Kluwer-Nijhoff Publishing.

Norris, N. (1993). *Understanding Educational Evaluation*. London: Kogan Page.

Parlett, M., Hamilton, D. (1977). Evaluation as Illumination: A New Approach to the Study of Innovative Programmes. V: Hamilton, D. et al. (eds). *Beyond the Numbers Game*. London: Macmillan.

Patton, M. (1981). *Creative Evaluation*. Beverly Hills: Sage.

Patton, M. (1987). *How to Use Qualitative Methods in Evaluation*. Newbury Park, London, New Delhi: Sage Publications.

Radaelli, C.M., Dente, B. (1996). Evaluation Strategies and Analysis of the Policy Process. *Evaluation*, 1, 51-66.

Rossi, H.P., Freeman, E.H. (1993). *Evaluation, a systematic approach*. Newbury Park: Sage Publications.

Sanders, J.R. (1994). *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs*. The Joint Committee on Standards for Educational Evaluation. Newbury Park, London, New Delhi: Sage Publications.

Scriven, M. (1993). Evaluation ideologies. V: Madaus, G.F., Stufflebeam, D., Scriven, M.S. (eds). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Boston, The Hague, Dordrecht, Lancaster: Kluwer-Nijhoff Publishing, 229-260.

Simons, H. (1987). *Getting to Know Schools in a Democracy*. London: Falmer Press.

Stake, R.E. (1976). *Evaluating Educational Programmes: The Need and the Response*. Washington: OECD Publications Centre.

Stake, R.E. (1993). Program Evaluation, Particularly Responsive Evaluation. V: Madaus, G.F., Stufflebeam, D., Scriven, M.S. (eds). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Boston, The Hague, Dordrecht, Lancaster: Kluwer-Nijhoff Publishing.

Toulemonde, J. (1995). Should Evaluation Be Freed from its Causal Links?" *Evaluation and Program Planning*, 2, 179-190.