

Standardi znanja v edukaciji#

*Darko Zupanc**
Državni izpitni center, Ljubljana

Povzetek: Sodobni edukacijski sistemi temeljijo na standardih znanja. Izobraževalni sistemi po svetu so različni, imajo različno tradicijo in možna so različna poimenovanja podobnih stvari. Standardi že v fiziki pomenijo nekaj empiričnega - merljivega in takšno vlogo morajo imeti tudi v edukaciji. Standardi znanja se lahko razdelijo na vsebinske oz. kurikularne standarde in na standarde dosežkov oz. dejavnosti, bistvo standardov pa je v slednjih. Opredelitev učnih ciljev oz. vsebinskih standardov je šele prvi korak v postopku merjenja znanja. Namesto izraza vsebinski standardi bi lahko uporabljali tradicionalni pojem - učni cilji. Če naj izobraževalni sistem temelji na standardih, je pri posameznih predmetih na različnih stopnjah šolanja potrebno (še) priti do standardov dosežkov.

Ključne besede: standardi dosežkov, standardi dejavnosti, vsebinski standardi, kurikularni standardi, znanje, preverjanje, ocenjevanje, merske lestvice

Standards in education

Darko Zupanc
National Examination Centre, Ljubljana, Slovenia

Abstract: Modern educational systems are based on standards. Systems in countries all over the world differ and have different tradition, so different terms are used for similar concepts. In education, standards should have a similar role as in physics, where they are empirical and measurable. Two types of standards can be defined: content or curriculum standards and performance standards, with the latter being essential. Setting educational goals or content standards is the first step in the construction of knowledge measurement procedures. In Slovenia, the traditional term 'performance objective' or 'instructional objective' can be used instead of the term 'content standards'. For each school subject on different levels in Slovenian educational system, performance standards have to be set in the future, if educational system is to be based on standards.

Key words: performance standards, content standards, curriculum standards, knowledge, assessment, grading, levels of measurement

CC = 3530

Stališča, zapisana v prispevku, izkazujejo avtorjev pogled na vprašanja, povezana s preverjanjem znanja, in ne odražajo stališč institucije, v kateri je avtor zaposlen.

* Naslov / address: mag. Darko Zupanc, Državni izpitni center, Ob železnici 16, tel.: 01 5484 616, email: darko.zupanc@guest.arnes.si

Glass (1978) je že skoraj pred tridesetimi leti zapisal, da je problem vzpostavitve standardov temeljni problem izobraževanja, ki je še danes nerešen, kot je bil že pred 20 leti. Torej se o problemu standardov piše in govori že najmanj pol stoletja. Britanska akademija je leta 1998 organizirala poseben simpozij na temo standardov v edukaciji. Med povabljenimi uglednimi akademiki in raziskovalci nihče ni ponudil enostavne rešitve problema definicije standardov znanja in spretnosti, merjenja le-teh, niti recepta za uradno šolsko politiko (Goldstein in Heath, 2000).

Cannell (1987) je prepričljivo vzbudil pozornost javnosti v ZDA, kar je znano kot efekt Lake Wobegon (Gipps, 2000; Koretz, 1988), ko je ugotovil, da so pravzaprav vse države v ZDA in večina okrožij objavljali, da njihovi učenci dosegajo nadpovprečne rezultate glede na nacionalne normative. Na osnovi svojih podatkovnih preglednic je zaključil, da učencem, staršem, šolskemu sistemu, zakonodajalcem in sredstvom javnega obveščanja prikazujejo napihnjena in zavajajoča poročila o doseženih nivojih pri standardiziranih, nacionalno normiranih testih dosežkov.

Standardi znanja so bili osrednja iniciativa v izobraževanju v ZDA že v Clintonovi administraciji. Tudi ključna oblika aktualne reforme vzgoje in izobraževanja v ZDA je prizadevanje za izdelavo standardov (Goals 2000: Educate America Act). S tem se na takšen ali drugačen način ukvarjajo tudi v drugih državah.

Večina držav centralne in vzhodne Evrope je imela v preteklosti zapisane t.i. izobraževalne standarde; šlo je za naslove in podnaslove vsebinskih področij in število ur, ko naj bi to obravnavali, redko pa je bilo napisano, kakšne dosežke naj bi učenci dosegli (West in Crighton, 1999). V Sloveniji se pojem 'standard znanja' uporablja zelo različno (Zupanc, 2005). V veljavnih katalogih znanja za srednje strokovne šole 'standardov znanja' pri strokovno-teoretičnih predmetih sploh (še) ni. V učnih načrtih za predmete v gimnaziji obstojajo zapisi o standardih znanja, med različnimi predmeti pa je očitno razumevanje le-tega zelo različno, celo znotraj posameznega predmeta je uporaba tega termina nekonsistentna. V učnih načrtih za osnovno šolo je besednjak o standardih bogatejši in celo šolski predpisi govorijo o njih. Razen redkih izjem termin 'standard znanja' nadomešča ali celo podvaja termin 'učni cilj' (Zupanc, 2005).

Če termin 'standard znanja' v izobraževanju v Sloveniji predstavlja novost, je za dobro umestitev le-tega potrebno pogledati vire. Pregledati je potrebno zapise v državah in vzgojno-izobraževalnih sistemih, kjer se ta izraz že dolgo uporablja. Gre predvsem za področja z angleško govorečim prebivalstvom: ZDA, Anglija, Kanada in Avstralija. Po drugi strani pa je preverjanje znanja področje merjenja v psihologiji in o standardih se eksaktno govori pri merjenjih v naravoslovju.

V različnih šolskih sistemih zaradi lastne zgodovine in tradicije podobne stvari različno poimenujejo. Tako kot je bilo v zgodovini znanosti, tehnike in nenazadnje trgovine in gospodarstva za pretok blaga, ljudi in idej pomembno natančno definiranje in prevajanje pojmov, je to pomembno tudi pri edukacijskih sistemih. Poimenovanje določenega izraza samo po sebi ni pomembno, pomembno je, da je razumevanje vsaj med strokovnjaki enako. Če različni strokovnjaki določen temeljni pojem razumejo in interpretirajo različno, je to slabo in povzroča nejasnosti.

Ob analizi tujih virov lahko ugotovimo, da 'standardi znanja' v edukaciji niso enovit pojem; ravno nasprotno, ob široki uporabi, razumevanju in interpretaciji tega pojma je nujno postaviti razmejitve na t. i. vsebinske ali kurikularne standarde znanja in na standarde dejavnosti ali dosežkov. Tudi postopki pri opredelitvi in postavitvi enih in drugih standardov znanja so različni.

Metoda

Pojem 'standard znanja' v edukaciji ni tako jasno opredeljen, kot je pri uporabi v naravoslovju. V Sloveniji je to najbrž še bolj izrazito, ker je uporaba standardov znanja v edukaciji terminologija, ki je včasih niso uporabljali. Vendar tudi v izobraževalnih sistemih v drugih državah ti pojmi niso enoznačni.

Namen prispevka je s primerjavo tujih in domačih virov določiti bolj enotno oz. manj dvoumno poimenovanje podobnih stvari s sorodnimi izrazi. Za dobro opredelitev standardov znanja v edukaciji je potrebna primerna umestitev le-teh v širši sklop psihološkega testiranja, ki se ukvarja z merjenjem znanja učencev. Ker imajo standardi korenine v naravoslovju, je osnova razumevanja proučitev in uporaba pojma(ov) v temeljni naravoslovni znanosti – fiziki – in preko razumevanja pojma(ov) na tem področju smiselna transformacija na področje edukacije.

Iz zgledeov z osnovnimi fizikalnimi količinami in merjenjem le-teh lahko z analogijo poskušamo opredeliti učne cilje oz. kurikularne (vsebinske) standarde in kot osnovo za merjenje znanja in spretnosti uporabimo standarde dosežka(ov).

Predstavljena študija je zasnovana na deskriptivni metodi. Za tehniko in instrumentarij je bila uporabljena analiza tuje literature z iskanjem podobnosti in razlik v pojmovanju in konceptualnih pristopov pri standardih znanja oz. vsebinsko sorodnih pojmi. Ob preučevanju fizikalnih pristopov v temeljni naravoslovni znanosti, poznavanju osnov merjenja v psihologiji in pregledu sistemov edukacije v različnih državah prispevek s primerjanjem ugotavlja podobnosti in analogije. V prispevku so poleg slovenskih izkušenj obravnavani še sistemi edukacije v ZDA, Angliji, Kanadi, Avstraliji in mednarodna raziskava znanja matematike in naravoslovja TIMSS 2003.

