

Marjan Šimenc

## **Standardi evalvacije**

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

Zdenko Lapajne je v članku "Psihometrične pripombe k pravdi o zunanjem preverjanju" predlagal, "da morajo vsi javni programi za zunanje vrednotenje izobraževalnih učinkov v Sloveniji, ki jih financira država, izpolnjevati vsaj tiste ameriške standarde za pedagoško in psihološko testiranje, ki so označeni kot primarni." (Lapajne 1993, 277) Njegov argument za ta predlog je bil, da je kot vodilo pri ravnanju (gre torej za področje etike in morale) na področju zunanjega vrednotenja znanja potrebno vzpostaviti določena merila. Teh pa ni treba nanovo izumljati, temveč je smiselno uporabiti izkušnje iz držav, kjer so se ta merila izkristilizirala in kjer je pri razvoju in uporabi testov pridobljenih največ izkušenj, se pravi v ZDA. Ne nazadnje je pomembno tudi to, da omenjene standarde ni vsilila država, temveč sta jih razvili dve stanovski organizaciji, to sta: *Ameriško društvo psihologov* in *Ameriško združenje za raziskave izobraževanja*, skupaj z *Državnim svetom za uporabo meritev v izobraževanju* in *Državnim združenjem za izobraževanje*. Standardi so se oblikovali na osnovi konkretnih izkušenj in oblikovali so jih praktiki, tako da gre v nekem smislu za samoregulacijo same prakse.

Podoben argument je mogoče uporabiti tudi na področju evalvacije. ZDA so namreč tudi država, od koder izhajajo prizadevanja izsledke in metode znanosti uporabiti za regulacijo družbe, zato je smiselno tudi na tem področju upoštevati njihove izkušnje. To pa ni edina povezava med Lapajnetovim opozorilom in standardi za evalvacijo. Lapajne namreč navaja hrvaški prevod ameriških standardov *Standardi za pedagoško i psihološko testiranje*.<sup>1</sup> Standardi za evalvacijo programov, *The Program Evaluation Standards*, pa so dejansko nastali kot stranski proizvod ene od verzij standardov za psihološko testiranje. Standarde se namreč na vsake toliko časa popravlja in ko je leta 1974 odbor, ki so ga imenovali *American Educational Research Association*, *American Psychological Association* in *National Council on Measurement in Education* končal revizijo *Standardov za izobraževalne in psihološke teste in priročnike* iz leta 1966, so se pri obravnavi vsebinske členitve soočili tudi z

---

<sup>1</sup> AERA, APA in NCME (1992). *Standardi za pedagoško i psihološko testiranje*. (Prevedel A. Kulenović). Zagreb: EDUCA (izvirnik 1985).

vprašanem rabe testov pri evalvaciji izobraževalnih programov. Po daljši razpravi se je odbor odločil, da vprašanje evalvacije zasluži lastne standarde in to tudi predlagal organizacijam in združenjem, ki so ga imenovali. Ustanovljen je nov odbor, ki je leta 1981 objavil *Standards for Evaluations of Educational Programs, Projects, and Materials*. Zanimivo je, da se je tudi pri delu tega odbora identificirala nova tema, ki jo je treba obravnavati ločeno, namreč standardi za evalvacijo osebja. Tako so leta 1988 objavili še tretji niz standardov, *The Personnel Evaluation Standards*.<sup>2</sup>

Ameriške standarde za evalvacijo programov je izdelal *Joint Comitee on Standards for Educational Evalvation*. Odbor je akreditirala *American National Standard Institute* (ANSI), ki skrbi za ameriške nacionalne standarde, s čimer so standardi postali uradni ameriški standardi. Posebno težo jim daje spisek 16 združenj, ki so sodelovala pri njihovem nastanku.<sup>3</sup> Vsebino standardov povzemam po izdaji *The Program Evaluation Standards, 2th Edition*.<sup>4</sup>

Standardi se členijo na štiri skupine, na standarde koristnosti (Utility Standards), izvedljivosti (Feasibility Standards), primernosti (Propriety Standards) in natančnosti (Accuracy Standards), s čimer so podani štirje osnovni atributi dobre evalvacije. Členitev na utility, feasibility, propriety in accuracy je na neki način postala tudi standard za neko vrsto standardov, saj jo najdemo tudi pri standardih za evalvacijo osebja (v zaporedju primernost, koristnost, izvedljivost in natančnost). Standardi koristnosti se nato členijo na sedem, standardi izvedljivosti na tri, standardi primernosti na osem in standardi natančnosti na dvanajst (pod)standardov.

