

Lesley Saunders:

## **Evalvacija kurikularne reforme v Sloveniji**

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

### **1.1. UVOD**

Pričujoče delo je dodatek k poročilu *Evalvacija kurikularne prenove v Sloveniji: poskus izhodišč (prvi osnutek)* in je bilo izdelano na prošnjo predsednika Nacionalne komisije za uvajanje in spremljanje programov v vzgoji in izobraževanju (NK). Želel je, da se obdelajo naslednje točke:

- opis glavnih alternativ pri načrtovanju metodologije evalvacije;
- opis evalvacijskih postopkov, metod in tehnik, ki so se po mnenju evalvatorjev kurikularne prenove (v Angliji in drugod) izkazali za učinkovite;
- podroben opis "kriterijev uspešnosti za dokazovanje vpliva kurikularne prenove", kot so bili opisani v prvem dokumentu.

Napisano obravnava prvi dve točki. Glede tretje točke NFER (National Foundation for Educational Research – Nacionalna ustanova za raziskovanje vzgoje in izobraževanja) sicer razume, da se zdi NK zelo zanimiva in pomembna, vendar je za nas bistvenega pomena, da se o kriterijih za uspešnost prenove pogovorijo, pogajajo in dogovorijo predstavniki "udeležencev" – bistveni del procesa vzpostavljanja evalvacije v izobraževalni reformi je namreč ta, da tisti, ki jih ocenjujemo, lahko načela in kriterije "vzamejo za svoje", ne pa da so

jim ti vsiljeni od zunaj ali od zgoraj. Zato smo namesto tega opisali primer ali dva, da bi pokazali, kaj razumemo pod "kriteriji uspešnosti".

Številni predlogi v tem besedilu se nanašajo na nekatera priporočila OECD (Reid et al. 1998), ker so bili uvedeni v prakso ali so jih vsaj upoštevali pri uvajanju reform. Nanje se zato tudi večkrat sklicujemo, seznam priporočil iz besedila pa je naveden v prilogi A.

## **1.2. ALTERNATIVNI NAČINI EVALVACIJE**

Najbolj očitnega ali najboljšega načina za evalviranje kurikularne prenove in njenega vpliva ni – s to težavo se spoprijemajo številne države. Ključni vprašanji za določitev najprimernejšega načina evalvacije v posameznem primeru bi se morali glasiti: "Kaj si najbolj želimo evalvirati?" in "Kaj želimo narediti z rezultati?" Glavne možnosti lahko razdelimo takole:

### ***1.2.1. Formalne konzultacije in konzultativni pregledi***

Vlada se lahko odloči za izvajanje konzultacij in konzultativnih pregledov, da zakonsko pristojno javnost povpraša o pogledih in izkušnjah v zvezi s predlaganimi reformami; ti pregledi so lahko ozko omejeni ali vseobsegajoči. Pri tem običajno upoštevajo, kaj se je zgodilo ali se še dogaja in se na ta način podučijo o vsem, kar jim bo v prihodnosti prišlo zelo prav. Izvedbo konzultacij pogosto prepustijo marketinškim ali drugim za to usposobljenim organizacijam.

### **1.2.2. Merjenje vpliva reform**

Ta oblika evalvacije je sveti graal evalvacij – po eni strani naj bi odgovorila na najpomembnejše vprašanje, in sicer, ali je prenova uspela, po drugi strani pa je v resnici precej težko določiti vpliv pobude vlade – še zlasti pri vseobsegajoči prenovi – celo tedaj, ko smo želene rezultate prenove določili vnaprej in vemo, da so izmerljivi in zbirljivi (kar pa pogosto ne drži!).

Do tega pride zato, ker število "vhodnih" spremenljivk, ki jih je treba upoštevati, ni zanemarljivo in obsega tako razlike v vodstvenih spretnostih pri različnih ravnateljih, ki vplivajo na to, kako dobro se bodo reforme izvajale, ali razlike v sposobnosti učiteljev, ki vplivajo na to, kako dobro bodo reforme podane, kot tudi razlike v socialno-ekonomskem izvoru učencev ali skupin učencev, ki vplivajo na njihovo uspešnost in prizadevanje. Izolacija vpliva reform na zgolj izpitne rezultate učencev ali na njihov prehod na trg delovne sile v nasprotju z vplivom katerekoli spremenljivke je dolgotrajno in tehnično zahtevno opravilo. Poleg tega tovrstnih evalvacij ni mogoče izvesti brez zadostnih "izhodiščnih točk" ali "osnovnih" informacij, s katerimi bi lahko merili spremembe "prej in potem". Tema je obdelana v referatu z naslovom *Measuring the Effects of Educational Interventions (Merjenje vpliva vzgojno-izobraževalnih posegov)* (Brooks et al. 1997), ki je na voljo pri NFER.

Kljub tem omejitvam obstaja vrsta tehnik, po katerih lahko posežemo, da izmerimo obseg vpliva. Podane so v skrajšani obliki in v padajočem vrstnem redu glede strogosti (povzeto po: Holtermann 1998):

- naključni kontrolirani preizkusi – dajejo jasne rezultate s trdnim pripisovanjem vpliva/učinkovitosti;
- raziskave "prej in potem" s kontrolno skupino – dajejo jasne rezultate, če je kontrolna skupina usklajena;
- raziskave "prej in potem" brez kontrolne skupine – ta metoda daje negotovo osnovo za pripisovanje vpliva;
- longitudinalne raziskave populacijskih vzorcev – ta metoda mora statistično preveriti vse preostale vplive na rezultate, preden lahko pokaže pripisovanje vpliva;
- raziskave tipičnih vzorcev – ta metoda se opira na predpostavke o prenosljivosti izsledkov z ene populacije na drugo;
- prikazi mnenj in odnosa – so negotova osnova za ocenjevanje učinkovitosti, čeprav so sami po sebi zanimivi;
- raziskave storitev in uporabnikov (procesov in praks) – vpogledi iz teh metod so bistvenega pomena za izdelavo izvedljivih in sprejemljivih storitev, vendar brez kontrolne skupine ne dajejo podatkov o učinkovitosti.

Zaradi tehtnih političnih razlogov je treba priznati, da večina vzgojno-izobraževalnih pobud, zlasti obsežnih reform, običajno ni zastavljena ali zasnovana tako, da bi dopuščala kontrolne skupine ali "uskrajene" populacije. To prav gotovo velja za nacionalni kurikulum in z njim povezane reforme v Angliji in Walesu. Zato je treba evalvacije po navadi izvajati po manj zanesljivih metodah.

Večina oblik evalvacije vzgojno-izobraževalnih sistemov ali pobud v praksi spada v eno ali več spodaj navedenih kategorij. Naj še opozorimo, da je kritična ocena teh pristopov, kot so jih razvrstili v Veliki Britaniji, podana v 3. poglavju.

### **1.2.3. Pregledi sedanjih raziskav**

Vlada ali njeni uradi lahko naročijo neodvisno identifikacijo, analizo in oceno sedanjih raziskav – ki jih financira vlada ali so neodvisne; objavljene in neobjavljene –, da preverijo uvajanje in vpliv kurikularne prenove. Tega se po navadi lotijo samo v odločilnih trenutkih (na primer pri tehtanju revizije v zakonodaji), da bi preskrbeli najrazličnejše informacije za vladni program spremljanja, predvsem pa vlado obvestili o problematiki in področjih, kjer sta potrebna nadaljnji razvoj ali prilagoditev.

### **1.2.4. Potekajoče spremljanje učinkovitosti sistema**

1.2.4.1 **Potekajoče strukturirano in nestrukturirano spremljanje:** izvajajo ga državni uradi z dejavnostmi, kot so na primer prikazi z vprašalniki in obsežnim terenskim delom (ki ga lahko opravijo neodvisne organizacije), obiski v šolah, razprave v fokusnih skupinah, povezave in sodelovanje z drugimi uradnimi organi – združenji učiteljev, subjektivnimi organizacijami, podeljevalnimi organi – in neuradne analize prostovoljnega dopisovanja in stikov.

1.2.4.2 **Merjenje uspešnosti učencev** na ključnih področjih znanja, vključno z dokazom o izboljšanju ali poslabšanju v času in relativnih razlikah, če obstajajo, med različnimi skupinami učencev. To lahko izvaja oddelek za statistiko pri šolskem ministrstvu ali vladni urad, lahko pa tudi neodvisne ekspertne ustanove. Če so bili sistemi vzpostavljeni pred reformo, je možno do nadaljnjega pripisovati spremembe v uspešnosti vplivu prenove, vendar samo, če pri tem lahko upoštevamo tudi druge relevantne spremenljivke, kot so na primer

demografske spremembe. Merjenje uspešnosti lahko, odvisno od namena, izvajamo z vzorčenjem ali na splošni osnovi (gl. 3. poglavje).

1.2.4.3 Različico zgornjega pristopa predstavljajo **mednarodne raziskave dosežka**, kot sta na primer Third International Mathematics and Science Study (TIMSS) in International Assessment of Educational Progress (IAEP), v katerih so sodelovali tudi slovenski šolarji (gl. npr. Šetinc 1999; gl. tudi Brown 1998 za kritičen pogled na mednarodne raziskave).

1.2.4.4 **Merjenje uspešnosti ustanov** glede ključnih kazalcev vključuje dokaz o izboljšanju ali poslabšanju v času. Tudi to lahko izvaja vlada ali njeni uradi ali pa relevantne ekspertne ustanove za zbiranje podatkov in analizo. Ta pristop zahteva splošen sistem ocenjevanja uspešnosti učencev. Če sistem dovoljuje zbiranje potrebnih podatkov, lahko dokaz vsebuje merjenje na "dodano vrednost" relativnega napredka, ki so ga naredile dane skupine učencev med vstopom in izstopom.

1.2.4.5 **Neodvisna inšpekcija šol** glede kvantitativnih in kvalitativnih kazalcev, ki jo izvaja vladni urad ali neodvisni timi glede na eksplicitne kriterije in je, če je možno, ne pa nujno, podprta z nasveti in napotki šolam, kako naj dosežejo izboljšanje. Ta pristop lahko oriše glavne spremembe, težko pa je zagotoviti natančno primerljivost izvajanja inšpekcijskih kriterijev v različnih okoliščinah.