Opredelitev pojma 'standardi znanja'

V slovarjih se standard¹ pojavlja kot norma, pravilo, merilo; enotna mera ali teža; vzorec, vzor, splošno veljaven tip; smernica ipd. V posebnem slovarju s področja preverjanja znanja v edukaciji so standardi navedeni kot dogovorjena vrednotenja, ki

¹ Pojem 'standard' se v angleških zapisih prvič pojavi leta 1138, v povezavi z bitko 'Battle of the Standard', med Angleži in Škoti, 22. avgusta 1138 (Aldrich, 2000). 'Standard' v angleškem jeziku pomeni prapor, posebno zastavo, ki se ji izkazuje lojalnost; kraljevska zastava ('the royal standard') plapolala, ko je kraljica v rezidenci. Figurativno 'standard' predstavlja prominentnega vodjo v določeni stvari. (Hornby, 1975).

jih uporabljamo pri merjenju kakovosti dejavnosti učenca, metod poučevanja ali kurikula (Glossary of Educational Assessment, 2003). Standardi so ključni pojem sodobne diskusije v edukaciji in če raziščemo kontekst, v katerem se pojem uporablja, ugotovimo, da je bistvo predvsem kvantitativno (Bartholomew, 2000). Ocenjevanje znanja v edukaciji je merjenje (Grgin, 1994). Vključuje vse nujne komponente, ki so pri vsakem merjenju. Vsako merjenje predpostavlja: predmet (vsebino, pomen) ali tisto, kar merimo, instrument, s katerim merimo, in tehniko merjenja. Merilo umerimo; navadno za to uporabimo natančnejše merilo. Tudi tega umerimo, navadno s še natančnejšim merilom. Ta veriga se konča pri dogovorjeni enoti, pri njeni izvedbi, etalonu ali standardu.

Izboljšave v edukaciji se morajo začeti z jasnimi pojmi, kaj se od učencev pričakuje, da se (na)učijo. To izhodišče je osnova za izboljšanje ameriškega sistema izobraževanja (Linn in Gronlund, 2000). Navedbe (zapisi) določajo, kaj naj se poučuje in kaj naj se učenci (na)učijo. Standardi določajo pričakovane cilje za učence, vendar ne opredeljujejo določenega učnega načrta, učbenika ali učnega pristopa. Lahko je veliko načinov, kako doseči cilje. Busheva zakonodaja v ZDA iz leta 2001, znana pod geslom "No Child Left Behind" (NCLB), zahteva od vseh držav v ZDA, da imajo teste znanja, ki temeljijo na standardih, in da so vsi učenci do dvanajstega leta šolanja testirani (Morgan in Perie, 2004). Pooblastilo o testiranju na osnovi standardov v več razredih in iz različnih vsebinskih področij ter določeno poročanje o dosežkih na več ravneh povečuje potrebo po trdno postavljenih procesih postavitve standardov, ki se jih lahko tudi uspešno zagovarja.

Standardi znanja so povezani s kriterijsko uporabo oz. interpretacijo preverjanja znanja (Nacionalna komisija za uvajanje in spremljanje novosti in programov v vzgoji in izobraževanju, 2000; American Educational Research Association, American Psychological Association in National Council on Measurement in Education, 1999), kjer je izbran kriterij tista kritična točka, ki razdeli preizkušance na one, ki so dosegli, in na tiste, ki niso dosegli določenih standardov znanja. Natančna opredelitev v učnem načrtu zapisanih operativnih ciljev je pri tem nujna. Za nacionalno oz. državno določene učne cilje se uporablja tudi izraz (vsebinski) standardi (Medveš in Svetlik, 2002; American Educational Research Association, American Psychological Association in National Council on Measurement in Education, 1999).

V strokovni literaturi s področja preverjanja v edukaciji (Cresswell, 2000; Linn in Gronlund, 2000; Morgan in Perie, 2004; International Center for Technology in Learning, 2004); American Educational Research Association, American Psychological Association in National Council on Measurement in Education, 1999; Taylor, 2002) se govori o dveh vrstah standardov, vsebinskih standardih in standardih dejavnosti oz. dosežkov. Vsebinski standardi navajajo, kaj je pomembno, da učenci znajo oz. naredijo, kakšne vrste veščin in katero znanje naj bi bili učenci sposobni prikazati. Standardi dejavnosti oz. dosežkov pa so vezani na empirično izmerjene vrednosti znanja in spretnosti. Standardi dosežkov kvantificirajo vsebinske standarde; koliko in v kakšni meri mora učenec znati in narediti glede na vsebinske standarde, tj. koliko in v kakšni meri učenec dosega vsebinske standarde, da se ga lahko prišteva med poznavalca oz. med tiste, ki so standard dosegli.

Vsebinski standardi (še) niso standardi dosežka(ov)

Svetovna banka za zgled vzgoje in izobraževanja, ki temelji na standardih, ponuja primere iz Avstralije, Kanade in ZDA (Standards & testing, 2001). V ameriški (Linn in Gronlund, 2000) in britanski (Cresswell, 2000) strokovni literaturi je nesporno, da morajo biti učinkoviti standardi postavljeni tako, da opredelijo dvoje:

- kaj se mora učenec (na)učiti oz. kaj se bo preverjalo in
- določiti, kdaj je standard dosežen, oz. določiti nivo dosežkov, ki so primerljivi s tistimi, ki so potrebni za vsako oceno pri drugih izpitih iz iste družine.

Za razumevanje izpitnih standardov je potrebno upoštevati oboje. Oba namena – opredelitev, kaj se mora učenec (na)učiti, in določitev, kdaj je standard dosežen – sta v povezavi z dvema vrstama standardov, ki jih običajno ločimo na vsebinske standarde in standarde dosežkov oz. dejavnosti. Vsebinski standardi določajo »kaj«, medtem ko standardi dejavnosti določajo »kako dobro«. Vsebinski standardi so javne navedbe (izjave, *angl. statements*), ki določajo, kaj naj bi učenec znal in bil sposoben narediti pri določeni vsebini ali šolskem predmetnem področju na določeni stopnji izobraževanja in usposabljanja (npr. pri branju v 4. razredu, pri matematiki v 8. razredu). Standardi dejavnosti (*angl. performance*) so odvisni od vsebinskih standardov, vezani so na empirično izvedena merjenja znanja in dodajo specifikacijo nivoja dejavnosti, ki jo pričakujemo od učencev v povezavi z vsebinskimi standardi. Z drugimi besedami, odgovarjajo na vprašanje: »Kako dobro je dovolj dobro?«

Nabor vsebinskih standardov razvijemo z namenom obsežnega opisa učenčeve dejavnosti z razmejitvijo vsebin in splošnih veščin, ki jih merimo. Vsebinskim standardom sledi razvoj opisov in empirično ugotovljenih (izmerjenih) ciljev ali nivojev dosežkov. Takšni standardi dosežkov imajo namen definirati potrebna znanja in veščine za vse različne kategorije znanja ali sposobnosti. (American Educational Research Association, American Psychological Association in National Council on Measurement in Education, 1999)

Pavešić (2005) v prevodu ameriškega članka o standardih razpravlja, da brez dobro izbranih vzorčnih nalog celo skrbno izdelani standardi ne zmorejo natančno posredovati različnih spoznavnih ravni, potrebnih za doseganje ustreznega matematičnega znanja. Samo besede, četudi jasne in pravilne, ne morejo v celoti opisati matematičnega razumevanja. Na primer, »reši sistem dveh linearnih enačb z dvema neznankama« je povsem jasna naloga, vendar ne pove nič o zahtevnosti, ki izvira iz tipa neznank ali koeficientov. Dobro premišljeni problemi lahko posredujejo to značilno enovitost matematike na način, ki ga naštevanje in opisovanje standardov nikoli ne more. Matematiki lahko najbolje pomagajo gibanju za standardizacijo znanja v preduniverzitetni matematiki s ponujanjem problemov – problemov, ki so izzivalni, ki presenečajo, ki širijo obzorja – torej z nadgradnjo opisov učnih ciljev oz. vsebinskih standardov h konkretnim empiričnim standardom dejavnosti (Pavešić, 2005).

Udeleženci mednarodnega posveta v okviru projekta JONAS (JOint National ASsessment project) so ugotovili, da obstajajo različna pojmovanja izrazov »cilj« in »standard« (National Examinations and their meaning for Quality Assessment, 2001, str. 5). Dogovorili so se, da bodo pojem »standard« uporabljali le za vrednotenje rezultatov preverjanja znanja. Vse vhodne »standarde« (*angl.* input standards), ki se v procesu edukacije pojavljajo pred konkretnim preverjanjem, naj bi poimenovali »izpitne cilje« (*angl.* examination objectives).

