

**Dr. Mojca Štraus**

## Pomen mednarodnih raziskav znanja in nacionalnega preverjanja znanja za vrednotenje dosežkov učencev

**Povzetek:** Nacionalno preverjanje znanja in mednarodne raziskave v izobraževanju dajejo možnosti za vrednotenje znanja učencev v širšem spektru primerjav, kot je to mogoče pri preverjanju in ocenjevanju znanja v šoli. Mednarodne raziskave v izobraževanju so se pojavile prav iz zanimanja več držav za pridobivanje dodatnih informacij o dosežkih učencev in posredno kakovosti njihovega izobraževalnega sistema ob primerjavi z drugimi državami. Splošni koraki v izvedbi raziskave znanja so, da se predstavniki sodelujočih držav dogovorijo o populaciji učencev, zajetih v raziskavi, področjih znanja, ki bodo merjena, ter o instrumentih za merjenje znanja in zajemanje podatkov o spremljajočih dejavnikih. Instrumente rešujejo reprezentativni vzorci učencev, navadno sočasno v vseh sodelujočih državah. Mednarodne raziskave znanja, ki se izvajajo v cikličnih zajemih podatkov, kot so PISA, TIMSS in PIRLS, in izsledki nacionalnega preverjanja znanja, ki se izvaja vsako leto v šestem in devetem razredu osnovne šole, postajajo čedalje pomembnejša podlaga za oblikovanje strategij razvoja izobraževalnega sistema tudi v Sloveniji. Vendar, preden sklepamo o slabih ali dobrih rezultatih glede na razvrstitev države v lestvici sodelujočih držav ali glede na kak drug enoštevilčni kazalnik, se je treba seznaniti tudi s posameznimi elementi raziskav, saj utegnejo vplivati na možnosti našega sklepanja in pojasnjevanja rezultatov. Tema prispevka so konceptualne podobnosti in razlike v zasnovah raziskav na primeru zgoraj omenjenih raziskav ter njihov pomen in uporabnost tako pri oblikovanju standardov znanja kot pri ugotavljanju dejavnikov, ki so povezani z uspešnostjo učencev v šoli. Pri tem je zelo pomembno stalno preverjati veljavnost izpeljanih sklepov iz podatkov in rezultatov analiz, da lahko izsledke posameznih raziskav smiselno posplošujemo na širše populacije in daljša časovna obdobja.

**Ključne besede:** mednarodne raziskave v izobraževanju, izobraževanje, kazalniki izobraževanja, dosežki v izobraževanju.

UDK: 371.26

Pregledni znanstveni prispevek

*Dr. Mojca Štraus, raziskovalka, Pedagoški inštitut; e-naslov: mojca.straus@pei.si*

## Uvod

V raziskavah znanja iščemo odgovore na vprašanja, kot so: Kako dobro se učenci učijo v izobraževalnem sistemu (glede na splošna pričakovanja, cilje kurikula ali pripravo na nadaljnje življenje)? Ali obstajajo podatki o posebno močnih ali šibkih področjih v znanju učencev? Kateri dejavniki v šolskem sistemu ali zunaj njega so povezani z uspešnostjo učencev? Temeljni namen raziskav znanja je prispevati k izboljševanju kakovosti izobraževanja v državi. V zadnjih desetletjih so ravni znanja učencev, dijakov in študentov (v nadaljevanju imenovani splošno z besedo učenci) postale pomemben kazalnik te kakovosti. Tudi v Sloveniji je eden od ciljev prenovljenega šolskega sistema »doseganje mednarodno primerljivih standardov znanja« (Bela knjiga, str. 92).

Slovenija se je v letu 2004 pridružila državam, ki so se pripravljale na izvedbo tretjega cikla Programa mednarodne primerjave dosežkov učencev PISA (Programme for International Student Assessment), ki poteka v sklopu Organizacije za ekonomsko sodelovanje in razvoj (OECD). Pred tem so bile znane tudi že Mednarodna raziskava trendov znanja matematike in naravoslovja (TIMSS, Trends in International Mathematics and Science Study), Mednarodna raziskava bralne pismenosti (PIRLS, Progress in International Reading Literacy Study), ki obe potekata v sklopu Mednarodnega združenja za evalvacijo dosežkov v izobraževanju IEA (International Association for Evaluation of Educational Achievement). Ob teh raziskavah smo se lahko seznanili s primerjavami dosežkov slovenskih učencev z dosežki učencev iz drugih držav. Obenem vsako šolsko leto učenci ob koncu drugega prostovoljno in ob koncu tretjega obdobja obvezno opravljajo nacionalno preverjanje znanja (NPZ), pri katerem s preizkusi, pripravljenimi na nacionalni ravni, zbiramo podatke o njihovem znanju slovenščine (oziroma italijanščine in madžarščine), matematike in tretjega predmeta.

