Zlatan Magajna, Amalija Žakelj

Primerjalna analiza zunanjega preverjanja znanja iz matematike ob koncu devetletke v Sloveniji med obdobjema 2002–2005 in 2006–2010

Povzetek: V prispevku primerjamo dve obdobji zunanjega preverjanja znanja iz matematike ob koncu devetletne osnovne šole, in sicer nacionalne preizkuse znanja v obdobju 2002–2005 in nacionalno preverjanje znanja v obdobju 2006–2010. V prvem obdobju je izid na zunanjem preverjanju vplival na šolsko oceno učencev in na njihovo všolanje v srednjo šolo, v drugem obdobju pa so zunanja preverjanja namenjena predvsem pridobivanju dodatnih informacij o matematičnem znanju le-teh. Povprečni dosežki na zunanjih preverjanjih iz matematike so v drugem obdobju bistveno nižji kot v prvem. Analiza je pokazala, da se v obdobju 2002–2010 glede na številne obravnavane parametre struktura matematičnega znanja učencev ni spremenila in da so preizkusi v drugem obdobju zahtevnejši kot v prvem v smislu, da so v njih glede na obravnavane parametre pogosteje zastopane naloge zahtevnejših kategorij. V prispevku ugotavljamo, da nižji dosežki učencev niso posledica upada matematičnega znanja le-teh, temveč so posledica drugih vzrokov – eden od njih je večja zahtevnost preizkusov v drugem obdobju.

Ključne besede: zunanje preverjanje, matematično znanje, nacionalni preizkusi znanja, nacionalna preverjanja znanja

UDK: 37.091.279.7

Pregledni znanstveni prispevek

Dr. Zlatan Magajna, docent, Univerza v Ljubljani, Pedagoška fakulteta, Kardeljeva ploščad 16, SI-1000 Ljubljana, Slovenija; e-naslov: zlatan.magajna@pef.uni-lj.si

Dr. Amalija Žakelj, docentka, Zavod RS za šolstvo, Oddelek področnih/predmetnih skupin, Poljanska 28, SI-1000 Ljubljana, Slovenija; e-naslov: amalija.zakelj@zrss.si

SODOBNA PEDAGOGIKA 2/2011, 134–156

Uvod

Zunanje preverjanje znanja ob koncu osnovne šole v Sloveniji izvajamo že dve desetletji, pri čemer sta se poimenovanje in namen le-tega spreminjala. V šolskem letu 1990/1991 je bilo uvedeno *skupinsko preverjanje znanja*, ki je bilo namenjeno razvrščanju učencev v primeru omejitve vpisa v posamezne srednje šole.[1] Skupinsko preverjanje znanja je imelo za učence, ki so se vpisovali na šolo z omejitvijo vpisa, predvsem selekcijsko vlogo, za ostale pa je pomenilo le povratno informacijo. Z uvedbo *nacionalnih preizkusov znanja* v šolskem letu 2001/2002 se je namen zunanjega preverjanja znanja razširil – kot inštrument za ugotavljanje doseganja standardov znanja je notranjim ocenam ob zaključku osnovne šole dodal še zunanje ocene. Natančneje: zaključne ocene iz predmetov, ki so bili ocenjevani na nacionalnem preverjanju, so bile sestavljene iz internih ocen učiteljev in rezultatov, doseženih na nacionalnih preizkusih. S tem naj bi uveljavili komplementarnost zunanjega in notranjega ocenjevanja znanja, kar je pomembno z vidika njunih prednosti in slabosti. Hkrati so bili rezultati zunanjega preverjanja eden od kriterijev za vpis učencev v srednje šole z omejitvijo vpisa. Z uvedbo *nacionalnega preverjanja znanja* v šolskem letu 2005/2006 se je vloga zunanjega preverjanja znanja ponovno spremenila. Nacionalno preverjanje znanja namreč nima več selekcijske vloge (razen v izjemnih primerih, kot bo pojasnjeno v nadaljevanju), ampak sta poudarjeni njegova informativna in formativna vloga – le-ti sta bili seveda posredno prisotni tudi pri nacionalnih preizkusih znanja. Dosežki na nacionalnem preverjanju znanja niso več del šolske ocene, niti tako rekoč ne vplivajo na vpis v srednje šole z omejitvijo le-tega, temveč so dodatna informacija o znanju učencev, ki je pomembna za učence same (osnovna šola starše pisno obvesti o dosežkih učenca na nacionalnem preverjanju znanja), za učitelje in šolo, ki jo učenec obiskuje, ter šolo, na katero se vpisuje; dosežki kot celota so poleg tega pomembni tudi za državo oziroma ustanove, ki skrbijo za izobraževanje na državni ravni. Število doseženih

[1] Dosežek na skupinskem preverjanju znanja iz matematike in maternega jezika je v primeru omejitve vpisa predstavljal delni kriterij v izbirnem postopku v posamezne srednje šole. Udeležba osmošolcev na skupinskem preverjanju znanja je bila prostovoljna, a nujna za tiste osmošolce, ki so se želeli vpisati na šole z omejitvami vpisa.

točk na nacionalnem preverjanju znanja je zapisano tudi v učenčevo spričevalo. Pomemben namen nacionalnega preverjanja znanja je prav tako ugotavljanje močnih in šibkih točk izvajanja pouka matematike ter ustreznost spremljajočih dejavnikov, kot so na primer učni načrt, učbeniki, izobraževanje in usposabljanje učiteljev. Zakonske podlage za izvedbo nacionalnih preizkusov znanja so urejene z Zakonom o osnovni šoli in podrejenimi predpisi (ZOsn 1996; ZOsn 2005), podlage za nacionalna preverjanja znanja pa z dopolnitvami Zakona o osnovni šoli iz leta 2006 (ZOsn 2006). V nadaljevanju bomo tako za »nacionalne preizkuse znanja« kot tudi za »nacionalno preverjanje znanja« uporabljali kratico NPZ, sama izraza »preverjanje« in »preizkus« pa bomo uporabljali v različnih pomenskih odtenkih: izraz preizkus bo označeval sam inštrument preverjanja, torej naloge in spremljajočo dokumentacijo, izraz (zunanje) preverjanje pa bo predstavljal s šolsko zakonodajo predviden postopek ugotavljanja znanja.

Nacionalno preverjanje znanja je v mnogih šolskih sistemih – poleg notranje evalvacije in mednarodnih primerjalnih raziskav, kot sta na primer PISA ali TIMSS – eden izmed mehanizmov za ugotavljanje kakovosti izobraževanja (Markelj in Majerič 2009). Z zunanjim preverjanjem znanja lahko pridobivamo različne informacije, npr. o dosežkih posameznika in povprečnih dosežkih učencev posameznih šol. Pomen in vpliv nacionalnih preverjanj znanja je večplasten in sega vse od oblikovanja šolske politike do razvrščanja učencev ob vstopu v šole, od preverjanja uspešnosti šolskega sistema pa do izboljševanja učne prakse (Zupanc 2005). Osnovni namen nacionalnih preverjanj in njihova izvedba sta prilagojena značilnostim in potrebam posameznih šolskih sistemov, vloga zunanjih preverjanj pa praviloma ni enopomenska. Glede na osnovni namen lahko nacionalna preverjanja znanja delimo v tri skupine (Nacionalno preverjanje … 2010, str. 19–20).

V prvi skupini so nacionalna preverjanja, ki so namenjena predvsem ocenjevanju učencev. Tovrstna preverjanja poznajo npr. v Združenem kraljestvu (na Škotskem), na Nizozemskem, Malti, v Luksemburgu, na Poljskem in v Belgiji (francoska skupnost). Uspeh na tovrstnih preverjanjih vpliva na učenčev šolski uspeh in na nadaljnjo izobraževalno pot učencev (prav tam).

V drugi skupini so nacionalna preverjanja, ki so namenjena zlasti vrednotenju šolskega sistema. Nekatere države, npr. Latvija, Madžarska, Avstrija, Združeno kraljestvo (Anglija), s preverjanji ugotavljajo predvsem izobraževalno uspešnost posameznih šol. Druge države, npr. Belgija (flamska skupnost), Estonija, Irska, Španija, Francija, Litva, Finska, pa tovrstna preverjanja uporabljajo predvsem kot inštrument za spremljanje celotnega šolskega sistema (prav tam).

V tretjo skupino uvrščamo nacionalna preverjanja znanja, katerih osnovni namen je podpora učnim procesom: na podlagi rezultatov preverjanja naj bi izobraževalne ustanove ugotovile pomanjkljivosti v svojem delovanju, učne potrebe učencev in možne izboljšave. Tovrstna podpora je seveda posredno prisotna tudi pri nacionalnih preverjanjih prve in druge skupine. V tretjo skupino sodijo nacionalna preverjanja, kot jih izvajajo npr. v Bolgariji, Italiji in Avstriji (prav tam).