#### **1.2.5. Posebne, lokalne in kontekstualne raziskave**

Omenjene oblike spremljanja in evalvacije pogosto dopolnjujejo naslednje:

1.2.5.1 **Cenovno ugodne raziskave** posebnih vidikov sistema, ki jih izvajajo neodvisni ocenjevalci. Tovrstne raziskave naj bi izvajali nacionalni organi, ki so za tovrstno delo primerno usposobljeni in jih najame vlada pri tržnih svetovalcih ipd. V praksi obstajajo bolj ali manj natančni načini določanja in ocenjevanja "cenovno ugodnega". Gre za razmeroma mlado področje evalvacije v vzgoji in izobraževanju, zato ne bi mogli reči, da je njena metodologija že dobro razvita.

1.2.5.2 **Programske evalvacije** posebnih programov ali pobud, ki jih po navadi po pogodbi z državnim ministrstvom za šolstvo izvajajo neodvisne raziskovalne ustanove. Njihov namen je podati dokaze – kvantitativne in kvalitativne – o vplivu posebnih financiranih programov; to se je običajno odvijalo pod njihovimi pogoji, tj. brez merjenja vpliva programa v primerjavi z drugim.

1.2.5.3 **"Občasna" poročila urejevalnih organov**, kot je inšpektorat za šolstvo, o določenih, trenutno zanimivih temah – na primer o razvoju in uporabi informacijske in komunikacijske tehnologije v šolah.

1.2.5.4 **Kvalitativne raziskave** procesa (ki pogosto vključujejo upravljanje sprememb) in zaželeni rezultati izbranih vidikov prenove, z namenom da bi raziskali problematiko, ki jih spodbudi dokaz o uspešnosti.

1.2.5.5 **Lokalne inšpekcije**, ki jih opravljajo izbrani timi upravnih enot šolske uprave, o dogovorjenih kazalcih uspešnosti, ki so posebno zanimivi za lokalne skupnosti.

1.2.5.6 **Neodvisne raziskave**, tj. take, ki jih financira kdo drug kot vlada in ki se jih lotijo univerzitetni timi ali drugi raziskovalni organi za raziskovanje vidikov prenove z različnih perspektiv.

1.2.5.7 **Interna evalvacija, neformalna evalvacija ali samoevalvacija v šolah**, ki lahko morda raziskuje procesne in izkustvene kazalce.

V naslednjem poglavju so navedeni primeri za vsakega od teh pristopov.

### **1.3. UČINKOVITI EVALVACIJSKI PRISTOPI**

#### **1.3.1. Evalvacijske metode v Veliki Britaniji**

V tem podpoglavju navajamo glavne evalvacijske metode, ki so jih nedavno uporabljali ali pa jih še uporabljajo v Veliki Britaniji (ali bolje rečeno v Angliji in Walesu), ne da bi dokončno ocenili njihovo učinkovitost. Predvsem skušamo povzeti njihove prednosti in pomanjkljivosti, kot so jih razložili drugi (avtorica osebno ni pristojna, da bi presojala nekatere tehnične vidike, povezane na primer z razvijanjem testov).

Če upoštevamo trditev, da je "evalvacija politična dejavnost" (Holt 1981), moramo v celotnem podpoglavju priznati, da sta tipično zastavljena ali oblikovana na podlagi ideologij najrazličnejših odtenkov, moči in ravni izdelanosti ne samo sprejeta zakonodaja sukcesivnih vlad, ampak tudi ustanovitev raznih evalvacijskih organov in metodoloških pristopov, ki jih ali so jih predstavljale ali podpirale. Tovrstne ideologije so med spreminjajočo se politično upravo in intelektualno modo neposredno tekmovala druga z drugo ali nasledile druga drugo. V tem dokumentu je nemogoče prikazati zapleteno socialno in politično preteklost evalvacije na področju vzgoje in izobraževanja, celo tisto iz zadnjega desetletja (za katero sta značilna

prepiranje in nasprotovanje). Na to temo so bili seveda napisani že številni članki in knjige; en ali dva izbrana primerka sta navedena v virih na koncu besedila.

### **1.3.1.1 Državno spremljanje in evalvacija nacionalnega kurikula in njegovo ocenjevanje**

Uradi, odgovorni za nacionalni kurikulum in njegovo ocenjevanje (NCC in SEAC, pozneje SCAA, zdaj QCA), so pooblaščen za potekajoče spremljanje in pregledovanje; to so delali na več načinov, tudi z obsežnimi neodvisnimi evalvacijami novih ocenjevalnih režimov, prikazov z vprašalniki in razgovori v šolah kot delom Projekta vzorčenja šol (Schools Sampling Project) ter naročenim pregledom sedanjih raziskav uvajanja in vpliva nacionalnega kurikula (gl. podpoglavje 3.1.8). Ta program spremljanja je povečal število poročil o različnih tematskih sklopih, od katerih pa mnoga niso bila objavljena. Evalvacijo uvajanja nacionalnega kurikula je finančno podprl Nacionalni kurikulumni svet (gl. npr. NCC 1991), lotil pa se je tudi Britanski inšpektorat (gl. podpoglavje 3.1.4). V postopen pregled za razvijanje ocenjevalnih instrumentov za merjenje dosežka učencev v jedrnih predmetih ob koncu "vzgojno-izobraževalnih obdobj" kurikula so bile vgrajene tudi široko zasnovane evalvacijske strategije (gl. podpoglavje 3.1.3). Občasno so izvajali tudi formalne preglede kurikula, ki so se jih lotili s široko zastavljenimi konzultacijami (gl. podpoglavje 3.1.6).

NFER bi bil pripravljen izdelati ustrezno dokumentacijo, ki bi bila na voljo po razpravah s slovensko NK, v katerih bi izostrili problem.

*Prednosti:* pomembno načelo je, da glavne reforme že od začetka spremlja in ocenjuje vlada; uporabljeni so bili še drugi evalvacijski pristopi, čeprav se – kot je razvidno iz 2. poglavja – z nobenim ni dalo izmeriti vpliva z uporabo modela "prej in potem".

*Pomanjkljivosti:* gledano za nazaj, je bilo težko odkriti (i) jasen pogled, ki bi povzel ugotovitve raznih evalvacij, k čemur še pripomore to, da večina evaluativnega gradiva sploh ni bila objavljena; (ii) kakšen je bil odziv na rezultate evalvacij in ali so bili sistematično uporabljeni v podporo uvajanju.

### **1.3.1.2 Merjenje uspešnosti učencev**

Če želimo imeti pregled dosežka učencev na nacionalni ravni na ključnih področjih kurikula, lahko uspešnost preverjamo s testiranjem *vzorca* populacije učencev; namen je omejiti porabo sredstev in nezaželeni "povratni" učinek na kurikulum. To je v Veliki Britaniji od sredine sedemdesetih let do 1990. leta izvajala Enota za ocenjevanje uspešnosti (Assessment of Performance Unit, APU), ki jo je ustanovil Urad za šolstvo in znanost (Department of Education and Science) za "pospeševanje razvoja ocenjevanja in spremljanja dosežka šolarjev in za ugotovitev razloga za pojavljanje slabih dosežkov" (Foxman et al. 1991). APU je najela raziskovalne time na petih kurikularnih področjih – na področju materinščine, matematike, naravoslovja, tujih jezikov ter oblikovanja in tehnologije –, da bi razvili in pripravili za uporabo primerne metode in instrumente za ocenjevanje uspešnosti učencev in ugotovili, katere so "pomembne razlike, povezane z okoliščinami, v katerih se otroci učijo" (Foxman et al. 1991). APU je izvedla približno triinštirideset obsežnih raziskav v programu, ki je pokrival šole v Angliji, Walesu in na Severnem Irskem med letoma 1978 in 1988. Ocenjevalne metode so presegle visoko formalizirane postopke za testiranje uspešnosti učencev, ki so bili v rabi prej, in razvile interaktivne, slušne in vidne ter časovno neomejene načine in tradicionalne časovno omejene, pisne in na spominu temelječe načine.

Enoto so razpustili leta 1990, njeno funkcijo pa je začasno prevzela Enota za evalvacijo in spremljanje (Evaluation and Monitoring Unit, EMU) pri Šolskem izpitnem in ocenjevalnem svetu (School Examination and Assessment Council, SEAC).

*Prednosti:* "Lahko vzorčenje" je šole minimalno bremenilo, a še vedno omogočalo obsežno obravnavanje uspešnosti učencev različnih starosti in v različnih obdobjih, dajalo podrobno sliko na kurikulum navezane uspešnosti na ključnih predmetnih področjih in določalo šolske dejavnike in dejavnike učencev, povezane z uspešnostjo. Razvili so primerne postopke vzorčenja in učinkovite delovne odnose z lokalnimi šolskimi upravami in šolami.

*Pomanjkljivosti:* Pojavila so se vprašanja o zanesljivosti in veljavnosti, postavil pa jih je Holt (1981), čeprav je zavzel skrajno stališče, da "ima nacionalno spremljanje ceneno privlačnost ... APU sama po sebi nima nikakršne izobraževalne vrednosti". Še zlasti je delo APU povzročilo mednarodno polemiko o t. i. Raschevem modelu. Ta model je bil zasnovan kot rešitev naslednje težave: ker testni primerki lahko zaradi kulturnih in družbenih sprememb zastarajo, so lahko longitudinalne raziskave uspešnosti resno ogrožene; model je med drugimi nepopustljivo kritiziral Goldstein (1979).

Če pa po drugi strani želimo dobiti sistematično sliko, kaj doseže prav vsak učenec in kako napreduje, potrebujemo splošen sistem testiranja in ocenjevanja. V Angliji in Walesu praviloma od leta 1988 dalje v nacionalnem kurikulumu določijo, kaj naj bi bili učenci sposobni storiti, potem pa zberejo dokaze, ali to tudi res lahko naredijo. Uspešnost učencev pri določenih ciljnih znanja nacionalnega kurikula ocenjujejo pri 7., 11., 14. in 16. letih (torej ob koncu vsakega "vzgojno-izobraževalnega obdobja" nacionalnega kurikula) in o tem tudi poročajo. Ocenjevanje je bilo kombinacija zunanjega preverjanja znanja s standardnimi preizkusi znanja (standard assessment tests, SAT) in ocenjevanja učiteljev (teacher

assessment, TA) pri temeljnih predmetih. Razvoj, preverjanje in izpopolnjevanje tega režima je bila zahtevna naloga, ki je potrebovala veliko finančnih in drugih virov in sredstev.