Kako nam torej te raziskave pomagajo odgovoriti na zgornja vprašanja? Vsaka raziskava se teh vprašanj loteva s svojimi metodološkimi pristopi. V

uvodu prispevka je dan pregled podobnosti in razlik v zasnovah omenjenih raziskav znanja, ki mu sledi razprava o njihovi pomembnosti kot podlaga pri sprejemanju odločitev na vseh ravneh izobraževalnega sistema. Tako so izsledki raziskav znanja pomembni na državni ravni, kjer se sprejemajo odločitve o razvoju izobraževalnega sistema kot celote, obenem pa so lahko pomembni tudi za samo delo v šolah in razredih. Izpostavljeni so nekateri mogoči problemi z veljavnostjo uporabe izsledkov raziskav, ki se lahko pojavijo v vseh elementih raziskave, od same zasnove prek izvedbe do izbire analiz za zbrane podatke in interpretacije rezultatov.

### Podobnosti in razlike v zasnovi raziskav

Preden lahko sestavimo lestvico držav, šol ali učencev po dosežkih, moramo pripraviti opredelitev o tem, katere dosežke merimo. Opredeliti moramo področje merjenja znanja, populacijo učencev, ki jo obravnavamo, kakšne instrumente in pristope bomo uporabili pri zbiranju podatkov o njihovem znanju in kako bomo podatke analizirali. Med raziskavami PISA, TIMSS, PIRLS in NPZ so v teh elementih razlike, ki so opisane v nadaljevanju.

Ciljna populacija v raziskavi PISA so petnajstletni udeleženci v šolskem sistemu, kar v Sloveniji večinoma predstavlja dijake prvih letnikov srednjih šol, v manjšem deležu pa tudi učenci zadnjih razredov osnovnih šol in udeleženci izobraževanja za odrasle. Ta populacija je izbrana, ker v večini držav OECD predstavlja udeležence, ki so tik pred ali tik po koncu obveznega dela izobraževanja. V raziskavi TIMSS je ciljnih populacij več. Temeljna razlika v primerjavi z raziskavo PISA je, da so v raziskavi TIMSS ciljne populacije učencev na splošno izbrane glede na razred, ki ga obiskujejo, oziroma glede na leta šolanja v kombinaciji s starostjo učencev. V večini držav je populacija nižjih razredov izbrana kot populacija učencev, ki obiskujejo četrti razred, populacija višjih razredov pa kot populacija učencev, ki obiskujejo osmi razred. Tudi v Sloveniji mednarodna opredelitev ciljnih populacij TIMSS ustreza četrtemu in osmemu razredu devetletnega programa. Posebej pa se v letu 2008 izvaja raziskava TIMSS za maturante, ki bo zajela populacijo dijakov v zadnjih letnikih gimnazij. Raziskava PIRLS ima eno ciljno populacijo, ki je vezana na leta šolanja v kombinaciji s starostjo. Mednarodno in v Sloveniji so to učenci četrtih razredov devetletnega programa. V preglednici 1 je predstavljena shema teh opredelitev ciljnih populacij.

PISA	TIMSS	PIRLS	NPZ
petnajstletniki v izobraževalnem sistemu (konec obveznega izobraževanja)	4. razred (mednarodno)	4. razred (mednarodno)	6. razred (prostovoljno)
	8. razred (mednarodno)		9. razred (obvezno)
	TIMSS za maturante: konec gimnazije		

Preglednica 1: Ciljne populacije v raziskavah znanja

Raziskave znanja so uporabne kot podlaga za odločanje o razvoju izobraževalnega sistema tako na ravni države kot na ravni posameznih šol. Podatki imajo še večji pomen, če potekajo raziskave v več zaporednih zajemih podatkov. Vse omenjene raziskave imajo večletne cikle zajemanja podatkov. Raziskava PISA poteka v triletnih, TIMSS v štiriletnih in PIRLS v petletnih ciklih. Časovni presledki med posameznimi zajemi podatkov v teh raziskavah so bili določeni glede na namene zbiranja podatkov in možnosti povezovanja z drugimi podatki. Poglavitno načelo v vseh raziskavah je, da vsakoletno zbiranje podatkov ne bi bilo smiselno, saj poteče od zbiranja podatkov do mednarodnega poročila o rezultatih analiz teh podatkov več kot leto dni. Hkrati lahko pričakujemo, da so izsledki, pridobljeni v nekem letu, posplošljivi tudi na nekaj več let. Vendar pa je pomemben interes držav, da čim pogosteje pridobivajo podatke za podlago pri oblikovanju odločitev v šolskem sistemu. V raziskavi PISA je bil cikel zajema podatkov zastavljen kot triletni, saj je to tako rekoč najmanjši razpon časa, ko je raziskave v tako velikem obsegu moč izvesti. V raziskavi TIMSS med posameznima zajemoma podatkov preteče štiri leta, ker je to čas, ko generacija iz populacije četrtilih razredov napreduje v osmi razred. S tem so omogočene ne samo primerjave v času med četrtilim razredom prvega zajema in četrtilim razredom drugega zajema podatkov, ampak tudi primerjave v isti generaciji učencev, od podatkov, ko so bili v četrtem razredu v prvem zajemu, do podatkov za isto generacijo učencev v osmem razredu v drugem zajemu podatkov. Preizkusi znanja in pristopi k analizam podatkov so zastavljeni tako, da so te primerjave na ravni celotnih generacij smiselne. V raziskavi PIRLS je cikel zajemanja podatkov nekaj daljši, saj ni zasnovana kot primerjave med generacijami. Nacionalno preverjanje znanja pa se izvaja vsako leto, saj je za razliko od mednarodnih raziskav eden od njegovih namenov ponuditi dodatno in primerljivo informacijo o znanju tudi posameznikom, ne le šolam in državam. Zato je pomembno, da ima vsak učenec ob končevanju 9. razreda možnost, da svoje znanje primerja z drugimi učenci devetega razreda v državi.