Če že je uporabljen izraz »vsebinski standardi«, pomeni le-ta široko dogovorjena pričakovanja, kaj naj bi učenec znal, razumel in bil sposoben narediti za določeno oceno pri določenem predmetu (International Center for Technology in Learning, 2004). Nekateri avtorji menijo, da ni nujno, da učni cilji vključujejo tudi standarde izvedbe oz. rezultata za zadostno oceno (Oosterhof, 2001). Drugi (npr. Mager, 1984) pa zagovarjajo, da bi morali vsi učni cilji vsebovati tudi standard izvedbe oz. definirano uspešnost izvedbe. Oosterhof (2001) meni, da to ni vedno možno niti ni nujno.

Preverjanje znanja lahko služi različnim namenom: certificiranju dosežkov, selekciji ali nadzoru doseganja izobraževalnih standardov (The Nature Of Public Examinations, 2002). Nadzor nad doseganjem standardov v vzgoji in izobraževanju najbolj dosežemo z nacionalnimi preverjanji (*angl.* national assessment). Gre za sistematično merjenje tipičnega nivoja dosežkov učencev oz. standarde dosežkov na nacionalni ravni. Pri nacionalnem preverjanju ugotavljamo doseganje nacionalnih vzgojno-izobraževalnih standardov. Običajno ga izvajamo na reprezentativnem vzorcu učencev (Glossary of Educational Assessment, 2003). Pri opredeljevanju ciljev na nacionalni ravni je potrebno imeti v mislih dvoje (National Educational Goals Panel, 1991):

- vrsto dokaza oz. evidence (npr. esej, matematični dokaz, znanstveni eksperiment, projekt, izpit ali kombinacijo naštetih), ki je potrebna za prikaz doseganja vsebinskega standarda, in
- kakovost opravljene dejavnosti učenca, ki bo pri presojanju sprejemljiva (s čim si učenec prisluži zadostno, dobro ali odlično oceno).

Standardi dosežkov v drugih državah in v Sloveniji

Sistem nacionalnega preverjanja znanja v šolstvu ZDA (NAEP – National Assessment of Educational Progress) je zelo znan in velja kot primer dobrega poročanja o rezultatih. V bogati dokumentaciji je podrobno zapisano, kaj se pri NAEP preverja pri določenem predmetu – t. i. vsebinski standardi. Z umerjenimi preizkusi znanja, ki jih dograjujejo in izvajajo že več kot trideset let, imajo dobro določene standarde dosežkov za posamezno starostno skupino na treh stopnjah oz. štirih, če se šteje še neuspešne (NCES - National Center for Education Statistics, 2004). S t. i. ravnmi dosežkov ('basic', 'proficient', 'advanced') so podrobno opisani standardi dosežkov, ki opredeljujejo matematično znanje učencev.

V Angliji in Walesu so v nacionalnem kurikulumu zapisani standardi znanja z ravnmi zahtevnosti, ki jih vsako leto preverjajo z nacionalnimi preizkusi v različnih starostnih obdobjih ('Key Stage 1' pri 7 letih, 'Key Stage 2' pri 11 letih, 'Key Stage 3' pri 14 letih). Za celotno obdobje obveznega šolanja so ravni znanja – standardi dosežkov – opredeljeni na 9 oz. 10-stopenjski lestvici (Statistics of Education, 2001).

V Kanadi izvajajo nacionalno preverjanje (SAIP – School Achievement Indicators Program) pri 13- in 16-letnikih (Council of Ministers of Education of Canada, 1996). Dosežki v znanju – standardi dosežkov – s posameznega predmetnega področja so razvrščeni na pet ravni, ki predstavljajo kontinuum znanja in spretnosti, ki jih dosegajo učenci v osnovni in srednji šoli (Council of Ministers of Education of Canada, 2002).

V Avstraliji so postavili nacionalne profile ravni znanja za obdobje obveznega šolanja učencev. Preverjanja izvajajo v 3., 5. in 7. letu šolanja. Z zgodnjim preverjanjem znanja si pridobijo informacije, ki so potrebne za ukrepanje, še posebej pri učencih, ki ne dosegajo minimalno potrebnih standardov znanja. Merijo napredek v znanju glede na nacionalne profile ravni in poročajo o dosežkih (Ministerial Council on Education, Employment Training and Youth Affairs, 2003).

Pri raziskovanju dosežkov učencev pri posameznem predmetu v mednarodnih raziskavah (npr. TIMSS) v poglobljenih analizah proučujejo in ločujejo načrtovani kurikulum (*angl.* intended curriculum) in izvedbeni kurikulum (*angl.* implemented curriculum) od doseženega kurikuluma (Mullis, Martin, Gonzales in Chrostowski, 2004; Štraus, 2005). Pri tem že izrazi poudarjajo, da načrtovane vsebine in učni cilji pri določenem predmetu v praksi še ne pomenijo, da se vse to pri poučevanju tudi izvaja. Kaj od načrtovanega in s strani učitelja izvedenega pa se učenci naučijo, pa se brez preverjanja znanja in spretnosti ne more ugotoviti.

V raziskavi o znanju matematike in naravoslovja TIMSS so raziskovalci kot enega od možnih načinov uporabili opisovanje znanja na štirih mejnikih, tako da so učence po znanju razdelili v pet skupin (Mullis, Martin, Gonzales, Chrostowski, 2004). Po večletnih pripravah in izvedbi preverjanja znanja v 51 državah in deželah po svetu so s kompleksno obdelavo rezultatov preverjanja znanja prišli do standardov – kaj učenci znajo in so sposobni narediti. Poleg opisov dosežkov znanja na vsaki od štirih ravni v TIMSS-u empirično prikažejo primere konkretnih nalog iz preverjanja, ki jih učenci na določeni ravni dosežkov tipično rešijo pravilno oz. jih tisti, ki ne dosegajo določenih standardov dosežkov, ne dosežejo. Gre za odgovore na vprašanja, kako dobro je dovolj dobro za doseganje določenega mejnika znanja. To razvijejo z metodo sidrnih nalog na različnih točkah zahtevnosti v celotnem območju merjenja znanja za vsak predmet in vsako starostno skupino posebej. Preko mednarodne raziskave množica držav pride do standardov dosežkov za učne cilje, ki jih takšna raziskava preverja. TIMSS je preverjal več učnih ciljev matematike in naravoslovja za učence po štirih in osmih letih šolanja.

Slovenija se je ob prenovi vzgoje in izobraževanja zavezala, da bo svoj sistem in spremembe v njem vrednotila. V Izhodiščih za evalvacijo kurikularne prenove vzgoje in izobraževanja v Republiki Sloveniji (1999) je na prvem mestu med problematiko, na

katero se bodo na vsebinskem področju znanja vezale evalvacijske študije, navedena mednarodna primerljivost standardov znanja. Slovenija se je vključila v raziskavo TIMSS in je s tem pridobila bogato zbirko informacij o doseženih standardih za matematiko in naravoslovje v dveh starostnih obdobjih učencev v osnovni šoli. Poleg tega so zbrani podatki o dosežkih v znanju matematike in slovenščine pri nacionalnem preverjanju znanja (NPZ). Dosežki v znanju za t. i. tretje predmete pri NPZ za obdelavo v smeri standardov dosežkov niso bili uporabni, ker jih niso pisali vsi učenci oz. primeren vzorec učencev osnovne šole. Pri skupinskem preverjanju znanja ob koncu osemletne osnovne šole se je opravljalo preverjanje iz matematike in slovenščine, pa vendar v ta ocenjevanja niso bili vključeni vsi učenci oz. dogovorjen vzorec celotne populacije. Podobno velja za srednješolsko izobraževanje: rezultati pri slovenščini kot obveznem predmetu na splošni in poklicni maturi bi se lahko uporabili za obdelave, ki bi pripeljale do standardnih dosežkov, pri splošni maturi pa velja enako tudi za matematiko in podobno za prvi tuj jezik. Pri poklicni maturi pri tretjem predmetu kandidati izbirajo med tujim jezikom in matematiko. Pri vseh drugih predmetih splošne mature po opravljenem preverjanju ni nikakršnih podatkov o mejnem, tipičnem ali povprečnem dosežku in znanju slovenskega srednješolca. V skladu s teorijo o postavljanju standardov in ob poudarku, da pojem standard pomeni merljiv standard dejavnosti oz. dosežka (*angl.* performance standard), v Sloveniji ni postavljenih nacionalnih standardov znanja.