Namen standardov koristnosti je zagotoviti, da bo evalvacija nameravanim uporabnikom evalvacije v resnici ponudila informacije, ki jih potrebujejo. S tem bo evalvacija izpolnila svoj primarni cilj, saj ima evalvacija izobraževalnih programov kot vsako načrtovano delovanje cilj, ta pa je pridobitev ustreznih informacij glede

---

<sup>2</sup> The Joint Committee on Standards for Educational Evaluation: *The Personnel Evaluation Standards*, Sage, London 1988.

<sup>3</sup> American Association of School Administrators, American Educational Research Association, American Evaluation Association, American Federation of Teachers, American Psychological Association, Association for Assessment in Counseling, Association for Supervision and Curriculum Development, Canadian Society for the Study of Education, Council of Chief State School Officers, Council on Postsecondary Accreditation, National Association of Elementary School Principals, National Association of Secondary School Principals, National Council on Measurement in Education, National Education Association in National School Boards Association.

<sup>4</sup> The Joint Committee on Standards for Educational Evaluation: *The Program Evaluation Standards, How to Assess Evaluations of Educational Programs, 2th Edition*, Sage, London 1994.

njegove ustreznosti. Na ustreznost informacije se navezuje četrta skupina standardov, standardi natančnosti. Ti naj bi zagotovili, da evalvacija "razkrije in posreduje tehnično ustrezne informacije o značilnostih, ki določajo vrednost oziroma zasluge programa, ki je v evalvaciji" (ibid, str. 125). Druga in tretja se ne navezujeta na rezultat, temveč na sam proces evalvacije. Standardi izvedljivosti skušajo zagotoviti, da bo evalvacija "realistična, preudarna, diplomatska in varčna" (ibid, str. 63). Upoštevanje standardov primernosti pa naj bi zagotovilo legalen in etičen potek evalvacije, to je tak, ki bi upošteval dobrobit vseh vpletenih v evalvacijo in tudi tistih, na katere bi vplivali njeni rezultati.

Da bi dobili boljši vpogled v vsebino standardov, na kratko povzemimo opredelitve sedmih "Standardov koristnosti".

1. Identifikacija vpletenih: ugotoviti je treba, kdo vse je osebno vključen v evalvacijo oziroma na koga bo evalvacija učinkovala. Tako bo mogoče nasloviti njihove potrebe. Navadno gre za organe, ki sprejemajo odločitve, v široki evalvaciji izobraževalnega sistema pa med vpletene seveda spada tudi širša javnost.
2. Kredibilnost evalvatorja: osebe, ki bodo evalvacijo izvajale, morajo biti vredne zaupanja in biti morajo ustrezno usposobljene. Le v tem primeru bo evalvacija kredibilna in bodo njeni rezultati sprejeti.
3. Obseg in selekcija informacij - izbira mora biti dovolj široka, da naslovi vsa pertinentna vprašanja, obenem pa se mora odzivati na potrebe naročnikov evalvacije in vseh drugih, ki jih evalvacija zadeva.
4. Identifikacija vrednot - perspektive, postopke in načela pri interpretaciji izsledkov je treba skrbno opisati, tako da so jasna izhodišča za vrednostne sodbe. Pri šolskem sistemu so osnova za vrednotenje verjetno že podane v zakonsko sprejetih ciljih in izhodiščih posameznih programov. Druga možnost je, da so tudi ti proklamirani cilji sami deležni kritičnega premisleka, le da je v tem primeru treba jasno predstaviti, kaj je osnova za njihovo vrednotenje.
5. Jasnost poročila - poročilo mora jasno opisati evalvirani program, kontekst, namen uporabljene procedure in seveda izsledke evalvacije, tako da so bistvene informacije zbrane in razumljivo predstavljene.
6. Pravočasnost in dostopnost poročila - pomembne preliminarne rezultate je treba pravočasno posredovati naročnikom evalvacije. Potreben je kompromis med hitrostjo in temeljitostjo evalvacije. Lahko so nekateri rezultati tako pomembni, da jih je treba posredovati čim prej.