Opredelitve merjenega področja so med raziskavami delno podobne in delno različne. V raziskavi PISA so v vsakem ciklu merjena tri področja pismenosti: bralna, matematična in naravoslovna pismenost. Pismenost v raziskavi PISA ni le ozek pojem znanja branja in pisanja, ampak je zasnovan širše kot izkazovanje kompetenc, ki naj bi jih petnajstletniki potrebovali za uspešno nadaljnje izobraževanje in vključevanje v družbo na področjih branja, matematike in naravoslovja. V raziskavi TIMSS sta merjeni področji matematika in naravoslovje, bralna pismenost pa je področje merjenja znanja v raziskavi PIRLS. Nacionalno preverjanje znanja ima kot nacionalno zasnovano zbiranje podatkov o znanju učencev prednost v tem, da je lahko natančno usmerjeno v področje merjenja znanja, za katero je pomembno pridobiti informacije v državi. Tako je nacionalno preverjanje znanja opredeljeno z okvirom, ki ga določa učni načrt posameznega predmeta. Zgornje opredelitve so predstavljene v preglednici 2.

<b>PISA</b>	<b>TIMSS</b>	<b>PIRLS</b>	<b>NPZ</b>
2000: bralna pismenost* (glavno področje) matematična pismenost naravoslovna pismenost	1995: matematika naravoslovje	(1991: raziskava Reading Literacy)	2005/06
2003: bralna pismenost matematična pismenost* naravoslovna pismenost	1999: matematika naravoslovje (samo 8. razred)	2001: bralna pismenost	2006/07
2006: bralna pismenost matematična pismenost naravoslovna pismenost*	2003: matematika naravoslovje	2006: bralna pismenost	2007/08
2009: bralna pismenost* matematična pismenost naravoslovna pismenost	2007: matematika naravoslovje		itn.
itn.			

*Preglednica 2: Cikli zajemanja podatkov in merjena področja v raziskavah znanja*

Potem ko je opredelitev, kaj želimo meriti, izdelana, je treba sestaviti naloge za preizkus znanja. Kot pomemben del usklajevanja v mednarodnih raziskavah, ki pri nacionalnem preverjanju znanja ni potreben, je usklajevanje pri oblikovanju mednarodnega preizkusa znanja. Nemogoče je pričakovati, da bo nabor nalog v preizkusu ustrezal vsem sodelujočim državam. V raziskavah IEA, ki pri sestavljanju preizkusa znanja upoštevajo kurikule sodelujočih držav, je cilj, da je mednarodni preizkus znanja enako »nepravičen« do vseh držav. Za druge raziskave (npr. OECD PISA in zaradi narave področja merjenja tudi IEA PIRLS) je preizkus znanja sestavljen neodvisno od kurikulumov sodelujočih držav na podlagi mnenja strokovnjakov o tem, kaj je znanje, ki je pomembno za učence v zajeti populaciji. Poleg nalog za merjenje znanja pa za pojasnjevanje razlik v dosežkih v raziskavah znanja pogosto zbiramo tudi vrsto podatkov o spremljajočih dejavnikih, kot so na primer podatki o spolu in starosti učenca, njegova stališča do merjenega področja in okoliščine, v katerih poteka pridobivanje znanja. Tako so na primer v vseh mednarodnih raziskavah pripravljeni vprašalniki za sodelujoče učence in vprašalniki za ravnatelje sodelujočih šol. V raziskavi TIMSS je zaradi povezave s kurikularnimi vsebinami, in ker je ciljna populacija učencev določena kot oddelek učencev, smiselna tudi uporaba vprašalnikov za učitelje, ki v sodelujočem razredu poučujejo matematiko oziroma naravoslovne predmete. V raziskavi PIRLS pa je zaradi nižje starosti učencev še posebnega pomena zbiranje podatkov z vprašalniki za starše, saj le ti pomagajo usklajevati podatke, ki jih navajajo učenci. Hkrati so za učence te starosti podatki, ki jih navajajo starši, pomemben pojasnjevalni dejavnik pri ugotavljanju razlik v dosežkih med učenci.

## Kazalniki iz raziskav znanja in oblikovanje standardov znanja

Rezultat v vsaki posamezni nalogi pri preizkusu znanja nam pove svojo zgodbo o znanju učencev. Rezultati pri preizkusu znanja z več nalogami nam povedo več zgodb. Vendar pa določene skupine nalog, ki so povezane glede znanja in spretnosti, ki jih merijo, lahko povedo podobne zgodbe. Temeljni zakon povzemanja podatkov je, da združevanje poudari skupne lastnosti v nalogah in hkrati zasenči lastnosti, ki so značilne le za posamezne naloge (Mislevy 1995). Zato je ključnega pomena v raziskavah znanja, kako sestavimo preizkus in kako bomo iz odgovorov učencev izpeljevali sklepe in jih posploševali na širšo populacijo ali daljša obdobja.