Ne samo, da v Sloveniji ne zbiramo podatkov v smeri, da bi lahko prišli do nacionalnih standardov dosežkov, ampak je bilo pri analizi učnih načrtov in katalogov znanja za osnovne in srednje šole ugotovljeno, da prihaja tudi do konceptualne zmede pri razumevanju pojma 'standard znanja' (Zupanc, 2005). Zmotno je mišljenje, da tisto, kar je o standardih zapisano in sprejeto z učnimi načrti in katalogi znanj, že uvršča Slovenijo med izobraževalne sisteme, kjer imamo standarde dosežkov v izobraževanju določene in lahko doseženo znanje posamezne generacije učencev ali dosežke v znanju in spretnostih posameznega učenca primerjamo z opredeljenimi nacionalnimi standardi dosežkov. Sistem tudi ni zastavljen tako, da bi lahko doseganje standardov znanja spremljali in vrednotili longitudinalno, med leti.

Države gredo skozi procese, ki spreminjajo razmere: liberalizacijo šolske politike in predpisov, kurikula, učbenikov in izobraževanja učiteljev ter usmerjanje k bolj centraliziranemu sistemu eksternega preverjanja z namenom najprej vzpostaviti in nato zaščititi nacionalne standarde (West in Crighton, 1999). Razvoj in uporaba nacionalnih standardov pri preverjanjih ob zaključkih šolanja neizogibno vodi do vzpostavitve nacionalnega ali osrednjega organa za nadzor nad standardi in priznavanjem (Crighton, 1993). Golo dejstvo je, da 'decentralizirani standardi' ne obsojajo, to je pojmovno protislovje. Določena stopnja nadzora in kontrole je neizogibna, če želimo zagotoviti kakovost.

Opredelitev učnih ciljev ('vsebinskih standardov') je (še) prvi korak v postopku merjenja znanja

Pri merjenju v družboslovju, psihologiji in edukaciji gre za gradnjo teoretičnih pojmov, konceptov, modelov tega, kar želimo meriti; kar se prepleta z izkustvom, opazovanjem in beleženjem. Empirična faza vpliva nazaj na potrditev, dopolnitev ali spremembe v konceptih in teoretičnih modelih. Med teoretičnim pojmom (konceptom), ki ga želimo meriti, in merskim postopkom (instrumentom) mora biti ustrezna usklajenost (Ferligoj, Leskošek in Kogovšek, 1995). Merjenje je proces, ki se začne z opredelitvijo teoretičnega pojma in nadaljuje z razvojem merskega postopka, s katerim je mogoče proučevani pojem operacionalizirati. Natančno je potrebno opredeliti predmet merjenja, to je konceptualizirati teoretični pojem, na drugi strani pa določiti merske metode, tehnike in postopke, s katerimi se bo mogoče temu teoretičnemu pojmu najlažje empirično približati (Bucik, 1997). V prvem koraku se z razjasnitvijo teoretičnega pojma (konceptualizacijo) opravi predvsem pregled različnih definicij pojma in izbor najustreznejše, če ni ustrezne definicije, pa izdelava le-te ter razgradnja pojma z določitvijo razsežnosti teoretičnega pojma (dimenzioniranjem) (Ferligoj, Leskošek in Kogovšek, 1995). V naslednjem koraku sledi operacionalizacija, kjer se za teoretične spremenljivke opredeli specifične operacije, merske instrumente ali postopke. Operacionalizacija povezuje teorijo z empiričnim merjenjem, z realnostjo, ki jo lahko opazujemo. Naslednji korak je izvajanje konkretnega merjenja, ko se dobi izmerjene spremenljivke. Sledijo analize empirično dobljenih meritev, popravki, razreševanje problemov z merskimi napakami, zanesljivostjo in veljavnostjo merjenja idr.

Ko se v edukaciji za učne vsebine pri vsakem predmetu razmišlja o preverjanju znanja in spretnosti, tj. o merjenju, je tudi potrebno prehoditi vse faze. Z natančno opredelitvijo učnih ciljev oz. vsebinskih standardov (»kaj«) je postavljen še le teoretični konstrukt – dimenzioniranje oz. razsežnost pojma znanja in spretnosti. Samo z zapisi, kaj naj bi učenec znal in bil sposoben narediti, v učnih načrtih in katalogih znanj, brez konkretnih nalog za preverjanje, postopkov in metod, ni opravljen niti drugi korak – operacionalizacija, da bi se preverjanje znanja obravnavalo kot merjenje. Od tu naprej je za postavitve standardov dosežkov (standardov znanja) še dolga pot, kjer se empiričnemu delu, tj. preverjanju znanja kandidatov s konkretnimi nalogami, ne da izogniti.

Nunnally in Bernstein (1994) merjenje opredelita s pravili pri določanju simbolov objektom s pripisovanjem številčnih vrednosti za določene lastnosti (lestvičenje, *angl. scaling*) ali z opredelitvijo, če se objekti razvrščajo v isto ali različne kategorije glede na določeno lastnost (razvrščanje, *angl. classification*). Pri tem je za merjenje lastnosti (*angl. attributes*) pomembno, da imajo objekti določene značilnosti (*angl. features*). Meri se lastnosti pri objektih, tj. oseb, ne pa oseb. Različne osebe morajo za merjenje določene lastnosti imeti različne značilnosti, da se le-te lahko razvrsti ali oštevilči v določeni merski lestvici.

Merske lestvice (nominalno, ordinalno, intervalno in razmernostno) lahko v zapisanem vrstnem redu razumemo nivojsko (Nunnally in Bernstein, 1994). Z izmerjenimi vrednostmi v lestvici, ki je uvrščena na višji merski nivo, so dovoljene zahtevnejše temeljne operacije (primerjanje, razvrščanje po velikosti, računanje razlik, razmerij ipd.), pretvarjanja in statistike. Če obravnavamo fenomen, ki ga lahko opisujemo z nominalno lestvico, gre samo za poimenovanje stvari, dogodkov ali pojmov in klasifikacijo le-teh na osnovi nekih lastnosti. Ne moremo opredeljevati, kaj je večje ali manjše, boljše ali slabše. Lestvica, ki omogoča vsaj rangiranje (večje – manjše, boljše – slabše ipd.), razlike med rangi pa niso določene, je ordinalna lestvica. Šele intervalna lestvica omogoča izražanje velikosti in enakosti intervalov ter računanje razlik med vrednostmi.

V edukaciji poskušamo meriti znanje in spretnosti učencev oz. doseganje opredeljenih učnih ciljev. Če naj bo preverjanje znanja merjenje, se je potrebno zavedati, da je nujno potrebno različne dosežke v znanju učencev razlikovati in jih enoznačno razvrstiti v določene kategorije znanja (standardne dosežke) ali določiti številčne vrednosti glede na doseženo znanje. V kateri od naštetih štirih merskih lestvic bi lahko merili znanje? Opisane učne cilje v učnih načrtih in katalogih znanj, pa četudi jih imenujemo (vsebinski) standardi in jih morda celo opredelimo za manj in bolj zahtevne, minimalne, temeljne, višje ipd., brez empiričnega merjenja znanja težko prikazujemo v merski lestvici, ki bi bila višje kot nominalna. Preverjanje znanja učencev naj bi potekalo v skladu z zapisanimi cilji (standardi) v učnih načrtih. Znanje geografije se pri učencu ob koncu osnovne šole preverja tako, da na primer »s pomočjo ustrezne karte opredeli lego Slovenije«, ali da »ovrednoti lego Slovenije z vidika odprtosti in prehodnosti v prometnem in gospodarskem smislu«, ali da »opiše lego Slovenije na stičišču Alp, Dinarskega gorstva, Panonske kotline in Jadranskega morja« (Ministrstvo za šolstvo in šport, 1998a). Pri slovenščini »učenec po poslušanju dramskega besedila z imenovanjem prvin zvočne opreme zvočnega posnetka pove, kako je to doživel« ali »po ogledu in branju prikaže doživljanje dramskega besedila, kar pokaže tako, da pisno primerja dramsko besedilo in uprizoritev« (Ministrstvo za šolstvo in šport, 1998b). Učitelji na osnovi množice takšnih zapisov v učnih načrtih lahko večrazsežnostni pojem – znanje neke učne snovi – dimenzionirajo in pristopijo k operacionalizaciji, tj. pripravijo naloge in načine preverjanja. Ko učenci pri opravljenem preverjanju prikažejo različne konkretne dosežke oz. izdelke, jih najverjetneje učitelji na osnovi takšnih zapisov v učnih načrtih težko enoznačno rangirajo na ordinalni merski lestvici. Kakšen opis lege Slovenije ali katero doživljanje dramskega besedila, ki je izraženo s pisno primerjavo prebranega in uprizorjenega besedila, je boljše in katero slabše? O tem takšni zapisi v učnih načrtih ne povedo nič. Takšni zapisi naj bi zagotavljali, da bi za vse šole, učitelje določenega predmeta in učence veljalo enako (*kaj se (na)učijo*), v učnih načrtih pa ni odgovorov na vprašanja o rangiranju dosežkov (na ordinalni lestvici) in na vprašanja: »Kako dobro je dovolj dobro za...?« Učitelj je bolj ali manj prepuščen samemu sebi, da si ustvari svojo »mersko lestvico«. Namen standardov znanja, še posebej v razsežnosti nacionalnih standardov dosežkov, ni dosežen.

