
Interno ocenjevanje pri slovenski maturi: velike razlike med šolami

Darko Zupanc, Gašper Cankar in Matevž Bren

V večini evropskih držav se spričevalo ob koncu srednje šole pridobi z neke vrste zaključnim izpitom. Gre za kombinacijo notranjega preverjanja znanja in zunanjega izpita. V večini držav se pri zaključnih izpitih znanje preverja tako pisno kot ustno (Eurydice, 2012: 166–168). V državah z eksternim preverjanjem znanja je način preverjanja v veliki meri izveden s pomočjo pisnih izpitov. V posameznih državah pa so pri nekaterih predmetih praktični deli izpitov (»coursework«), ki jih kandidati opravljajo na šoli v času priprave na izpit. Ta je lahko usten, lahko je laboratorijsko ali projektno delo ali seminarska, projektna naloga idr. Interni deli izpita so lahko moderirani (Eurydice, 2012: 168), tako da učitelji ocenjujejo znanje in dosežke dijakov na osnovi vnaprej postavljenih in usklajenih kriterijev. V praksi pa to vedno ni tako, v Sloveniji interno ocenjevanje pri maturi ni eksterno moderirano (Brejc, Sardoč in Zupanc, 2011). Z internim ocenjevanjem praktičnih nalog se lahko preverja doseganje tistih ciljev in standardov v učnih načrtih in katalogih znanj, ki jih z internim ocenjevanjem lažje preverjamo kot s pisnimi preizkusi oz. jih je pisno praktično nemogoče preverjati. Vseh taksonomskih stopenj znanja in vsebin namreč ni možno enako kakovostno preveriti le z eno vrsto izpita (Državna komisija za splošno maturo, 2010)

V Sloveniji so pri ocenjevanju učencev in dijakov velika odstopanja v porazdelitvah izmerjenih dosežkov. Porazdelitve dosežkov prikazujejo višje doseženo znanje, če v šolah ocenjujejo učitelji, kot če je znanje izmerjeno eksterno (Pučnik Ozimič, 2010; Starc, 2011; Zupanc, 2010a; Zupanc in Bren, 2010a). O razlikah pri izmerjenem znanju poročajo tudi študije na podlagi mednarodnih meritev znanja, npr. TIMSS (Japelj Pavešič idr., 2004; Japelj Pavešič idr., 2009; Mullis idr., 2004; Zupanc, 2005). Te razlike v dosežkih, tako v večletnem ča-

sovnem obdobju (Zupanc, Urank in Bren, 2009), kakor tudi v primerjavi z mednarodnimi preizkusi znanja (Zupanc, 2005), za strokovnjake in praktike predstavlja izzive, soočenje s problemi in motive za iskanje rešitev (Townsend, 2007). Stalno analiziranje izkazanega znanja omogoča odkrivanje možnih sistematičnih napak, ki jih navaja že Deming (1982; Marolt in Gomišček, 2005), na različnih nivojih šolskega sistema in predstavlja izhodišče za popravke in izboljšave (Allal, Mottier in Lopez, 2005).

V članku so prikazana izjemno visoka povprečja doseženih točk pri internem ocenjevanju praktičnih nalog na maturi, kakor tudi trendi vsakoletnega naraščanja. Z grafi ordinalne dominantnosti so za posamezne šole v primerjavi s celotno Slovenijo prikazani primeri razlik pri treh parih porazdelitev dosežkov: za šolske ocene pred maturo, za eksterno ocenjen in za interno ocenjen praktični del maturitetnega izpita. Ocenjevanje ob zaključku gimnazije mora tudi diskriminirati in prikazati razlike med maturanti po izkazanem znanju, ne pa da praviloma odlične dosežke eksterne preverjanja nižajo slabše interne ocene in obratno slabše eksterne ocene višajo visoko ocenjeni interni deli. Ker se s praktičnimi nalogami preverja pomembne učne cilje, ki se jih s pisnimi preizkusi ne more preveriti, bi bilo ukinjanje tega dela maturitetnega izpita neprimerno. Zato je podan predlog, da bi tudi interno ocenjevanje praktičnih nalog eksterno moderirali.

Tako klasična testna teorija (Nunnally in Bernstein, 1994) kot modernejši pristopi k testni teoriji (McDonald, 1999) poudarjajo, da mora preizkus odražati konstruktno veljavnost. Med drugim je del konstruktne veljavnosti tudi interna konsistentnost merjenega konstrukta (npr. znanje posameznega predmeta na maturi). Čeprav različni deli izpita tipično zaradi učinkovitosti merijo različna znanja, pa vseeno razhajanja med njimi ne smejo biti prevelika, sicer nismo upravičeni do seštevanja v skupen rezultat. Iz tega izhaja, da si morajo biti tudi porazdelitve dosežkov na različnih delih izpita podobne.

Izhodiščna hipoteza glede razhajanj med eksterno in interno ocenjenimi deli izpitov bi tako bila, da sta porazdelitvi dosežkov obeh delov izpita enaki. Ker gre za relativno enostavne primerjave, bo metoda raziskovanja primerjava osnovnih deskriptivnih statistik za omenjene podatke. Od uporabljenih statističnih primerjav lahko posebej izpostavimo graf ordinalne dominantnosti, ki omogoča tako pregledno vizualno primerjavo kot izpeljavo statistik (ρ), ki govorijo o velikosti razlik med dvema porazdelitvama.

Ustno preverjanje znanja in ocenjevanje praktičnih nalog

V strokovni literaturi, še posebej v knjigah, ki se ukvarjajo z merjenjem znanja v šoli oz. razredu (Linn, Gronlund, 2000; Oosterhof, 2001), se

v pregledu različnih oblik testov¹ in načinov preverjanja in ocenjevanja znanja ustnega spraševanja ne najde. Nastopi učencev, ki so seveda tudi ustni, pa se navajajo med ocenjevanjem dejavnosti učencev (ang. *performance assessment*).

Z ocenjevanjem dejavnosti se lahko meri kompleksne cilje. Lahko se ocenjuje dejavnost pri laboratorijskem ali terenskem delu, pri izdelavi in predstavitvi seminarских nalog, govornih nastopih ipd. Pri ocenjevanju takšnih spretnosti mora imeti kandidat možnost, da po svoji presoji predstavi in uporabi prej naučene informacije, koncepte in pravila, da reši nalogo, ki jo dobi pri preverjanju. Običajni pisni preizkusi ne omogočajo te fleksibilnosti in niso primerni za ta namen. Pri internem ocenjevanju dejavnosti pa je več subjektivnosti. Če bi želeli z različnimi dejavnostmi oceniti znanje in spretnosti, so študije pokazale (Taylor, 2003), da bi bilo za zanesljivost preverjanja treba oceniti več različnih dejavnosti. V Washingtonu (Taylor, 2003) so se pri vključevanju ocenjevanja dejavnosti iz razreda v sistem zunanjega preverjanja velikega števila dijakov pokazali naslednji problemi:

- usposobljenost učiteljev in priprava za ocenjevanje v razredu,
- izbor dejavnosti, ki se jih ocenjuje, in pomanjkljivi zapiski preverjanja – zmanjšuje se primerljivost rezultatov kandidatov,
- zanesljivost točkovanja dejavnosti in veljavnost rezultatov zaradi učinkov, ki niso povezani z znanjem in spretnostmi kandidatov – neprijemna pomoč učitelja, sošolcev, staršev ipd.

Pri oblikah preverjanja znanja, ki ne temeljijo na strokovno pripravljene pisnem preizkusu, je veliko težje zagotoviti pomembne merske značilnosti, kot so objektivnost, zanesljivost in veljavnost. Black in William (v Sentočnik, 2004: 72) učenje in preverjanje naučenega z netestnimi oblikami priporočata za formativno preverjanje. Dokler se v omejenem obsegu netestne oblike uporabljajo v formativni, diagnostični ali v manjšem, skrbno načrtovanem deležu, sumativni funkciji (Zupanc, 2004) preverjanja znanja, lahko prednosti prevladajo nad slabostmi.

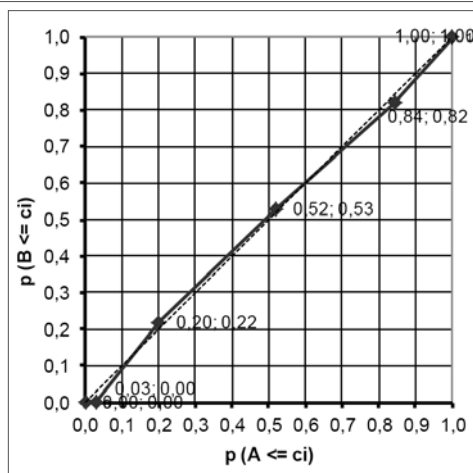
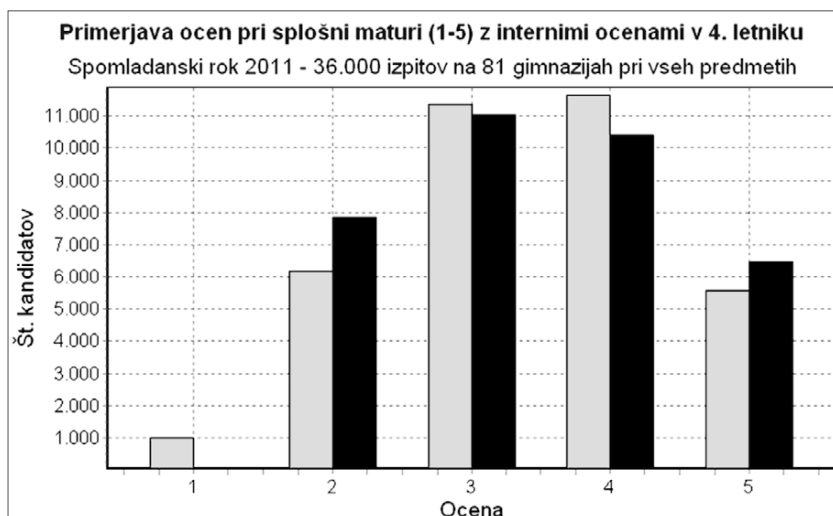
Interno in eksterno ocenjevanje v gimnazijah in pri maturi

V slovenskih šolah je tradicija, da učitelji ob koncu šolskega leta z zaključno oceno ovrednotijo znanje svojih učencev. Učiteljeve interne ocene odločajo o uspehu učenca in njegovem napredovanju. Gimnazijsko izobraževanje pa se ne zaključuje z uspešno opravljenim zaključnim (četrtim) letnikom, ampak je le-ta pogoj za pristop k splošni maturi, ki je v večjem delu ocenjena ekster-

1 V strokovni literaturi v angleškem jeziku se uporablja izraz *test* za različne oblike merjenja znanja. Pri citiranju splošnih virov je tako tudi v članku ohranjen izraz *test*, medtem ko se pri navajanju zunanjih preverjanj znanja v slovenskem prostoru uporablja izraz *preizkus* znanja.

no. Vsak maturant opravlja tri obvezne predmete, slovenščino oz. italijanščino ali madžarščino, matematiko in prvi tuji jezik; poleg tega izbere še dva, lahko tudi tri izbirne predmete. Maturanti, njihovi učitelji in ravnatelji šol spremljajo dosežene ocene na maturi, ki so v pretežnem delu dodeljene eksterno, in jih primerjajo z interno dodeljenimi šolskimi ocenami (Semen, 2011b; Semen, 2012). Za vsakega maturanta so pri internih ocenah v zaključnem letniku (na Sliki 1) upoštevane ocene iz petih oz. šestih maturitetnih predmetov.