Kompromisov, ki jih je treba skleniti pri razvijanju nacionalnega sistema – na primer med predstavljanjem idealnih standardov in odsevom resnične prakse, med natančnostjo in jasnostjo, med izčrpnostjo in prilagodljivostjo ali pa med zanesljivostjo in skladnostjo s pravim poučevanjem v učilnici –, je neverjetno veliko; poročil in polemik o teh zadevah in prilagoditvah režima, ki sledijo, je preveč in so prepomembni, da bi tukaj skušali narediti kakršenkoli povzetek. Primera ocenitvenih raziskav sta Sainsbury et al. (1992) in Ashby et al. (1995). Koristni pregledi tovrstne tematike vključujejo Blacka (1998) in Whettona (1997). Podrobna razprava o nekaterih ključnih teoretičnih področjih je predstavljena v Sainsburyju in Sizmurju (1998).

Hkrati so predlogi za ta sistem – ki jih je zbrala Delovna skupina za ocenjevanje in preverjanje (Task Group on Assessment and Testing; gl. GB.DES/WO 1987, 1988) – vključevali predpise za uporabo ocenjevanja nacionalnega kurikula za ocenjevanje in primerjavo uspešnosti *šol*. To je bilo eno od sredstev, ki jih je sprejela vlada, da bi postale šole javno odgovorne za svoje delo, ker jim je zanj dodelila znatne finančne in upravne odgovornosti (gl. podpoglavje 3.1.3). Poleg tega so še trdili (ali vsaj domnevali), da bodo starši glede na rezultate izbrali šolo in da bo to pomagalo dvigniti standarde na trgu izobraževanja. Ocenjevanje uspešnosti učencev v Angliji in Walesu je tako v zadnjem obdobju prenove bolj ali manj nerazdružljivo povezano tako z odgovornostjo kot s tržno vrednostjo – gledano za nazaj, je to verjetno zveza, ki ne obeta nič dobrega.

Rezultati z zakonom predpisanih preizkusov niso objavljeni za posamezne učence, šola pa jih sporoči staršem (in tudi učencem samim); rezultate, zbrane na šolski ravni, vsako leto objavi

DfEE v razpredelnicah uspešnosti. Tudi rezultate, zbrane na lokalni ravni, objavi DfEE in jih uporabijo za presojo uspešnosti lokalnih šolskih uprav. Rezultate, zbrane na nacionalni ravni, objavi prav tako DfEE; uporabijo jih za spremljanje napredovanja glede na nacionalne cilje.

*Prednosti:* Zaradi zmešnjave ali kombinacije različnih načinov dela je težko izločiti prave prednosti in pomanjkljivosti nacionalnega sistema ocenjevanja. Upravičeno pa lahko trdimo, da je sodelovanje mnogih učiteljev in drugih strokovnjakov pri ustvarjanju skupnega razumevanja namena in pomena ocenjevanja, ne glede na omejitve sistema, kakršen je zdaj, zelo učinkovito v obsežnem programu nacionalnega sestavljanja testov in evalvacije. Spremenila se je tudi podlaga: dandanašnji bi bilo težko prepričljivo zagovarjati trditev, da dosežka in napredka učencev ni treba ocenjevati in da ni treba iskati vzrokov za slabe dosežke.

*Pomanjkljivosti:* Poleg dvomov o veljavnosti in zanesljivosti, ki se nanašajo na vse ocenjevalne režime, ima angleški model še druge pomanjkljivosti. Ker sistem velja za vse in ni delan z "lahkim vzorčenjem" in ker se rezultati učencev uporabljajo tudi za dajanje šol in vedno bolj tudi uspešnosti učiteljev pod drobnogled, obstaja resna nevarnost "učenja za teste". To pomeni, da postane uspešnost pri testih to, kar v resnici v večini učilnic žene kurikulum in njegovo poučevanje, namesto da bi bilo obratno. Prišlo je do poplave testov za testiranje za neobvezne letnike, ker si učitelji prizadevajo "napovedati" prihodnji uspeh iz tekočega. Največja nevarnost ali strah je, da bi kurikulum postal osiromašen, čeprav to še ni čisto dokazano. Poleg tega pa nekateri razlagalci dokazujejo, da ocenjevalni sistem, ki naj bi služil več namenom, za katere se izkaže, da si nasprotujejo, nobenemu od njih ne služi tako, kot bi moral.

### 1.3.1.3 Merjenje institucionalne uspešnosti

V podpoglavju 3.1.2 je podan povzetek o tem, kako so vpeljali merjenje institucionalne uspešnosti. Ta zadeva je postala razvpita tudi zato, ker so šolski rezultati v obliki razvrščenih dosežkov ali "razrednih razpredelnic" prišli na dan z veliko publicitete – objavljena so bila imena, pa tudi osramotili so jih – v lokalnem in nacionalnem tisku. Obstaja pa tudi zanimiv članek (Downes 1998), v katerem avtor, ravnatelj, dokazuje, da "je odpiranje šol javnemu interesu dobrodošel, a zakasnel korak v razvoju".

Originalne predloge (ki jih je izdelal TGAT), da bi s pomočjo uspešnosti učencev ocenjevali šole, sta že kmalu kritizirala Goldstein in Cuttance (1988), češ da bi bil ta sistem krivičen tako do šol kot tudi do skrbnikov in staršev ter t. i. selektorjev kakovosti v izobraževanju. Dokazovala sta, da bi sistem, ki primerja šole na podlagi ocenjevanja šolarjev v okoliščinah, kjer obstajajo bodisi družbene in izobrazbene razlike med učenci neke šole – pri čemer se predvideva, da v resničnem življenju take razlike seveda so –, prikrikl pravo stopnjo napredka učencev različnih šol, ne pa standardov, ki bi jih dosegli. Ta trditev, ki so jo od tedaj že obdelali v številnih člankih in poročilih, je bila bistveni nasprotni udarec tedanjemu vladnemu zavzemanju za "razredne razpredelnice" uspešnosti, ki temeljijo na "neobdelanih" rezultatih.

Zaposleni na področju vzgoje in izobraževanja so razdeljeni v dva tabora glede razrednih razpredelnic. Prvi menijo, da jih je treba odpraviti, čeprav bi obenem najraje še dodali, s čim bi jih nadomestili, da bi zadostili zakonitim zahtevam po odgovornosti. Drugi pa zahtevajo, da bi morali razviti mere uspešnosti na "dodano vrednost", torej mere, ki upoštevajo statistična poročila o spremenljivkah v ozadju, za katere je znano, da so v korelaciji z

dosežkom. To je smer prejšnje konzervativne vlade, nadaljuje pa jo laburistična državna uprava. Kljub vsemu je ta pot preprejena s tehničnimi in operativnimi težavami, še zlasti glede izoblikovanja kazalcev, ki bi jih brez težav lahko razumelo "nepoučeno" občinstvo, tj. prav tisti ljudje, za katere so bile razpredelnice uspešnosti pravzaprav izdelane – starši in javnost torej. Pregled razvoja in vprašanj na temo "dodane vrednosti" najdemo v Foxmanu (1997) in Saundersu (1999).

To področje še vedno razvijajo DfEE, QCA in OFSTED; sprejem najnovejšega sistema kazalcev uspešnosti, imenovanega "jesenski paket", je bil pri učiteljih in višjih upraviteljih zelo različen. Treba bo dokazati, da se prizadevanja pri zbiranju, analiziranju, poročanju in širjenju podatkov kažejo tudi v izboljšavah šol.

*Prednosti:* Zaradi razpredelnic uspešnosti so vrata šol kot še nikoli doslej na široko odprta očem javnosti, večji del polemike, ki so jo povzročile na področju vzgoje in izobraževanja in zunaj njega, pa je upravičen in omenja temeljna načela in zadeve tehnične narave, postranske zadeve in "zaščitena območja". Ena od posledic je ta, da šolsko osebje bolje razume prednosti, uporabo in omejitve statističnih podatkov, česar si še pred desetletjem ni bilo mogoče zamisliti.

*Pomanjkljivosti:* Obstaja nevarnost, da bodo osredotočanje na merjenje uspešnosti in analize na dodano vrednost ustvarili "premestitveno dejavnost" za šolsko osebje, in to prav v času izjemno obsežnih dejavnosti na področju vzgoje in izobraževanja, ki jih prinaša zakonodaja laburistične vlade. Še večja nevarnost pa je ta, da je veliko prizadevanj na področju politike, raziskovanja in prakse neke dejavnosti, ki morda na koncu koncev sploh ne bodo sposobna

javnosti zrcaliti kompleksne realnosti uspešnosti šol, po drugi strani pa ne bodo niti zanesljiv način dvigovanja standardov v vseh šolah.

#### **1.3.1.4 Neodvisno preverjanje ustanov**

Britanski inšpektorat so ustanovili pred več kot 150 leti z nalogo, da obvešča vlado in ji na podlagi inšpekcije svetuje o izobraževalnih storitvah. Ko je bila v poznih sedemdesetih letih v ospredju domnevna povezava izobraževanja s stanjem gospodarstva, je bila vlada bolj kot kdajkoli prej zainteresirana, da bi v šolah dosegali standarde; v zvezi s tem, kar so poučevali in se učili, v zvezi z zdravim upravljanjem šol in lokalnih šolskih uprav ter v zvezi s financiranjem, sredstvi in "cenovnimi ugodnostmi" – informacije o tem bi lahko pridobili z inšpekcijami (Bolton 1998). Inšpekcije so opravljale lokalne šolske uprave za šole na svojem območju, vendar pa niso bile v skladu s splošnimi in strogimi kriteriji. Inšpektorat je bil zato tista osrednja ustanova, ki je posredovala podrobne najnovejše informacije o rezultatih in delovanju izobraževalnega sistema.

Od leta 1983 dalje je treba inšpekcijska poročila objavljati; pojavilo se je pričakovanje, da bodo vsa poročila inšpektorata – tudi poročila o posebnih vladnih pobudah, o izpopolnjevanju učiteljev, usposabljanju mladine in porabi lokalnih šolskih uprav – postala javna. Od leta 1988 objavljajo letna poročila, ki so povzetki šolske inšpekcije (gl. npr. GB.DES.HMI 1989, 1990, 1991, 1992). V zvezi s tem referatom je pomembno, da se je vsako poročilo v veliki meri osredotočalo na zadeve, povezane z uvajanjem nacionalnega kurikula.