Optimalne ravni povzemanja informacije iz odgovorov učencev na preizkus znanja se lahko razlikujejo za različne namene poročanja rezultatov. Vse raziskave, o katerih govorimo v tem prispevku, imajo med temeljnimi cilji poročanje povprečnega dosežka v državi. V mednarodnih raziskavah je dodatni poudarek na primerjavah povprečnih dosežkov med državami. Primerjave so pomembne tako na celotnem merjenem področju kot tudi po posameznih, posebej opredeljenih podpodročjih (npr. dosežki v geometriji kot podpodročju matematike). V večletnih ciklih zbiranja podatkov iz kazalnikov iz posameznega cikla sestavljamo kazalnike trendov razvoja dosežkov na merjenih področjih.

Vendar pa zgolj iz številčnega kazalnika o dosežkih učencev ne moremo sklepati o vsebini znanja, ki ga učenci izkazujejo. Kaj znajo učenci z višjimi dosežki, česar učenci z nižjimi dosežki ne znajo? Odgovor na to vprašanje poskušamo izluščiti s posebno analizo številčnih podatkov o dosežkih v povezavi z nalogami, ki so bile vključene v preizkus znanja. V tej analizi najprej lestvico dosežkov razdelimo na več območij, od nižjih do višjih dosežkov, ali določimo konkretne točke na lestvici, ki jih imenujemo mejniki med nižjimi, srednjimi in višjimi dosežki. Potem pogledamo, katere naloge so učenci z dosežki v spodnjih območjih reševali uspešno. S pomočjo vsebine teh nalog lahko vsebinsko opišemo znanje, ki ga učenci s temi dosežki izkazujejo. Nadalje na območjih višjih dosežkov poiščemo tiste naloge, ki so jih učenci s temi dosežki dodatno uspešno reševali, poleg nalog, ki so jih reševali uspešno učenci tako z višjimi kot z nižjimi dosežki. Iz vsebine dodatno uspešno rešenih nalog lahko izluščimo znanje, ki ga izkazujejo učenci z višjimi dosežki, in ne učenci z nižjimi dosežki. Tako postopoma dobimo vsebinske opise znanja po posameznih območjih dosežkov na lestvici, ki opredeljuje znanje, ki smo ga merili v raziskavi. Če nam to omogoča zasnova raziskave v smislu, da je v raziskavi sodelovalo zadostno število učencev in je bilo v preizkus znanja vključenih dovolj nalog, lahko lestvico dosežkov v raziskavi opišemo precej natančno. S tem na neki način oblikujemo zunanje standardne znanja, saj sicer absolutnih standardov v izobraževanju ni. Standarde znanja iz podatkov o dosežkih učencev je navadno mogoče oblikovati v mednarodnih raziskavah, kjer so preizkusi znanja zasnovani z več delovnimi zvezki, tako da s širšim naborom nalog zajamemo več področja, ki ga merimo. In še več. Zasnova mednarodnih raziskav, o katerih govorimo v tem prispevku, omogoča ohranjanje lestvice dosežkov na posameznem področju v več ciklih zbi-

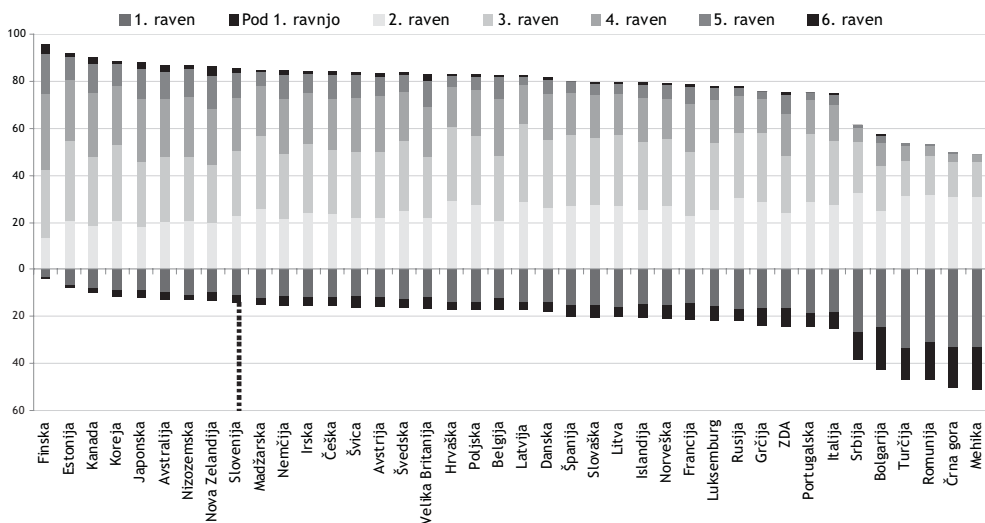
ranja podatkov in tako lahko ugotavljamo tudi trende v doseganju standardov znanja, ki se ohranjajo v cikličnih zbiranjih podatkov. S tem dobimo izhodišče za ugotavljanje, ali se znanje generacij učencev z leti in spremembe v šolskih sistemih zvečujejo, ostaja enaki ali nazadujejo. V preglednici 3 je primer vsebinskega opisa ravni znanja na lestvici naravoslovnih dosežkov v raziskavi PISA 2006.