7. Vpliv evalvacije - evalvacija mora biti načrtovana, izvedena in izsledki predstavljeni tako, da spodbujajo nadaljnje aktivnosti in da se poveča verjetnost uporabe njenih izsledkov.<sup>5</sup>

Če poskušamo opredeliti status standardov evalvacije, se odpira nekaj bistvenih vprašanj. Predvsem to, kakšna je vloga standardov pri evalvaciji. Iz zadnjega standarda natančnosti, iz refleksivnega standarda, ta govori o metaevalvaciji, je očitno, da so to standardi, ki se jih lahko uporabi za evalvacijo evalvacije. Ko že imamo neki model evalvacije in ga uporabimo v praksi, lahko njegovo kakovost merimo s temi standardi. Tudi iz opisa nastanka je razvidno, da so se standardi razvili iz že obstoječih praks evalvacije in so izhajali iz tako pridobljenih izkušenj. Vendar se odbor oziroma komite, ki je kolektivni avtor standardov, ne omeji na to. V predgovoru kot svoj cilj postavijo: "Standardi so napor priskrbeti vodstvo (guidance) za učinkovito evalvacijo. Sami ne morejo jamčiti ali zagotoviti kakovosti nobene evalvacije. Dobra evalvacija zahteva uporabo profesionalne razsodnosti, da bi se adekvatno uporabilo standarde na posebno evalvacijsko konstelacijo. Standardi naj bi izražali najboljšo prakso v času, ko so bili napisani." (Ibid, xviii) Standardi so torej vodič do dobre evalvacije. Kdor še ne ve, kaj dobra evalvacija je, ga vodijo do nje. Iz tega bi lahko sklepali, da so standardi uporabni tudi tam, kjer se šele vpeljuje kultura evalvacije. Niso samo merila za evalvacijo evalvacije, temveč tudi napotki za njeno vpeljavo. Vendar standardi zahtevajo smiselno uporabo, za katero je potrebna presodnost. Kako je presodnost (se pravi pravo mero in občutek, ki se kaže v delovanju) mogoče pridobiti? Zanj ni standardov, ni dodatnih standardov, ki nam bi povedali, kako te prve standarde uporabiti. Sposobnost presoje nekdo pač ima ali pa je nima. Seveda ne gre za stvar osebnostnih značilnosti oziroma vrojenih kvalit, temveč za nekaj, kar izhaja iz dolgoletne ustrezne prakse, se pravi iz že uveljavljene prakse evalvacije. Standardi so se razvili iz prakse, in tako s svojim faktičnim in konceptualnim nastankom predpostavljajo obstoj razširjene prakse evalvacije. To sicer ni ovira: če želimo preprosto prenesti prakso, ki je nekje že razvita (in ocenimo, da gre za

---

<sup>5</sup> Ker na tem mestu za daljšo obravnavo standardov ni prostora, členitve ostalih treh standardov navajam brez razlage. Standardi izvedljivosti: praktičnost procedur, politična primernost, učinkovitost glede na stroške. Standardi primernosti: usmerjenost na uslugo, potrebnost formalnega pristanka, spoštovanje pravic vključenih subjektov, spoštovanje človeškega dostojanstva pri interakciji, popolnost raziskave in poštenost ocene, javnost izsledkov, poštena obravnava konflikta interesov, finančna odgovornost. Standardi natančnosti: natančna dokumentacija evalviranega programa, upoštevanje konteksta analize, detajlni opis ciljev in procedur, izčrpen opis virov informacij, veljavnost informacij, zanesljivost informacij, sistematičnost informacij, analiza kvantitativnih informacij, analiza kvalitativnih informacij, utemeljitev sklepov, nepristrano poročanje, metaevalvacija.

tehnologijo, ki je nevtralna in ki je tudi dokaj neobčutljiva za okoliščine), drugam, potem lahko standarde uporabljamo tudi tam, kjer prakse ni - ker pač parazitiramo na praksi od drugod. Večji problem je, da standardi to prakso predpostavljajo tudi pri svoji rabi: standarde bo z ustreznim občutkom uporabljal samo tisti, ki že ima dovolj nakopičenih izkušenj z evalvacijo in je torej ta občutek že postopno pridobil. To pomeni, da bi za neposredno uporabo standardov morali poleg modelov za aplikacijo uvoziti tudi strokovnjake (in strokovno javnost), ki so si ustrezen občutek že razvili: ob predpostavki seveda, da bo ta občutek uspešno deloval tudi v spremenjenih okoliščinah.