Slika 1: Primerjava porazdelitve vseh ocen (v lestvici od 1 do 5) pri spomladanskem roku splošne mature 2011 (sivi stolpci v histogramu) z internimi ocenami v zaključnem, četrtem letniku gimnazije (črni stolpci v histogramu), spodaj graf ordinalne dominantnosti primerjave obeh porazdelitev, $N = 36.000$ izpitov.



Na Sliki 1 sta prikazani porazdelitvi ocen (v lestvici od 1 do 5) pri spomladanskem roku splošne mature 2011 (sivo) in interne ocene v zaključnem – četrtem letniku gimnazije (črno). V spodnjem delu slike je z grafom ordinalne dominantnosti² prikazana primerjava obeh porazdelitev (Zupanc, Bren, 2010b). Porazdelitvi sta si zelo podobni, med njima ni pomembnih razlik ($\rho = 0,501^2$; $d = 0,001$; Zupanc, 2010b; Vargha in Delaney, 2000). Tudi povprečni oceni sta praktično enaki: povprečna maturitetna ocena pri vseh predmetih, ki so jih dijaki pisali pri maturi, to je pri več kot petintrideset tisoč izpitih pri več kot 30 predmetih je 3,41; povprečna interna ocena teh dijakov pri istih predmetih v četrtem letniku pa je 3,43.

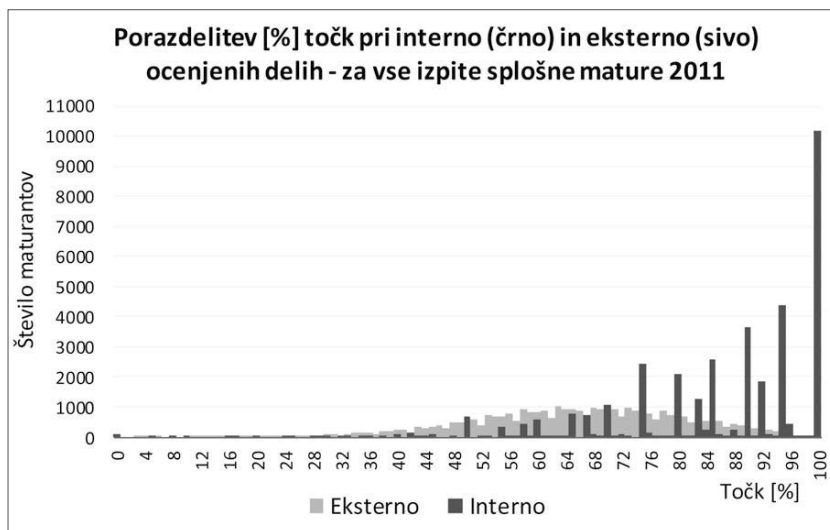
Pri slovenski maturi se izvajajo izpiti iz več kot 30 predmetov (Zupanc, 2001). Vsi materni in tuji jeziki ter matematika imajo pri maturi interni ustni del, ki k skupni oceni prispeva 20 %. Pri zgodovini umetnosti in likovni teoriji ustni del prispeva 30 %. Naravoslovni predmeti, geografija in zgodovina imajo v internem ocenjevanju laboratorijske vaje, eksperimentalno delo, terensko delo, ekskurzije, ki k skupni oceni prispevajo 20 %. Filozofija, psihologija in sociologija imajo 25 % vredno interno ocenjeno seminarsko nalogo. Laboratorijske vaje, eksperimentalno delo, terensko delo, ekskurzije, seminarske naloge in projektne naloge kandidati opravijo v zaključnem letniku. Mentorji – praviloma njihovi učitelji – jih ocenijo pred zaključkom pouka. Ustni izpiti pa se izvajajo po zaključku pouka pred šolsko komisijo.

Doseženo število odstotnih točk pri vseh oblikah internega ocenjevanja je pri vseh predmetih in vsa leta mnogo višje kot pri eksterni ocenjenih delih (Pučnik Ozimič, 2010; Starc, 2011; Zupanc, 2001; Zupanc in Bren, 2010a). Na Sliki 2 je prikazana porazdelitev odstotnih točk za vse izpite na spomladanskem roku splošne mature 2011 – posebej za eksterno (sivo) in posebej za interno ocenjeni del (črno). Pri več kot deset tisoč izpitih, skoraj tretjini, maturanti pri internem delu maturitetnega izpita dobijo 100 %. Verjetnost, da je imel naključno izbrani maturant v spomladanskem roku splošne mature 2011 ($N = 36.000$) višje odstotne točke pri internem kot pri eksternem ocenjevanju je 0,84 (upoštevana je tudi polovica verjetnosti, da sta bila oba dela ocenjena enako).

Na razkorak med internim in eksternim ocenjevanjem so različni avtorji opozarjali že pred leti. Šimenc in Slavec (1999) sta pri filozofiji na maturi poročala, da v povprečju vsako leto kandidati pri seminarskih nalogah dosežejo bistveno boljše rezultate kot pri pisnem eksternem delu izpita. Tako

2 Indeks ρ se uporablja kot mera različnosti dveh ordinalnih porazdelitev. Vrednost ρ je tudi ploščina nad (pod) grafom in je enaka verjetnosti, da je naključno izbrani kandidat iz ene skupine po dosežku uvrščen višje kot naključno izbrani kandidat iz druge skupine; prišteta je še polovica verjetnosti, da sta oba naključno izbrana kandidata po primerjanih dosežkih rangirana enako (Zupanc, Bren, 2010b).

Slika 2: Porazdelitev odstotnih točk za vse izpite v spomladanskem roku splošne mature 2011 – posebej za eksterno ocenjeni del (sivo) in posebej za interno ocenjeni del (črno).



izboljšajo tudi svojo skupno oceno pri tem izbirnem predmetu. Pri maturi iz sociologije se je leta 1998 povečalo število tistih šol, na katerih so bile zelo velike razlike med notranjo in zunanjo oceno. Počkar in Slavec (1999: 25–29) sta ugotovila, da na nobeni šole niso kandidati pri pisnem delu izpita dobili boljše ocene kot pri maturitetni nalogi. Pri psihologiji sta Skarza-Žerovnik in Slavec (1999) povzela podobno ugotovitev: v večini so bile šole z 20,1 do 30,0 % razliko med interno in eksterno oceno. Na petih šolah so razlike presegale 40 %, največja razlika pa je bila 75,5 %.

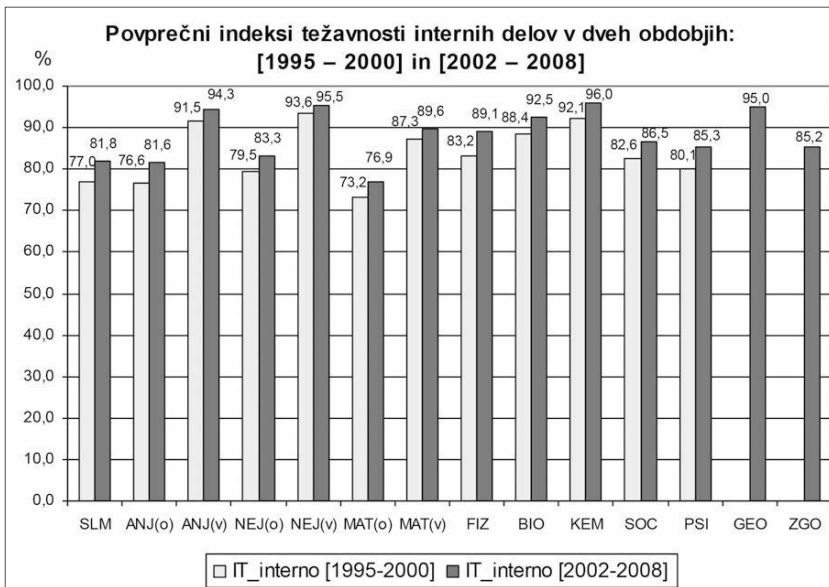
Rezultati internega ocenjevanja laboratorijskega dela pri fiziki so si iz leta v leto zelo podobni. Glavni ocenjevalec in koordinator pri maturitetnem predmetu Trampuš in Kranjčevič (1998: 15) v svoji analizi ugotavljata, da kandidati v glavnem dosega visoko število točk – razlika med povprečnim dosežkom in največjim možnim številom točk je vse manjša. Analiza internih ocen za leto 1997 je pokazala, da je stanje zaskrbljujoče, in avtorja domnevata (ibid.), da gre pri internem ocenjevanju za subjektivno nagrajevanje kandidatov. Razkorak med obema ocenama je največji ravno na šolah z najnižjimi rezultati pri pisnem eksternem delu maturitetnega izpita. Da notranja ocena vključuje tudi nagrajevanje truda kandidatov in potemtakem ni objektivna, sta Osredkar in Kranjčevič (1999) ugotavljala tudi pri kemiji.

Državna komisija za splošno maturo je v zadnjih letih razpravljala o prevelikih razlikah pri dodeljenih točkah pri notranjem oziroma zunanjem delu izpita (Starc, 2011) in o različnih interpretacijah za te razlike v notra-

njem in v zunanjem delu izpita, ki jih izražajo predvsem posamezne državne predmetne komisije in tudi šolske maturitetne komisije (Pučnik Ozimič, 2010).

Zupanc in Bren (2010a) sta objavila analizo spreminjanja (višanja) indeksa težavnosti³ interno ocenjenih delov izpitov pri več predmetih splošne mature v letih 2002 do 2008 v primerjavi z leti od 1995 do 2000 (Slika 3). Pri vseh predmetih so se indeksi težavnosti zvišali, kar pomeni, da so interno ocenjeni deli izpitov v zadnjih letih »lažji« kot pred leti; maturanti pri tem v povprečju dosegaajo 80 % oz. tudi prek 90 % možnih točk.

Slika 3: Povprečni indeksi težavnosti pri predmetih v spomladanskem roku splošne mature; posebej za obdobje 1995–2000, N = 49.046 kandidatov – sivo, in posebej za obdobje 2002–2008, N = 55.131 kandidatov – črno (vir: Zupanc, Bren, 2010a).