V prvem obdobju izvajanja NPZ-ja v Sloveniji, torej v letih 2002–2005, so bile bolj poudarjene značilnosti prve skupine nacionalnih preverjanj. »Nacionalni preizkusi znanja glede na svojo naravo in namen ne nastopajo v vlogi formativnega

preverjanja, temveč pomenijo sumativno ocenjevanje znanja.« (Razdevšek-Pučko 2002, str. 5) Vendar pa so ti preizkusi vsebovali pomembne elemente formativnega preverjanja znanja – temu so bila namenjena predvsem poskusna preverjanja, ki so se izvajala nekaj mesecev pred nacionalnimi preizkusi: »Ker poskusne preizkuse analizirajo učitelji skupaj z učenci, imajo eni in drugi možnost, da pridobijo formativno povratno informacijo.« (Prav tam, str. 6) V drugem obdobju, torej v letih 2006–2010, zasledimo posamezne elemente tako druge kot tudi tretje skupine preverjanj: poleg tega, da NPZ-ji merijo izkazano (matematično) znanje učencev, učiteljem, šolam in drugim institucijam nudijo informacijo o dosežkih njihovih učencev in raznih vidikih njihovega znanja (Cankar 2008). Zakon o osnovni šoli sicer izrecno prepoveduje uporabo rezultatov NPZ-ja za razvrščanje šol (ZOsn 2006); slednje izvedo le za porazdelitev dosežkov svojih učencev ter kako se povprečni dosežek le-teh umešča med povprečne dosežke učencev drugih šol.

S to analizo, ki je bila narejena v okviru projekta »Priprava konceptualnih osnov za dograjevanje sistema vzgoje in izobraževanja«, želimo podrobneje osvetliti razlike med zadnjima dvema obdobjema zunanjega preverjanja znanja matematike ob koncu devetletne osnovne šole, torej med nacionalnimi preizkusi znanja v obdobju 2002–2005 in nacionalnimi preverjanji znanja v obdobju 2006–2010. V nadaljevanju se bomo zato omejili zgolj na zunanja preverjanja matematike ob koncu devetletke. Tovrstna analiza je pomembna iz dveh razlogov. Prvič, ker predstavlja nujen predpogoj za primerjavo kakovosti matematičnega znanja med generacijami učencev na osnovi rezultatov zunanjih preverjanj, in drugič, ker je pomembno ugotoviti, kako se različni zasnovi in namena preverjanj odražajo na rezultatih preverjanj.

Opis problema

Čeprav se osnovni namen NPZ-ja v prvem obdobju razlikuje od namena v drugem obdobju, ne moremo govoriti o povsem različnih preverjanjih. Tako v namenu preverjanja znanja kot v vlogi samih preverjanj in njihovi izvedbi so med prvim in drugim obdobjem določene razlike in podobnosti, ki jih predstavljamo v nadaljevanju.

Osnovni namen NPZ-ja v prvem obdobju je zagotavljanje doseganja standardov znanja ob koncu devetletne osnovne šole. Iz 64. člena Zakona o osnovni šoli iz leta 1996 je razvidno, da vsi učenci devetletne osnovne šole opravljajo NPZ iz matematike[2], iz 72. in 73. člena pa je razvidno, da je pozitivna (izpitna) ocena iz tega predmeta pogoj za uspešno zaključen deveti razred osnovne šole (navedene so izjeme, kjer zadostuje 80 ali 90 % točk, potrebnih za pozitivno oceno) (Zosn 1996). Po Zakonu o osnovni šoli iz leta 2006 NPZ iz matematike prav tako opravljajo vsi učenci ob zaključku devetletne osnovne šole (64. člen), vendar so dosežki

[2] Učenci, ki so bili vključeni v poskusno uvajanje devetletne osnovne šole, so v skladu s prehodnimi določbami NPZ opravljali prostovoljno (ZOsn 1996, 108. člen). Kot je razvidno iz letnih poročil o izvedbi NPZ-ja, je v prvem obdobju izvajanja le-tega udeležba devetošolcev na preverjanju nihala med 91 in 93 % (Letna poročila … 2011).

nacionalnega preverjanja le dodatna informacija o znanju učencev in niso pogoj za uspešno dokončanje devetega razreda osnovne šole (64. in 72. člen) (ZOsn 2006). V enem in drugem obdobju pa so učenci seznanjeni s svojimi dosežki. V drugem obdobju je bil osnovni namen NPZ-ja torej pridobivanje informacij, ki bi učiteljem, šolam in državi omogočale izboljšavo pouka in posredno tudi šolskega sistema. Kot rečeno, se namena NPZ-ja v prvem in drugem obdobju ne izključujeta; tudi v prvem obdobju so se zbirale in obdelovale informacije, pomembne za izboljšavo poučevanja, in tako kot v prvem je bil v drugem obdobju dosežek posameznega učenca sporočen učencu, njegovim staršem in učitelju, posredno (iz učenčevega spričevala) pa je bil dostopen tudi šoli, na kateri je učenec nadaljeval izobraževanje.

Pomembna razlika med obdobjema se nanaša na upoštevanje rezultatov NPZ-ja v primeru omejitve opisa v posamezne srednje šole. Zakonske odredbe so v prvem obdobju določale, da sta pri vpisnih merilih v srednjo šolo rezultata NPZ-ja iz slovenščine in matematike skupaj pomenila toliko kot šolske ocene iz zadnjega triletja. V drugem obdobju je selekcijska vloga NPZ-ja pri vpisu v srednje šole zelo omejena. Objavljena merila[3] iz leta 2008 določajo, da so v primeru omejitve vpisa v posamezno srednjo šolo pri izbiri kandidatov rezultati NPZ-ja upoštevani le kot sekundarni kriterij, kadar je na kateri šoli na spodnji meji več kandidatov z enakim številom točk iz šolskih ocen in uspeha.

V prvem obdobju so se NPZ-ji izvajali le tam, kjer se je poskusno uvajala devetletna osnovna šola, v drugem obdobju pa na vseh osnovnih šolah. V prvem obdobju je število učencev, ki so se v posameznem letu udeležili NPZ-ja (ob koncu devetletne šole), naraščalo od 794 do 1.788, v drugem obdobju pa se je zaradi upadanja generacije zmanjševalo od 20.892 do 17.924. Številčnost obravnavane populacije, to je devetošolcev, ki so opravljali NPZ, je v obdobjih sicer različna, pomembno pa je, da je bila šolska pot učencev v prvem in drugem obdobju podobna: prvih pet let šolanja so obiskovali osemletno, zadnja tri leta pa devetletno osnovno šolo (izjema so le učenci, ki so se udeležili NPZ-ja v drugem obdobju in so se všolali v šole, ki so devetletko v poskusnem obdobju uvajale že od prvega razreda).

Omenimo naj še eno razliko med NPZ-ji v prvem in drugem obdobju. Zakon o osnovni šoli iz leta 1996 v svojem 64. členu določa, da se (ob koncu devetletke) »znanje učencev obvezno preverja z nacionalnimi preizkusi znanja, s katerimi se preverjajo minimalni standardi znanja učencev« (ZOsn 1996, 64. člen). Čeprav je tudi ustavno sodišče v eni od svojih odločb[4] v zvezi z NPZ-jem ob zaključku devetletke med drugim zapisalo, da je »smisel in namen tega preverjanja, da se preverjajo minimalni standardi znanja učencev«, je bila že v prvem obdobju izvajanja NPZ-ja navedena dikcija razumljena širše. Žagar je na primer zakonsko rešitev razumel takole: pri predmetih, kjer se je izvajal nivojski pouk, naj bi preizkusi vključevali »najmanj 50 % minimalnih standardov znanja, okoli 20 % temeljnih ciljev za nižjo nivojsko skupino, 20 % temeljnih ciljev za srednjo nivojsko skupino

[3] Sklep Ministrstva za šolstvo in šport RS o merilih za izbiro kandidatov v primeru omejitve vpisa v programe srednjega poklicnega izobraževanja, srednjega strokovnega in tehniškega izobraževanja ter gimnazij (Merila za izbiro kandidatov … 2010).

[4] Odločba o ugotovitvi, da četrti odstavek 40. člena in peti odstavek 64. člena Zakona o osnovni šoli nista v neskladju z ustavo (Odločba o … 2000).

in 10 % zahtevnejših ciljev za višjo nivojsko skupino« (Žagar 2002, str. 19). Zaradi normativnih nejasnosti v zvezi s 64. členom Zakona o osnovni šoli so bile komisije (vsaj komisija za matematiko) do njegovega uradnega tolmačenja kljub strokovnim priporočilom, naj se preverjajo vsi standardi, zadržane pri vključevanju neminimalnih standardov v preizkus.

V drugem obdobju pomislekov glede standardov, ki naj bi jih preverjali, ni bilo, saj Zakon o osnovni šoli iz leta 2005 v 64. členu v zvezi z NPZ-jem govori le o »preverjanju standardov znanja« (ZOsn 2005, 64. člen), Zakon o osnovni šoli iz leta 2006 pa v 64. členu določa, da se (ob koncu devetletke) »znanje učencev preverja z nacionalnim preverjanjem znanja, s katerim se preverjajo standardi znanja, določeni z učnim načrtom« (ZOsn 2006, 64. člen).

Čeprav so med NPZ-ji v prvem in drugem obdobju razlike v namenu preverjanja in uporabi rezultatov preverjanja, pa se je oblika njihovega izvajanja med obdobjema v vseh pomembnih ozirih ohranila. Preizkuse so pisali v enakem obdobju šolskega leta, nanašali so se na isti učni načrt, čas pisanja se ni spremenil, tipično število nalog in točk je ostalo enako, pa tudi v načinu izdelave in ocenjevanja preizkusov ni prišlo do bistvenih sprememb.