Tako v državah, ki se ne morejo pohvaliti z dolgoletno prakso evalvacije, in evalvacijo šele uvajajo ter iščejo uspešen model oziroma modele, naletimo na paradokšno situacijo: standardi za izvedbo evalvacije so na eni strani na moč potrebni, saj še ne obstaja zbir preteklih bolečih izkušenj in si je treba čim več učenja na lastnih napakah prihraniti, na drugi strani pa imajo omejen domet, saj je za njihovo ustrezno rabo potreben 'občutek' oziroma sposobnost presodnosti, slednja pa je rezultat dolgoletne prakse evalvacije. Standardi so rezultat kulture evalvacije in v samoumevnosti neke kulture evalvacije tudi najboljše delujejo. Pri uvajanju le-te pa so sicer zelo potrebni, vendar brez uveljavljene rabe pogosto delujejo kot (pre)abstraktna pravila za naknadno evalviranje samega poteka evalvacije, s katerimi si je težko pomagati. Na samem začetku je namreč potreben razmislek o evalvaciji kot taki, zato bi bil subjektom, ki so se znašli v takem zagatnem položaju, dobrodošel model evalvacije, ravno tega pa standardi ne ponujajo in zaradi svoje narave ne smejo ponuditi. Vendar je ambicija snovalcev standardov vendarle večja, ne gre samo za metaevalvacijo, temveč je že iz prejšnjega citata razvidno, da hočejo ponuditi vodič za evalvacijo. Včasih formulacije avtorjev res opisujejo minimalen cilj evalvacije, po katerem naj bi standardi, opredeljeni kot splošno dogovorjena načela, ki, kadar so spoštovana, povečajo kvaliteto in poštenost prakse, za katero gre,<sup>6</sup> zagovarjali "odgovorno izvajanje in rabo evalvacij izobraževalnih programov, projektov in materialov" (ibid, str. 2). Vendar vedno znova naletimo na formulacije maksimalnega programa:

"Programe izobraževanja in usposabljanja se evalvira, da bi se določilo njihovo kvaliteto in dobilo smer za njihovo izboljšanje. V odsotnosti vsake jasne definicije, kaj vzpostavlja razumno evalvacijo izobraževalnih programov, je skupni odbor kompilirjal vednost o evalvaciji programov, pridobljeno iz profesionalne literature in dolgoletnih

---

<sup>6</sup> "A standard is a principle mutually agreed to by the people engaged in a professional practice, that, if met, will enhance the quality and fairness of that professional practice, for example, evaluation." (Ibid, str. 2)

izkušenj edukatorjev in specialistov za evalvacijo. Ta vednost je organizirana in predstavljena v tej knjigi kot standardi za prakso evalvacije izobraževalni programov (...) Standardi vodijo oblikovanje, uporabo in ocenjevanje evalvacij izobraževalnih programov, projektov in gradiv." (Ibid, str. 1)

V ZDA sicer obstaja dolgoletna praksa evalvacije, vendar pa to ni pripeljalo do njene enotne definicije, zato je treba opredelitve zbirati in se opirati na izkušnje. Za formulacije avtorjev se sicer lahko zdi, da nihajo med minimalnim in maksimalnim programom: ali nam standardi že nudijo model evalvacije (jih torej lahko uporabimo pri uvajanju evalvacije) ali pa nastopijo naknadno, ko model že imamo in ga uporabljamo, za evalvacijo kvalitete same evalvacije. Vendar premislek pokaže, da je tako postavljena dilema napačna. Seveda nam standardi lahko pomagajo pri oblikovanju evalvacije, vendar sami na sebi za oblikovanje modela evalvacije ne zadoščajo. Prav tako standardi sami na sebi ne vzpostavijo prakse evalvacije, da moramo vpeljati prakso evalvacije (niti nam ne narekujejo, da jo moramo), šele po odločitvi zanjo nam povedo, kako evalvirati, če hočemo, da bo praksa kvalitetna in poštena. Imajo torej status hipotetičnega imperativa: če smo se že odločili za evalvacijo, potem nam povedo, kaj storiti. Ne povedo nam torej, da se moramo odločiti za evalvacijo, niti ne, kakšen model natančno moramo izbrati. Ponujajo pa nam merila za izbiro med različnimi modeli, če so le-ti že dani.