Pri postavitvi standardov znanja v edukaciji se ne moremo izogniti nekaterim ključnim aktivnostim, ki so značilna za vsako merjenje. Postopek izdelave psihološkega testa kot merskega instrumenta se lahko strne v šest aktivnosti (Bucik, 1997): (i) opredelitev testa, (ii) izbiro vrste in ravni lestvice, (iii) sestavo testnih nalog in zbiranje podatkov za delovno verzijo testa, (iv) analizo reševanih testnih nalog, (v) na podlagi opravljenih analiz popravek testa in (vi) izdajo testa. Pred izdajo končne verzije testa je potrebno testiranje izvesti na reprezentativnem vzorcu posameznikov in v tem delu govorimo o določanju standardov.

Količine, merske enote in standardi v fiziki

Cresswell (2000) se zaveda kočljivosti termina 'standardi' v edukaciji in primerja izpitne in izobraževalne standarde s standardnimi merami v fiziki.

Zgodovina fizike je tudi zgodovina merjenja in razprav o fizikalnih količinah in enoznačnih definicijah le-teh. Včasih so imeli v vsakem okrožju, na razdaljah nekaj deset ali sto kilometrov, drugačne enote (mere) za iste količine (dolžino, maso, prostornino). Mednarodni sistem merskih enot (SI), ki ga uporabljajo skoraj vse države sveta, je vpeljal osnovne fizikalne *količine* in njim pripadajoče osnovne *merske enote* (Bueche, 1986). Enake merske enote omogočajo enostavno sporazumevanje med ljudmi po vsem svetu. To velja za vsa področja, od znanosti in tehnike do trgovine. Z državami, ki uporabljajo druge sisteme merskih enot, moramo pri prenašanju podatkov poznati pretvornike med enotami.

Zanimivo je, da se samo iz sedmih fizikalnih količin in njim pripadajočih merskih enot lahko izpelje vse ostale. To je razlog, da se tem sedmim reče osnovne fizikalne količine, ki imajo svoje osnovne merske enote. To so: dolžina z mersko enoto meter [m], čas z enoto sekunda [s], masa z enoto kilogram [kg], temperatura z enoto kelvin [K], električni tok z enoto amper [A], množina snovi z enoto mol [mol] in fizikalna količina svetilnost z mersko enoto kandela [cd]. Merske lestvice fizikalnih količin so vsaj intervalne ali celo razmernostne, kjer ima vrednost nič svoj absolutni pomen. V izrazu 'merska enota' v fiziki ni poudarek na enoti kot enici in merski lestvici kot taki, ampak na dogovoru in postopku za enoznačno merjenje npr. metra, kilograma, sekunde itd. Z mersko enoto je enoznačno opredeljena fizikalna količina, tako da navedba fizikalne količine ni več nujna. Če izmerimo neko vrednost in jo zapišemo z mersko enoto, npr. 1,75 [m], 73 [kg] in 45 [s], je nesporno, da gre v prvem primeru za fizikalno količino dolžino, v drugem za maso in v tretjem za čas. Pomembno je natančno in dobro opredeliti in definirati vsako fizikalno količino, da je vsebina in lastnost dobro opisana ter lahko natančno merljiva. Pri opredelitvi fizikalne količine ali vpeljavi nove je potrebno znati odgovoriti na vprašanja vrste »Kaj?«.

Ko se ukvarjamo z merljivostjo fizikalne količine, mislimo na njeno mersko enoto. Ko poročamo o meritvi, poročamo o obeh skupaj. Izmerjeno vrednost katerekoli količine moramo vedno zapisati z mersko enoto. Za vsako mersko enoto, še posebej

pomembno je to za osnovnih sedem, je opredeljen in definiran poseben (primarni) standard. Pri opredelitvi standardov za merske enote fizikalnih količin je potrebno znati ponovljivo izmeriti in odgovoriti na vprašanja vrste »Koliko?«.

Za merjenje časa so v zgodovini znanosti vedno izbrali periodičen pojav: vrtenje Zemlje, kroženje Lune ali vrtenje Zemlje okrog lastne osi (Strnad, 1985). Že stari antični narodi so za merjenje časa vpeljali merske enote: leto, mesec, dan. Za osnovno časovno mersko enoto je bila kasneje izbrana sekunda, kot 86.400-ti del povprečnega sončnega dne. Merska enota za čas, standard za 1 [s], je sedaj določen z 9.192.631.770 nihajnimi časi sevanja enobarvne svetlobe atomov cezija (${}_{55}\text{Cs}^{133}$).

Za dolžino so včasih ljudje uporabljali različne merske enote: ura hoda, čevljev, palec ... Prvi pravi mednarodni standard je leta 1790 določila Francoska akademija znanosti in sicer so za meter izbrali 1/10.000.000 dolžine pariškega poldnevnikarja od ekvatorja do pola. Izdelali so traverzo iz platine in iridija in vanjo vrezali dve črtici (Giancoli, 1991). Razdalja med njima pri temperaturi tališča ledu je predstavljala dolžino 1 [m] – prameter. Primarni standard je bil osnova za sekundarne, ki so bili shranjeni v večjih državah sveta. Od leta 1983 je 1 [m] definiran s potjo, ki jo svetloba prepotuje v vakuumu v 1/299.792.458 delu sekunde.

Za merjenje mase so v Franciji leta 1889 izdelali valj iz platine in iridija – prakilogram, ki še danes velja kot primarni standard (Zupanc, 1995). Kot izhodišče so vzeli maso 1 litra vode pri 4 °C. Hkrati so izdelali več identičnih valjev iz platine in iridija, ki so shranjeni v različnih državah sveta in veljajo kot sekundarni standardi za merjenje mase.

Pri merjenju električnega toka se fizika osredotoči na učinke usmerjenega gibanja nosilcev električnega naboja (Breuer, 1993). V okolici vodnikov, po katerih teče električni tok, nastaja magnetno polje, ko teče električni tok skozi upornike, se zaradi njega sprošča toplota, v elektrolitih pa električni tok povzroča elektrolizo. Standard za električni tok 1 [A] je določen kot konstanten tok, ki teče po dveh zelo dolgih ravnih vzporednih vodnikih z majhnim presekom, ki sta v razmiku 1 [m], če delujeta vodnika drug na drugega s privlačno oziroma odbojno silo $2 \cdot 10^{-7}$ [N] na vsak meter dolžine.

Temperaturo se lahko vpelje kot količino, s katero je sorazmerna prostornina razredčenega plina ali s katero je sorazmerna prenesena toplota pri idealnem toplotnem stroju ali s katero je sorazmerna povprečna kinetična energija molekule v razredčenem plinu. Merjenje v kelvinih so vpeljali, ko so ugotovili, da se ničesar ne da ohladiti pod določeno temperaturo. Tej temperaturni točki se reče 0 [K] = -273,15 °C. Temperatura v kelvinih je določena s temperaturo trojne točke vode $T_{\text{trojne točke}} = 273,15$ [K]. Voda je pri temperaturi $T_{\text{trojne točke}}$ in pri tlaku $p_{\text{trojne točke}} = 6,1$ [mbar] v vseh treh agregatnih stanjih hkrati: plin, kapljevina in trdno.

Fizikalno količino priredimo lastnosti telesa ali posebnosti pri pojavu ali stanju telesa, ki ga želimo opazovati (Strnad, 1985). Za vsako količino je potrebno navesti postopek za merjenje in standard za mersko enoto. Zanimiva je zgodovina vsakega standarda, kako se je spreminjal z razvojem znanosti, tehnike, metod in natančnosti

merjenja. Že sv. Avguštin (Bartholomew, 2000) naj bi o vprašanju časa govoril: »Kaj je čas? Če nihče ne vpraša o tem, potem vem, če pa moram na vprašanje, kaj je to, nekomu odgovoriti, pa ne vem.« V podobni zadregi smo pri vsem znanih fizikalnih količinah: masi, električnem toku, temperaturi, toploti, energiji ... in v nič manjši zadregi nismo, če želimo definirati učne cilje oz. kurikularne standarde znanja. V fiziki opisovanje posamezne (fizikalne) količine prav tako ni (bila) preprosta stvar. Vsekakor pa količine ne enačimo s standardom za mersko enoto, s katerim jo izmerimo, npr. dolžine z metrom, mase s kilogramom, temperature s kelvinom.