Prišlo pa je do večje spremembe v uspešnosti reševanja nalog na NPZ-ju. V prvem obdobju je povprečno število doseženih odstotnih točk na NPZ-ju iz matematike nihalo med 60 in 76 %, v drugem obdobju pa med 50 in 55 % (podrobnejši podatki so v nadaljevanju). Obstaja veliko dejavnikov, ki bi lahko vplivali na uspešnost učencev na NPZ-ju iz matematike, na primer motiviranost le-teh, selekcijska vloga preverjanj, obseg priprav v okviru šolskega dela, uporabljeni učbeniki in druga gradiva, izkušenost učiteljev v poučevanju po spremenjenem učnem načrtu. Mi se bomo ukvarjali le z enim dejavnikom – izpitnimi nalogami.

Namen raziskave

Zastavili smo si naslednja vprašanja:

– Kakšne so razlike v vsebinski strukturi preizkusov NPZ-ja iz matematike med obdobjema?
– Ali so morebitne razlike v vsebinski strukturi preizkusov eden od vzrokov za izkazane nižje povprečne dosežke v drugem obdobju izvajanja NPZ-ja iz matematike?
– Ali rezultati NPZ-ja iz matematike izkazujejo spremembe v strukturi matematičnega znanja devetošolcev med obdobjema?

Pri tem pod izrazom vsebinska struktura preizkusov razumemo deleže nalog, ki se nanašajo na posamezne vsebinske sklope, taksonomske ravni in druge v analizo zajete parametre.

Pod strukturo matematičnega znanja razumemo razmerje med izkazanim znanjem učencev na različnih vsebinskih področjih, taksonomskih ravneh in drugih kategorizacijah, ki bodo podrobneje opisane v razdelku o metodologiji. Do

spremembe v strukturi znanja bi na primer prišlo, če bi v prvem obdobju bolje reševali naloge iz geometrije kot pa naloge iz obdelave podatkov, v drugem pa ravno obratno, ali pa če bi na primer naloge, ki vsebinsko sodijo v 7. razred, v primerjavi z ostalimi nalogami v prvem obdobju reševali bistveno bolje kot v drugem.

Odgovori na vprašanja zahtevajo podrobno analizo nalog in same strukture preizkusov NPZ-ja od leta 2002 do 2010 ter analizo uspešnosti reševanja nalog.

Metodologija

Izhodišče obravnave so predstavljale naloge na prvih spomladanskih rokih NPZ-ja od leta 2002 do 2010 in uspešnost učencev pri njihovem reševanju. Pridobili smo tudi vso dostopno dokumentacijo o sestavi posameznih preizkusov (mrežne diagrame), ki so jih izdelale komisije za pripravo le-teh. Dokumentacijo nam je posredoval Državni izpitni center.

Kot rečeno, so posamezni preizkusi NPZ-ja iz matematike vsebovali 14–16 nalog, ki so bile skupaj ovrednotene s 53–60 točkami. Komisije so vsaki točki pripisale ključne parametre (npr. vsebinski sklop, taksonomsko raven), za vsako točko pa je bil tudi posebej izračunan indeks težavnosti (IT), torej delež učencev, ki so dosegli določeno točko. Zato smo kot *enote analize vzeli posamezne točke preizkusa* (*item*) in ne posameznih nalog (*task*). V nadaljevanju bomo izraz naloga uporabljali v obeh pomenih: bodisi kot zadana naloga na preizkusu bodisi kot ocenjevana enota, tj. enota analize.

Strukturo preizkusov NPZ-ja in uspešnost reševanja nalog smo analizirali z vidika naslednjih parametrov:

– *Razred obravnave*: glede na razred obravnave tematike smo naloge uvrstili v *6., 7., 8., ali 9. razred*. Naloge, ki so se nanašale na znanja od prvega do petega razreda, smo uvrstili v 6. razred.
– *Standard znanja:* komisije za pripravo preizkusov so posameznim enotam analize (točkam) določile, ali ugotavljajo doseganje *minimalnega*, *temeljnega* ali *zahtevnejšega* standarda (v skladu z učnim načrtom). Pri nekaterih nalogah iz prvega obdobja NPZ-ja v pridobljenih mrežnih diagramih standard ni bil zapisan (naloge so seveda preverjale učne cilje iz učnega načrta). V takem primeru smo standard, ki ga je preverjala naloga, določili avtorji prispevka.
– *Tip naloge (tj. tip zahtevanega odgovora):* v tem pogledu smo razlikovali naloge *zaprtega tipa* (npr. izbirni odgovori, povezovanje), *polodprtega tipa* (npr. vpisovanje besed ali podatka na določeno mesto) in *odprte naloge* (prost zapis rešitve ali utemeljitve).
– *Taksonomska raven:* v obeh obdobjih izvajanja NPZ-ja in v tej raziskavi je bila uporabljena taksonomija, ki izhaja iz ameriških nacionalnih preverjanj znanja NAEP (Allen idr. 1997, str. 25). Upoštevali smo štiri taksonomske ravni: *konceptualno*, *proceduralno*, *kompleksno proceduralno* in *problemsko*. *Konceptualna raven* se nanaša na zmožnost, da pri razmišljanju uporabljamo

pojme, njihove definicije, povezave med njimi in njihove različne predstavitve. *Proceduralna raven* predstavlja zmožnost izvajanja (preprostih, nesestavljenih) postopkov, poznavanje in učinkovito obvladovanje algoritmov, procedur, pravil in obrazcev. *Kompleksna proceduralna raven* se nanaša na zmožnost zaporednega povezovanja (matematičnih) postopkov, izvedbo kompleksnih postopkov ter zmožnost izbire, preverjanja in utemeljevanja pravilnosti postopkov. *Četrta, tj. problemska raven,* pa se nanaša na zmožnost povezovanja in uporabe konceptualnega in proceduralnega znanja v novih situacijah, torej na zmožnost prepoznavanja in formuliranja problemov, obravnavo zadostnosti in konsistentnosti podatkov, uporabo strategij, kritično presojo smiselnosti in ustreznosti rešitve. Posamezne ravni sta podrobno opisali M. Cotič in A. Žakelj (Cotič in Žakelj 2004).

– *Vrsta konteksta:* v tem pogledu smo upoštevali tri kategorije: *simbolno-matematično* (naloga je podana v matematičnih simbolih, npr. kot enačba ali zapisan račun), *besedilno-matematično* (besedilo v matematičnem kontekstu je treba prevesti/zapisati v matematično simboliko), *besedilno-nematematično* (besedilo iz nematematičnega konteksta je treba povezati z matematičnimi pojmi in pretvorbo zapisati v matematični simboliki).
– *Vsebinsko področje:* v skladu z učnim načrtom smo vključili naslednje vsebine: *aritmetiko in algebro*, *geometrijo in merjenje*, *podatke in verjetnost*. Pri nekoliko natančnejši obdelavi in analizi preizkusov smo vsebine delili na*: števila, merjenje in količine, ravninsk*o geometrijo, prostorsko geometrijo, odnose med količinami, algebro ter podatke in verjetnost.
– *Povezovalnost:* ocenjevano enoto smo ocenili kot *povezovalno*, če pomembno zajema dve ali več matematičnih tem.

Razlike v strukturi preizkusov med obdobjema 2002–2005 in 2006–2010 bomo predstavili glede na posamezne parametre. Odvisnost uspešnosti reševanja nalog glede na obdobje preizkusov in posamezne zgoraj navedene parametre smo ugotavljali z dvosmerno analizo variance.

Analiza

V razdelku predstavljamo dosežke učencev pri NPZ-ju iz matematike v posameznih letih, temu sledi obravnava razlik v strukturi preizkusov glede na posamezne preučevane parametre in hkratna obravnava uspešnosti reševanja nalog glede na obdobje in kategorije posameznih parametrov.

Analiza dosežkov pri NPZ-ju iz matematike po letih

Preglednica 1 nazorno kaže več razlik med preverjanji v prvem in v drugem obdobju izvajanja NPZ-ja. Razlika v številu udeležencev na preverjanju je posledica dejstva, da so v prvem obdobju NPZ izvajali le na šolah, kjer so preizkusno uvajali nove učne načrte. Le v prvem obdobju so bili dosežki posameznih učencev

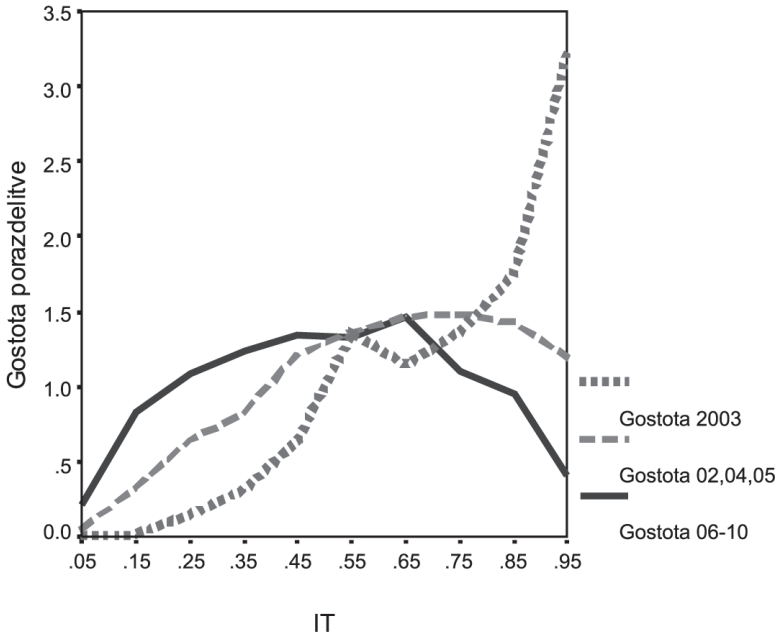
prevedeni v šolsko oceno, posebej pa je bila določena tudi meja za pozitivno oceno. V drugem obdobju izvajanja NPZ-ja dosežki niso bili vključeni v šolsko oceno, zato tovrstne meje ni.