Raven	Kompetence učencev z dosežki na posamezni ravni
6. raven spodnja meja: 707,9 točke	<p>Učenci z dosežki na 6. ravni zanesljivo prepoznavajo, razlagajo in uporabljajo naravoslovno znanje ter znanje o naravoslovnih znanostih v različnih zapletenih življenjskih situacijah. Lahko povezujejo različne vire informacij in razlag ter podatke iz teh virov uporabljajo za utemeljevanje svojih odločitev. Jasno in zanesljivo izkazujejo napredno naravoslovnoznanstveno mišljenje ter utemeljevanje in so pripravljeni svoje razumevanje uporabiti za podporo rešitvam v novih naravoslovnoznanstvenih ter tehnoloških situacijah. Učenci z dosežki na tej ravni lahko uporabljajo naravoslovno znanje in izpeljujejo utemeljitve za podporo priporočilom in odločitvam, ki so osredotočene na osebne, družbene in globalne situacije.</p> <p>To raven dosega 2,2 % slovenskih učencev in povprečno 1,3 % učencev iz držav OECD.</p>
5. raven spodnja meja: 633,3 točke	<p>Učenci z dosežki na 5. ravni lahko prepoznavajo naravoslovnoznanstvene komponente različnih zapletenih življenjskih situacij. V njih uporabljajo tako naravoslovne pojme kot znanje o naravoslovnih znanostih. Lahko primerjajo, izbirajo in vrednotijo ustrezne naravoslovnoznanstvene dokaze, podatke in preverjena dejstva za odzivanje v življenjskih situacijah. Učenci z dosežki na tej ravni uporabljajo visoko razvite sposobnosti raziskovanja, ustrezno povezujejo svoje znanje in prinašajo kritične vpoglede v te situacije. Temelječ na kritični analizi dokazov, podatkov in preverjenih dejstev lahko sestavijo razlage ter utemeljitve.</p> <p>To raven dosega 12,9 % slovenskih učencev in v povprečju 9,0 % učencev iz držav OECD.</p>
4. raven spodnja meja: 558,7 točke	<p>Učenci z dosežki na četrti ravni lahko učinkovito delujejo v situacijah in pri vprašanjih, ki lahko vključujejo pojave, pri katerih morajo sklepati o vlogi naravoslovnih znanosti in tehnologije. Lahko izbirajo in združujejo razlage iz različnih naravoslovnih ali tehnoloških disciplin ter te razlage neposredno povezujejo z življenjskimi situacijami. Učenci z dosežki na tej ravni lahko razmišljajo o svojih dejanjih in sporočajo svoje odločitve z uporabo naravoslovnega znanja, dokazov in podatkov.</p> <p>To raven dosega 35,4 % slovenskih učencev in v povprečju 29,3 % učencev iz držav OECD.</p>
3. raven spodnja meja: 484,1 točke	<p>Učenci z dosežki na tretji ravni lahko prepoznavajo jasno opisana naravoslovna vprašanja v različnih kontekstih. Lahko izbirajo dejstva in znanje za razlage pojavov ter uporabijo preproste modele ali raziskovalne pristope. Učenci z dosežki na tej ravni lahko interpretirajo naravoslovne pojme iz različnih disciplin in jih neposredno uporabljajo. Z uporabo podatkov lahko sestavijo kratke izjave in sprejmejo odločitev na podlagi naravoslovnega znanja.</p> <p>To raven dosega 63,0 % slovenskih učencev in v povprečju 56,7% učencev iz držav OECD.</p>

2. raven spodnja meja: 409,5 točke	<p>Učenci z dosežki na druge ravni imajo zadovoljivo naravoslovno znanje, da lahko ponudijo mogoče razlage v kontekstih, ki so jim blizu, ali izpeljejo sklepe iz preprostih raziskav. Lahko neposredno utemeljujejo in dobesedno interpretirajo rezultate naravoslovno-znanstvenih raziskav ali reševanja tehnoloških problemov.</p> <p>To raven dosega 86,1 % slovenskih učencev in v povprečju 80,8 % učencev iz držav OECD.</p>
1. raven spodnja meja: 334,9 točke	<p>Učenci z dosežki na prvi ravni imajo zelo omejeno naravoslovno znanje, ki ga uporabljajo le v maloštevilnih situacijah, ki so jim blizu. Lahko ponudijo znanstvene razlage, ki so očitne in ki neposredno izhajajo iz danih podatkov oziroma dokazov.</p> <p>To raven dosega 97,2 % slovenskih učencev in v povprečju 94,8 % učencev iz držav OECD.</p>

Preglednica 3: Opis šestih ravni na lestvici naravoslovnih dosežkov v raziskavi PISA 2006