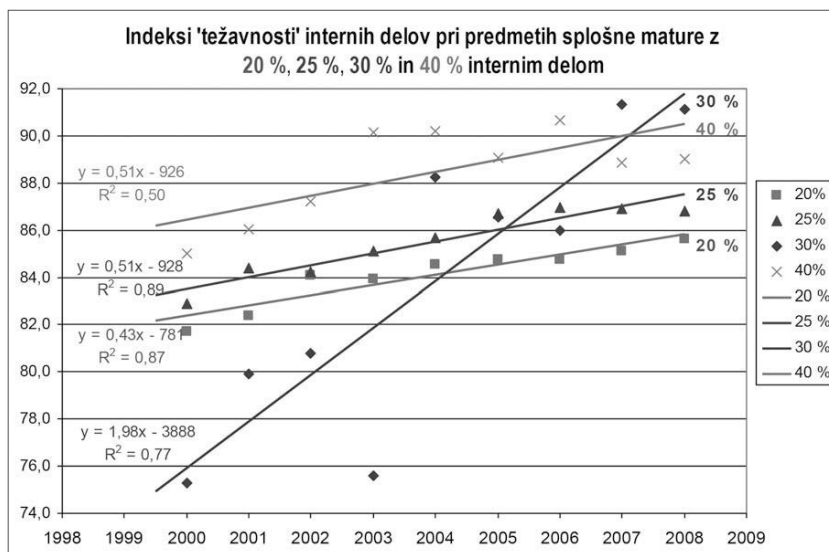


Kako visoko so glede na točke ocenjeni dijaki pri internih maturitetnih delih posameznih predmetov, je po skupinah predmetov prikazano na Sliki 4 (Zupanc, Bren, 2010a). Za vsako leto od leta 2000 do 2008 je na sliki označeno povprečno število točk, doseženih pri internem delu izpita za posamezno skupino predmetov – indeks težavnosti. S kvadrati so označene točke za skupino maturitetnih predmetov z 20-odstotnim internim delom, s trikotniki so označene točke za skupino s 25-odstotnim internim delom, z rombi so označene točke za skupino s 30-odstotnim internim delom

3 Lažja naloga ima višji indeks, težja naloga pa nižjega. V angleščini se uporablja izraz indeks »lahkosti« naloge (*facility index*).

in s križci točke za skupino s 40-odstotnim internim delom. S slike je razvidno, na kakšni odstotkovni višini so bile posamezne skupine leta 2000 in kako visoko so leta 2008. Za vsako skupino predmetov so dodane premice, ki kažejo trende naraščanja indeksov »težavnosti« v zadnjih letih. Smerni koeficienti za vse štiri skupine so statistično pomembni (Zupanc, Bren, 2010a).

Slika 4: Dosežene povprečne odstotne točke pri internih delih maturitetnih izpitov (indeksi »težavnosti«) po različnih skupinah predmetov splošne mature od leta 2000 do 2008 (vir: Zupanc, Bren, 2010a).



Glede na vsako leto višje dodeljene točke pri internem ocenjevanju praktičnih nalog pri slovenski maturi pa ni dokazov, ki bi kazali na iz leta v leto boljše znanje in doseganje zahtevnejših učnih ciljev ter standardov znanja (Zupanc, Bren, 2010a). Mnenja ravnateljev o aktualnem internem ocenjevanju ustnih in praktičnih delov izpitov pri splošni maturi so pozitivna (Semen, 2011a: 24). Tri četrt jih trdi, da so notranji deli izpitov potrebni, saj preverjajo znanje in spretnosti, ki jih pri zunanem delu ni mogoče preveriti, malo manj kot četrtnina pa jih meni, da notranji deli izpitov niso potrebni, saj so dijaki notranje ocenjeni skozi štiri leta gimnazijskega izobraževanja. Z argumenti za nujnost preverjanja učnih ciljev, ki se jih s pisnimi preizkusi ne da preverjati, je mogoče soglašati, nujne pa so spremembe izvedbe in ocenjevanja do sedaj interno ocenjenih ustnih delov izpitov in praktičnih nalog. V šolah spremembam niso naklonjeni: več kot dve tretjini (67 %) ravnateljev zagovarja dosedanja izvedbo, le slaba tretjina (31 %) jih zagovarja, da bi bili ustni izpiti ocenjeni že med šolskim letom (Semen, 2011a: 23). Ravnatelji so bili vprašani tudi za podporo spremembi, da bi interno ocenjene praktične

dele izpitov zunanje kontrolno ocenili. Nasprotovanje je bilo izrazito: s spremembo bi soglašala četrtnina (24 %) ravnateljev, tri četrtine (75 %) jih je bilo proti (Semen, 2011a: 22).

Različne porazdelitve ocen pri internem in eksternem ocenjevanju – primer fizike

Navkljub podatkom o visoko dodeljenih točkah (v povprečju približno 90 %) pri interno ocenjenih praktičnih nalogah in trendu vsakoletnega povečevanja v strokovni literaturi ni zaslediti ugotovitev, da se je zvišal nivo znanja, ki se ga na ta način preverja. Zasledimo pa lahko trend, o katerem je pisal Kodelja (2000), ko učitelji vedo, da je za nadaljnje šolanje učencev odločujoča skupna ocena, sestavljena iz interne in eksterne ocene, in jim namenoma dvigujejo interne ocene (ali pa znižajo kriterije ocenjevanja) ter tako vnaprej kompenzirajo njihov predvideni slabši uspeh na eksternem preverjanju. Kritična opozarjanja (Osredkar in Kranjčević, 1999; Počkar in Slavec, 1999; Skarza-Žerovnik in Slavec, 1999; Šimenc in Slavec, 1999; Trampuš in Kranjčević, 1998; Zupanc, 2001; Zupanc in Bren, 2010a) na previsoko ocenjene interne dele doslej niso obrodila sadov.

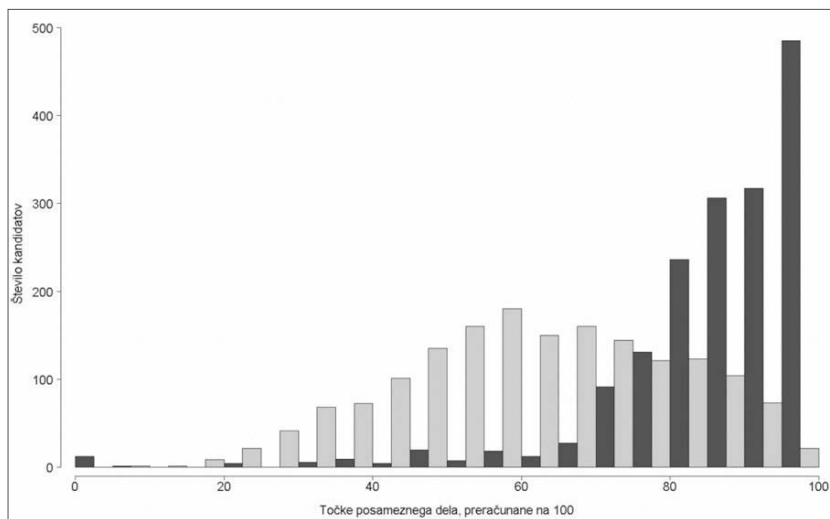
Ric in Državna komisija za splošno matura v Sloveniji že od leta 1995 opravljata analize kakovosti eksternih delov maturitetnih izpitov na podlagi rezultatov sistematične edukometrične analize testnih nizov ali posameznih izpitnih vprašanj. Za pisne izpite, eksterne ocenjene dele mature, se poleg veljavnosti sistematično analizira zanesljivost, diskriminativnost in objektivnost maturitetnih delov izpita za vsak predmet v vsakem izpitnem roku. Merske značilnosti so javno objavljene (navedba zadnjih dveh letnih poročil: Pučnik Ozimič, 2010; Starc, 2011) tudi za spremljanje kakovosti maturitetnih izpitov s strani zunanjih strokovnjakov, posebej pa se s povratno informacijo kot izhodiščem za delo vnaprej ukvarjajo državne predmetne komisije. Na ta način maturitetni sistem v Sloveniji zagotavlja transparentno, preverljivo in poglobljeno edukometrično kakovost za eksterne ocenjene dele mature; za interno ocenjene dele in šolske ocene v Sloveniji pa sistematičnih analiz na takem strokovnem nivoju ni. Ob odsotnosti analiz, ki bi sedanje dosežke na interno ocenjenih delih izpita potrdile kot veljaven odraz znanja kandidatov, se lahko naslonimo le na objektivnost eksterne ocenjenih delov izpita in jih jemljemo kot veljavno mero znanja kandidatov.

Na tej osnovi se merskim podatkom pri zunanje ocenjenih delih mature da posebno referenčno mesto in se podatke iz drugih načinov internega ocenjevanja primerja s prvimi. Ali so visoka (približno 90 %) povprečja in izrazito v desno zamaknjene porazdelitve točk pri interno ocenjenih praktičnih nalogah dovolj prepričljiv argument, da gre pri tem za sistemsko napako pri merjenju znanja? Državni izpitni center ima za vsakega maturanta v

okviru mature za vsak predmet različne podatke o dosežkih v znanju. Porazdelitve dosežkov interno in eksterno ocenjenih delov izpitov (v surovih ali odstotnih točkah) so si od šole do šole v veliki večini podobne, kar se odraža tudi na porazdelitvah za posamezni predmet na državnem nivoju (Pučnik Ozimič, 2010; Starc, 2011). Porazdelitve eksterno ocenjenih delov izpitov so praviloma zvonaste oblike z vrhovi okrog povprečnih dosežkov (60 % točk); deleži na obeh repih (maturanti z najnižjimi in najvišjimi dosežki) se bližajo ničli. Porazdelitve točk pri interno ocenjenih praktičnih delih maturitetnih izpitov pa so praviloma izrazito pomaknjene v desno s povprečnimi dosežki preko 80 % ali celo preko 90 % možnih točk in desnega repa sploh nimajo, saj so vrhovi porazdelitev pri najvišjih vrednostih.

Zanimiva je tudi drugačna primerjava, in sicer že omenjenih porazdelitev točk pri maturi in internih šolskih ocen maturantov ob zaključku zadnjega letnika gimnazije pred pristopom k maturi. Če se pri maturi z eksternim in internim delom (lahko) preverja drugačne učne cilje, pa se mora pri pouku v šolah ocenjevati celoten spekter znanj pri določenem predmetu – znanja, ki se preverjajo s pisnimi preizkusi, z ustnimi nastopi, s praktičnimi nalogami ipd. Ugotovljeno je bilo, da pri vseh gimnazijah pri vseh maturitetnih izpitih v Sloveniji med porazdelitvami šolskih ocen in maturitetnih ocen, kjer se v ustreznem deležu eksterne in interne točke seštejejo, ni pomembnih odstopanj (Slika 1).