Ker je bila potreba po inšpekcijskih dokazih vedno večja in prav tako zahteva po večji odgovornosti v sistemu, se je struktura inšpektorata morala spremeniti. Na voljo so bile štiri

izvedljive možnosti (Bolton 1998), leta 1992 pa so izbrali model, po katerem so neodvisni inšpektorati prevzeli delo inšpekcije. Tega opravljajo po ustaljenih nacionalnih kriterijih in protokolih, kakovost njihovega dela pa nadzoruje Britanski inšpektorat, in sicer iz popolnoma neodvisne Pisarne za standarde v vzgoji in izobraževanju (Office for Standards in Education, OFSTED).

Rezultati dela so po eni strani tako kot prej individualna poročila o šolah, na katerih izsledke se morajo zdaj odzvati upravni organi šol z objavljenim načrtom dela in roki za izvedbo, po drugi strani pa letna poročila glavnega inšpektorja. Zdaj je težava v tem, da nasveti in smernice niso več del poročil o šolah, zaradi česar so nekatere šole izpostavljene, češ da imajo "resne šibkosti", in ne omogočajo učinkovite podpore za njihovo odpravljanje. Letnih poročil, ki vendarle ponujajo splošna načela in napotke, niso pisali in jih tudi ne morejo pisati ljudje, ki opravljajo inšpekcijske preglede, na katerih ta poročila temeljijo.

*Prednosti:* Sedanji sistem temelji na jasnem izhodišču kriterijev, ki pokrivajo precejšen del kurikularnih, vzgojnih in družbenih ukrepov, ki bi jih morale izvajati šole, zadevajo pa vprašanje zmogljivosti in učinkovitosti na področju porabe javnih sredstev. Zamisel o objavi tako pozitivnih kot negativnih vidikov šol (o večini šolskih poročil pišejo v lokalnem tisku), ker s tem za slabo prakso ni "nobenega skrivališča", se zdi večini staršev privlačna.

*Pomanjkljivosti:* OFSTED je v svoji sedanji podobi pritegnila toliko nasprotujočih si mnenj – nestalnost inšpekcijskih timov in vprašljiva natančnost njihovih dognanj ob polemičnih izjavah –, da je skrajno težko razločiti te pomanjkljivosti od onih manj opaznih. Za obvladovanje nestalnosti timov, pravi Bolton (1998), morajo biti inšpekcijska izhodišča strogo preskriptivne narave in se tako ne morejo uporabljati kot instrument opazovanja prave

raznovernosti in kompleksnosti šol. Med teoretiki na področju vzgoje in izobraževanja vlada trdno prepričanje, da bi se moral inšpektorat spet ukvarjati s svetovanjem in usmerjanjem. Obstaja pa tudi vprašanje glede meja odgovornosti inšpektorata v primerjavi z DfEE.

### **1.3.1.5 Programske evalvacije**

Državna uprava je poleg nacionalnega kurikula in njegovega ocenjevanja v zgodnjih osemdesetih letih uvedla številne posamezne politične programe, namenjene prenovi ukrepov na drugačne načine. Številni programi so se ukvarjali s prilagajanjem izobraževanja – zlasti poklicnega izobraževanja in usposabljanja (vocational education and training, VET) – potrebam britanskega gospodarstva; na primer Program usposabljanja mladine (Youth Training Scheme), Pobuda za tehniško in poklicno izobraževanje (Technical and Vocational Education Initiative, TVEI) in Compacts. Poznejši programi, kot sta programa nacionalnega opismenjevanja in računanja, so usmerjeni na zgodnejša leta.

Ker tovrstne programe znatno finančno in z drugimi sredstvi podpira vlada, se od njih pričakuje, da bodo prepričljivo vplivali na dosežke mladine; njihovo evalvacijo finančno podpira vlada, ker se ji zdi potrebno, da oceni njihov vpliv. Nekatere zgodnejše programske evalvacije so bile pomembne za razvoj kazalcev učinkovitosti, po katerih presojava vpliv. Ker pa se programi pogosto uvajajo postopoma, vlade običajno želijo formativen vnos poleg sumativnih evalvacij. Potemtakem je torej pomembno, da so med vlado in evalvatorji dobri delovni odnosi in da tako eni kot drugi razumejo in spoštujejo poklicne prioritete. Zgodi se tudi, da je znotraj široke nacionalne strategije in izhodišč usmerjanja odgovornost za programsko strukturo, vsebino in podajanje na ramenih lokalnih upraviteljev, in sicer zaradi odzivnosti na lokalne gospodarske in izobraževalne potrebe. Evalvacije morajo upoštevati različne modele uvajanja in procesa, da lahko osvetlijo potencialno diferencialne rezultate,

tako da znanje in spretnosti, ki se zahtevajo od evalvatorjev, presegajo tradicionalno psihometrično raziskovanje (gl. Parlett in Hamilton 1987). Nekateri razlagalci trdijo, da so programski evalvatorji, ker jih financirajo pobudniki programov, nagnjeni k temu, da se postopoma poistovetijo, s sabo ali z drugimi na nepriznan način (posebno s tistimi, ki jih evalvirajo), s politično perspektivo; to škodi nepristranskosti evalvacije (gl. Kushner in MacDonald 1987).

*Prednosti:* Programske evalvacije se izjemno razlikujejo, vendar pa je kot posplošitev pošteno povedati, da spodbujajo razpravo in razvoj primernih kvantitativnih in kvalitativnih lokalnih in nacionalnih kazalcev uspešnosti. Mnoge so tudi poudarile pomembnost razumevanja upravljanja sprememb in interaktivne narave inovacij na področju vzgoje in izobraževanja, tako da sta na primer koncepta "dodana vrednost" in "pomnožitveni efekt" nadomestila ali vsaj dopolnila pojem "vpliva".

*Pomanjkljivosti:* Splošna težava je v tem, da programske evalvacije omejujejo lastno področje raziskovanja in običajno ne vključuje (i) evalvacije same politike, ampak zgolj njeno uvajanje; (ii) omejuje tudi kakršnokoli primerjanje informacij, da bi osvetlile razmerje med stroški in učinkovitostjo med dvema programoma, ki se ukvarjata s podobnimi zadevami.

#### **1.3.1.6 Formalne konzultacije in konzultativni pregledi**

Pošteno je povedati, da sta nacionalni kurikulum in z njim povezana prenova pri mnogih zaposlenih v vzgoji in izobraževanju zlasti na začetku naletela na nenaklonjenost in celo sovražnost, deloma zaradi hitrosti in premalo posvetovanj pri uvajanju, pa tudi zaradi zahtev, ki so postavljene šolskim upraviteljem in učiteljem v razredu. Kot smo že omenili, ni bilo

popolnoma razvidno, da so evalvacijo in programe spremljanja uporabljali na odprt, povezan in dosleden način za prilagajanje prenove. Šele poznejše državne uprave so spoznale, da je za vzdrževanje izvedljivosti in verodostojnosti prenove v očeh javnosti potrebna obsežna revizija kurikula in ukrepov, ki ga podpirajo. Tedanja Uprava za šolski kurikulum in ocenjevanje (School Curriculum and Assessment Authority) je imela vrsto konzultacij, kar je povzročilo objavo high-profile, poljudno napisanih poročil, ki so jih spremljala priporočila: ti pregledi so postali znani pod imenom predsednika, sira Rona Dearinga (gl. Dearing 1993, 1994). Grobo rečeno, njihova posledica so bile racionalizacije in poenostavitve kurikula in njegovega ocenjevanja. Čeprav so bile sprejete z odprtimi rokami, je treba ujeti ravnotežje in se izogniti nenehnemu "krpanju" in prekinitvam.

Pobudo za zadnje poročilo je dala na začetku leta 1998 laburistična vlada – nova državna uprava, ki se želi posvečati vzgoji in izobraževanju kot konstitutivnemu delu svoje politike, njeno stališče glede kurikula pa je nekoliko drugačno od konzervativne vlade. Njen namen je, da bi imeli od septembra 2000 v vseh šolah pripravljen pregledan in popravljen kurikulum, ki bi zmanjšal prekinitve in se osredotočil na dvig standardov. Ministrovi predlogi za pregledana in popravljena izhodišča nacionalnega kurikula ne obsegajo le nadaljnjih poenostavitev, ampak tudi večjo prilagodljivost in bolj jasno osnovo vrednot (gl. prejšnji referat slovenske NFER). Organ, odgovoren za svetovanje in praktične smernice, je Uprava za kvalifikacije in kurikulum (Qualification and Curriculum Authority); v kratkem se bo lotila novih konzultacij, katerih izvajanje bo prepustila raziskovalni organizaciji. Po določitih za konzultacije bodo predloge poslali vsem šolam in lokalnim šolskim upravam, poklicnim organizacijam in skupinam, ki podpirajo njihovo delo, združenjem učiteljev, poslovnim in proizvodnim podjetjem. Konzultacijski vprašalnik bodo poslali vzorcu šol, drugim pa bo na voljo, če ga bodo zahtevali; o analizi pregleda bo izdelano tudi poročilo.

### **1.3.1.7 Neodvisne raziskave**

V Angliji in Walesu so poleg državnih ministrstev in uradov številni organi, ki bodisi stalno ali zgolj občasno financirajo raziskave na področju vzgoje in izobraževanja. Je pa tudi veliko ustanov, od univerzitetnih oddelkov za izobraževanje do zasebnih svetovalcev za upravljanje, ki izvajajo raziskave različnih in nasprotujočih si perspektiv, od postmodernih do upravnih. Tako je torej nemogoče podati pregled neodvisno financiranih raziskav h kurikularni prenovi v osemdesetih in devetdesetih letih, čeprav je zasilno izhodišče pregled, omenjen v podpoglavju 3.1.9. Primer uporabne in obsežne empirične raziskave, ki ga v tem poročilu ne navajamo, je Weston et al. (1992), ki je preučeval upravljanje celotnega kurikula za učence od 5 do 16 let in evalviral zahteve Zakona o izobraževalni reformi (Education Reform Act) v luči prizadevanj šol in lokalnih šolskih uprav za vzdrževanje, ustvarjanje ali razvijanje skladnosti v kurikulu.