Po drugi strani pa lahko čisto resno beremo poudarke, da so standardi tudi vodilo za načrtovanje evalvacije, in ne samo vodič za njeno ocenjevanje - da imajo tako formativno kot sumativno vlogo. V uvodu je tako razvitih pet korakov pri aplikaciji standardov. Prvi z lepo samonanašalno gesto pravi, da se moramo najprej seznaniti s standardi evalvacije. Drugi zahteva, da nam postane jasen namen evalvacije. V tretjem si moramo razjasniti kontekst evalvacije. V četrtem uporabimo vsak standard glede na namen in kontekst evalvacije. Peti terja, da se odločimo, kaj bomo naredili z rezultati. Standardi za evalvacijo tako vključujejo standarde za uporabo samih sebe - tako pri oblikovanju programa evalvacije kot pri njegovi evalvaciji.

Poleg tega ima predstavitev vsakega posameznega standarda preišljeno strukturo, ki naj bi kar najbolj premostila prepad med dejansko prakso evalvacije in abstraktnimi kriteriji, v katero jo skuša vsak seznam standardov prevesti. Vsak od tridesetih standardov je najprej opisan s kratkim pravilom, ki naj bi ga evalvacija upoštevala. To je formulacija standarda. Nato sledijo štiri rubrike. Najprej kratka razlaga vsebine

standarda (Overview), nato navodila (Guidelines), ki so v bistvu seznam čisto konkretnih stvari, ki jih je treba narediti. Sledi spisek običajnih napak (Common errors), se pravi spet navajanje čisto konkretnih primerov možnih napak, ki se jim je treba izogniti. Na koncu sta dodana dva konkretna primera kršitve standarda (Illustrative case), najprej opis konkretnega poteka evalvacije, nato analiza, zakaj je do napake prišlo. Standard je tako opredeljen pozitivno, v svoji abstraktni obliki, nato je uporabljen in konkretiziran, na koncu pa je z navajanjem kršitev standarda opredeljen še negativno.

Prednost standardov je prav v odprtosti: združljivi so z različnimi modeli evalvacije in z različnimi konkretnimi programi. To pa je tudi njihova šibka točka, saj skrbna sistematizacija včasih zakrije, da evalvacija nastopa v družbi, ki ni tako kot standardi urejena v enoten in pregleden sistem. Kolikor ta razčlemba vidikov, ki jih je treba pri evalvaciji upoštevati, razkriva logiko področja evalvacije, jo hkrati tudi zakriva. Naštevane samo identificira elemente, ne pokaže pa njihovih medsebojnih povezav, predvsem pa ne možnih konfliktov med njimi. Pa ne samo med različnimi dejavniki, tudi znotraj enega samega standarda so možne konfliktne situacije. Če si ogledamo samo prvo postavko znotraj standardov primernosti, usmerjenost na uslugo. Komu naj služi evalvacija: kateri skupini, kateremu interesu? To je po svoje temeljno vprašanje evalvacije in v tem standardu se nanj opozori, a istočasno se ga uvrsti v sistem in se tako dilematičnost vprašanja zabriše. To je razumljivo, saj so standardi sestavljeni v čisti sistem, ki zaradi svoje forme ne dopušča razcepov in razkolov. Taka pregledna in sistematična forma standardov skuša urejati evalvacijski govor o realnosti, v kateri pa so interesi, ki se na tak ali drugačen način začnejo artikulirati ob evalvacijskih posegih vanjo, lahko zelo konfliktni in ne dopuščajo pomiritve in konsenza.<sup>7</sup> Zdi se, da standardi pravzaprav izpostavijo vse točke, ki so sporne in zaradi katerih je evalvacija že vnaprej obsojena na neuspeh oziroma vsaj na to, da se ji bo neuspešnost očitala. V tem smislu standardi implicitno opozarjajo na nemožnost konsenza o evalvaciji. Po drugi strani pa standardi ravno so uspeli konsenz, zato je spisek organizacij, ki so sodelovale pri njihovem nastanku, tako pomemben. In prav lahko je spregledati, da je morda pomembnejše od njihove vsebine dejstvo, da ti standardi sploh obstajajo. To priča, da je zanje obstajal močan interes in da je konsenz (vsaj glede nekaterih stvari) možen.