Leto	Število udeležencev	Povprečni delež doseženih točk (%)	St. odklon povpr. deleža doseženih točk (%)	Delež učencev nad 50 % točk (%)	Meja (%) za pozitivno oceno	Delež negativno ocenjenih učencev (%)
2002	794	64,9	22,4	74,1	20,0	2,8
2003	1104	75,8	19,0	88,8	36,0	2,8
2004	1241	60,5	23,1	67,1	20,0	3,6
2005	1788	62,9	21,8	71,3	20,0	2,6
2006	20892	54,9	22,0	57,8	–	–
2007	19311	51,5	25,5	53,0	–	–
2008	18652	53,0	23,9	55,4	–	–
2009	18574	51,3	22,3	53,5	–	–
2010	17924	50,1	22,3	51,4	–	–

Preglednica 1: *Uspešnost reševanja nalog NPZ-ja iz matematike po letih*

Povprečni deleži doseženih točk so z leti padali, kar pomeni, da so učenci na preverjanju v povprečju dosegali nižje število točk. Najvišji povprečni delež, ki je leta 2003 znašal 0,76, se je znižal do 0,50 v letu 2010. Vsi doseženi povprečni deleži točk iz obdobja 2002–2005 so očitno višji od vseh povprečnih deležev točk iz obdobja 2006–2010 (preglednica 1). V obdobju 2002–2005 se le-ti gibljejo od 0,60 do 0,76, v obdobju 2006–2010 pa od 0,50 do 0,55. Meja za pozitivno oceno v obdobju 2002–2005 je znašala 20 %, višja je bila le leta 2003, ko je bil povprečni indeks težavnosti preizkusa 0,76. Poudarimo naj, da v prvem obdobju rezultati NPZ-ja pri predmetu matematika v nobenem letu niso odstopali od rezultatov pri ostalih predmetih. Pri predmetu slovenščina, ki je bil poleg matematike edini obvezen za vse učence, so bili na primer povprečni dosežki v letih 2002 do 2005 zaporedoma 66,9, 73,9, 70,6 in 69,3 %, prav tako pa je bila pri tem predmetu v prvem obdobju povprečna meja za pozitivno oceno (na desetstopenjski lestvici) 34, za najvišjo oceno pa 92 % točk (Letna poročila … 2011).

Med prvim in drugim obdobjem so izrazite razlike v porazdelitvi doseženih točk. Porazdelitve dosežkov so znotraj posameznih obdobij v posameznih letih med seboj podobne, izstopa le porazdelitev v letu 2003. Slika 1 prikazuje shematski prikaz povprečne porazdelitve v prvem in v drugem obdobju ter posebej za leto 2003. Porazdelitve v prvem obdobju odražajo poudarjanje občutljivosti preizkusa na spodnjem delu lestvice. Pri tem so se kazale težave z uskladitvijo z drugimi zahtevami, kot so preverjanje široke palete ciljev in standardov ter ocenjevanje učencev na zgornjem delu lestvice.



Slika 1: *Shematski prikaz porazdelitev odstotnih točk učencev na NPZ-jih iz matematike v letih od 2002 do 2010*

Razlike v strukturi in zahtevnosti nalog med prvim in drugim obdobjem izvajanja NPZ-ja

V obeh obdobjih so preizkusi vsebovali 14–16 nalog, ovrednotenih s 53–60 točkami. Naloge so zastavljene zelo raznoliko, čeprav lahko preverjajo zelo podobne cilje. Le naloge, ki preverjajo izračun številskih izrazov, algebrskih izrazov in reševanje enačb, so v različnih letih večkrat zastavljene na enak ali skoraj enak način – te naloge skupaj predstavljajo največ 20 % točk.

Kot smo že omenili, smo razlike v strukturi in zahtevnosti nalog med prvim in drugim obdobjem izvajanja NPZ-ja analizirali glede na naslednje parametre: razred obravnave, standard znanja, tip naloge, taksonomska raven, vrsta konteksta, vsebinsko področje, povezovalnost. Za posamezen parameter smo najprej za obe obdobji analizirali sestavo preizkusov glede nanj. Na primer: za parameter »razred obravnave« smo ugotovili, koliko nalog (točk) je v posameznem obdobju sodilo v 6., 7., 8. oz. 9. razred.

Raziskali smo tudi, kako dobro so učenci v posameznem obdobju reševali naloge glede na posamezne parametre. Za vsako obdobje smo torej za naloge iz posameznih kategorij (danega parametra) izračunali uspešnost učencev pri reševanju. Pomembnost razlik pri dobljenih rezultatih smo ugotavljali z dvosmerno analizo variance.

Parameter »razred obravnave«

V drugem obdobju izvajanja NPZ-ja je bil nekoliko večji delež nalog, ki so zajemale vsebine od 1. do 6. in 9. razreda. Dvosmerna analiza variance kaže, da obstajajo statistično pomembne razlike med težavnostjo nalog med obdobjema ($p = 0{,}000$), prav tako pa obstajajo statistično pomembne razlike med težavnostjo nalog po razredih ($p = 0{,}000$). Učenci nasploh najbolje rešujejo naloge, ki zajemajo snov 8. razreda, najslabše pa naloge, ki zajemajo snov 9. razreda. Verjetna razlaga za to je dejstvo, da snov 9. razreda zajema zahtevnejše vsebine iz algebre in prostorske geometrije, snov 8. razreda pa je pri učencih že utrjena in se tudi tesno navezuje na snov 9. razreda. Pri vsebinah 6. in 7. razreda so bili učenci uspešni tam, kjer se snov močno navezuje na vsebine poznejših let (npr. ulomki, decimalna in naravna števila), manj uspešni pa so bili pri nekaj ostalih vsebinah, kjer je prišlo do izraza pozabljanje (npr. zahtevnejša pretvarjanja merskih količin, geometrijske konstrukcije). Interaktivnega učinka med faktorjema nismo zaznali ($p = 0{,}115$).

Parameter »standard znanja«

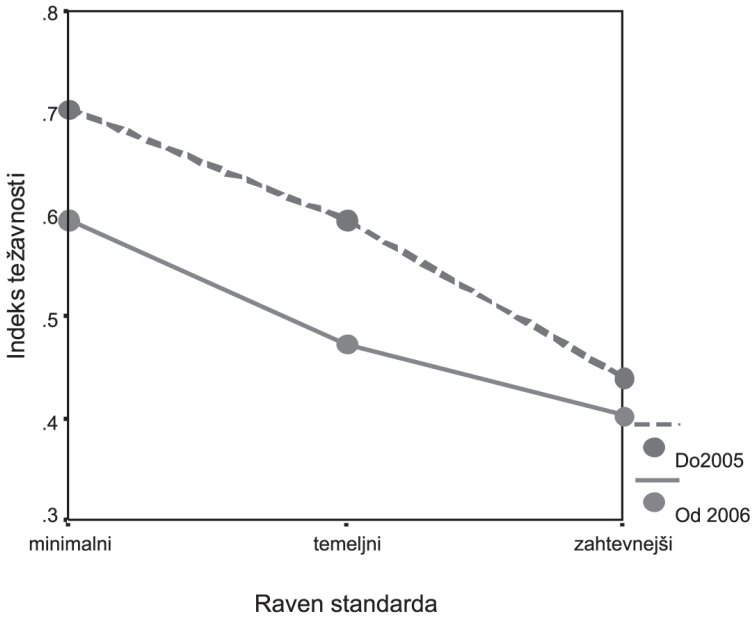
V prvem obdobju izvajanja NPZ-ja je bilo bolj poudarjeno ugotavljanje doseganja minimalnih standardov (preglednica 2).

		Raven standarda Število točk in delež v obdobju			Skupaj
		Minimalni	Temeljni	Zahtevnejši	
Obdobje	Do 2005	157 70,1 %	63 28,1 %	4 1,8 %	224 100,0 %
	Od 2006	118 42,8 %	141 51,1 %	17 6,2 %	276 100,0 %
Skupaj		275 55,0 %	204 40,8 %	21 4,2 %	500 100,0 %

Preglednica 2: Struktura preizkusov iz matematike v dveh obdobjih izvajanja NPZ-ja glede na raven standarda

Uspešnost reševanja nalog po obdobjih glede na raven standarda prikazuje slika 2, iz katere je razvidna (samoumevna) ugotovitev, da učenci najbolje rešujejo naloge, ki ugotavljajo minimalne, najslabše pa naloge, ki ugotavljajo zahtevnejše standarde znanja.