Uporabnost vsebinskih opisov dosežkov na lestvicah v raziskavah znanja je tudi v primerjavah med državami. Države se med seboj primerjajo po deležih učencev, ki dosežajo posamezne ravni znanja (glej sliko 1). Uspešnejše države bodo praviloma imele večje deleže učencev z dosežki na višjih ravneh. Vendar pa z ugotavljanjem doseganja posameznih ravni znanja lahko izluščimo tudi informacije o razpršenosti dosežkov v državi. Tako na primer imata dve državi lahko podobne srednje povprečne dosežke, pri čemer v prvi državi velika večina učencev dosega srednje dosežke, v drugi državi pa je povprečje dobljeno iz združitve zelo nizkih dosežkov nekega deleža učencev z na drugi strani zelo visokimi dosežki drugega deleža učencev. Za ti dve državi je ta informacija ključna pri



Države so razvrščene po padajočih odstotkih učencev in učenk z dosežki na 2., 3., 4. in 5. ravni.  
Vir: OECD PISA 2006, Tabela 2.1b

Slika 1: Odstotki učencev po ravneh na lestvici naravoslovnih dosežkov v raziskavi PISA 2006



ugotavljanju, kateri pristopi bi bili primerni pri poskusih zviševanja ravni znanja učencev teh držav. V prvi državi bi bilo smiselno pristope oblikovati glede na glavnino, ki dosega srednje znanje, v drugi državi pa glede skupino učencev z nizkimi dosežki.

### **Raziskovanje dejavnikov, ki vplivajo na uspešnost učencev in izobraževalnega sistema v celoti**

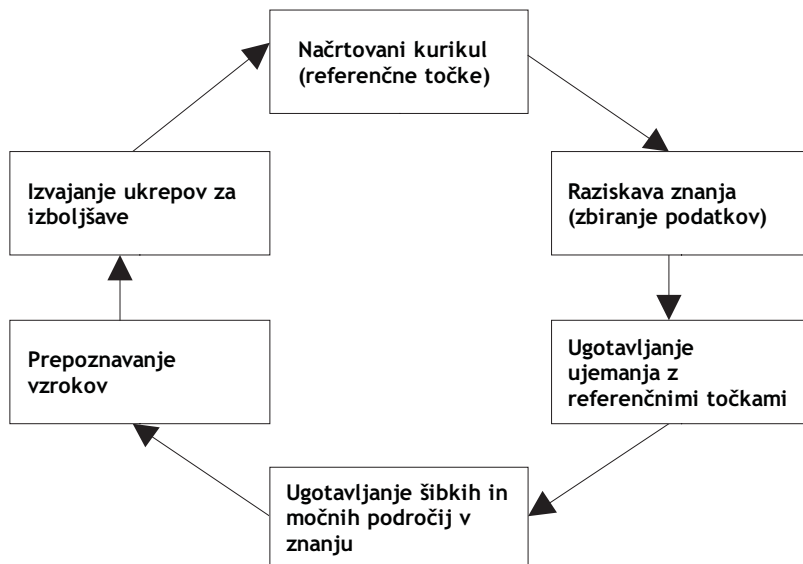
Ob prepoznavanju, da preproste primerjave dosežkov pri preizkusu znanja ne dajo veliko informacij za izboljševanje kakovosti izobraževanja, uporabniki podatkov iz raziskav znanja, tako nacionalnih kot mednarodnih, iščejo bolj praktične nasvete. Kakšna naj bo nadaljnja politika? Katere naj bodo metode in pristopi k poučevanju v šolah? Mednarodne raziskave se zdijo dobro izhodišče za iskanje odgovorov na ta vprašanja. Raziskave zbirajo podatke o spremljajočih dejavnikih, kot so šolski program, karakteristike učencev, podatke o okoliščinah doma in v šoli in drugih dejavnikih, kot so motivacija, stališča in želje za prihodnost.

Vendar pa rezultati analiz omogočajo ugotavljanje zgolj povezav med posameznimi dejavniki in uspešnostjo, ne moremo pa iz njih sklepati neposredno o vzrokih za uspešnost v šoli oziroma pri mednarodnem preizkusu znanja. Vzročne vplive bi lahko raziskovali, če bi lahko raziskave znanja izvajali kot znanstvene poskuse s slučajnostnim razvrščanjem opazovanih enot v kontrolno in eksperimentalno skupino. Tega pa pri raziskavah v izobraževanju seveda ne moremo izvesti. Vendar to ne pomeni, da je raziskovanje povezav med dosežki in spremljajočimi dejavniki povsem nesmiselno. Občasno lahko poenotimo in primerjamo skupine učencev tudi, ko niso slučajnostno razporejeni v primerjalne skupine, in uporabimo statistične pristope, kot je regresija, da izločimo »moteče« vplive. S temi pristopi seveda izločimo le del takšnih motenj. Ugotavljanje močnih povezav spremljajočih dejavnikov z dosežki nikoli povsem ne izloči možnosti, da so opazovane povezave še vedno posledica neopazovanih dejavnikov. Torej je raziskovanje povezav med spremljajočimi dejavniki in dosežki uporabno kot posredna podpora za hipoteze o vplivih na dosežke učencev. Slika 1 ponazarja krog dejavnosti, ki jih izvajamo za spremljanje in izboljševanje učinkovitosti šolskega sistema.