### **1.3.1.8 Institucionalna samoevalvacija**

V prvem poročilu NFER Pedagoškemu inštitutu je bilo omenjeno, da je bila večina pred kratkim opravljenega dela ministrstva za izobraževanje in zaposlovanje v Angliji osredotočena na razvijanje sistematičnega načina samoevalvacije šol kot del vladnega dvigovanja ravni dosežka. Bela knjiga z naslovom *Odličnost v šolah* in z njo povezana poročila o napredku v šoli, zastavljanju ciljev in vzpostavljanju instrumentov za ocenjevanje so predstavljali smernice za petstopenjski cikel revidiranja in zastavljanja ciljev, ki so ga izvajale šole. Jasno je, da je samoevalvacija pomemben vidik zagotavljanja kakovosti v izobraževanju, čeprav – kot je bilo omenjeno v prejšnjem referatu NFER – obstaja vrsta zadev, katerih kontekst in upravljanje je treba še razložiti.

Ocenjevalna komisija, OFSTED in DfEE so dali pobudo za pomoč lokalnim šolskim upravam pri podpiranju sistematične samoevalvacije v šolah.

*Prednosti:* Samoevalvacija lahko doseže dele življenja in prakse v ustanovah, kamor večina drugih oblik evalvacije ne seže; ob primerni uporabi lahko poveže osebje in učence na "ustvarjalni" in poklicno ter osebno bogat način razumevanja samih sebe in svoje vloge.

*Pomanjkljivosti:* Nekateri oblikovalci politike, ki bi se jim zdelo idealno, da bi bila samoevalvacija oblika ponotranjene inšpekcije, vidijo v njej "mehko opcijo". Poleg tega je težko dati splošno oceno vpliva samoevalvacije na standarde, čeprav je z mislijo na to v letih 1997/98 NFER pri DfEE naročil obsežno kvalitativno raziskavo.

### **1.3.1.9 Izčrpni pregledi sedanjih raziskav**

Edini avtorici znan primer izčrpne raziskave na področju evalvacije nacionalnega kurikula, tj. raziskave, ki naj bi zajela vse od vlade objavljene in neobjavljene raziskave, in vse verodostojne neodvisne raziskave, je poročilo, ki ga je naročila tedanja Uprava za šolski kurikulum in ocenjevanje pri nekem angleškem univerzitetnem oddelku za izobraževanje, a le-to pozneje ni bilo objavljeno (Ford et al. 1998). Zaradi tega je status poročila dvomljive narave, izvlečke iz gradiva, ki sledijo, pa bi morali obravnavati kot zaupne. Podrobnejši pregled seznama virov pri tem poročilu nam daje slutiti, da pregled pravzaprav sploh ni izčrpen. Kakorkoli pa je ta dokument že zaradi preglednega pristopa izjemen; ključne točke so pomembne za Slovenijo. Izstopajo naslednje:

### *Izobraževalni nameni*

- "Koncept državljanstva, ki je opredeljen tako, da otroke pripravlja na življenje dejavnih pripadnikov demokratične skupnosti, se v literaturi nenehno pojavlja zato, da bi to postalo osrednje načelo kateregakoli pregledanega in popravljenega kurikula."<sup>1</sup>

### *Kurikularna struktura*

- "V nekaterih državah sistemi kurikularne evalvacije upoštevajo širši obseg kurikularnih ciljev in so manj odvisni od splošnega testiranja /.../ kot v Angliji in Walesu. V številnih državah uporabljajo metode vzorčenja, podobne nekdanji Enoti za ocenjevanje uspešnosti" [gl. zgoraj].
- "Premalo je bilo raziskav o vplivu kurikula na učenje in dosežke učencev ter različnih načinov njegove organizacije."
- "Naraščajoče poznavanje kognitivnega razvoja učencev še ni vplivalo na strukturo kurikula ali predmetov. Na tem področju bo še potrebno več dela."
- "V literaturi vlada napetost med konceptom 'prilagodljivosti' ali lokalne determinacije in osrednjimi predpisi; glavna skrb je ta, da lahko prilagodljivost, če ni usmerjana, vodi k še manj uravnoteženemu kurikulu, kot je sedanji."

### *Medpredmetne teme*

- "Vedno bolj narašča zanimanje za določanje učnih rezultatov glede na medpredmetne teme, na primer družbene raziskave, državljska vzgoja ter osebni in družbeni razvoj in njihovo povezovanje z učinkovitim sistemom ocenjevanja, tako da bo to področje kurikula dobilo status, ki mu pripada."

---

<sup>1</sup> To priporočilo je del ministrovih predlogov za pregledana in popravljena kurikularna izhodišča, kot je omenjeno v NFER-ovem prejšnjem poročilu in zgoraj.

## *Življenjsko učenje*

- "'Življenjsko učenje' in 'učeca se družba' sta koncepta, ki se borita za svoj prostor pod soncem; njune številne različice trenutno tekmujejo v bitki za pozornost. Nekatere poudarjajo gospodarsko tekmovalnost, druge družbeno vezljivost, tretje naučiti se učiti, nekatere pa so kar kombinacija vseh treh."
- "V literaturi vlada silna podpora mnenju, da je treba bolj poudariti razumevanje poteka učenja, ter načina, kako pridobimo znanje, spretnosti in vedenje, potrebne za razvoj ljudi, ki se učijo vse življenje. Pedagogika ostaja izjemno slabo raziskano področje."

## *Standardi*

- "Obstajajo številne izrazite težave glede [zaželene] veljavnosti in zanesljivosti testov, tako da bodo morali še dokazati, koliko so v resnici vredni kot primerne mere standardov."
- "Nemogoče je nedvoumno zatrditi, da so se standardi uspešnosti izboljšali ali poslabšali zaradi nacionalne kurikularne preнове, in tako ostaja polemika o standardih, tako mednarodna kot tudi v Angliji in Walesu, polna nesoglasij."

## *Ocenjevanje*

- "Obstajajo dokazi o izboljšanju dela v razredu v prvem vzgojno-izobraževalnem obdobju [od 1 do 3 let], ki so ga narekovale zahteve ocenjevanja."
- "Nekateri dokazi kažejo, da sta še zlasti pri slabših učencih dobro mnenje o sebi in nagnjenost k učenju prizadeta zaradi nekaterih ocenjevalnih praks."

- "Starši v glavnem sprejemajo z zakonom predpisano ocenjevanje in z njim povezane poročevalne ukrepe. Vendar pa to v celoti ne poteši njihove želje po podrobnejših informacijah glede učenja in razvoja njihovih otrok."

### *Obvladovanje izvajanja*

- "Premalo pozornosti je bilo posvečene procesu spreminjanja z zakonom predpisanega vodenja v kurikularno vsebino, ki jo podajajo učitelji, doživljajo pa učenci. Zaradi tega so se ponekod preveč zanašali na učbenike in objavljene programe, dajali prevelik poudarek podajanju vsebine in premalo pozornosti procesu poučevanja in učenja."
- "Ne da se izračunati, koliko časa in energije so učitelji zaman vložili v svoje delo, to pa zaradi pogostega spreminjanja nacionalnega kurikula. To je poslabšalo njihovo dobro voljo in pustilo posledice pri nadaljnjem razvoju."

### *Postopki v razredu*

- "Učitelji se pri uvajanju vidikov reševanja problemov, preiskovanja in raziskovanja v kurikulu soočajo s težavami. Kaže, da je to deloma zaradi omejitev učiteljev pri konceptualnem razumevanju predmetov, zaradi prevlade določenih slogov poučevanja in zaradi omejenega obsega interakcije v razredu, ki se ji le težko odpovejo."
- "Pomemben del nacionalnega kurikula, Priporočila, se ne izvaja, ker posvečajo premalo pozornosti podpori učiteljev pri spreminjanju dela v razredu, posebno tam, ko to vključuje pomoč učencem pri boljšem razumevanju lastnega učenja."

### *Razvoj učiteljev*

- "Številne raziskave kažejo, da sta prepričanje in predstava učiteljev o poučevanju močna posredovalna dejavnika pri izvajanju nacionalnega kurikula. Premalo pozornosti je posvečene spreminjanju ali pa izkoriščanju prepričanja učiteljev o poučevanju in učenju z omejenim raziskovanjem učinka in refleksijo."
- "Razvoj učiteljev kot ključni dejavnik pri učinkovitem spreminjanju so v glavnem spregledali, kurikularne spremembe pa so obravnavali kot zadevo tehnične, ne pa človeške narave."

NFER ni uspelo dognati, zakaj pregled ni bil objavljen. Čeprav izsledki morda niso popoln in natančen pogled na položaj v Angliji in Walesu, pa je treba zgornje točke resno upoštevati kot "zgodnje opozorilo" za druge sisteme.

### ***1.3.2. Nov generični razvoj na področju evalvacije v Veliki Britaniji***

Kot posledica novih pritiskov na državna ministrstva, ki morajo natančneje utemeljiti porabo sredstev iz javnega sektorja, morata zdaj raziskovanje in evalvacija na področju vzgoje in izobraževanja zadostiti strožjim zahtevam politike, dokazati, da se izplača, in dati boljši, bolj sistematičen in usklajen dokaz, "kaj deluje".

#### **1.3.2.1 Ocenjevanje cenovno ugodnega**

V razpravah o "ekonomiki izobraževanja" se običajno pojavlja tudi vprašanje, kakšni so stroški izobraževanja, tj. odstotek bruto nacionalnega dohodka, in vprašanje o rezultatih izobraževanja, to so spretnosti, osebno bogastvo in produktivnost. "Cenovno ugodno" na

področju izobraževanja pa ni samo po sebi razvidno načelo ali kriterij: v hipu se nam porodi vprašanje, kakšna vrednost in za katere zainteresirane skupine.