Dvosmerna analiza variance kaže, da obstajajo statistično pomembne razlike med težavnostjo nalog med obdobjema ($p = 0{,}021$), prav tako pa obstajajo statistično pomembne razlike med težavnostjo nalog glede na raven standarda ($p = 0{,}000$), pri čemer interaktivnega učinka med faktorjema ni ($p = 0{,}750$).



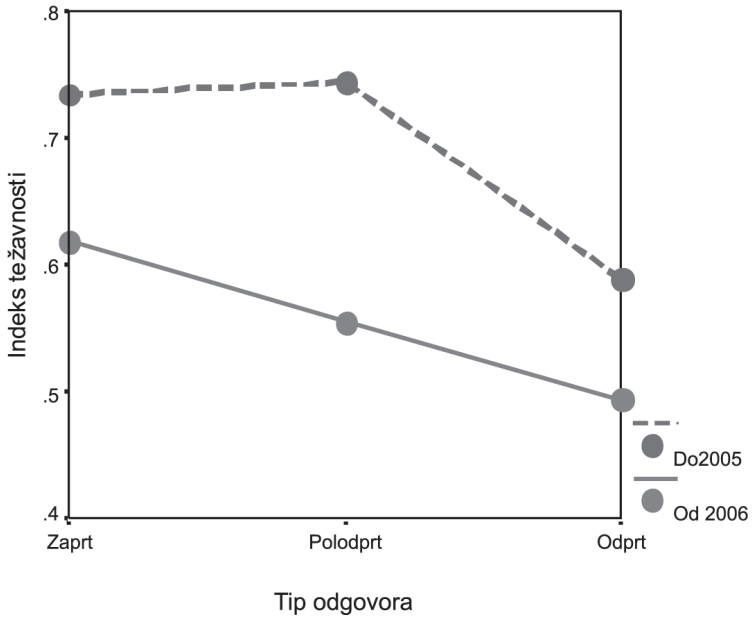
Slika 2: *Uspešnost reševanja nalog NPZ-ja v obdobjih 2002–2005 in 2006–2010 glede na standard naloge*

Parameter »tip naloge«

Preglednica 3 razkriva, da je bilo v drugem obdobju več nalog odprtega tipa. Uspešnost reševanja posameznih tipov nalog v enem in drugem obdobju prikazuje slika 3.

		Tip naloge Število in delež točk			Skupaj
		Zaprt	Polodprt	Odprt	
Obdobje	Do 2005	96 42,9 %	26 11,6 %	102 45,5 %	224 100,0 %
	Od 2006	41 14,9 %	41 14,9 %	194 70,3 %	276 100,0 %
Skupaj		137 27,4 %	67 13,4 %	296 59,2 %	500 100,0 %

Preglednica 3: Struktura preizkusov iz matematike v dveh obdobjih izvajanja NPZ-ja glede na tip naloge



Slika 3: *Uspešnost reševanja nalog NPZ-ja v obdobjih 2002–2005 in 2006–2010 glede na tip naloge*

Dvosmerna analiza variance kaže statistično pomembne razlike v težavnosti nalog glede na obdobje NPZ-ja (p = 0,000), prav tako pa so razvidne statistično pomembne razlike med težavnostjo nalog glede na tip le-te (p = 0,000). Najtežje so naloge odprtega tipa (skupni povprečni IT = 0,53), sledijo polodprte naloge (skupni povprečni IT = 0,62), najbolje pa so rešene naloge zaprtega tipa (skupni povprečni IT = 0,69). Interakcija med parametroma »obdobje« in »tip naloge« ni statistično pomembna (p = 0,229).

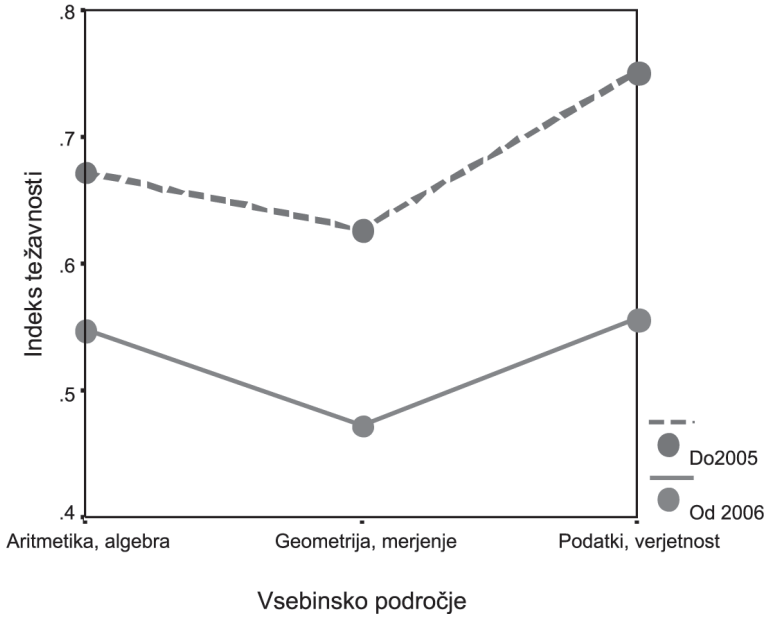
Tako v prvem kot v drugem obdobju so bile v povprečju slabše rešene naloge odprtega kot pa polodprtega in zaprtega tipa. Mnenja o tem, da je indeks težavnosti naloge, ki preverja dani cilj, odvisen od tipa naloge, so med strokovnjaki deljena (Nitko 1996, str. 141). J. Dudaite navaja več raziskav, pri čemer nekatere ugotavljajo, da so vsebinsko enakovredne naloge bolje rešene, če so zastavljene kot odprte ali polodprte, druge ugotavljajo večjo uspešnost reševanja pri zaprtih formulacijah, tretje pa niso zaznale vpliva načina formulacije na uspešnost reševanja (Dudaite 2007). Avtorica je ugotovila, da so litvanski učenci bolje reševali naloge zaprtega tipa iz raziskave TIMSS, kot če so jim bile naloge zastavljene kot (ekvivalentne) (pol)odprte. Tega, da bi odprtost nalog lahko bila dejavnik, ki vpliva na uspešnost njihovega reševanja pri NPZ-ju, ne moremo trditi. Prepričljivejša je razlaga, da v preizkusih NPZ-ja v obeh obdobjih naloge posameznih taksonomskih ravni vsebujejo različne deleže nalog odprtega tipa: pri nalogah konceptualne in proceduralne ravni je delež odprtih nalog pod, pri nalogah kompleksne proceduralne in problemske ravni pa nad povprečjem.

Izvor	SS	df	MS	F	p
Obdobje	1,422	1	1,422	36,044	0,000
Tip naloge	1,720	2	0,860	21,796	0,000
Obdobje * tip naloge	0,117	2	0,058	1,478	0,229

Preglednica 4: *Učinek tipa naloge in obdobja izvedbe preverjanja na indeks težavnosti nalog NPZ-ja*

Parameter »vsebina«

Deleži nalog s področja aritmetike in algebre, geometrije in merjenja ter podatkov in verjetnosti so v obeh obdobjih podobni, in sicer približno enaki 50, 35 in 15 %. Profil zahtevnosti nalog glede na vsebinsko področje je v obeh obdobjih podoben: najzahtevnejše so naloge s področja geometrije (skupni povprečni IT = 0,54), sledijo naloge s področja aritmetike in algebre (skupni povprečni IT = 0,60), najlažje pa so naloge s področja podatkov in verjetnosti (skupni povprečni IT = 0,65). V drugem obdobju so po posameznih področjih indeksi težavnosti sistematično nižji (p = 0,000), interakcijskih učinkov med vsebinskim področjem in obdobjem izvajanja NPZ-ja pa ni (p = 0,400). Tudi pri natančnejši razčlenitvi vsebinskih področij pridemo do podobnih rezultatov.



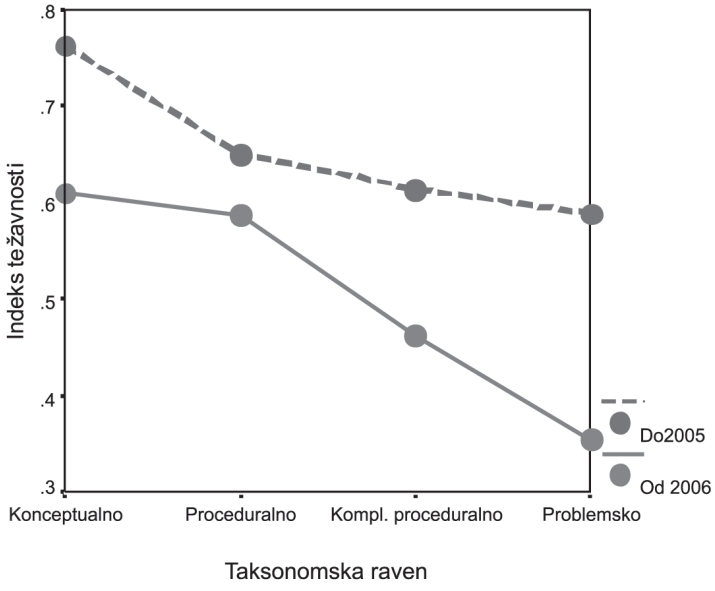
Slika 4: *Indeks težavnosti nalog v obdobjih 2002–2005 in 2006–2010 glede na vsebinsko področje*

Parameter »taksonomska raven«

Med obravnavanima obdobjema NPZ-jev so razvidne razlike v zastopanosti posameznih taksonomskih ravni v preizkusih (preglednica 5). V drugem obdobju se je na račun proceduralnih povečal delež problemskih nalog. V obeh obdobjih so najbolje rešene naloge, ki ugotavljajo konceptualno znanje (poznavanje dejstev in razumevanje pojmov, skupni povprečni IT = 0,68), sledijo proceduralne (skupni povprečni IT = 0,62), kompleksne proceduralne (skupni povprečni IT = 0,53) in problemske naloge (skupni povprečni IT = 0,42). Dvosmerna analiza variance pokaže, da so v drugem obdobju indeksi težavnosti sistematično nižji, pri čemer je bil zaznan interakcijski vpliv med obdobjem izvajanja in taksonomsko ravnijo (p = 0,019). V drugem obdobju je bil torej indeks težavnosti problemskih nalog v primerjavi z ostalimi nalogami nekoliko nižji kot v prvem obdobju.