Kot izhodišče za spremljavo učinkovitosti izobraževalnega sistema si postavimo referenčne točke, kot so na primer standardi v učnem načrtu ali uspešnost učencev v drugih, primerljivih državah. Z raziskavo znanja zberemo podatke o znanju učencev v državi oziroma v vseh sodelujočih državah v mednarodni raziskavi. Iz ugotovljenih dosežkov ugotavljamo ujemanje rezultatov z zastavljenimi referenčnimi točkami in torej ugotavljamo, kje so močna in šibka področja v znanju učencev. Glede na ugotovitve poskušamo z dodatnimi študijami izluščiti vzroke za nastalo stanje in oblikovati ukrepe za izboljšanje. Ko so ti ukrepi izpeljani, lahko znova oblikujemo izhodišča za ugotavljanje učinkovitosti sistema in druge dejavnosti, s katerimi poskušamo to učinkovitost spremljati.

### **Veljavnost izpeljevanja in posploševanja sklepov o znanju učencev**

Za uspešno oblikovanje strategij razvoja izobraževalnega sistema je še posebno pomembna veljavnost sklepov o stanju v sistemu in potrebnih spremembah v



Slika 2: Krog dejavnosti spremljanja in izboljševanja učinkovitosti izobraževalnega sistema

njem, ki jih izpeljemo iz rezultatov analiz zbranih podatkov v raziskavah znanja. Kako natančno lahko iz skupnih dosežkov učencev pri nacionalnem ali mednarodnem preizkusu znanja sklepamo o njihovem znanju celotnega področja, o znanju celotne populacije ali celo prejšnjih in prihodnjih generacij? Veljavnosti izpeljanih sklepov o tem znanju učencev ne moremo neizpodbitno dokazati, lahko pa se s skrbno premišljeno zasnovo in natančno izvedbo raziskave izognemo mnogim nevarnostim, ki to veljavnost ogrožajo. Konceptualni okvir pri posploševanju podatkov in rezultatov, zbranih neposredno z raziskavami znanja, je oblikovan v Mislevy, Wilson, Erickan in Chudowsky (2003).

V tem okviru so v posploševanju dosežkov na konkretnih nalogah na znanje učencev predstavljene štiri stopnje: naloge, s katerimi oblikujemo situacije, v katerih naj učenec izkaže svoje znanje; vrednotenje odgovorov učencev na te naloge; statistični ali merski model, s katerim iz posameznih odgovorov oblikujemo enoštevilčni kazalnik o znanju učenca; in model, kako si predstavljamo znanje, o katerem želimo pridobiti informacije, čeprav ga ne moremo neposredno meriti. Posploševanje dosežkov pomeni, kako povežemo naloge, ob katerih naj bi učenci izkazovali svoje znanje, z modelom znanja, ki ga želimo meriti. Na veljavnost posploševanja in izpeljevanja sklepov moramo biti pozorni na vsakem koraku povezovanja med temi stopnjami. Prvi korak je povezava med nalogami in postopkom vrednotenja. Vrednotenje je postopek, ki nam omogoča, da se osredotočimo na vidike učenčevega odgovora, ki so povezani z znanjem, ki ga želimo meriti. Odgovore lahko razporedimo v urejene kategorije, ki nakazujejo naraščanje ravni znanja. Najpreprostejše vrednotenje razvrsti odgovore zgolj v dve kategoriji, pravilne in nepravilne. Lahko pa odgovore vrednotimo tudi kot delno pravilne, celo v več stopnjah. Še posebno v tem primeru je pomembno, da vsi ocenjevalci sledijo enakemu

modelu vrednotenja na osnovi primerov odgovorov, ki imajo obrazložene dodeljene ocene ali kode.

Naslednji korak je povezati tako dobljene ocene z znanjem, kot smo ga zasnovali v konceptualnem okviru raziskave. Kako, s katero vrednostjo naj te ocene povežemo z našim konceptom znanja? Povezavo oblikujemo prek merskega ali statističnega modela merjenja. Pri tem predpostavljamo, da je znanje učencev neposredno vplivalo na to, kako so učenci odgovarjali na vprašanja v preizkusu znanja, in torej na ocene, ki smo jih dobili iz vrednotenja. Z drugimi besedami, kar učenec zna in zmore, vpliva na to, koliko verjetno bo pravilno odgovoril vsako posamezno nalogo, ki je zastavljena v preizkusu znanja. Znanje v našem konceptualnem okviru se lahko torej izkazuje na eni ali več nalogah, ena naloga pa je lahko povezana tudi z več različnimi vidiki znanja. Tako na primer lahko znanje naravoslovja merimo z več različnimi nalogami, hkrati pa ena naloga lahko meri tako znanje naravoslovja kot znanje branja in razumevanja besedila.

Veljavnost posploševanja podatkov iz raziskav znanja predstavlja, koliko naloge v našem preizkusu znanja v resnici dajejo smiselne podatke o znanju učencev iz našega konceptualnega okvirja. Ali obstajajo tudi alternativne razlage za višje ali nižje dosežke na nalogah? Z dobro zasnovano preverjanja znanja prek povezav med štirimi stopnjami v posploševanju podatkov o dosežkih učencev se lahko izognemo situaciji, ko bi alternativne razlage izhajale iz nenadzorovanih dejavnikov pri preverjanju znanja.