Slika 5: Porazdelitev točk pri eksterno ocenjenih maturitetnih izpitih (sivo) in porazdelitev točk pri interno ocenjenih laboratorijskih vajah (črno) za fiziko na spomladanskem roku mature 2011. Obakrat so dosežki preračunani na 100 % (pri maturi interno ocenjeni del zajema 20 % skupnega dosežka).



Na primeru enega predmeta – fizike – bomo primerjali porazdelitve treh merjenj znanja: interne učiteljeve ocene ob koncu pouka v zaključnem letniku, eksterne točke in še interne točke z maturitetnega izpita. Slika 5 predstavlja porazdelitvi točk pri eksterno ocenjenih maturitetnih izpitih (sivo) in porazdelitev točk pri interno ocenjenih laboratorijskih vajah za fiziko na spomladanskem roku mature 2011 (črno).

Vidimo lahko, da sta porazdelitvi dosežkov na eksternem in internem delu maturitetnega izpita iz fizike za vse kandidate v Sloveniji zelo različni. Medtem ko je porazdelitev na eksternih maturitetnih preizkusih zvonasta in zelo dobro razvršča kandidate, je porazdelitev dosežkov na internem delu popolnoma asimetrična, saj ima največ kandidatov najvišje možno število točk.

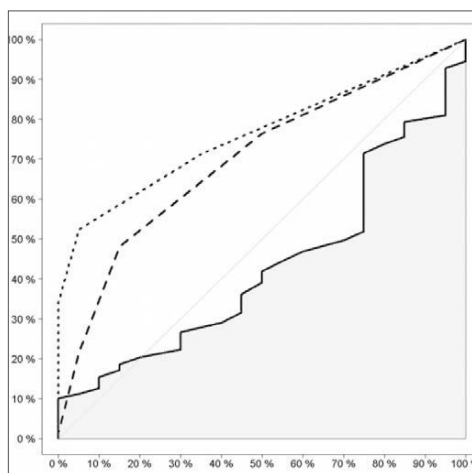
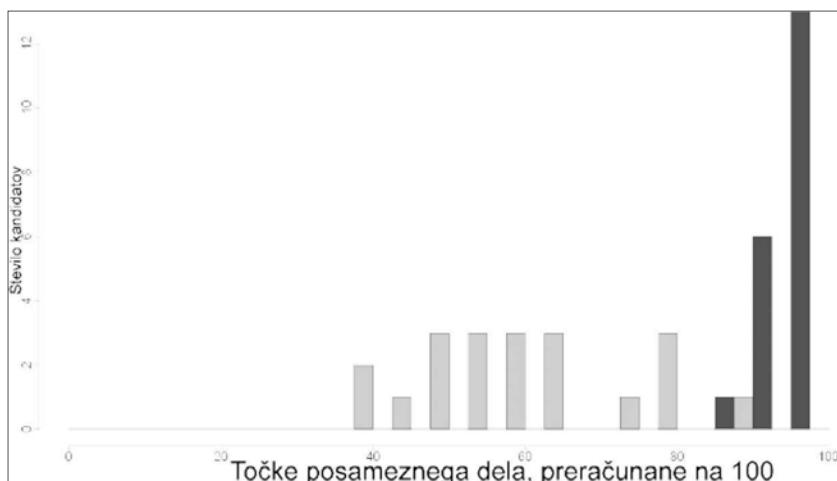
V nadaljevanju bosta prikazana primera dveh gimnazij – A in B – z zelo različnim načinom ocenjevanja dosežkov dijakov pri eksternem in internem preizkusu ter šolskem ocenjevanju. Gimnazija A, kjer je na spomladanskem roku 2011 maturo iz fizike opravljalo 20 kandidatov – prikazana je (zgoraj na Sliki 6) porazdelitev točk pri eksterno ocenjenih maturitetnih preizkusih (sivo); kandidati na šoli so dobili od približno 40 % do 90 % možnih točk. Na isti sliki (črno) je za iste maturante prikazana tudi porazdelitev točk pri interno ocenjenih laboratorijskih vajah; najslabše ocenjeni kandidat je dobil 85 %, največ pa jih je dobilo 100 % možnih točk.

Za posamezno šolo želimo hkrati prikazati primerjavo dosežkov na eksterno ocenjenih preizkusih z istovrstnimi dosežki vseh kandidatov v Sloveniji in še primerjavo interno ocenjenih delov izpita njihove šole z dosežki vseh kandidatov. Zato so na Sliki 6 spodaj prikazani grafi ordinalne dominantnosti, ki prikazujejo primerjavo dosežkov na šoli z dosežki vseh na istem maturitetnem izpitu. Razlika med dosežki eksterno ocenjenih preizkusov in interno ocenjenih laboratorijskih vaj na Sliki 5 je velika in zanimiva. Na grafih ordinalne dominantnosti pa so prikazana odstopanja ocenjevanja dosežkov na šoli A v primerjavi z ocenjevanjem na vseh šolah.

Vse krivulje, prikazane na Sliki 6 spodaj, predstavljajo primerjavo porazdelitev točk oz. ocen med izbrano šolo A (x os na grafu) in vsemi maturanti iz fizike v Sloveniji (y os na grafu).⁴ Posamezna krivulja predstav-

4 Graf ordinalne dominantnosti poteka od spodnjega levega oglišča, ki primerja dosežke kandidatov z najnižjim številom točk (ocen), in se zaključuje v zgornjem desnem oglišču, ki primerja dosežke kandidatov z najvišjim številom točk (ocen). Enostavna interpretacija grafa ordinalne dominantnosti je, da krivulja »po diagonali« razdeli kvadrat na dva dela in se primerja ploščini spodaj desno pod krivuljo (v tem primeru za konkretno šolo) s ploščino zgoraj levo nad krivuljo (v tem primeru za Slovenijo). Če graf (krivulja) skoraj razpolovi kvadrat po diagonali, sta ploščini skoraj enaki in porazdelitvi nista pomembno različni. Če pa sta ploščini (pod in nad krivuljo) različni, dosežki, ki pripadajo skupini z večjo ploščino, dominirajo nad dosežki skupine, s katero se primerjajo. Ploščina nad (pod) grafom je enaka verjetnosti, da je naključ-

Slika 6: Primer porazdelitve točk pri interno (črno) in eksterno (sivo) ocenjenem delu mature iz fizike spomladi 2011, $N = 20$ kandidatov - histogram zgoraj. Spodaj so z grafi ordinalne dominantnosti za isto šolo prikazane razlike pri treh parih porazdelitev dosežkov²: za šolske ocene pred maturo (črtkano), za eksterno ocenjen (polno) in interno ocenjen praktični del (pikčasto) maturitetnega izpita v primerjavi s celotno Slovenijo.



lja primerjavo para (dveh) porazdelitev: točke oz. ocene za izbrano šolo v primerjavi s točkami oz. ocenami za vse maturante v Sloveniji. Na črtkano označenem grafu se za ti dve skupini maturantov primerja interne, s stra-

no izbrani kandidat iz ene skupine po dosežku uvrščen višje kot naključno izbrani kandidat iz druge skupine; prišteta je še polovica verjetnosti, da sta oba naključno izbrana kandidata po primerjanih dosežkih rangirana enako (Zupanc, Bren, 2010b).

ni učiteljev dodeljene, šolske ocene iz fizike v zaključnem letniku. S šolskimi ocenami so bili na gimnaziji A maturanti v zaključnem letniku s strani učitelja ocenjeni višje kot vrstniki v Sloveniji (črtkana krivulja je v celoti izrazito nad diagonalo). Krivulja s polno črto predstavlja primerjavo porazdelitev točk, doseženih pri eksterno ocenjenem pisnem delu mature iz fizike. Graf s polno črto je najbolj oddaljen od črtkanega. Graf s polno črto je najbližji diagonalni,⁵ maturanti iz predstavljene šole so pri eksternem delu mature iz fizike dosegli nekoliko slabše dosežke kot njihovi vrstniki v Sloveniji ($\rho = 0,43^2$; razlika je šibka; Zupanc, 2010b; Vargha in Deleney, 2000). Podrobnejši pogled pove še, da so po dosežkih slabši maturanti izbrane šole pri eksternem delu mature iz fizike dosegli nekoliko višji rezultat kot vrstniki v Sloveniji (krivulja s polno črto je v spodnjem delu nad diagonalo), maturanti z nekoliko boljšimi dosežki pa vse do najboljših na šoli pa so se pri eksterno ocenjenih preizkusih izkazali slabše znanje kot vrstniki v Sloveniji (krivulja s polno črto je v zgornjem delu pod diagonalo).

Če se pogleda še primerjavo porazdelitve točk pri interno ocenjenih laboratorijskih vajah (pikčasta krivulja na Sliki 6), je očitno, da so bili z internimi točkami za praktični del na konkretni šoli maturanti s strani učitelja ocenjeni še izrazito višje kot vrstniki v Sloveniji (pikčasta krivulja v celotnem definicijskem območju razdeli kvadrat na dve izrazito neenaki ploščini) ($\rho = 0,76^2$; razlika je izrazita; Zupanc, 2010b; Vargha in Deleney, 2000). Učitelj(i) na izbrani šoli je (so) po dosežkih slabše in srednje dobre maturante pri laboratorijskem delu ocenjeval(i) še višje kot z zaključnimi ocenami v zadnjem letniku gimnazije (razvidno iz razmika med črtkano in pikčasto krivuljo na Sliki 6 – spodaj); svoje najboljše maturante je (so) za laboratorijsko delo ocenil(i) podobno kot z zaključno oceno v zadnjem letniku (vse primerjave so z maturanti iz fizike v Sloveniji).⁶

Težko je zagovarjati, da je s prikazanimi razlikami pri treh parih porazdelitev – črtkani (šolske ocene), polni (eksterne točke na maturi) in pikčasti (interne točke na maturi) – vse v redu, torej, da so razlike posledice izkazane ga znanja, ne pa ocenjevanja. Krivulja s polno črto kaže, da je bilo znanje maturantov izbrane šole pri eksterno ocenjenih delih mature iz fizike približno enako oz. nekoliko slabše od vrstnikov v Sloveniji. Kako bi poskušali razložiti povsem drugačno porazdelitev točk pri interno ocenjenih laboratorijskih

5 Graf ordinalne dominantnosti ima veliko prednost, da na enem mestu lahko primerjamo pare porazdelitev z različnimi lestvicami dosežkov; nekje so lahko dosežki v »surovih« točkah v lestvici od 0 do 20, drugje v odstotnih točkah od 0 % do 100 %, v tretjem primeru so to lahko šolske ocene od 1 do 5.