---

<sup>7</sup> Zdi se, da je ena od predpostavk standardov - ob temeljni pragmatični usmerjenosti seveda, ki pa je nemara značilnost vsake evalvacije - univerzalna skupnost komunikacije, v kateri na koncu le pride do racionalnega soglasja vseh.

Opisani standardi seveda niso edini možni, a ustrezno izpostavijo vidikov, za katere je ameriška praksa pokazala, da jih je treba upoštevati. Zdi se, da pregled standardov opozori na odločitve, ki jih je treba sprejeti, preden se evalvacije sploh lotimo. In te odločitve niso samo metodološke, temveč zadevajo predvsem vprašanje, kaj se od evalvacije pričakuje. Natančna definicija cilja in pričakovanj lahko prepreči številne nesporazume pri intepretaciji izsledkov.

Ko so pomembne odločitve glede evalvacije sprejete, so to standardi korektnosti in imajo poleg vloge pri metaevalvaciji vseskozi lahko tudi status regulativnih idej, ki nam pomagajo voditi proces.<sup>8</sup> Ne kot zadnja beseda, a kot nujna opora. Drugače se lahko zgodi, da dolgim načelnim razpravam sledi hitra in površna izvedba. Standardi izvedbe evalvacije (in evalvacija pravzaprav pomeni izvedbo evalvacije, tako da je evalvacija na neki način po svojem pojmu praktična in pragmatično usmerjena) so varovalka, ki lahko to prepreči.

## Literatura

Lapajne, Zdenko: "Psihometrične pripombe k pravdi o zunanjem preverjanju", *Sodobna pedagogika* 1993/5-6, str. 258-278.

Murphy, Roger; Torrance, Harry (ur.): *Evaluating Education: Issues and Methods*, Paul Chapman, London 1987.

Noris, Nigel: *Understanding Educational Evaluation*, Kogan Page, London 1990.

*Standardi za pedagoško i psihološko teastiranje*, Educa, Zagreb 1992.

*The Pesonnel Evaluation Standards*, Sage, Newbury Park 1988.

*The Program Evaluation Standards*, Sage, Thousand Oaks 1994.

---

<sup>8</sup> "Standardi so vodila, ne mehanična pravila. Vsebujejo opomine in opozorila proti možnim napakam, izločijo prakse, za katere obstaja splošno strinjanje, da so sprejemljive oziroma nesprijemljive, in predlagajo vodila, ki izražajo trenutno najboljšo prakso." (Ibid, str. 8) Ta citat dobro prikazuje dvojnost, ki opredeljuje status standardov: na eni strani so standardi v središču evalvacije, saj so vodila, ki nam kažejo pot. Ne dajo sicer konkretnih odgovorov, vendar nas opozarjajo na bistvena vprašanja, ki si jih moramo zastaviti, in nam povedo, kakšne morajo biti značilnosti ustreznega odgovora. Na drugi strani pa je njihov status obrobni, zgolj negativen: samo opozarjajo nas na nekatere napake, ki bi jih lahko naredili, a če jih nismo, so standardi za nas pravzaprav čisto brez pomena. V nekem smislu nobena postavka iz standardov nima vnaprejšnje veljavnosti. Vsaka je pač rezultat dolgoletnih izkušenj, ki so jih poskusili strniti vanje. A če katere postavke ne upoštevamo, tudi tega, da je ni smiselno upoštevati, ne moremo reči vnaprej, temveč je tak odgovor lahko samo rezultat konkretnega premisleka.