## Sklep

Zakaj raziskave znanja? Znanje v današnji družbi je kritični element uspešnosti v globalni gospodarski konkurenci (Kellaghan in Greany 2001). Potrebe po informacijah in podatkih o kazalnikih te uspešnosti so tako vedno večje. Eden od načinov pridobivanja teh podatkov je izvajanje nacionalnih in mednarodnih primerjalnih raziskav znanja. Te raziskave dopolnjujejo podatke o ravneh znanja učencev, ki jih dajejo ocene učiteljev in rezultati državnih preizkusov znanja. V prispevku so tri mednarodne raziskave, PISA, TIMSS in PIRLS, ter nacionalno preverjanje znanja primerjani po posameznih karakteristikah zasnove in pomenu za pridobivanje podatkov o učinkovitosti slovenskega šolskega sistema. Še tako dobro izvedeno preverjanje znanja bo omejeno uporabno, če uporabniki rezultatov in interpretacij niso dobro seznanjeni s potekom preverjanja znanja, postopki in nalogami, ki so bili uporabljeni. Nepoznavanje teh elementov lahko vodi do napačnih sklepanj iz rezultatov. Podobno lahko neprimerna pričakovanja za dosežke učencev vodijo do napačnih sklepov. Standardi v mnogih državah imajo pomembno vlogo pri poučevanju in učenju in pogosto so zapisani v učnih načrtih. Če so previsoki, so lahko dosežki učencev (napačno) interpretirani kot slabi, če so prenizki, so lahko dosežki učencev (napačno) interpretirani kot dobri. Evalvacija standardov in preverjanje načinov njihove uporabe imata zato pomembno vlogo pri interpretiranju rezultatov preverjanj znanja, med njimi tudi rezultatov mednarodnih raziskav.

Čeprav so primerjave dosežkov med državami in nacionalno preverjanje znanja zaželeni, pa se izobraževalni sistemi in dosežki učencev v mnogočem razlikujejo in ni mogoče zahtevati, da bi bilo raziskovanje tako kompleksnih sistemov, kot so izobraževalni sistemi z vsemi svojimi karakteristikami in okoliščinami, preprosto. En sam kazalnik ne more opisati vseh značilnosti procesa izobraževanja in dosežkov učencev v državi. Mednarodne primerjave dosežkov na primer pripomorejo k razumevanju in določanju standardov v državi. Kazalniki raziskav znanja ponujajo sicer nepopolno, vendar relativno poceni informacijo o poteku izobraževanja in rezultatih učenja v državi. Še vedno pa so pomembni tudi drugi načini zbiranja podatkov in informacij o učinkovitosti šolskega sistema, navsezadnje neposredno vrednotenje in ocenjevanje znanja v razredih in šolah. Šele povezave med različnimi viri podatkov o znanju učencev bodo ponudile jasnejšo sliko resničnega stanja o znanju učencev.

## Literatura

- Bela knjiga o vzgoji in izobraževanju v Republiki Sloveniji (1996). Ministrstvo za šolstvo in šport. Ljubljana.
- Kellaghan, T. in Greaney, V. (2001). Using assessment to improve quality of education. Pariz: UNESCO International Institute for Educational Planning.
- Mislevy, R. J. (1995). What can we learn from international assessments? Educational evaluation and policy analysis, 17, str. 419–437.
- Mislevy, R. J., Wilson, M. R., Erickson, K. in Chudowsky, N. (2003). Psychometric principles in student assessment. V: Kellaghan, T., Stufflebeam, D. L. in Wingate, L. A. (ur.), International handbook of educational evaluation. Dordrecht: Kluwer Academic, str. 489–532.

Dr Mojca ŠTRAUS

### THE SIGNIFICANCE OF INTERNATIONAL SURVEYS ON KNOWLEDGE AND THE NATIONAL EXAMINATION OF KNOWLEDGE FOR EVALUATING PUPILS' ACHIEVEMENTS

**Abstract:** National assessment and international studies in education offer a larger spectrum of comparisons of student achievement than assessment practices at school. International studies in education emerged from the interest of many countries to obtain additional information about the educational achievement of their students and consequently about the quality of their education systems. General steps in the study are that representatives from participating countries agree on the target population of students, curriculum areas to be assessed and on the instruments to be used. The instruments are solved by the representative samples of students, usually at the same time in all participating countries. International studies in education conducted in cycles as for example PISA, TIMSS and PIRLS, are increasingly important in educational policy-making also in Slovenia. Before we can draw conclusions from the league tables of countries or from any other single number indicator one needs to be thoroughly familiar with individual procedures in the study that could impede the generalizability of the conclusion or interpretations of the results. The theme of this paper are the conceptual similarities and differences in the designs of the studies and their importance for the educational standard setting as well as eliciting the factors that are related to student achievement at school. At the same time it is extremely important to constantly examine the validity of conclusions from the results of data analyses for making sensible generalizations of the individual studies over wider populations and over longer time.

**Key words:** international studies in education, education, educational indicators, educational achievement.