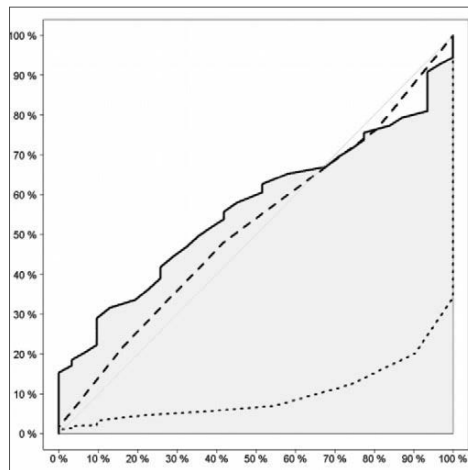
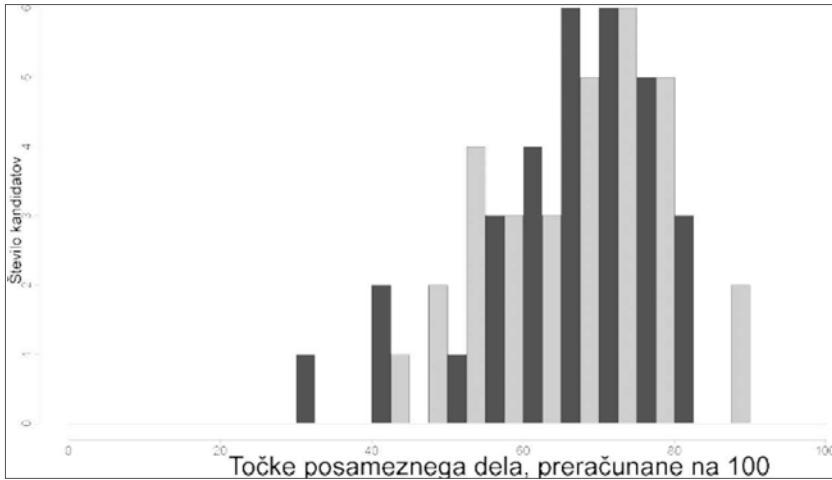
6 Ena možnih interpretacij je, da učitelj po dosežkih slabšim kandidatom pri interno ocenjenih praktičnih nalogah dodelijo neobjektivno veliko število točk (točke internega in eksternega dela mature se za končno maturitetno oceno seštejejo), da tudi ob pričakovano slabšem dosežku na eksterno ocenjenem delu mature na koncu ne bi bili neuspešni.

vajah (pikčasta krivulja)? Da se v tem delu preverja povsem drugačne učne cilje, v doseganju katerih so maturanti šole A (pač) izrazito boljši od vrstnikov v Sloveniji? Razliko med obema grafoma je težko zagovarjati, če upoštevamo še tretji (črtkani) graf – razlike v porazdelitvah šolskih ocen iz fizike iz zaključnega letnika ($\rho = 0,70^2$; razlika je srednja; Zupanc, 2010b; Vargha in Deleney, 2000). V šolskih ocenah bi se skladno z učnimi načrti za posamezni predmet morale pokazati doseganje celotne palete učnih ciljev, tistih, ki se jih tako pri pouku kot na maturi preverja pisno, in tudi preverjanje praktičnih nalog, pa so učiteljeve ocene na izbrani šoli mnogo boljše, kot so jih dobili vrstniki v Sloveniji. Ob prikazanih razlikah pri treh parih porazdelitev je verjetno, da gre v tem primeru za neobjektivno visoko interno ocenjevanje (Kodelja, 2000; Starc, 2011; Zupanc in Bren, 2010a).

Drugačen pa je primer gimnazije B, ki je bila med analizami razlik pri ocenjevanju fizike med gimnazijami izbrana kot nasprotni primer glede na gimnazijo A; primer gimnazije B je prikazan na Sliki 7. Na tej gimnaziji je na spomladanskem roku 2011 maturo iz fizike opravljalo 31 kandidatov. Histogram (Slika 7 zgoraj) kaže porazdelitev točk pri eksterno ocenjenih maturitetnih preizkusih (sivo) in za iste maturante porazdelitev točk pri interno ocenjenih laboratorijskih vajah (črno). Porazdelitvi se prekrivata. Kandidate je (so) učitelj(i) v šoli ocenil(i) v razponu od približno 30 % do 80 % možnih točk, pri eksterne delu pa so dobili od 40 % do 90 % možnih točk.

Na grafu ordinalne dominantnosti (Slika 7 spodaj) so za iste maturante prikazane primerjave treh porazdelitev dosežkov s porazdelitvijo dosežkov pri fiziki za vse slovenske maturante istega izpitnega roka^{2,3}. Graf s polno črto se približno prilega diagonalni, maturanti iz gimnazije B so pri eksterne delu mature iz fizike dosegli približno takšne dosežke kot njihovi vrstniki v Sloveniji ($\rho = 0,56^2$; razlika ni pomembna; Zupanc, 2010b; Vargha in Deleney, 2000). Podrobnejši pogled pove še, da so po dosežkih nižji in dobri maturanti konkretne gimnazije dosegli nekoliko višji rezultat kot vrstniki v Sloveniji (krivulja s polno črto je v dveh tretjinah nad diagonalo), najboljši maturanti na gimnaziji pa so pri eksterne ocenjenih preizkusih izkazali nekoliko nižje znanje kot vrstniki v Sloveniji (krivulja s polno črto je v zadnji tretjini pod diagonalo). Na črtkano označenem grafu se za isti skupini maturantov² (iz gimnazije B in za vse maturante, ki so v spomladanskem izpitnem roku 2011 v Sloveniji izbrali fiziko) primerja interne, s strani učiteljev dodeljene, šolske ocene iz fizike v zaključnem letniku – pred maturo. Črtkani graf na Sliki 7 spodaj skoraj idealno razpolovi kvadrat po diagonalni ($\rho = 0,52^2$; razlika ni pomembna; Zupanc, 2010b; Vargha in Deleney, 2000) in je po obliki podoben grafu s polno črto; učiteljeve interne ocene ob zaključku pouka v zadnjem letniku gimnazije so dobro prekrivajo z maturitetnimi rezultati. Tudi pri učiteljevih šolskih ocenah se je na gimnaziji B izkazalo, da

Slika 7: Primer porazdelitve točk na drugi šoli pri interno (črno) in eksterno (sivo) ocenjenem delu mature iz fizike spomladi 2011, $N = 31$ kandidatov – histogram zgoraj. Spodaj so z grafi ordinalne dominantnosti prikazane razlike pri treh parih porazdelitev dosežkov: za šolske ocene pred maturo (črtkano), za eksterno ocenjen (polno) in interno ocenjen praktični del (pikčasto) maturitetnega izpita v primerjavi s celotno Slovenijo.



so bili po dosežkih slabši maturanti v gimnaziji s strani učitelja v zaključnem letniku ocenjeni nekoliko bolje kot vrstniki v Sloveniji (črtkana krivulja je v dveh tretjinah nad diagonalo), najboljši maturanti na gimnaziji pa so bili s strani učitelja ocenjeni nekoliko slabše kot vrstniki v Sloveniji (črtkana krivulja je v zadnji tretjini pod diagonalo). Razlika v porazdelitvah ocen med gimnazijo B in Slovenijo je pri ocenah v zaključnem letniku manj izra-

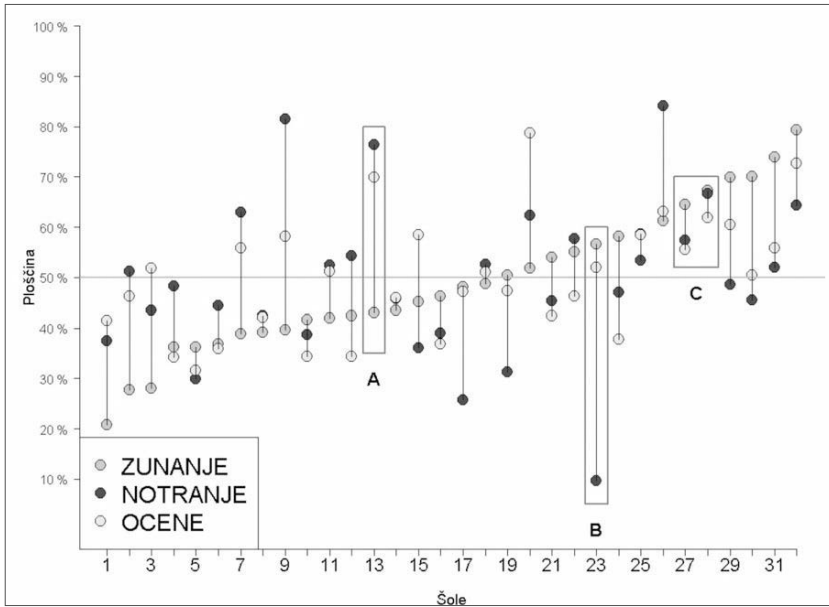
zita, kot je bila razlika, izmerjena pri eksterno ocenjenem delu mature (črtkana krivulja poteka bližje diagonalni kot polna krivulja).

Če se s porazdelitvijo točk za vse gimnazijce pri fiziki v Sloveniji pogleda še primerjavo porazdelitve točk pri interno ocenjenih laboratorijskih vajah v gimnaziji B (pikčasta krivulja na Sliki 7), je očitno, da so bili maturanti na tej šoli z internimi točkami za praktični del s strani učitelja(ev) ocenjeni izrazito nižje ($\rho = 0,10^2$, razlika je izrazita; Zupanc, 2010b; Vargha in Deleney, 2000) kot vrstniki v Sloveniji (ploščina pod pikčasto krivuljo je približno 10 %, nad krivuljo pa 90 % celotne ploščine kvadrata). Verjetnost, da je naključno izbrani kandidat, ki je izbral fiziko za maturo v Sloveniji, pri laboratorijskih vajah od svojega učitelja(ev) na šoli dobil večje ali enako število točk kot maturant na gimnaziji B, je 0,90. Pri eksterno ocenjenem delu mature iz fizike pa so gimnazijci te šole, glede na vrstnike v Sloveniji, izkazali nadpovprečen uspeh (graf s polno črto). Učitelj(i) na gimnaziji B je (so) s točkami za laboratorijsko delo svoje maturante ocenil(i) v širokem razponu (vse točke so preračunane na enako lestvico od 0 do 100 %), podobno širokem, kot so se kasneje razlikovale njihove dosežene točke pri eksterno ocenjenem delu mature. Ob predpostavki, da je (so) učitelj(i) fizike na gimnaziji B svoje dijake pri praktičnih nalogah ocenil(i) strokovno ustrezno, pa je ob pogledu na prikazane razlike treh parov porazdelitev težko spregledati, da (so) jih je primerjalno z drugimi gimnazijami in vrstniki v Sloveniji »oškodoval(i)«. Očitno je »normalna« zvonasta porazdelitev dodeljenih točk za praktični del maturitetnega izpita v Sloveniji (črni stolpci v histogramu na Sliki 7) atipična (glej Sliko 2) in po neodvisno izmerjenih maturitetnih dosežkih skupini maturantov gimnazije B »v škodo«.