Včasih se izmenično uporabljajo izrazi "razmerje med stroški in učinkovitostjo", "cenovno ugodno", "najboljša vrednost", "analiza stroškov in koristi", "analiza obrestovanja", "varčevanje, učinkovitost in uspešnost", čeprav ne pomenijo isto. V veliko pomoč nam je opis razlike razmerij med stroški in učinkovitostjo ter analizo stroškov in koristi, ki ga je podala Holtermannova (1998): "Raziskave razmerja med stroški in učinkovitostjo običajno merijo rezultate s kvantitativnimi pojmi, vendar ne merijo vedno vseh rezultatov in ne skušajo dati rezultatom, ki niso opisani z monetarnimi izrazi, monetarne vrednosti." Nasprotno pa analiza stroškov in koristi "meri vse koristi in stroške in daje monetarno vrednost vsem koristim in tudi vsem stroškom". Nadaljuje: "Vendarle pa je včasih možno in sprejemljivo, da merimo vse koristi z monetarnimi izrazi, zato je analiza stroškov in koristi v tem formalnem pomenu prava redkost."

Holtermannova pravi, da "analiza stroškov in koristi želi odgovoriti na vprašanje, kaj se najbolj obnese glede na stroške, vendar to ni mogoče, dokler najprej ne odgovorimo na vprašanje, kaj je najboljše". Ta zadeva je tesno povezana s težnjo sedanje vlade po ustanovitvi trdnejše, z dokazi podprte osnove za politiko (gl. podpoglavje 3.2.2) in ima daljnosežne posledice na oblikovanje pobud in njihovo evalvacijo.

V Angliji in Walesu so leta 1983 ustanovili Ocenjevalno komisijo. Njeno področje dela naj bi zajemalo "ocenjevanje porabe ne le zaradi poštenosti in pravičnosti, ampak tudi zaradi cenovne ugodnosti" pri lokalnih (šolskih) upravah. Komisija je v začetku desetletja objavila vrsto poročil o šolskem upravljanju finančnih sredstev; naslov se glasi *Seštevanje (Adding Up*

*the Sums*, 1993 a in b, 1994, 1996). Ta poročila so ocenila, ali šole uspešno upravljajo svoj delež proračunskega denarja, in izdelala primerjalne vzorce porabe in sredstev, da so lahko ravnatelji in upravitelji primerjali, kako njihova šola uporablja sredstva v primerjavi z drugimi šolami.

Iz tega dela so razvidne tri točke: (i) potreba po novih trendih, da se bo lahko ocenilo "cenovno ugodno"; (ii) potreba po natančni statistični analizi za določitev, ali so ti trendi pomembni; (iii) zagotovitev, da so bile primerjave narejene na podlagi primerjave enakega z enakim in posledično odstranitev iz podatkov vse "zunanje dogodke", zaradi katerih primerjave ne bi bile veljavne.

Razmerje med stroški in učinkovitostjo je v Veliki Britaniji še premalo razvito področje evalvacije v vzgoji in izobraževanju. Zato namerava DfEE tri leta finančno podpirati raziskovalno središče, ki se bo ukvarjalo s tem področjem in ki naj bi razvilo program interdisciplinarnega raziskovanja ter na ta način skušalo najti odgovore na vprašanja, kot so:

- kako so lahko izobraževalni posegi povezani s spretnostmi učencev;
- kakšen vpliv ima vlaganje v izobraževanje v različnih fazah na dolgoročne rezultate;
- kako se na področju vzgoje in izobraževanja primerjajo politične opcije.

Raziskovalno središče se bo usmerilo v razvijanje metodologije za izvajanje analize stroškov in koristi na področju izobraževalnih posegov.

#### 1.3.2.2 Z dokazi podprta izobraževalna politika

Kot smo že omenili, je pristop, ki žene "z dokazi podprto" izobraževalno gibanje, tesno povezan z ustanavljanjem cenovne ugodnosti, in sicer tako, da mora ta predvidevati sistematično poznavanje, "kaj se obnese", utemeljeno na najboljših možnih dokazih raziskav. V zadnjem času se pojavlja precej kritik glede pomanjkljive koristnosti večine raziskovanja na področju vzgoje in izobraževanja, nesistematičnega ponavljanja (kar ni isto kot replikacija) prejšnjega dela ali pa preprosto neupoštevanje drugega dela, včasih dvomljive metodologije, nerazloženih temeljnih načel, nepodprtih izsledkov, zlasti pa pomanjkanja tehtnosti za politiko in za delo v razredu (za analitičen pregled na področju vzgoje in izobraževanja gl. Hillage et al. 1998; za bolj polemičen pogled gl. Hargreaves 1998).

Mednarodno gibanje za z dokazi podprto izobraževanje, ki se v veliki meri zgleduje po načelih, povzetih po modelu varovanja zdravja, si hitro pridobiva zaupanje med raziskovalci in oblikovalci politike: za pregled problematike in pristopov gl. npr. Davies (v pripravi), Harlen (1996), Sebba (1999), Slavin (1986). Glavna področja, ki se jih je treba lotiti, so:

- metodologija za izvajanje izčrpnih in usklajenih pregledov raziskav;
- izvajanje novih pregledov raziskav na ključnih področjih;
- zagotavljanje kakovosti in nadzor empiričnega raziskovanja (vključno s kvalitativnim);
- jasna analiza in poročanje o raziskovalnih študijah;
- razširjanje raziskovanja z oblikovalci politike in praktiki.

Danes je še prezgodaj, da bi napovedali, kakšni bodo praktični rezultati gibanja. Vladni predlogi vključujejo ustanovitev "skupnega raziskovalnega vira za izobraževanje in zaposlovanje", katerega namen bi bil:

- "s pridobivanjem ustreznih, zanesljivih, najnovejših dokazov, ki zmanjšujejo pristranost, povečati gotovost ustreznosti dokazov na danem področju;
- koordinirati in podpreti pregledovanje in primerjanje raziskovanja na področju vzgoje in izobraževanja;
- poiskati dostopen dokaz za obveščanje na področju raziskovanja, politike in prakse ter tako prispevati k razvoju celotne strategije razširjanja".

Predlagani vir bi te cilje dosegel z:

- "razvijanjem izčrpne baze podatkov tekočih in objavljenih raziskav;
- ustanovitvijo skupnih raziskovalnih skupin, ki bi se lotile revizij na določenih področjih in prevzele odgovornost za njihovo posodabljanje. Članstvo bi zajemalo raziskovalce, oblikovalce politike, praktike in druge, ki želijo imeti dostop do zanesljivih dokazov" [po domače so znani pod imenom Cochranove revizije po modelu, vzpostavljenem na področju varovanja zdravja].

Če se bo ta predlog ali pa vsaj kakšen podoben uresničil, bo pomenil ogromno naložbo v raziskovanje, hkrati pa bo prinesel tudi številne izzive raziskovalnim ustanovam in posameznikom. Mednje sodijo strinjanje z metodološko problematiko pri pregledovanju raziskovanja v vzgoji in izobraževanju, izoblikovanje vprašanj tipa "Cochranovih revizij", opis možnih modelov in izpeljava tega, kar bodo nemara daljnosežne posledice za raziskovalno prakso na področju vzgoje in izobraževanja (vključno z usposabljanjem in viri).

### **1.3.3. Kaj je še treba upoštevati**

1.3.3.1 Iz zgornje razprave izvirajo trije dodatni problemi, ki jih bo moral raziskati NK pri izvajanju evalvacije v Sloveniji. Prvi je odnos med vlado in neodvisnimi raziskovalnimi ustanovami, ker se nanaša na kakovost raziskovalnega upravljanja vlade; tu gre še zlasti za financiranje strateških raziskovalnih programov, specifikacijo raziskovalnih ponudb za izvedbo dela in preskrbovanje raziskovalnih pogodb. To področje je prezapleteno, da bi ga lahko obdelali v tem referatu, NFER pa bi bil pripravljen ločeno NK podrobneje seznaniti s tem, upoštevaje sorazmerno zrelost prakse ministrstva za izobraževanje in zaposlovanje v Angliji.

1.3.3.2 Drugi problem zajema razvoj primernih instrumentov, ki obsegajo vse, od testov in drugih pripomočkov za ocenjevanje uspešnosti prek vprašalnikov, razgovorov in opazovalnih razporedov do instrumentov za odprte etnografske raziskave. Tudi to je preveč raznoliko in tehnično zahtevno področje, da bi se z njim ukvarjali tukaj; domnevamo, da bo imel NK dostop, med drugim tudi prek Pedagoškega inštituta, do zahtevanih informacij. NFER pa bi po želji lahko ponudil tehnično vodstvo.

1.3.3.3 Tretji pa se kvalitativno razlikuje od prejšnjih dveh, a je izjemno pomemben za razvoj racionalnega sistema evalvacije. Zadeva preobilico nacionalnih uradov, ki se ukvarjajo z evalvacijo, spremljanjem, testiranjem, inšpekcijo in podobno. Holt (1981) to zelo primerno zapiše:

Za pluralistične družbe – in še zlasti za načrtovanje izobraževanja – je značilno, da je razmeroma preprosto dovoliti nastanek novih organizacij, skoraj nemogoče pa se jih je znebiti. Takoj ko se institucionalizirajo, njihovi pripadniki postanejo politično jedro, in to, za kar se je zdelo, da

je nastalo zgolj po naključju, se nenadoma zdi, da izpolnjuje neustavljivi namen.

V Angliji so se v zadnjih letih v vzgoji in izobraževanju zgodile racionalizacije posameznih kvaziavtonomnih nevladnih organizacij. Tako sta se leta 1993 Nacionalni kurikularni svet in Šolski izpitni in ocenjevalni svet, ustanovljena 1988. leta z Zakonom o izobraževalni reformi, združila v Šolsko kurikularno in ocenjevalno upravo. Ta organ se je leta 1997 z Nacionalnim svetom za poklicne kvalifikacije, ustanovljenim 1986. leta, združil v Upravo za kvalifikacije in kurikulum (Qualifications and Curriculum Authority, QCA). DfEE je po drugi svetovni vojni zaživel kot Odbor za izobraževanje, pozneje ministrstvo za izobraževanje in znanost; leta 1992 se je preimenoval v ministrstvo za izobraževanje in se potem združil z ministrstvom za zaposlovanje (ki je imelo prav tako pisano zgodovino v kontekstu povojnih potreb po "delovni sili" in skrbi za usposabljanje) in leta 1995 postal DfEE.