		Taksonomska raven Število in delež točk				Skupaj
		Konceptualno	Proceduralno	Kompleksno proceduralno	Problemsko	
Obdobje	Do 2005	63 28,1 %	92 41,1 %	51 22,8 %	18 8,0 %	224 100,0 %
	Od 2006	74 26,8 %	85 30,8 %	68 24,6 %	49 17,8 %	276 100,0 %
Skupaj		137 27,4 %	177 35,4 %	119 23,8 %	67 13,4 %	500 100,0 %

Preglednica 5: *Struktura preizkusov iz matematike v dveh obdobjih izvajanja NPZ-ja glede na taksonomsko raven*



Slika 5: *Indeks težavnosti nalog v obdobjih 2002–2005 in 2006–2010 glede na taksonomsko raven*

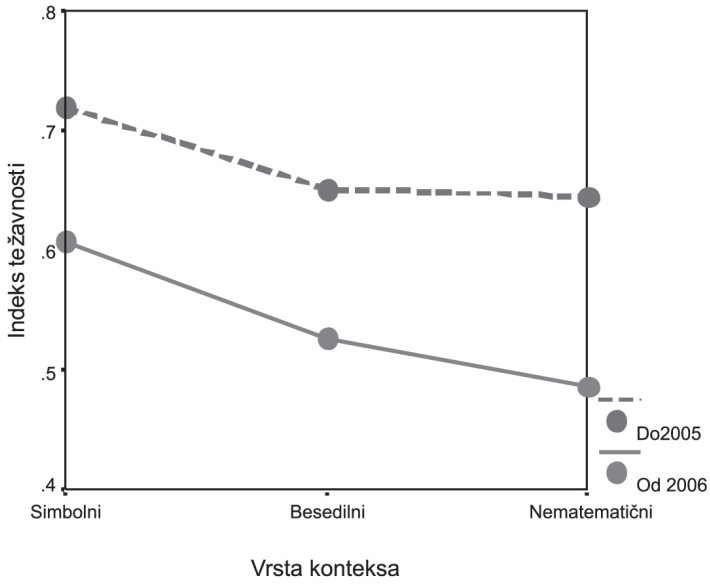
Parameter »vrsta konteksta«

Med obravnavanima obdobjema izvajanja NPZ-ja so razlike tudi v pogledu konteksta v nalogah (preglednica 6). V drugem obdobju je v primerjavi s prvim izrazito večji delež nalog podan v besedilno-nematematičnem kontekstu in izrazito manjši delež v simbolno-matematični obliki.

		Vrsta konteksta (število točk)			Skupaj
		Simbolni	Besedilno-matematični	Besedilno-nematematični	
Obdobje	Do 2005	65 29,0 %	87 38,8 %	72 32,1 %	224 100,0 %
	Od 2006	42 15,2 %	113 40,9 %	121 43,8 %	276 100,0 %
Skupaj		107 21,4 %	200 40,0 %	193 38,6 %	500 100,0 %

Preglednica 6: *Struktura preizkusov iz matematike v dveh obdobjih izvajanja NPZ-ja glede na vrsto konteksta*

V preizkusih NPZ-ja so bile naloge, ki so podane v simbolnem matematičnem zapisu (v nalogi je na primer treba zapisan računski ali algebrski izraz izračunati, poenostaviti, rešiti in grafično predstaviti), pomembno bolje rešene kot besedilne naloge matematične vsebine in bolje kot besedilne naloge, ki so zadane v nematematičnem kontekstu in zahtevajo pretvorbo v matematični jezik. Skupni povprečni indeksi zahtevnosti nalog navedenih kategorij so bili 0,68 za naloge, podane v matematičnem simbolnem jeziku, 0,58 za besedilne naloge matematične



Slika 6: *Indeks težavnosti nalog v obdobjih 2002–2005 in 2006–2010 glede na vrsto konteksta*

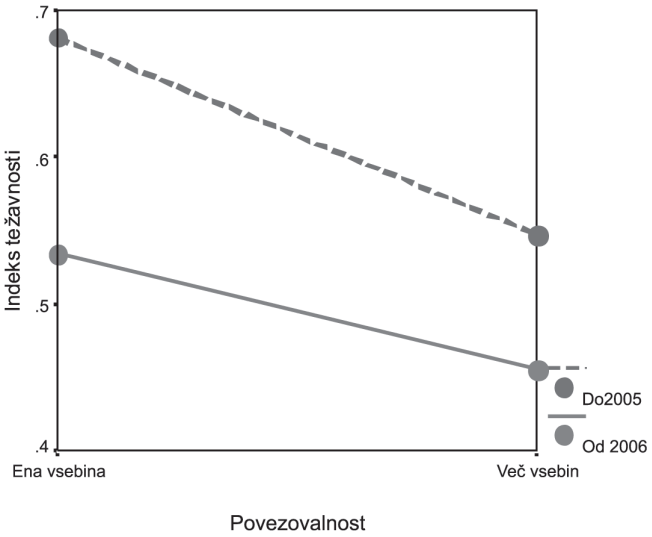
vsebine in 0,54 za besedilne naloge v nematematičnem kontekstu. Tako kot pri večini ostalih parametrov so v drugem obdobju pri vseh kategorijah konteksta indeksi težavnosti sistematično nižji (slika 6), vendar ni zaznati interakcijskega vpliva obdobja izvajanja na vrsto konteksta (p = 0,580).

Parameter »povezovalnost«

Povezovalnost se na preizkusih NPZ-ja »obnaša« podobno kot mnogi drugi parametri. V drugem obdobju izvajanja NPZ-ja je bil delež povezovalnih točk pomembno višji kot v prvem obdobju. Pri tem naj poudarimo, da se povezovalnost nanaša na enote obravnave, tj. posamezne ocenjevalne točke, in ne naloge iz preizkusa. Posamezne naloge so tipično zajemale 3–6 točk; nedvomno bi bil torej delež povezovalnih nalog večji. Glede zahtevnosti so se povezovalne naloge v obeh obdobjih izkazale kot zahtevnejše glede na nepovezovalne (p = 0,000), interakcijskega vpliva med povezovalnostjo in obdobjem izvajanja na indeks težavnosti pa nismo zaznali (p = 0,337).

		Povezovalnost (število točk)		Skupaj
		Ena vsebina	Več vsebin	
Obdobje	Do 2005	203 90,6 %	21 9,4 %	224 100,0 %
	Od 2006	230 83,3 %	46 16,7 %	276 100,0 %
Skupaj		433 86,6 %	67 13,4 %	500 100,0 %

Preglednica 7: *Struktura preizkusov iz matematike v dveh obdobjih izvajanja NPZ-ja glede na povezovalnost*



Slika 7: *Indeks težavnosti nalog v obdobjih 2002–2005 in 2006–2010 glede na povezovalnost*

Izvor	SS	df	MS	F	p
Obdobje	0,726	1	0,726	17,395	0,000
Povezovalnost	0,578	1	0,578	13,846	0,000
Obdobje * povezovalnost	0,04	1	0,04	0,925	0,337

Preglednica 8: *Učinek povezovalnosti naloge in obdobja izvedbe preverjanja na indeks težavnosti nalog NPZ-ja*

Sklepne ugotovitve

Doseženi povprečni deleži točk na NPZ-ju iz matematike so se gledano za celotno obdobje od 2002 do 2010 spreminjali – najvišji je znašal 0,76 (leta 2003), najnižji pa 0,50 (leta 2010). Povsem očitna je razlika v doseženih deležih točk med obdobjema 2002–2005 in 2006–2010. V prvem obdobju so povprečni deleži točk očitno višji kot v drugem (preglednica 1); v prvem se tako gibljejo od 0,62 do 0,75, v drugem pa od 0,50 do 0,55.

Preizkusi se glede na veliko večino obravnavanih parametrov med obdobjema statistično pomembno strukturno razlikujejo. V drugem obdobju so glede na preverjan standard znanja, tip odgovora, taksonomsko raven, vrsto konteksta in povezovalnost v večjem deležu zastopane kategorije, ki so v celotnem obravnavanem obdobju povezane z večjo zahtevnostjo (nižjim IT-jem) nalog (preglednica 9). Glede na preostala obravnavana parametra, razred obravnave in vsebinsko področje, je struktura zastopanosti nalog v prvem in drugem obdobju enaka.