Del ploščine, ki jo graf ordinalne dominantnosti v enotskem kvadratu razmeji, je uporabna mera za primerjavo skupin večjega števila šol in prikaz različnih situacij. Graf ordinalne dominantnosti je zaradi svoje podrobnosti in natančnosti zelo primeren za primerjavo dveh skupin. Ploščina, ki jo na grafu razmejuje lomljena črta, pa je celovit in smiseln povzetek informacije celotnega grafa z eno samo mero. Kadar želimo primerjati več skupin, to tako najlažje storimo s primerjavo omenjenih ploščin. Na Sliki 8 so prikazani deleži ploščin enotskega kvadrata, ki jih dobimo s primerjavami eksterno ocenjevanega dela izpita (sivi krožci), interno ocenjevanega dela maturitetnega izpita (črni krožci) in šolskih ocen (votli krožci) – vsakič v primerjavi z vsemi kandidati, ki so se udeležili maturitetnega izpita iz fizike v spomladanskem roku 2011. Velik razkorak med tremi deleži ploščin za posamezno šolo kaže na neusklajenost različnih načinov ocenjevanja (istih kandidatov pri istem predmetu).

Prikazanih je 32 šol, kjer je fiziko na maturi v tem izpitnem roku izbralo vsaj 20 gimnazijcev.⁷

Slika 8: Deleži plosččin, ki jih dobimo s primerjavami eksterno ocenjevanega dela izpita (sivi krožci – polne črte v grafih ordinalne dominantnosti), notranje ocenjevanega dela maturitetnega izpita (črni krožci – pikčaste črte) in šolskih ocen pred maturo (prazni krožci – črtkane črte) – vsakič v primerjavi z vsemi kandidati, ki so se udeležili maturitetnega izpita iz fizike v spomladanskem roku 2011.



Vseh šol, kjer so kandidati v tem roku izbrali fiziko, je bilo 65; z zgornjim kriterijem smo se izognili velikemu številu izračunanih statistik po šolah, ki bi bile zaradi majhnega števila kandidatov podvržene večji naključni napaki in manj uporabne za oblikovanje zaključkov.

Za lažjo primerjavo so šole na sliki urejene po velikosti deleža plosčine eksterno ocenjevanega dela mature. Izhodišče je, da so merski podatki pri eksternem maturitetnem ocenjevanju najbolj zanesljivi in objektivni (izmed dostopnih treh merjenj znanja). Pred tem smo že pojasnili, da maturitetni sistem v Sloveniji zagotavlja preverljivo edukometrično kakovost za eksterne ocenjene dele mature (Pučnik Ozimič, 2010; Starc, 2011); za interno ocenjene dele in za šolske ocene v Sloveniji pa ni sistematičnih analiz na takem strokovnem nivoju.

7 Za primerjanje ordinalnih porazdelitev in ugotavljanje razlik med njimi smo uporabljali ne-parametrične statistike. Mann-Whitney-Wilcoxonov U-test ima omejitve pri številu kandidatov, ki je večje ali enako 20 (Zupanc, 2010a). Tudi edukometrične analize v letnih poročilih se v Sloveniji delajo za predmete z 20 ali več kandidati (Pučnik Ozimič, 2010; Starc, 2011).

Gimnaziji, ki sta bili podrobneje predstavljeni z rezultati na slikah 6 in 7, sta na Sliki 8 označeni s črkama A in B. Maturanti na obeh gimnazijah so pri eksternem ocenjevanju znanja fizike na maturi v povprečju dosegli rezultat bližju povprečja, gimnazija A na Sliki 6 nekoliko pod, gimnazija B na Sliki 7 pa nekoliko nad povprečjem Slovenije (ploščini sta pod oz. nad 50 %). Na Sliki 8 pa vidimo, da so odstopanja pri dosežkih na posameznih gimnazijah zelo različna. Sivi krožci, ki predstavljajo dosežke maturantov na gimnazijah pri zunanje ocenjenem delu mature iz fizike glede na povprečje v Sloveniji, so od leve proti desni na grafu razporejeni po višini, od šole z najnižjim dosežkom na skrajni levi do gimnazije z najvišjim dosežkom na skrajni desni. Na sredini so tako prikazane gimnazije (sivi krožci) s povprečnimi maturitetnimi dosežki. Levo in desno od sredine so razlike od povprečja v dosežkih manjše, število gimnazij v tem območju je veliko; na obeh robovih pa je manjše število gimnazij, kjer pa so odstopanja od povprečja večja.

Nekatere gimnazije (npr. gimnaziji, označeni s črko C) imajo zelo majhne razlike med posameznim načinom ocenjevanja oziroma so različni načini ocenjevanja pri maturi in v gimnaziji dokaj usklajeni. Pri večini gimnazij pa so odstopanja večja, tako v smeri, ko so kandidati na internem ocenjevanju glede na dosežke eksternega dela dobili boljše ocene od njihovih vrstnikov (gimnazija A), kot tudi v drugi smeri, ko predvsem gimnazije z najboljšimi dosežki na eksternem delu svoje kandidate ocenjujejo strožje, kakor je to tipično za celotno populacijo (gimnazija B).

S Slike 8 je razvidno, da ima učiteljevo oz. notranje ocenjevanje v posamezni gimnaziji tendenco, da po višini nadomesti – »uravnoteži« – ocene svojih gimnazijcev glede na izmerjeno znanje s strani neodvisnih zunanjih ocenjevalcev pri maturi. Za dodeljene točke pri notranje ocenjenih delih izpita (črni krožci) in tudi za ocene v zaključnem letniku gimnazije pred maturo (votli krožci) velja, da so na gimnazijah s slabšimi maturitetnimi dosežki praviloma boljše, in obratno, na gimnazijah z nadpovprečnimi maturitetnimi dosežki praviloma slabše. Točke pri laboratorijskih vajah, kjer gimnazijce pri notranje ocenjenem delu maturitetnega izpita oceni njihov učitelj (črni krožci), in učiteljeve ocene ob zaključku gimnazije (votli krožci) so v povprečju na gimnazijah s podpovprečnimi dosežki – na levi strani grafa na Sliki 8 – nad, na gimnazijah z nadpovprečnimi dosežki – na desni strani grafa na Sliki 8 – pa pod krivuljo objektivno (eksterno) izmerjenega znanja. Razvidno je, da so gimnazije pri ocenjevanju s šolskimi ocenami bližje slovenski povprečni porazdelitvi, kot je to pri ocenjevanju maturitetnega izpita. V grafih ordinalne dominantnosti se to kaže skozi boljše prileganje črtkane črte diagonali, kar pomeni, da je ocenjevanje s šolskimi ocenami znotraj gimnazij »normalizirano«. Porazdelitve šolskih ocen v posameznih gimnazijah so si tako veliko bolj podobne kot porazdelitve dosežkov na eksterno ocenje-

nih preizkusih, kar lahko razložimo s preprostim dejstvom, da učitelji tako v slabših kot v boljših razredih oz. šolah svoje učence ocenjujejo po »relativnem« kriteriju: najslabšim pripišejo oceno ena, najboljšim oceno pet in tako izkoristijo ves razpon ocen. Različni avtorji navajajo, da učitelji ne ocenjujejo (samo) izkazanega znanja, ampak tudi prizadevnost in napredek (McMillan, 2002; Public Agenda, 2000), da so pristranski glede na socialno-ekonomski status učencev (Makarovič, 1984; Toličič in Zorman, 1977) in da prilagajajo svoja merila ocenjevanja skupini, ki jo poučujejo (Jurman, 1989). Jurman (ibid.) je že pred več kot dvajsetimi leti pisal, da kakovost znanja na šolah v Sloveniji pada že več let zapored, tako vsaj zatrjujejo učitelji v šolah. Učitelji pri ocenjevanju znanja v vseh oddelkih ne morejo uveljavljati istih standardov za ocene, temveč v »boljših« oddelkih zahtevnost dvigajo, v »slabših« jo spuščajo (ibid.). Pojavijo se pritiski na učitelje in ravnatelje, naj znižajo kriterije pri ocenjevanju znanja, olajšajo se zakonska določila v šolski zakonodaji za napredovanje učencev ali pa se celo menjajo lestvice ocenjevanja. Čeprav je »cilj« sicer dosežen, pa je ocenjevanje podvrženo inflaciji ocen (ibid.: 129). Prihaja do sprememb v oblikovanju standardov za posamezne ocene ter s tem do spremenjenega grupiranja ocen. Mejne vrednosti potrebnega znanja se vedno bolj pomikajo proti levi – znižujejo se nivoji doseženega znanja –, število učencev, ki dobivajo vedno višje ocene, pa proti desni. Na ta način se kandidate med sabo izenačuje, čeprav njihovi dosežki na eksternem delu kažejo drugače. Ob novi Beli knjigi (Krek, Metljak, 2011: 200) je bilo med drugim postavljeno vprašanje: »/A/li menite, da so učitelji, ki poučujejo na vaši gimnaziji, v zadnjih desetih letih pri ocenjevanju znižali kriterije zahtevnosti?« Samo 27 % ravnateljev meni, da ne; med učitelji gimnazij pa je takih celo samo 11 %.

Kališnik, Drole in Urank (1998: 37) so v okvirih teorije testiranja za slovensko maturo poudarili, da deli izpita, ki imajo nizko »težavnost«, ki jih z veliko verjetnostjo (npr. 0,80, 0,90 ali celo več) uspešno reši velika večina kandidatov, niso primerni. Taki deli izpita zaradi visokega indeksa »težavnosti« ne diskriminirajo med kandidati, ne ločujejo boljših od slabših in znanja od neznanja. Maturitetni izpiti za gimnazijce v Sloveniji izpolnijo normative za uvrstitev v kategorijo državnih zunanjih izpitov na osnovi kurikula (Bishop, 1998) (CBEES – curriculum-based external examination system). Znotraj OECD argumentirajo pomen CBEES ne samo za formalne zaključke izobraževanja na določeni stopnji in uporabo rezultatov za nadaljnje šolanje, ampak tudi za razvoj kakovosti izobraževanja v posamezni državi (OECD, 2010). En izmed kriterijev za »uvrstitev« izpitov v CBEES je, da poročanje o dosežkih kandidate razlikuje po doseženem znanju, izpitni rezultati morajo biti prikazani na več ravneh, v večstopenjski lestvici (Bishop, 1998). Menimo, da ni pravično, da večina kandidatov pri internem ocenjeva-

nju dobi (podoben) visok dosežek in se s tem izgubi ločljivost tega dela izpita. Čeprav lahko na eksterno vrednotenem delu izpita med kandidati zaznamo razlike v izkazanem znanju, dobijo na internem delu vsi zelo podobne ocene oz. celo na eksternem preverjanju boljši dobijo slabše interne ocene in obratno. Tako se ločljivost zmanjšuje, s tem pa se zmanjšuje tudi uporabnost maturitetnih rezultatov za povratno informacijo in razporejanje kandidatov glede na njihov dosežek, kar je zelo pomembna funkcija mature (Bishop, 1998; OECD, 2010).