Najbrž je v primerjavi s fiziko v edukaciji fenomene oz. količine težje definirati. Če v fiziki merimo vsaj na intervalni, če ne na razmernostni lestvici, je pri opredelitvah merjenja znanja pogosto težko zadostiti pogojem za uporabo ordinalne, kaj šele intervalne lestvice. Težko časovno absolutno vpeljemo mero za določeno količino, npr. znanje nečesa, ki je odvisno od stanja v družbeni(ih) sredini(ah) in se z razvojem družbe z generacijami spreminja. Pa vendarle bi morali operativne učne cilje v edukaciji poskušati opredeliti tako, da bi bili merljivi vsaj na ordinalnih lestvicah, in za njihovo merjenje pristopiti k izdelavi merskih instrumentov, ki vodijo do standardov znanja.

Učni cilji kot merljive količine, standardi dosežkov kot merske enote

Izraz preverjanje znanja velikokrat uporabljamo le kot stilistično različico merjenja znanja (Zupanc, 2004b). Iz opisa fizikalnih količin in merjenja le-teh v fiziki lahko prikažemo mnogo podobnosti z učnimi cilji in standardi znanja v vzgoji in izobraževanju. Za definicijo *količine* moramo specificirati pravilo ali postopek, kar se v naravoslovju imenuje 'operativna definicija' (*angl.* operational definition) (Giancoli, 1991). Podobno morajo biti operativni *učni cilji* za preverjanje (merjenje) znanja podrobno opisani tako, da so merljivi na lestvici, ki je višja od zgolj nominalne. Če se izraz standard znanja (že) uporablja za te namene, je najbolje uporabiti besedno zvezo 'vsebinski standard' oz. 'kurikularni standard' znanja (Linn in Gronlund, 2000).

Merjenje vsake fizikalne količine je v resnici *relativno* – je primerjanje izmerjene vrednosti z dogovorjenim standardom za določeno mersko enoto (Giancoli, 1991). Včasih je osnova za *dogovor* neka naravna lastnost, stalna vrednost ali pojav: standard za dolžino [m] je definiran z zvezo s hitrostjo svetlobe v vakuumu, ki je nespremenljiva; standard za čas [s] je določen z nihajnimi časi sevanja svetlobe atomov cezija ipd. V zgodovini fizike pa ni bilo vedno tako in še danes za nekatere količine veljajo dogovori, katerih glavna vrednost je, da vzdržijo primerljivost v času in/ali na različnih mestih. Primera takih dogovorov med osnovnimi merskimi enotami sta enota za električni tok – amper – in, morda za sodobni čas še najbolj »starinska«, enota za maso – (pra)kilogram. Pomembno dejstvo, ki je razpoznavno tudi iz standardov za merske enote v fiziki, je, da so lahko standardi določeni *posredno* – preko lastnosti

snovi, učinkov pojava ipd. Standard za dolžino je določen kot pot, ki jo prepotuje svetloba v vakuumu v določenem času. Standard za celzijevo temperaturno lestvico je bil določen ob upoštevanju dejstva, da je voda v različnih agregatnih stanjih, različna temperatur pa je bila določena z merjenjem raztezanja snovi pri segrevanju. Tudi v naravoslovnih znanostih so nekateri standardi za merske enote določeni oz. (še) veljajo zaradi *tradicije*. Standard za maso je najbrž dober primer preprostega in tradicionalnega standarda, tudi stopinje celzija se še uporabljajo za merjenje temperature, čeprav danes zaradi kasneje odkrite naravne zakonitosti v znanosti uporabljamo kelvine. Pri odločitvah za postavitev (novih) standardov igra pomembno vlogo zanesljiva izvedba meritve, ki je *ponovljiva* in natančna. Za katero lastnost, pojav ali učinek se odločimo, katero posredno meritev uporabimo za določitev standarda, je v naravoslovju odvisno od matematično eksaktnega zapisa posledic pojava. Zveze med fizikalnimi količinami zapišemo s *funkcijskimi zvezami* – matematičnimi formulami. Za merjenje dolžine ob znani stalni hitrosti svetlobe v vakuumu in določenem času je funkcijska zveza preprosta – dolžino poti izračunamo s produktom hitrosti in časa. Pri računanju sile med dvema vzporednima vodnikoma, po katerih teče električni tok, v fiziki poznamo funkcijsko zvezo med električnim tokom in silo med vodnikoma. Zaželeno je izbira takšnih posrednih pojavov, da so funkcijske zveze, ki te pojave opisujejo, matematično »enostavne«, najbolje *premosorazmerne*. Ker so merske lestvice v fiziki (vsaj) *intervalne*, lahko uporabljamo funkcijske zapise (npr. linearne funkcije), računamo razlike vrednosti, povprečja ipd. Pri merjenju znanja pa se je potrebno pri ordinalnih lestvicah zadovoljiti z razvrščanjem, računanjem rangov, median ipd. (Nunnally in Bernstein, 1994).

Tako kot ima vsaka fizikalna *količina* svojo *mersko enoto*, bi morali operativni

Tabela 1. *Učni cilji kot merljive količine, standardi dosežkov kot merske enote - analogije med učnimi cilji in fizikalnimi količinami ter standardi znanja in merskimi enotami*

fizika	edukacija
fizikalne količine	učni cilji oz. vsebinski ali kurikularni standardi
operativna definicija – podroben opis pravila ali postopka	operativni učni cilj – podroben in merljiv opis vsebine znanja ali spretnosti – cilja
merske enote	standardi dosežkov (dejavnosti)
relativno merjenje količine glede na standard	relativno merjenje dosežkov glede na standard
dogovori	dogovori
posredne definicije	posredne definicije
tradicija	tradicija
ponovljivost	ponovljivost
funkcijske zveze - (premo)sorazmernost	opisi stopenj, ravni, lestvice ocen
merjenje na intervalnih ali razmernostnih lestvicah	merjenje na ordinalnih in (morda) intervalnih lestvicah

učni cilji oz. vsebinski ali kurikularni standardi imeti svoje *standarde dosežkov* oz. dejavnosti.

Če preverjanje znanja in spretnosti obravnavamo kot merjenje znanja, so očitno meritve takšne, da se tisti, ki meri, pri interpretaciji rezultatov sklicuje na nekaj - najbolj pogosto na primerljivo skupino drugih učencev ali na vnaprej postavljene kriterije. V ameriški literaturi se uporabljajo izrazi 'norm-referenced measurement' ali 'criterion-referenced measurement' (Linn in Gronlund, 2000). Gre za *relativno* meritev, primerjanje dobljenih vrednosti znanja z dosežki drugih oz. drugimi dosežki. Čeprav se včasih za razliko od interpretacije rezultatov s sklicevanjem na norme interpretacija s sklicevanjem na kriterije omenja kot absolutna (Nacionalna komisija za uvajanje in spremljanje novosti in programov v vzgoji in izobraževanju, 2000), je kot absolutni mišljen nek apriorno postavljen kriterij. Pri vrednotenju znanja nekega učenca je tudi v tem primeru meritev relativna, ker pri interpretaciji rezultatov dosežek učenca primerjamo z nekim vnaprej postavljenim kriterijem.

Pri preverjanju znanja so že učni cilji postavljeni z *dogovorom*, prav tako je stopnja doseženih ciljev dogovor. V določenem času, na določeni stopnji se ljudje v družbi ustrezno dogovorijo glede standardov znanja pri doseganju učnih ciljev. Naravnih standardov znanja ni, tudi apriorni kriteriji morajo biti dogovorjeni. Dull (po Westu in Crightonu, 1999) je zabeležil trditev učenca, kako je včasih kakšen učitelj postavil »standard«: »Bog zna za 5, jaz znam za 4, učenci pa lahko znate največ za 3. Mogoče boste lahko tudi vi znali za 4, če boste znali to, kar vam bom jaz povedal.« Takšni časi postavljanja »standardov« so najbrž minili.