Znotraj DfEE obstaja polavtonomna enota, ustanovljena leta 1997, imenovana Enota za standarde in učinkovitost. Odgovorna je za ključna področja, ki jih je v Beli knjigi z naslovom *Odličnost v šolah*, uvedla nova laburistična državna uprava. Mednje sodijo učinkovitost šol in lokalnih šolskih uprav ter nacionalni programi za opismenjevanje in računanje, poleg tega pa še številni ciljno usmerjeni posegi, kot so območja šolskega udejstvovanja (Education Action Zones).

Britanski inšpektorat v primerjavi z DfEE obstaja zelo dolgo, saj je bil ustanovljen že davnega leta 1839 za inšpekcijo izobraževalnih zahtev v Angliji, čeprav je moral glavni inšpektor do leta 1988 izdelovati letna poročila, v katerih je povzel dognanja iz programov inšpekcij posameznih šol in kolidžev. V osemdesetih letih je začel inšpektorat obveščati vladno politiko, zlasti glede neprijetih ali spornih dejanj, ki se jih je lotila vlada zaradi

kritičnih poročil inšpektorata. Po drugi strani pa je bila kot posledica Zakona o izobraževalni reformi iz leta 1988 večina stvari, o katerih je poročal inšpektorat, resnično, ne pa tudi nujno, kritična do vidikov reform ali pa vsaj njihovega uvajanja in nenameravanih posledic (Bolton 1998). Tako oblikovalcem politike kot praktikom se je zdela njegova neodvisnost sporna, čeprav iz nasprotnih razlogov. Inšpektorat je leta 1992 nadomestil Urad za standarde v izobraževanju (Office for Standards in Education, OFSTED), ki je bil ustanovljen z namenom, da poveča odgovornost vlade z izhodišči za razširjen in zelo oglaševan program inšpekcij – vsako šolo je treba popolnoma pregledati vsaka štiri leta – in za usposabljanje za zasebne time inšpektorjev, ki so jih morali najeti za izvajanje tega programa (gl. Bolton 1998 za pogled na ta nedavni razvoj).

(Področje poobveznega izobraževanja in usposabljanja, tj. izobraževanja odraslih in visokošolskega izobraževanja in usposabljanja ob delu, ima še bolj pestro zgodovino, še posebno glede mehanizmov financiranja in odgovornosti. Pregled na tem mestu ni predviden! Med koristne vire sodita Young in Spours 1998 ter Trow 1998).

To, da za isto področje obstaja več organov z zakonitimi interesi, včasih povzroči "ozemeljske spopade", v katerih želijo ljudje zaščititi svoje organizacijske pridobljene pravice, povzroči pa tudi "mešana sporočila", ki jih dobivajo šole in javnost. Znak stopnje racionalizacije, ki jo je še treba doseči, celo znotraj razmeroma zaprtega področja, kot je objava podatkov o uspešnosti, je ta, da je še vedno mogoče, da QCA, OFSTED in DfEE pridobivajo statistike o uspešnosti šol in učencev, ki imajo namen pomagati šolam in lokalnim šolskim upravam pri spremljanju in samoevalvaciji, vendar so izračunane na različnih osnovah, upoštevajo različne spremenljivke in dajejo različne nacionalne norme.

Zato je morda v pomoč, da upoštevamo eno od priporočil OECD, in sicer:

- Struktura ustanov, vključenih v prenavo srednjega izobraževanja, je vsaj približno primerna za sedanjo fazo prenove, še posebno pomembna pa je osrednja vloga Nacionalnega kurikularnega sveta. Vendar bi moralo ministrstvo za izobraževanje in znanost revidirati strukturo ustanov, da jo poenostavi, kadar je to primerno.

#### **1.4. NEKAJ PRIMEROV "KRITERIJEV USPEŠNOSTI"**

Kot smo že povedali v uvodu, se NFER zdi pomembno, da bi bili ključni zainteresirani vključeni pri določanju in merjenju kriterijev uspešnosti. Hkrati se strinjamo z enim glavnih priporočil OECD (Reid et al. 1998), ki se glasi:

Sedanji razvoj bi moral temeljiti na jasnem in storilnostno usmerjenem sistemu, v katerem so rezultati vseh delov izobraževalnega sistema izpostavljeni rednemu merjenju. Obstajati bi morala hierarhija mer, in sicer na nacionalni, regionalni, lokalni in institucionalni ravni, ki bi prispevala k splošnemu vrednotenju učinkovitosti sistema in njegove cenovne ugodnosti.

Izhodišča, ki jih je NFER izdelal v svojem prvem poročilu, so bila zasnovana z mislijo na to zahtevo. V pomoč pri razvijanju kriterijev uspešnosti, smo opisali enega ali dva obdelana primera.

## 1.5. PRILOGA A PRIPOROČILA IZ REVIZIJSKEGA POROČILA OECD O SLOVENIJI

Navedeni so izvlečki priporočil iz obsežnejšega besedila.

### Razvoj kurikula

- Vsebino kurikula v osnovnem in srednjem izobraževanju bi morali skrčiti. Kurikulum mora temeljiti na znanju, spretnostih in vrednotah, pomembnih za slovensko družbo in za vse zainteresirane na področju izobraževanja, ne samo za učence in tiste, ki so pomembni s tradicionalnega akademskega stališča.
- Kurikularni razvoj mora zlasti na osnovnošolski ravni povezovati različne šolske predmete, v Izhodiščih kurikularne prenove pa mora biti več interdisciplinarnih in medpredmetnih tematskih sklopov ali tem, z namenom spodbujanja šol k uvajanju takih učnih programov in dejavnosti.
- Priporočamo, da bi nacionalna strategija mestnim skupnostim in šolam posredovala ideje in pristope za spopadanje s tradicionalnimi metodami poučevanja in usposabljanja, pri katerih je v središču učitelj, pri čemer naj upošteva razumevanje o izboljšanju šol, pridobljeno v mednarodni skupnosti.
- Nacionalni kurikularni svet je pravilno ravnal, ko je povabil vse učitelje v vseh šolah k razpravljanju in komentiranju dokumentov, vključenih v Izhodišča kurikularne prenove, in to bi se moralo nadaljevati. Pomembno je povabiti tudi SVIZ, da se udeleži razprave o oblikovanju novega kurikula, ki pa naj jo vodijo državni organi.
- Raziskovanje na področju vzgoje in izobraževanja mora uvajati inovacije v izobraževalni sistem in preverjati učinke procesa prenove v šolah in v sistemu. Ministrstvo za šolstvo in šport ter ministrstvo za znanost in tehnologijo morata izoblikovati jasnejšo strategijo za

raziskovanje na področju vzgoje in izobraževanja v Sloveniji, ki podpira prizadevanja nacionalne prenovе v vzgoji in izobraževanju, v kar naj vključi tudi Pedagoški inštitut in ljubljansko univerzo.

- Potekajočo kurikularno prenovо v Sloveniji bi lahko pospešili s skupnim razumevanjem, kaj pomeni kurikulum. To med drugim zahteva tudi vključevanje raziskovalcev s področja vzgoje in izobraževanja, usposabljanje ravnateljev ob delu in poudarjanje vloge učiteljev pri načrtovanju kurikula.
- Na področju poklicnega izobraževanja mora biti razvoj kurikulumov tesno povezan s potrebami na trgu delovne sile, zato morajo biti tudi delodajalci popolnoma vključeni v proces razvijanja, da po potrebi lahko podprejo mrežo delodajalcev.

## **Standardi rezultatov, ocenjevanje, spričevala in kakovost**

### **I. Osnovno izobraževanje**

- V osnovnem izobraževanju se zahtevajo standardi, usmerjeni na rezultate, kamor sodijo tudi medpredmetne spretnosti, kurikularne komisije pa bi si morale prizadevati za izdelavo dokumentacije na tem temelju. Ti standardi rezultatov bi morali biti osredotočeni na temeljne spretnosti in sposobnosti, ne pa na znanje, specifično za posamezne predmete.
- Instrumente ocenjevanja, ki učiteljem omogočajo, da redno preverjajo, ali njihovo poučevanje dosega zahtevane standarde, bi morali razviti centralno.
- Treba bi bilo pretehtati, ali naj bo v mehanizmu izbiranja učencev za prestižnejše šole takšen poudarek na zaključnih testih v osnovni šoli.
- Razviti in uvajati bi morali programe za usposabljanje učiteljev, ki bi jim pomagali nadomestiti tradicionalno preverjanje znanja z alternativnimi oblikami, določati standarde in jih zanesljivo izvajati.

- Ministrstvo za šolstvo in šport bi si moralo prizadevati za izoblikovanje instrumentov za ocenjevanje nacionalne uspešnosti na področju vzgoje in izobraževanja.
- Izobraževalni proces bi morali evalvirati in razmisliti o vlogi inšpektorata, ki naj bi bil v prihodnje neodvisen od ministrstva.

## **Standardi rezultatov, ocenjevanje, spričevala in kakovost**

### **II. Srednje in posrednje izobraževanje**

- Tako kot pri osnovnem izobraževanju se tudi pri srednjem zahtevajo standardi, usmerjeni na rezultate, prizadevanje po njih pa bi moralo imeti veliko prednost.
- Prehajanje iz razreda v razred ali iz programa v program v srednjem izobraževanju, kjer trenutno potekajo spremembe, bi moralo biti bolj prilagodljivo.
- Državni izpitni center bi morali prositi, da ob sedanji izpitni strukturi za prihodnost razvije specializirane instrumente ocenjevanja, pri čemer naj zaščiti pozitivne vidike mature.
- Struktura ustanov, vključenih v prenovu srednjega izobraževanja, je za trenutno fazo prenove povsem primerna, še posebno pomembna pa je osrednja vloga Nacionalnega kurikularnega sveta. Vendar pa bi moralo ministrstvo za šolstvo in šport revidirati strukturo ustanov, da bi jo po potrebi lahko poenostavilo
- Razvoj poklicnih standardov je še vedno v povojih, popoln razvoj sistema standardov, ki temelji na sposobnosti, pa bi moral imeti prednost. Njegovo razmerje do potreb trga delovne sile mora biti jasno, ustvarjati pa mora hierarhijo ali družino standardov, ki jih lahko uporabijo proizvodna podjetja ali poklici.
- Treba bi bilo skrbno raziskati in evalvirati učinek maturitetnih izpitov na poučevanje in učenje.