Zaključek

Interno ocenjevanje praktičnih nalog pri predmetih slovenske mature kaže očitna odstopanja, anomalije, ki jih ni mogoče pojasniti kot naključne. Gre za sistemsko napako. Razlogi za preverjanje in ocenjevanje ustnih izpitov in praktičnih nalog pa so zaradi možnosti preverjanja pomembnih učnih ciljev, ki jih s pisnimi preizkusi ni mogoče preveriti, tako umestni, da bi se bilo pri maturitetnem izpitu praktičnim nastopom škodljivo odpovedati. Tako se mora pri notranjem šolskem ocenjevanju vrednotiti tudi znanje, izkazano z ustnim preverjanjem in praktičnim delom.

Za vse gimnazijce v Sloveniji pri vseh maturitetnih predmetih porazdelitev šolskih ocen ne odstopa od porazdelitev ocen na maturi. Učitelji so lahko dosledni in zahtevajo izkazane standarde znanja tudi pri ocenjevanju v gimnazijah. Problem pa je, da so med šolami in učitelji na šolah glede tega velike razlike; večina notranje dodeljene ocene prilagaja skupini dijakov, ki jo poučuje. Ne glede na dosežene standarde znanja tako učitelji ob zaključku pouka kot pri internem ocenjevanju praktičnih nalog porazdelijo ocene med skupinami in šolami podobno. Ko se takšna porazdelitev notranjih ocen za skupine dijakov primerja na eni strani med šolami in na drugi strani z izkazanim znanjem pri bolj objektivnem zunanem ocenjevanju na maturi, so notranje ocene med skupinami maturantov v različnih gimnazijah porazdeljene nepravilno. Ni pravična podobna (enaka) porazdelitev notranjih ocen maturantov po šolah, če je njihovo izkazano znanje (zelo) različno. Za merjenje znanja in ocenjevanje niso pomembne samo veljavnost, zanesljivost in objektivnost – ob zaključku gimnazije, za vstop v terciarno izobraževanje, morajo ocene tudi diskriminirati. Ne samo, da je diskriminativnost v sami naravi ocenjevanja znanja, ob zaključkih programa je ta pomembna in smiselna za dijaka in za nadaljnjo življenjsko pot mladega človeka, pomembna in smiselna je tudi za njegove učitelje in starše, za šolo in družbo kot celoto (Bishop, 1998; OECD, 2010). Da se prikažejo razlike (diskriminativnost) med gimnazijci po izkazanem znanju, je tudi edino prav(ično).

Za Kodeljo (2000) je ocenjevanje pravično, če je v skladu z retributivnim načelom pravičnosti, ki se glasi: za enako izkazano znanje enaka ocena.

Če se ga krši, je nepravilno. Pri tem je vseeno, ali je ocena previsoka ali prenizka. V obeh primerih je krivična. Pomembno je, da veljajo isti kriteriji pri ocenjevanju vseh učencev. Če na primer nek učitelj uporablja zelo visoke kriterije in drastično kaznuje najmanjšo napako, potem zanj sicer lahko rečemo, da je prestrog, nikakor pa, da je nepravilno. In obratno: če na primer učitelj pri ocenjevanju učenca, ki ima težave z nekim učnim predmetom, a se hkrati zelo trudi, uporabi milejši kriterij kot pri ostalih učencih, da bi ga rešil popravnega izpita ali ponavljanja razreda, potem je učitelj ravnal krivično, četudi obenem dobrosrčno. Kodelja (2000) opozarja, da ne samo, da so razlike med šolami, ampak imajo učitelji tudi znotraj iste šole različne kriterije pri ocenjevanju posameznih učnih predmetov in tudi istega predmeta. Kodelja je že pred leti opisal primer, ko učitelji vedo, da je za nadaljnje šolanje učencev odločujoča skupna ocena, sestavljena iz interne in eksterne ocene, in jim namenoma dvigujejo interne ocene (ali pa znižajo kriterije ocenjevanja) ter tako vnaprej kompenzirajo njihov predvideni slabši uspeh na eksternem preverjanju. Kljub samovoljnemu zvišanju ocen (ali znižanju kriterijev ocenjevanja) je ocenjevanje na ravni šole še vedno pravično, če učitelji pri zviševanju ocen ne delajo razlik med učenci in recimo vse učence ocenijo z eno oceno višje, kot bi jih, če bi upoštevali le njihovo znanje. Nepravilno pa je tako delovanje do učencev na šolah, kjer učitelji tega ne počnejo (Kodelja, 2000). Če bi kriterije ocenjevanja znižali vsi učitelji, bi bili učinki nični. Iz tega sledi, da zniževanja kriterijev ocenjevanja kot subjektivnega vodila za domnevno pravično ravnanje učiteljev ni mogoče univerzalizirati, ne da bi hkrati izničili prav tiste učinke, ki naj bi jih na ta način dosegli. V vsakem primeru pa je tako ravnanje učiteljev v nasprotju z retributivnim načelom pravičnosti, ki pravi, da je za enako znanje treba dobiti enako oceno. (Kodelja, 2000.)

Pri maturi v Sloveniji je problem v združevanju internega in eksternega ocenjevanja (Pučnik Ozimič, 2010; Starc, 2011; Zupanc, Bren, 2010a). Učitelji od šole do šole zelo različno ovrednotijo znanje in spretnosti svojih dijakov pri internem delu mature. Utemeljeni dvomi o inflaciji ocen pri internem ocenjevanju praktičnih nalog (Jurman, 1989; Zupanc, Bren, 2010a) mečejo slabo luč na preverjanje znanja s praktičnimi nalogami in dolgoročno škodujejo pomenu učnih ciljev, ki se preverjajo z ustnimi izpiti in s praktičnim delom. Interno ocenjeni deli mature so sicer v deležu končne ocene manjši kot eksterni, pa vendarle k oceni predmeta na maturi lahko pomembno prispevajo.

Če se pri maturi želi obdržati preverjanje delov izpitov z ustnim spraševanjem in s praktičnimi nalogami (merjenje drugih učnih ciljev), bodo potrebne spremembe. Rešitve v drugih sistemih (Anglija, mednarodna matura ...) gredo v smeri eksterne moderacije sicer interno ocenjenih delov izpitov. V primeru, ko se z eksterno moderacijo ugotovi pomembno odstopanje pri in-

ternem ocenjevanju, so predvideni učinkoviti mehanizmi, da se interne ocene popravijo in se v takih primerih razreši posamične primere neobjektivnega ocenjevanja. Dolgoročno pa že sama prisotnost takega mehanizma, skupaj z izobraževanjem učiteljev za ocenjevanje praktičnih nalog, vpliva na bolj zanesljivo in pravično ocenjevanje le-teh.

Druga, slabša možnost, saj ne odpravlja nepravilnosti pri ocenjevanju, je, da se interno ocenjevanje praktičnih nalog pri maturi ohrani, da pa se točke iz eksterno ocenjenega dela maturitetnega izpita in interno ocenjenega dela ne seštejejo in se na ta način ne dobi ena ocena za maturitetno oceno iz predmeta. Eksterno ocenjen izpit iz predmeta bi tako vodil do samostojne ocene, neodvisno od prve pa bi pri posameznem predmetu prišli do ocene interno ovrednotenega ustnega izpita oz. praktične naloge. O takšni rešitvi bi bile še potrebne analize in diskusije v okviru gimnazij in nenazadnje tudi s pristojnimi v Sloveniji, ki so odgovorni za vpis maturantov v terciarno izobraževanje.

Literatura

- Allal, L., Mottier Lopez, L. (2005). Formative Assessment of Learning: A Review of Publications in French. V: *Formative assessment – Improving learning in secondary classrooms*. Centre for Educational Research and Innovation. OECD.
- Bishop, J. (1998). The effect of curriculum-based external exit exams on student achievement. *Journal of Economic Education*, 29(2), 171–182.
- Brejc, M., Sardoč, M., Zupanc, D. (2011). *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes COUNTRY BACKGROUND REPORT SLOVENIA*. [Http://www.oecd.org/education/preschoolandschool/48853911.pdf](http://www.oecd.org/education/preschoolandschool/48853911.pdf) (10. 11. 2012).
- Deming, W. E. (1982). *Out of the Crises*. Cambridge: Massachusetts Institute of Technology, Center for Advanced Engineering Study.
- Državna komisija za splošno maturo (2010). *Maturitetni izpitni katalog za splošno mature 2012*. Ljubljana: Državni izpitni center.
- Eurydice, Eurostat (2012). *Key Data on Education in Europe – 2012*. Brussels: Eurydice.
- Japelj Pavešić, B., Brečko, B. N., Čuček, M., Vidmar M. (2004). *TIMSS 2003 – Mednarodna raziskava trendov znanja matematike in naravoslovja – Povzetek izsledkov*. Ljubljana: Pedagoški inštitut. [Http://www.pei.si/UserFilesUpload/timss2003-prviizsledki.pdf](http://www.pei.si/UserFilesUpload/timss2003-prviizsledki.pdf) (10. 11. 2012).
- Japelj Pavešić, B., Svetlik, K., Kozina, A., Rožman, M. (2009). *Znanje matematike in fizike med maturanti v Sloveniji in po svetu*. Ljubljana: JRZ Pedagoški inštitut. [Http://www.pei.si/UserFilesUpload/file/razisko](http://www.pei.si/UserFilesUpload/file/razisko)