Pri preverjanju znanja in spretnosti je očitno, da ju ne moremo meriti neposredno. Ne moremo »pogledati v glave« učencev in izmeriti njihovega znanja. Merjenje znanja in/ali spretnosti vedno poteka *posredno*: preko neke dejavnosti učencev, reševanja nalog pri testih, odgovarjanja na zastavljena vprašanja, nastopov ipd. V državah z uglednimi in urejenimi šolskimi sistemi pri učnih ciljih na določeni stopnji izobraževanja in pri doseganju standardov znanja pomembno vlogo igra *tradicija*. Čeprav se časi spreminjajo in z njimi tudi vzgoja in izobraževanje, se zelo pogosto postavlja vprašanje zahtevnosti šole v časovni dimenziji in pogosto se pojavljajo vprašanja, kot so: »Kako je s *ponovljivostjo*?« in »Ali standardi padajo?« Anglija je primer, kjer ob razglasitvi rezultatov eksternih izpitov vsako leto potekajo javne diskusije o tem. A. Paterson (2003) razumevanje pojma standard ponazori z dialogom med učencem in ravnateljem v angleški šoli.

Učenec: »Ravnatelj, ali ste kdaj pomislili, da so vaši standardi morda malo 'za časom'?«

Ravnatelj: »Seveda so 'za časom'. Standardi so vedno 'za časom'. To je tisto, kar jih naredi za standarde.«

Ko preverjamo doseganje učnih ciljev, moramo odgovore učencev objektivno vrednotiti. Na voljo moramo imeti enoznačna pravila za presojanje in ocenjevanje

(Bucik, 1997). Boljše znanje pri učencih naj bi pri preverjanju pomenilo višji dosežek, rezultati pri izmerjenem znanju naj bi bili enoznačno razvrščeni v *ravni* ali *lestvice ocen*, morda celo v (so)razmerju z doseženimi učnimi cilji. Težko je zagotoviti funkcijski zapis oz. (premo)sorazmerje, pa vendarle z mejniki, opisi stopenj ali lestvico ocen lahko vsaj na *ordinalnih lestvicah* izrazimo različne ravni doseženega znanja.

Pri preverjanju znanja bi bilo bolje, če bi t. i. 'vsebinske standarde' oz. 'kurikularne standarde' poimenovali drugače, brez besede 'standard', ker gre za opisovanje nekega znanja oz. količine, ki jo merimo pri izpitih, in je to šele prvi korak na poti do merjenja znanja. V Sloveniji in njeni šolski tradiciji bi bil najbližji izraz 'merljivi oz. operativni učni cilji'. Za 'standarde' pri merjenju rezultatov vzgoje in izobraževanja pa naj se uporabljajo t. i. 'standardi dejavnosti' (*angl.* performance standards), ki so merljivi in jih lahko kvantificiramo ali razvrščamo (Nunnally in Bernstein, 1994). Na primer, učenec za določeno raven znanja (oceno) pri testu z 'določenimi' nalogami doseže 'določen' rezultat.

Merilni instrumenti

Čeprav se o standardih znanja veliko govori, ni enostavnega merilnega instrumenta, ki bi ga uporabili za šolo ali za otroka in potem enostavno z njega prebrali določeno vrednost (Bartholomew, 2000). Prav tako ni nobene 'naravne' enote za merjenje.

V fiziki se z definiranjem standarda oz. merske enote za določeno količino določi merilni postopek in gradi merilni instrument. Podobno se pri merjenju znanja in spretnosti s testnim ali drugim postopkom določa preverjanja znanja in gradi merilni instrument za merjenje doseganja učnih ciljev. Vsak merilni instrument je potrebno umeriti. Če je primarni standard neke vrste dogovor, se mora vse druge merilne instrumente, sekundarne in etalone, umeriti po primarnem oz. bolj zanesljivem. Standardi dosežkov pri merjenju znanja morajo dobro opredeliti kategorije ali vrednosti, ker predstavljajo neke vrste standardno enoto oz. mero pri preverjanju doseganja določenih učnih ciljev. Različni testi znanja morajo biti umerjeni s standardnimi dosežki, če se uporabijo v različnih časih oz. na različnih mestih. Tako v fiziki in psihologiji kakor tudi v edukaciji se je potrebno zavedati, v katerih pogojih je merilni instrument uporaben. Če pogoji niso izpolnjeni, meritve z merilnim instrumentom niso veljavne oz. uporabne. Obseg delovanja določenega instrumenta ima svoje meje. Kakor isti termometer ni uporaben za merjenje temperature pri zelo nizkih in zelo visokih temperaturah in v različnih drugih pogojih, tudi instrument za preverjanje določenega znanja običajno ni uporaben za preverjanje zelo različnih stopenj znanja ali za preverjanje v zelo različnih sredinah. Vsako merjenje s seboj nosi tudi napake merjenja. Vsak merilni instrument ima določene tolerance oz. odstopanja, kar je potrebno pri zanesljivosti merjenja upoštevati.

Razprava

Ali je oz. naj bo preverjanje znanja in spretnosti v šolstvu merjenje? Med nekaterimi strokovnjaki in praktiki s šolskega področja v Sloveniji je najbrž še dvom, če standarde znanja in standarde dosežkov sploh potrebujemo. Učiteljem preverjanje kot merjenje, ocenjevanje in vse, kar je povezano s standardi, pogosto ni najbolj všeč. Mihajlovska (2003) pravi, da »... učitelj sprašuje in ugotavlja le napredek pri svojih urah. Ocene niso primerljive, ker vsak učitelj podaja učno snov in ocenjuje znanje drugače, eni s srcem, veselo in bogato, drugi suhoparno, le s predpisanimi učnimi standardi« (str. 5).

Če že enkrat presežemo dilemo, ali je preverjanje in ocenjevanje znanja sestavni del poučevanja in učiteljevega dela, in prenehamo s stalnim odpiranjem vprašanja o tem, ali so standardi znanja v šolstvu potrebni ali ne, moramo vseeno ugotoviti, da zapisani učni cilji oz. vsebinski standardi o tem, kaj naj bi se v šoli (na)učili, še niso standardi (dosežkov).

Bolje bi bilo, če t. i. vsebinskim oz. kurikularnim standardom ne bi rekli 'standardi', ampak kako drugače, morda 'merljivi učni cilji'. Postopek opredeljevanja le-teh je bolj podoben določanju učnih ciljev, razvrščanju in dimenzioniranju večrazsežnostnega pojma znanja. V teh postopkih ni konkretnih testnih nalog, (še) ni opravljenih meritev znanja in spretnosti in ni podatkov, kako so kandidati reševali konkretne naloge, ni opredeljeno, kaj je boljše, kaj slabše. O standardih naj se govori šele na ravni merljive dejavnosti oz. dosežkov kandidatov po konkretnem empiričnem merjenju znanja. Četudi iz raznih razlogov kurikularni standardi v terminologiji ostanejo, naj se imenujejo kakorkoli že, ostaja dejstvo, da je določitev merljivih učnih ciljev, pa četudi se jih proglasi za kurikularne oz. vsebinske standarde, proces, ki s tem še ni zaključen. Z zamenjavo imena učnih ciljev v učnih načrtih za (vsebinske) standarde ni bilo narejeno (še) nič ključnega na poti merjenja znanja. Še vedno je potrebno priti do standardov dejavnosti oz. dosežkov. To v Sloveniji (še) ni bilo narejeno. Da bi bilo preverjanje znanja v edukaciji *merjenje* znanja, je potrebno še veliko postoriti.

Postavlja se tudi vprašanje, če sta dva procesa lahko ostro ločena. Ali je lahko določanje merljivih učnih ciljev oz. vsebinskih standardov neodvisno od določanja standardov dejavnosti oz. dosežkov? Najbrž ne. Pri določanju standardov dejavnosti oz. dosežkov se pri empiričnih ugotovitvah v praksi lahko pokaže marsikaj, tako da to lahko vpliva na odločitve o nekaterih konkretnih učnih ciljih. To je običajen proces pri graditvi vsakega merskega instrumenta, ko se je po že obdelanih empiričnih podatkih in analizah potrebno vrniti na začetek, preveriti izhodiščne opredelitve in narediti popravke in dopolnitve.