- Treba bi bilo pretehtati, kako bi lahko pospešili in poglobili ustreznost univerzitetnih programov, kar je v interesu zagotavljanja kakovosti.

### **Regionalizacija in decentralizacija**

- Decentralizacija znotraj izobraževalnega sistema je pomemben cilj, za katerega bi si morali prizadevati, tako da bi bili tudi zainteresirani tesneje vključeni v odločanje na regionalni in lokalni ravni.
- Ministrstvo za šolstvo in šport bi moralo premisliti, ali sedanji način financiranja šol daje prava znamenja znotraj decentraliziranega sistema.
- Decentralizacija je še posebno pomembna za poklicno izobraževanje. Zdi se, da na regionalni ravni na tem področju ni jasnega žarišča, zato bi morala vlada pretehtati predlog o "odborih za razvoj delovne sile" na regionalni ravni, kjer bi imela bodisi koordinacijske odgovornosti ali pa moč razdeljevanja virov.
- Ministrstvo bi moralo pazljivo preučiti, ali res obstaja izobraževalna potreba po tretji univerzi v Sloveniji, kaj to pomeni za javno porabo na področju vzgoje in izobraževanja in oportunitetne stroške za druge dele sistema, ali pa je bolje ustanoviti več posrednješolskih poklicnih in strokovnih ustanov.

### **Učinkovitost in viri v sistemu**

- Vlada bi morala pretehtati, ali bi morala vzpostaviti medresorno povezovanje za zavarovanje doslednega načina razvoja delovne sile čez vsa ministrstva.
- Sedanje delo o informacijah s področja trga delovne sile in poklicnega vodstva bi morali okrepiti, tako da bi se podjetja, posamezniki in vlada lahko odločali na podlagi trdno utemeljenih in pravočasnih informacij.

- Vlada bi morala preučiti dolgoročne, zlasti pa kratkoročne učinke demografskega razvoja v Sloveniji na prihodnjo obliko in strukturo sistema vzgoje in izobraževanja, učinek na osnovne šole (zlasti na oddaljenih območjih), vpliv politike in virov pa bi morali pokazati jasno, kot zadevo, ki bi jo morala vlada še preučiti.
- Ministrstvo za šolstvo in šport bi moralo pretehtati, ali je nedavno sprejeta visokošolska zakonodaja primerna, da se bodo univerze lahko prilagodile novemu položaju, s katerim se sooča Slovenija.
- Sedanji razvoj bi moral temeljiti na jasnem in storilnostno usmerjenem sistemu, v katerem so rezultati vseh delov izobraževalnega sistema izpostavljeni rednemu merjenju. Obstajati bi morala hierarhija mer, in sicer na nacionalni, regionalni, lokalni in institucionalni ravni, ki bi prispevala k splošnemu vrednotenju učinkovitosti sistema in njegove cenovne ugodnosti.

## 1.6. VIRI

ASHBY, J., BURLEY, J., HARGREAVES, E. and McCULLOCH, K. with SAINSBURY, M. and SCHAGEN, I. (1995). National Curriculum Assessment 1995, Key Stage 1: Evaluation in the Core Subjects. Unpublished report. Slough: NFER.

AUDIT COMMISSION (1993a). *Adding up the Sums: Schools' Management of their Finances*. London: HMSO.

AUDIT COMMISSION (1993b). *Adding up the Sums 2: Comparative Information for Schools*. London: HMSO.

AUDIT COMMISSION (1994). *Adding up the Sums 3: Comparative Information for Schools – 1994*. London: HMSO.

AUDIT COMMISSION (1996). *Adding up the Sums 4: Comparative Information for Schools – 1995/96*. London: HMSO.

BLACK, P. (1998). 'Learning, league tables and national assessment: opportunity lost or hope deferred?', *Oxford Review of Education*, **24**, 1, 57-68.

BOLTON, E. (1998). 'HMI – the Thatcher years', *Oxford Review of Education*, **24**, 1, 45-55.

BROOKS, G., HAGUES, N., HUTCHISON, D., MILNE, A., MINNIS, M., MORRIS, M., SAINSBURY, M., SCHAGEN, I., SCHAGEN, S., SIZMUR, S. and WHETTON, C. (1997). 'Measuring the effects of educational interventions'. Paper presented at the British Educational Research Association annual conference, University of York, 11-13 September.

BROWN, M. (1998). 'The tyranny of the international horse-race'. In: SLEE, R. and WEINER, G., with TOMLINSON, S. (Eds) *School Effectiveness for Whom? Challenges to the School Effectiveness and School Improvement Movements*. London: Falmer Press.

DAVIES, P. (forthcoming). 'Evidence-based healthcare/education', *British Journal of Educational Studies*.

DEARING, R. (1993). *The National Curriculum and its Assessment: Interim Report*. London: SCAA.

DEARING, R. (1994). *The Review of the National Curriculum: a Report on the 1994 Consultation*. London: SCAA.

DOWNES, P. (1998). 'The head's perspective', *Oxford Review of Education*, **24**, 1, 25-33.

FORD, K., CLARK, J., LEAT, D. and MILLER, J. (1998). An Analysis of Research into the Impact of the National Curriculum and the Implications for Teachers and Schools. Unpublished report. Newcastle: University of Newcastle upon Tyne.

FOXMAN, D. (1997). *Educational League Tables: For Promotion or Relegation?* London: Association of Teachers and Lecturers.

FOXMAN, D., HUTCHISON, D. and BLOOMFIELD, B. (1991). *The APU Experience 1977 – 1990*. London: School Examination and Assessment Council.

GOLDSTEIN, H. (1979). 'Consequences of using the Rasch model for educational assessment', *British Journal of Educational Research*, **5**, 211-20.

GREAT BRITAIN. DEPARTMENT OF EDUCATION AND SCIENCE. HER MAJESTY'S INSPECTORATE. (1989). *The Implementation of the National Curriculum in Primary Schools. A Report by HM Inspectors, Summer 1989 (HMI Report 332/89)*. London: DES.

GREAT BRITAIN. DEPARTMENT OF EDUCATION AND SCIENCE. HER MAJESTY'S INSPECTORATE. (1990). *Standards in Education 1988-89. The Annual Report of HM Senior Chief Inspector of Schools*. London: DES.

GREAT BRITAIN. DEPARTMENT OF EDUCATION AND SCIENCE. HER MAJESTY'S INSPECTORATE. (1991). *Standards in Education 1989-90. The Annual Report of HM Senior Chief Inspector of Schools*. London: DES.

GREAT BRITAIN. DEPARTMENT OF EDUCATION AND SCIENCE. HER MAJESTY'S INSPECTORATE. (1992). *Education in England 1990-91. The Annual Report of HM Senior Chief Inspector of Schools*. London: DES.

GREAT BRITAIN. DEPARTMENT OF EDUCATION AND SCIENCE AND WELSH OFFICE. (1987). *Task Group on Assessment and Testing: a Report*. London: DES and Welsh Office

GREAT BRITAIN. DEPARTMENT OF EDUCATION AND SCIENCE AND WELSH OFFICE. (1988). *Task Group on Assessment and Testing: Three Supplementary Reports*. London: DES and Welsh Office

HARGREAVES, D. (1998). *Creative Professionalism: the Role of Teachers in the Knowledge Society*. London: Demos.

- HARLEN, W. (1996). 'Educational research and educational reform.' In: HEGARTY, S. (Ed) *The Role of Research in Mature Education Systems*. Slough: NFER.
- HILLAGE, J., PEARSON, R., ANDERSON, A. and TAMKIN, P. (1998). *Excellence in Research on Schools*. Research Report RR74. London: DfEE.
- HOLT, M. (1981). *Evaluating the Evaluators*. London: Hodder and Stoughton.
- HOLTERMANN, S. (1998). *Weighing It Up: Applying Economic Evaluations to Social Welfare Programmes*. London: Joseph Rowntree Foundation.
- KUSHNER, S. and MacDONALD, B. (1987). 'The limitations of programme evaluation'. In: MURPHY, R. and TORRANCE, H. (Eds) *Evaluating Education: Issues and Methods*. London: Harper and Row.
- PARLETT, M. and HAMILTON, D. (1987). 'Evaluation as illumination: a new approach to the study of innovatory programmes'. In: MURPHY, R. and TORRANCE, H. (Eds) *Evaluating Education: Issues and Methods*. London: Harper and Row.
- REID, G. *et al.* (1998). *Reviews of National Policies for Education: Slovenia. Examiners' Report*. Paris: OECD.
- SAINSBURY, M. and SIZMUR, S. (1998). 'Level descriptions in the national curriculum: what kind of criterion referencing is this?', *Oxford Review of Education*, **24**, 2, 181-93.
- SAINSBURY, M., WHETTON, C., ASHBY, J., SCHAGEN, I. and SIZMUR, S. (1992). *National Curriculum Assessment at Key Stage 1: 1992 Evaluation*. London: SEAC.
- SAUNDERS, L. (1999). *Value Added Measurement of School Effectiveness: a Critical Review*. Slough: NFER.
- SEBBA, J. (1999). 'Educational research: developing and implementing the Government's action plan', *Research Intelligence*, **67**, 19-20.
- ŠETINC, M. (1999). 'International comparative studies of science achievement from the time perspective: the Third International Mathematics and Science Study versus the International Assessment of Educational Progress', *Assessment in Education*, **6**, 1, 43-56.
- SLAVIN, R.E. (1986). 'Best-evidence synthesis: an alternative to meta-analytic and traditional reviews', *Educational Researcher*, **15**, 9, 5-11.
- TROW, M. (1998). 'American perspectives on British higher education under Thatcher and Major', *Oxford Review of Education*, **24**, 1, 111-29.
- WESTON, P. and BARRETT, E. with JAMISON, J. (1992). *The Quest for Coherence: Managing the Whole Curriculum 5 – 16*. Slough: NFER.
- WHETTON, C. (1997). 'The psychometric enterprise or the assessment aspiration'. In: HEGARTY, S. (Ed) *The Role of Research in Mature Education Systems: Proceedings of the NFER International Jubilee Conference, Oakley Court, Windsor, 2-4 December 1996*. Slough: NFER.
- YOUNG, M. and SPOURS, K. (1998). '14-19 education: legacy, opportunities and challenges', *Oxford Review of Education*, **24**, 1, 83-97.

Prevedla:

Nika Schlamberger