- valna_dejavnost/TIMSS/TIMSSAdvanced/TIMSS_A_nacPorVSE.pdf (10. 11. 2012).
- Jurman, B. (1989). Ocenjevanje znanja, Selekcija ali orientacija učencev. *Pedagoška obzorja*. Ljubljana: Državna založba Slovenije.
- Kodelja, Z. (2000). *Pravičnost in ocenjevanje. Problemi ocenjevanja znanja in devetletna osnovna šola*. Ljubljana: Pedagoška fakulteta, Univerza v Ljubljani in Zavod RS za šolstvo.
- Kališnik, M., Drole, R., in Urank, M. (1998). Edukometrična analiza ocenjevanja. V: Brešar, F. (ur.), *Notranje vrednotenje mature 1995–1997*. Ljubljana, Državni izpitni center, 1–120.
- Krek, J., in Metljak, M. (ur.) (2011). *Bela knjiga v vzgoji in izobraževanju v Republiki Sloveniji 201*. Ljubljana: Zavod RS za šolstvo, 1–518.
- Linn, R. L., Gronlund, N. E. (2000). *Measurement and Assessment in teaching* (8th ed.). New York: Merrill/Prentice Hall.
- Makarovič, J. (1984). *Družbena neenakost, šolanje in talenti*. Maribor: Založba Obzorja.
- Marolt, J., Gomišček, B. (2005). *Management kakovosti*. Kranj: Moderna organizacija.
- McDonald, R. P. (1999). *Test Theory: A Unified Treatment* (1st ed.). Lawrence Erlbaum.
- McMillan, J. H., Myran, S., in Workman, D. (2002, March/April). Elementary teachers' classroom assessment & grading practices. *The Journal of Educational Research*, 95(4), 203–213.
- Mullis, I. V. S., Martin, M. O., Gonzales, E. J., Chrostowski, S. J. (2004). *TIMSS 2003 International Mathematics Report. International Association for the Evaluation of Educational Achievement – IEA. TIMSS & PIRLS International Study Center*, 1–437.
- Nunnally, J., in Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- OECD (2010). PISA 2009 Results: What Makes a School Successful? – Resources, Policies and Practices (Volume IV). [Http://www.oecd-ilibrary.org/education/pisa-2009-results-what-makes-a-school-successful_9789264091559-en](http://www.oecd-ilibrary.org/education/pisa-2009-results-what-makes-a-school-successful_9789264091559-en) (2. 9. 2012).
- Oosterhof, A. (2001). *Classroom applications of Educational measurement – Third ed.* New Jersey: Merrill Prentice Hall, Inc.
- Osredkar, M., Kranjčević, F. (1999). *Analiza izpitnih nalog z mature 1998. Kemija*. Ljubljana: Državni izpitni center, 22–24.
- Počkar, M., Slavec, A. (1999). *Analiza izpitnih nalog z mature 1998. Sociologija*. Ljubljana: Državni izpitni center, 25–29.
- Public Agenda. (2000, 16. februar). Reality check 2000 [Special Report]. *Educational Week*, S1–S8.

- Pučnik Ozimič, I. (ur.). (2010). *Letno poročilo – splošna matura 2010*. Ljubljana. Državni izpitni center.
- Semen, E. (2011a). *Analiza anketnega vprašalnika za ravnatelje – Analiza končnega poročila o izvedbi splošne mature 2011*. Ljubljana: Državni izpitni center. [Http://www.ric.si/splosna_matura/statisticni_podatki/](http://www.ric.si/splosna_matura/statisticni_podatki/) (30. 12. 2011).
- Semen, E. (2011b). *Analiza anketnega vprašalnika za dijake 2011*. Ljubljana: Državni izpitni center. [Http://www.ric.si/splosna_matura/statisticni_podatki/](http://www.ric.si/splosna_matura/statisticni_podatki/) (30. 12. 2011).
- Semen, E. (2012). *Analiza anketnega vprašalnika za ravnatelje – Analiza končnega poročila o izvedbi splošne mature 2012*. Ljubljana: Državni izpitni center (neobjavljeno gradivo).
- Sentočnik, S. (2004). Portfelj kot alternativna oblika vrednotenja učencevega napredka – možnosti njegove uporabe v slovenskih šolah. *Sodobna pedagogika*, 55(1), 70–91.
- Skarza - Žerovnik, M., Slavec, A. (1999). *Analiza izpitnih nalog z mature 1998. Psihologija*. Ljubljana: Državni izpitni center, 22–25.
- Šimenc, M., Slavec, A. (1999). *Analiza izpitnih nalog z mature 1998. Filozofija*. Ljubljana: Državni izpitni center, 31–33.
- Taylor, C. S. (2003, oktober). Incorporating Classroom Based Assessment into Large Scale Assessment Programs. Neobjavljeno gradivo: 29th Conference of International Association for Educational Assessment, Societies' goals and assessment. Manchester, Anglija, 1–43.
- Starc, S. (Ur.). (2011). *Letno poročilo – splošna matura 2011*. Ljubljana. Državni izpitni center.
- Toličič, I., in Zorman, L. (1977). *Okolje in uspešnost učencev*. Ljubljana: DZS.
- Townsend, T. (Ed.). (2007). *International Handbook of School Effectiveness and Improvement. Review, Reflection and Reframing*. Springer, 1–998.
- Trampuš, M., Kranjčević, F. (1998). *Analiza izpitnih nalog z mature 1997. Fizika*. Ljubljana: Državni izpitni center.
- Vargha, A., Deleney, H. D. (2000). A Critique and Improvement of the CL Common Language Effect Size Statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics*, 25(2), 101–132.
- Zupanc, D. (2001). Interno ocenjevanje pri maturi. *Sodobna pedagogika*, 52(3), 98–109.
- Zupanc, D. (2004). Funkcije preverjanja znanja, interpretacija rezultatov, poročanje o dosežkih. *Sodobna pedagogika*, 55(4), 148–167.
- Zupanc, D. (2005). Ravni znanja pri (inter)nacionalnih preverjanjih. *Šolsko polje*, XVI(5/6), 161–195.
- Zupanc, D. (2010a). Šolsko ocenjevanje v Sloveniji. *Didactica Slovenica – Pedagoška obzorja*, 25(3–4), 157–169.

- Zupanc, D. (2010b). Razlike v dosežkih dijakov pri zunanjih preverjanjih znanja pred vpisom in ob zaključku gimnazijskih in drugih srednješolskih programov. *Sodobna pedagogika*, 61(2), 142–163.
- Zupanc, D., Bren, M. (2010a). Inflacija pri internem ocenjevanju v Sloveniji. *Sodobna pedagogika*, 61(3), 208–228.
- Zupanc, D., Bren, M. (2010b). Izbira predmetov pri maturi in splošni uspeh – ali naravoslovne predmete izberejo po uspehu boljši? *Šolsko polje*, XXI(3–4), 107–135.
- Zupanc, D., Urank, M., Bren, M. (2009). Variability analysis for effectiveness and improvement in classrooms and schools in upper secondary education in Slovenia: assessment of/for learning analytic tool. *School Effectiveness and School Improvement*, 20(1), 89–122.

From teaching *about* the European union to learning *for* European citizenship: an analysis of the discourse in the didactic materials for teaching the European union

Today, the incorporation of the EU content in education is important because contemporary transnational and global processes affect the lives of citizens, whose identity, political affiliation and participation in previous centuries were related particularly to the national state. In recent years, the content of the EU has been included in the Slovenian national curriculum. Therefore, the teachers have an obligation to teach about the EU and also an opportunity to broaden pupil's understanding of European citizenship mainly due to the number of didactic materials and projects aimed at teaching EU in schools. Slovenia's entry into the EU has increased the transnationalisation and globalisation of education, which is a positive change. Education for European citizenship is becoming an important component in citizenship education. The purpose of this article is to critically examine the dominant discourse of the EU content in the teaching materials, assess its impact and place it in a wider theoretical context of global or transnational citizenship. This paper argues that the EU's self-presentation discourse, which is predominant in the teaching materials, is paradoxically one of the main obstacles that prevent teachers and pupils from developing the transnational citizenship perspective. Therefore, the article calls for a new form of European citizenship education, which is not so much teaching *about* the EU, but teaching *for* the EU and learning for European citizenship.

Key words: European Union, education, discourse, didactic materials.

Darko Zupanc, Gašper Cankar in Matevž Bren

Interno ocenjevanje pri slovenski maturi: velike razlike med šolami

V Sloveniji obstajajo velika odstopanja v porazdelitvah dosežkov, ko je znanje na pisnem delu maturitetnega izpita ocenjeno eksterno ali ko v šolah v okviru mature ustne izpite oz. praktične naloge učitelji ocenjujejo interno. Prikazana so izjemno visoka povprečja doseženih točk pri internem ocenjevanju praktičnih nalog na maturi in trendi letnega naraščanja. Z grafi ordinalne dominantnosti so za posamezne šole v primerjavi s celotno Slovenijo prikazani primeri razlik pri treh parih porazdelitev dosežkov: za šolske ocene pred maturo, za eksterno ocenjen in za interno ocenjen praktični del maturitetnega izpita. Zaskrbljujoče in nepravilne so zelo različne porazdelitve pri eksternem in internem ocenjevanju v šolah. Ocenjevanje ob zaključku gimnazije mora tudi diskriminirati in prikazati razlike med maturanti po izkazanem znanju, ne pa, da praviloma odlične dosežke eksternega prever-

janja nižajo slabše interne ocene in obratno: slabše eksterne ocene višajo visoko ocenjeni interni deli. Ker se s praktičnimi nalogami preverja pomembne učne cilje, ki se jih s pisnimi preizkusi ne more preveriti, je podan predlog, da bi tudi interno ocenjevanje praktičnih nalog eksterno moderirali.

Ključne besede: dosežki učenca, interno in eksterno ocenjevanje, praktične naloge, primerjava porazdelitev, matura v Sloveniji.

Internal grading of Matura: large differences between schools

There are stark variations in the achievement distributions in Slovenia between paper and pencil Matura exams graded externally compared to Matura coursework and oral exams in schools graded internally. The extremely high average scores for internal grading of Matura subjects' coursework and upward annual trends are presented in this paper. The differences in the distribution of three pairs of achievements for some schools in comparison to the entire country are also shown with an ordinal dominance graphs: for teachers' grades before Matura exams, for externally graded paper and pencil exams and internally graded Matura subjects' coursework. Very different distributions for external and internal grading in schools are disturbing and unjust. Grading in the end of Gimnazija program must also discriminate and show differences among Matura candidates according to achieved goals and standards. It is shown that as a rule, high achievement for the external parts is followed by low achievements for the internal parts and vice-versa. Since there some important learning goals that cannot be assessed using paper and pencil tests but can however be assessed using coursework, the possibility for change is outlined in this paper, suggesting that internally graded coursework and oral exams should also be externally moderated.

Key words: student achievement, internal and external grading, coursework, distributions comparison, Matura in Slovenia.

Janez Justin, Maja Zupančič in Mojca Rožman

Spoznavna učinkovitost učbeniškega besedila

V kvaziexperimentalni raziskavi smo revidirali štiri vsebinsko zaokrožena besedila, vzeta iz štirih učbenikov geografije za deveti razred osnovne šole. Revizija naj bi s posegi v štiri razsežnosti besedil (kohezivnost, argumentacija, premišljenost in koherentnost) povečala njihovo spoznavno učinkovitost. To so štiri razsežnosti didaktičnih besedil, ki so bile v teoriji (besediloslovju, spoznavnem jezikoslovju, psiholingvistiki, pragmatiki) dokaj natančno opisane. Najprej smo v kritični analizi besedil identificirali njihove pomanjkljivosti. Potem smo izdelali domnevno – grafično in jezikovno – izboljšane različice besedil. Revizija besedil je bila torej nadzorovana; spremembe, vne-