Ugotovili smo, da so bili preizkusi v obdobju 2006–2010 zahtevnejši kot v obdobju 2002–2005, saj so bile v drugem obdobju v večji meri zastopane naloge, ki so se nanašale na kategorije, ki so jih učenci v celotnem obdobju izvajanja NPZ-ja slabše reševali. Manjše povprečno število doseženih točk na preverjanjih v drugem obdobju zato samo po sebi ne pomeni, da je bilo matematično znanje učencev ob času izvajanja NPZ-ja v drugem obdobju slabše kot v prvem.

Nadalje pri skoraj vseh obravnavanih parametrih opazimo enak vzorec: uspešnost reševanja nalog je v drugem obdobju statistično pomembno in sistematično nižje. Profil zahtevnosti posameznih kategorij je zelo podoben, to pomeni, da interakcijskega vpliva med obdobjem izvajanja NPZ-ja in posameznimi parametri ni (izjema je parameter »taksonomska raven«, kjer je povprečen indeks težavnosti v drugem obdobju glede na ostale naloge nižji kot v prvem obdobju). Analiza preizkusov in dosežkov udeležencev je torej pokazala, da se v obravnavanem obdobju struktura matematičnega znanja ni spremenila – v smislu, da se glede na vse obravnavane parametre (z izjemo taksonomske ravni) relativni povprečni dosežki po kategorijah med enim in drugim obdobjem niso pomembno razlikovali.

Preprosto bi lahko rekli, da so rezultati NPZ-jev v drugem obdobju posledica dveh neodvisnih učinkov: a) enake strukture znanja kot v prvem obdobju izvajanja NPZ-ja in b) učinkov, ki so povzročili sistematično znižanje dosežkov na preverjanjih. Med razlogi za sistematično znižanje dosežkov na NPZ-jih v drugem obdobju glede na prvo je gotovo pomembno dejstvo, da so preizkusi v drugem obdobju zahtevnejši (v zgoraj navedenem smislu). O drugih možnih razlogih za znižanje

dosežkov, predvsem o morebitnem znižanju ravni matematičnega znanja učencev in morebitni manjši motiviranosti učencev pri pripravah na NPZ, iz razpoložljivih podatkov ne moremo sklepati. O teh dveh možnih razlogih za znižanje dosežkov na NPZ-ju bomo razpravljali v interpretaciji rezultatov.

Parameter	Statistična pomembnost razlik v uspešnosti reševanja		Statistična pomembnost interakcijskega vpliva med parametrom in obdobjem reševanja na IT nalog	Opombe
	Glede na parameter	Glede na obdobja		
Razred obravnave	0,000	0,000	0,115	
Standardi	0,000	0,021	0,750	Zahtevnejši standardi pomenijo nižji IT. V prvem obdobju so prevladovali minimalni standardi, zahtevnejših pa skorajda ni bilo.
Tip odgovora	0,000	0,000	0,229	Odprti odgovori imajo nižji IT. V prvem obdobju je bil manjši delež nalog odprtega tipa.
Vsebinsko področje	0,000	0,000	0,400	Glede vsebinskih področij se struktura preizkusov ni spreminjala.
Taksonomska raven	0,000	0,000	0,019	Problemske naloge so zahtevnejše. V drugem obdobju je bilo tovrstnih nalog več, poleg tega pa so bile tudi relativno zahtevnejše.
Vrsta konteksta	0,001	0,000	0,580	Naloge v simbolno-matematičnem kontekstu imajo višji IT. V prvem obdobju je bil delež tovrstnih nalog večji.
Povezovalnost	0,000	0,000	0,337	Povezovalne naloge imajo nižji IT. V prvem obdobju je bil delež tovrstnih nalog manjši.

Preglednica 9: *Pregled razlik v strukturi nalog in dosežkih med obdobjema 2002–2005 in 2006–2010*

Interpretacija

Z analizo nalog na NPZ-ju smo skušali odkriti možne razloge za nižje dosežke v drugem obdobju izvajanja NPZ-ja. Eden od razlogov so ugotovljene razlike v vsebinski strukturi izpitov med prvim in drugim obdobjem: preizkusi v drugem obdobju so bili zahtevnejši kot v prvem, saj so bile v njih v večji meri zastopane naloge, ki so se nanašale na tiste standarde znanja, tipe nalog, taksonomske ravni, vrste konteksta in način povezovalnosti, ki so jih učenci v celotnem obdobju izvajanja NPZ-ja slabše reševali. Vzroki nižjih dosežkov pa so lahko tudi drugje.

Obravnavali bomo dva (najočitnejša) možna vzroka: 1. dejanski upad matematičnega znanja učencev ob koncu osnovne šole v drugem obdobju in 2. manjša neposredna pripravljenost učencev na NPZ-ja v drugem obdobju. Pokazali bomo, da v drugem obdobju ne moremo govoriti o upadu matematičnega znanja, manjša neposredna pripravljenost učencev (ki je posledica manjše motivacije zaradi spremenjene vloge NPZ-ja) pa ima omejen vpliv.

Da matematično znanje devetošolcev v obdobju med prvim in drugim obdobjem izvajanja NPZ-ja ni upadlo, najbolj prepričljivo kaže raziskava TIMSS, ki med drugim posebej preučuje trende spreminjanja matematičnega znanja v posameznih državah. Za Slovenijo je izkazan (statistično pomemben) trend rasti matematičnega znanja v letu 2007, in sicer tako glede na leto 2003 kot tudi glede na leto 1995 (Japelj Pavešić idr. 2008, str. 28). Slovenski osmošolci so v raziskavi TIMSS 2007 v povprečju dosegli 501 točko, v letu 2003 pa so testirani učenci višjih razredov v povprečju dosegli 493 točk. Če upoštevamo, da so bili v raziskavo TIMSS 2003 zajeti tako učenci sedmih razredov osemletke kot tudi učenci osmih razredov devetletke in da so slednji v povprečju dosegli 481 točk (prav tam, str. 27), nas to napeljuje k sklepu, da matematično znanje devetošolcev v drugem obdobju ni slabše kot v prvem.

Glede na to, da nacionalno preverjanje znanja od šolskega leta 2005/2006 neposredno ne vpliva na šolsko oceno iz predmeta matematika in nima več pomembne selekcijske vloge za učence, ki se vpisujejo na šolo z omejitvijo vpisa, se je v drugem obdobju verjetno spremenil odnos udeležencev do NPZ-ja: to pomeni manj priprav nanj v okviru šolskega in domačega dela učencev, morda manjšo motiviranost le-teh za čim boljše dosežke na njem in posledično tudi manjšo pripravljenost ob preverjanju. Medtem ko so na primer v letnem poročilu NPZ-ja za leto 2002/2003 opozarjali na pojav pretiranega pripravljanja učencev na NPZ (Letna poročila … 2011), pa v drugem obdobju takih opozoril ni. Raziskava o NPZ-ju v drugem obdobju (Žakelj in Grmek Ivanuš 2010, str. 83) med drugim ugotavlja obseg priprav učencev na NPZ v šolskem letu 2007/2008 v okviru dela v šoli in samostojnega dela doma. Glede obsega *dodatnih* priprav v šoli na NPZ iz matematike (poleg rednega pouka) so učenci navedli, da so le-te potekale samo zadnji teden (34 %), samo zadnji mesec (40 %), vse leto (15 %) oziroma da priprav ni bilo (11 %). Samostojno se učenci po svojih izjavah na NPZ pripravljajo še manj: 43 % se jih sploh ne pripravlja, 35 % se jih pripravlja le zadnji teden pred izpitom, ostali pa večinoma en mesec prej. Korelacija med dodatnim pripravami v šoli in dosežki na NPZ-ju je neznatna (predvidevamo, da je tako tudi zato, ker učitelji priprave na bolj ali manj opazen način vgrajujejo tudi v redni pouk). Po drugi strani je po našem mnenju vpliv obsega priprav na izid preverjanja omejen predvsem zato, ker preizkusi merijo znanje, ki ga učenci pridobivajo v daljšem obdobju, torej znanje, ki se ga ni mogoče naučiti v kratkem času, ampak ga je treba pri pouku postopoma razvijati in nadgrajevati. Preizkusi namreč preverjajo vse taksonomske ravni znanja, tudi višje, prav tako pa preverjajo zmožnost reševanja kompleksnejših problemov in problemov, ki povezujejo več matematičnih področij, vsega tega pa se ni moč dobro naučiti zgolj s posebnimi pripravami, ki potekajo zadnji mesec. V zvezi s tem naj omenimo raziskavo o primerjavi dosežkov učencev v raziskavi TIMSS in na NPZ-ju

(Mastnak 2010). V tej raziskavi je izvedena primerjava med nalogami NPZ-ja v obdobju 2006–2009 in nalogami raziskave TIMSS 2007 za osmošolce. Avtorica ugotavlja, da na posameznih vsebinskih področjih med indeksi težavnosti nalog v eni in drugi raziskavi ni pomembnih razlik – izjema je področje algebre, kar je vsebina 9. razreda. Naloge s področja algebre so na raziskavi TIMSS, ki se izvaja v 8. razredu, slabše rešene kot na NPZ-ju, ki se izvaja v 9. razredu. Med nalogami je bilo iz ene in druge raziskave identificiranih 59 parov nalog, ki preverjajo isti vsebinski cilj. V primerih, ko sta nalogi z NPZ-ja in TIMSS-a ugotavljali isti cilj in sta se ujemali v ključnih parametrih (npr. taksonomska raven, kompleksnost naloge), so bili indeksi težavnosti nalog podobni. V nasprotnem primeru je bilo v skoraj vseh primerih razliko v dosežkih moč razložiti z različno taksonomsko ravnijo nalog, neobravnavanostjo algebrske vsebine ali drugimi parametri. Tudi to je neke vrste potrditev veljavnosti preizkusov NPZ-ja kot instrumenta merjenja matematičnih znanj, pridobljenih v daljšem obdobju šolanja.

Šolska politika je tista, ki določa, na kakšen način naj bodo uporabljeni rezultati NPZ-ja na ravni posameznih udeležencev in širše, ter posledično določa tudi, kako naj bodo preverjanja zasnovana. V tem pogledu obstajajo precejšnje razlike med prvim in drugim obdobjem izvajanja NPZ-ja. V prvem obdobju so NPZ-ji bolj vplivali na šolsko oceno, poleg tega pa je bila bolj poudarjena njihova selekcijska vloga za učence, ki so se vpisovali na šole z omejitvijo vpisa. V drugem obdobju je bolj poudarjen formativni namen preverjanja: bolj kot dodatna informacija o znanju posameznih učencev je po našem mnenju pomembna informacija, ki jo NPZ daje vsem udeležencem (učencem, staršem, učiteljem, šolam in drugim ustanovam, ki skrbijo za izobraževanje) o pričakovanem znanju in doseganju ciljev. Iz rezultatov NPZ-ja lahko izvemo, kaj so močne in šibke točke matematičnega znanja na ravni razreda, šole in širše. Ob upoštevanju drugih dejavnikov (npr. socialne strukture populacije, deleža učencev, ki jim slovenščina ni materni jezik, deleža učencev s posebnimi potrebami) so dosežki na NPZ-ju lahko eden od kazalnikov (ne)ustreznosti pogojev za pouk (učni načrti, učbeniki, usposobljenost učiteljev idr.), kar je prvi korak na poti izboljšanja učenja in poučevanja. V prvem obdobju so bili NPZ-ji bistveno občutljivejši na spodnjem delu ocenjevalne lestvice, v drugem obdobju pa je občutljivost enakomernejša, vendar v obeh obdobjih izkazujejo enako strukturo matematičnega znanja (v smislu razmerja med izkazanim znanjem učencev na različnih vsebinskih področjih, taksonomskih ravneh in drugih obravnavanih kategorizacijah). Med razlogi za nižje povprečne dosežke v drugem obdobju je zagotovo povečana zahtevnost preizkusov, v določeni meri pa tudi manjši obseg priprav na preverjanje (manjša motivacija učencev za dodatno pripravljanje na preverjanje). NPZ-ji iz matematike so v obeh obdobjih izvajanja tudi pomembno vplivali na uvajanje novosti v pouk matematike (npr. delo s podatki, poudarjanje razumevanja in utemeljevanja). S kurikularnega vidika so NPZ-ji pomembni za postavljanje standardov znanja, saj nudijo informacije o tem, kakšna znanja na posameznih vsebinskih področjih ob koncu devetletke tipično izkazujejo učenci s skromnimi, povprečnimi, dobrimi in zelo dobrimi dosežki na NPZ-ju.

Literatura in viri

Allen, N. L., Jenkins, F., Kulick, E. in Zelenak, C. A. (1997). *Technical report of the NAEP 1996 state assessment program in mathematics*. Washington: National Center for Education Statistics.

Cankar, G. (2008). Nacionalno preverjanje znanja kot pripomoček za izboljšanje poučevanja. *Sodobna pedagogika*, 59, posebna izdaja, str. 130–141.

Cotič, M. in Žakelj, A. (2004). Gagnejeva taksonomija pri preverjanju in ocenjevanju matematičnega znanja. *Sodobna pedagogika*, 55, št. 1, str. 182–192.

Letna poročila o izvedbi NPZ (za posamezna leta 2001–2010). (2011). Ljubljana: Državni izpitni center. Dostopno na: http://www.ric.si/preverjanje_znanja/statisticni_podatki/ (pridobljeno 11. 1.2011).

Dudaite, J. (2007). The influence of item-stem format and item-answer format on the results for selected items in TIMSS 2003. V: P. Wagemaker (ur.). *Proceedings of the IEA IRC-2006. Vol 1.* Amsterdam: International Association for the Evaluation of Educational Achievements, str. 83–88.

Interno gradivo Državnega izpitnega centra o nacionalnih preizkusih znanja in o nacionalnih preverjanjih znanja. (2010). Ljubljana: RIC.

Japelj Pavešić, B., Brečko, B., Bezgovšek, H., Čuček, M., Krevh, A., Lipovec, A., Magajna, Z., Perat, Z. in Vidmar, M. (2005). *Slovenija v raziskavi TIMSS 2003*. Ljubljana: Pedagoški inštitut.

Japelj Pavešić, B., Svetlik, K., Rožman, M. in Kozina, A. (2008). *Matematični dosežki Slovenije v raziskavi TIMSS 2007*. Ljubljana: Pedagoški inštitut.

Markelj, N. in Majerič, M. (2009). Kaj nam povedo mednarodne raziskave in nacionalno preverjanje znanja? *Šolsko polje*, 20, št. 1–2, str. 55–81.

Mastnak, A. (2010). TIMSS in nacionalno preverjanje znanja: primerjava pisnih preizkusov iz matematike. *Preverjanje in ocenjevanje*, 7, št. 3/4, str. 29–40.

Merila za izbiro kandidatov v primeru omejitve vpisa v programe srednjega poklicnega izobraževanja, srednjega strokovnega in tehniškega izobraževanja ter gimnazij. (2010). Sklep Ministrstva za šolstvo in šport št. 6035-118/2008 z dne 1. julija 2008. Dostopno na: http://www.mss.gov.si/si/solstvo/srednjesolsko_izobrazevanje/vpis_v_srednje_sole/vpis_201011 (pridobljeno 7. 2. 2011).

Nacionalno preverjanje znanja učencev v Evropi: namen, organiziranje in uporaba rezultatov. (2010). Ljubljana: Ministrstvo za šolstvo in šport.

Nitko, A. J. (1996). *Educational Assessment of Students*. Englewood Cliffs: Prentice-Hall.

Odločba o ugotovitvi, da četrti odstavek 40. člena in peti odstavek 64. člena zakona o osnovni šoli nista v neskladju z ustavo. (2000). Uradni list RS, št. 54/2000, str. 7158–7160.

Razdevšek-Pučko, C. (2002). Nacionalni preizkusi znanja in sodobna pojmovanja preverjanja in ocenjevanja znanja. *Vzgoja in izobraževanje*, 33, št. 2, str. 2–10.

Zupanc, D. (2005). Ravni znanja pri (inter)nacionalnih preverjanjih. *Šolsko polje*, 16, št. 5–6, str. 161–195.

Zakon o osnovni šoli (ZOsn) (1996). Uradni list RS, št. 12/1996, z dne 29. 2. 1996.

Zakon o osnovni šoli (ZOsn) (2005). Uradni list RS, št. 23/2005, z dne 10. 3. 2005.

Zakon o osnovni šoli (ZOsn) (2006). Uradni list RS, št. 81/2006, z dne 31. 7. 2006.

Žagar, D. (2002). Napotki za pripravo pisnih preizkusov znanja v devetletni osnovni šoli. *Vzgoja in izobraževanje*, 33, št. 2, str. 18–21.

Žakelj, A. in Grmek Ivanuš, M. (2010). *Povezanost rezultatov pri nacionalnem preverjanju s socialno-kulturnim okoljem učencev, poukom in domačimi nalogami*. Ljubljana: Zavod RS za šolstvo.

Zlatan MAGAJNA, Ph.D. (University of Ljubljana, Slovenia)
Amalija ŽAKELJ, Ph.D. (The National Education Institute, Slovenia)

A COMPARATIVE ANALYSIS OF SLOVENIAN NATIONAL MATHEMATIC ASSESSMENTS (YEAR 9) FROM THE PERIODS 2002–2005 AND 2006–2010

Abstract: In the comparative analysis, we consider the Slovenian national assessments of mathematical knowledge for pupils at the end of compulsory education (Year 9) from the periods of 2002-2005 and 2006-2010. In the first period, the pupils' assessment results were included in their final school grades and also served as criteria for enrolment into secondary schools, while in the second period the results had, above all, an informative function for both pupils and stakeholders. The average assessment results in the second period were significantly lower than those in the first period. In the analysis, we demonstrate that the assessment results for the whole period the structure of the pupils' mathematical knowledge of pupils had not changed. Furthermore, we argue that the reason for the lower achievement during the second period was not a decrease in pupils' level of mathematical knowledge but the increased difficulty of the assessment tasks given in the second period.

Keywords: external assessment, mathematical knowledge, national assessment