Literatura

- American Educational Research Association, American Psychological Association in National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Bartholomew, D. J. (2000). The measurement of standards. V H. Goldstein in A. Heath (ur.), *Educational Standards* (str. 121–149). London: The British Academy.
- Breuer, H. (1993). *Atlas klasične in moderne fizike* (J. Strnad, prevod in priredba) [*Atlas of classical and modern Physics*]. Ljubljana: Državna založba Slovenije. (originalno delo objavljeno 1987)
- Bucik, V. (1997). *Osnove psihološkega testiranja [Essentials of psychological testing]*. Ljubljana: Filozofska fakulteta.
- Bueche, F. J. (1986). *Introduction to physics for scientists and engineers*. New York: McGraw-Hill Book Company.
- Council of Ministers of Education of Canada (1996). *School Achievement Indicators Program. Science Assessment. Framework and Criteria*. Toronto, Ontario: Avtor.
- Council of Ministers of Education of Canada (CMEC) (2002). *SAIP Mathematics III Assessment 2001: The Public Report* (str. 1–137). Toronto, Ontario: Avtor.
- Cannell, J. J. (1987). *Nationally normed elementary achievement testing in America's public schools: How all 50 states are above the national average* (2nd ed.). Daniels, WV: Friends of Education.
- Cresswell, M. (2000). The Role of Public Examinations in Defining and Monitoring Standards. V H. Goldstein in A. Heath (ur.), *Educational Standards* (str. 69–104). London: The British Academy.
- Crighton, J. (1993). *Poland – National Standards in Education: the creation of a national assessment and examination service - a proposal*. Washington D.C.: World Bank.
- Ferligoj, A., Leskošek, K. in Kogovšek, T. (1995). *Zanesljivost in veljavnost merjenja [The measurement reliability and validity]*. Ljubljana: Fakulteta za družbene vede.
- Giancoli, D. C. (1991). *Physics – Principles with Applications*. London: Prentice-Hall International.
- Gipps, C. V. (2000). *Beyond Testing: Towards a Theory of Educational Assessment*. London – Washington, D.C.: The Falmer Press.
- Glass, G. V. (1978). Standards and criteria. *Journal of Educational Measurement*, 15, 237–261.
- Glossary of Educational Assessment, Examination and Testing Terms* (2003). Posneto 22.6.2003 s strani Open Society Institute, Budapest, Education Support Program: http://194.149.10.80/bin/common/course.pl?course_id=_15_1&frame=top
- Goals 2000: Educate America Act* (2000). Posneto 4.10.2004 s <http://www.ed.gov/legislation/GOALS2000/TheAct/index.html>
- Goldstein, H. in Heath, A. (ur.) (2000). *Educational Standards. Proceedings of The British Academy – 102*. London: The British Academy.
- Grgin, T. (1994). *Školska dokimologija: procenjivanje i mjerenje znanja* (2. izd.) [*School docimology: assessment and measurement of knowledge*]. Jastrebarsko: Naklada Slap.

- Hornby, A. S. (1975). *Oxford Advanced Learner's Dictionary of Current English*. London: Oxford University Press.
- International Center for Technology in Learning. (2004). *PALS guide - Glossary*. Posneto 4.10.2004 s <http://pals.sri.com/pals/guide/glossary.html>
- Koretz, D. (1988). Arriving in Lake Wobegon: Are standardized tests exaggerating achievement and distorting instruction? *American Educator*, 12(2), 8–15.
- Linn, R. L. in Gronlund, N. E. (2000). *Measurement and Assessment in teaching* (8. izd.). New York: Upper Saddle River, Merrill/Prentice Hall.
- Mager, R. (1984). *Preparing instructional objectives* (2nd ed.). Belmont, CA: David S. Lake.
- Medveš, Z. in Svetlik, I. (2002). O splošni izobrazbi v poklicnem in strokovnem izobraževanju [General education in vocational and professional schools]. *Sodobna pedagogika*, 53 (2), 182–191.
- Mihajlovska, J. (2003). Zgodba strahov in primerjanj; Nacionalni preizkusi znanja [The story about fear and comparison; National assessment]. *Šolski razgledi*, 54 (11), 5.
- Ministrstvo za šolstvo in šport (1998a). *Geografija [Geography]*. Posneto 13.10.2004 s http://www.mszs.si/slo/solstvo/razvoj_solstva/viprogrami/os/9letna/ucni_nacrti/pdf/geo.pdf
- Ministrstvo za šolstvo in šport (1998b). *Slovenščina [Slovenian language]*. Posneto 13.10.2004 s http://www.mszs.si/slo/solstvo/razvoj_solstva/viprogrami/os/9letna/ucni_nacrti/pdf/sl.pdf
- Morgan, D. L. in Perie, M. (2004). *Setting Standards in Education: Choosing the Best Method for Your Assessment and Population*. (Neobjavljeno gradivo). Prispevek na konferenci International Association for Educational Assessment - 30th Annual Conference, Assessment in the Service of Learning, junij 2004, Philadelphia, Pennsylvania, USA.
- Nacionalna komisija za uvajanje in spremljanje novosti in programov v vzgoji in izobraževanju. (2000). *Izhodišča za pripravo nacionalnih preizkusov znanja v devetletni osnovni šoli [Starting-point for national assessment in the nine-grade primary school]*. (2000). Ljubljana: Avtor.
- Nunnally, J. C. in Bernstein, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Mullis, I. V. S., Martin, M. O., Gonzales, E. J. in Chrostowski, S. J. (2004). *TIMSS 2003 International Mathematics Report*. International Association for the Evaluation of Educational Achievement – IEA. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.
- National Educational Goals Panel. (1991). *The national educational goals report, 1991: Building a nation of learners. Executive summary* (str. 22). Washington, DC: Avtor.
- National Examinations and their meaning for Quality Assessment, Quality Assessment in Primary and Secondary Education*. (2001). Report on the third meeting of JONAS project. Posneto 15.8.2003 s http://download.citogroep.nl/pub/e_cens/JONAS_reports/JONAS_Report_2001-3_Ljubljana.pdf
- Ministerial Council on Education, Employment Training and Youth Affairs. (2003). *National Report on Schooling in Australia: National Benchmark Results: Reading, Writing and Numeracy Years 3, 5 and 7 – Preliminary Paper 2003*. Carlton South: Avtor
- Oosterhof, A. (2001). *Classroom applications of Educational measurement* (3. izd.). New Jersey: Merrill Prentice Hall, Inc.
- Paterson, A. (2003). Standards. V K. Tattersall (ur.), *Setting the Standard. A century of*

- public examining by AQA and its parent boards* (str. 139–160). Manchester: The Assessment and Qualifications Alliance.
- Pavešić, P. (prev.) (2005). Kako lahko matematiki prispevamo k preduniverzitetnemu izobraževanju. *Obzornik za matematiko in fiziko*, 52 (3), 94–96.
- Standards & testing*. (2001). Posneto 4.10.2004 s strani Public Examination System, The World Bank Group: <http://www1.worldbank.org/education/exams/standards.asp#Australia>
- Statistics of Education: National Curriculum Assessment of 7, 11 and 14 year olds in England – 2000* (2001). Posneto 15.1.2002 s <http://www.dfes.gov.uk/statistics/DB/SBU/b0262/index/html>
- Strnad, J. (ur.). (1985). *Fizika – leksikon [Physics – lexicon]*. Ljubljana: Cankarjeva založba.
- Štraus, M. (2005). Izvedbeni in doseženi kurikulum za matematiko v višjih razredih osnovne šole med 1995 in 2003 [Implementational and achieved curriculum for mathematics in higher grades of primary school between 1995 and 2003]. *Šolsko polje*, XVI (3/4), 19–40.
- Taylor, A. R. (2002). Setting Curriculum and Performance Standards for Gains in Achievement and in the Meaning of Results. Prispevek na konferenci 28-th Conference of International Association for Educational Assessment, Reforming Educational Assessment to Meet Changing Needs, September 2002, Hong Kong SAR, Kitajska.
- NCES – National Center for Education Statistics. (2004) *The NAEP Mathematics Achievement Levels*. Posneto 13.10.2004 s <http://nces.ed.gov/nationsreportcard/mathematics/achieve.asp>
- The Nature Of Public Examinations* (2002). Posneto 4.10.2004 s strani Public Examination System, The World Bank Group: <http://www1.worldbank.org/education/exams/nature/>
- Zupanc, D. (1995). *Fizika – interaktivni učbenik [Physics – electronic 'textbook']* [CD]. Ljubljana: Gekko.
- Zupanc, D. (2004a). Nekateri dileme šolskega ocenjevanja v Sloveniji [Certain dilemmas of assessment in Slovenian schools]. *Sodobna pedagogika*, 55 (1), 92–111.
- Zupanc, D. (2004b). Funkcije preverjanja znanja, interpretacija rezultatov, poročanje o dosežkih [The functions of knowledge assessment, interpretation of the results, reporting on achievements]. *Sodobna pedagogika*, 55 (4), 148–167.
- Zupanc, D. (2005). Standardi znanja v slovenski šoli [Knowledge standards in the Slovenian schools]. *Sodobna pedagogika*, 56 (1), 124–147.
- West, R. in Crighton, J. (1999). Examination Reform in Central and Eastern Europe: issues and trends. *Assessment in Education*, 6 (2), 271–289.

Prispelo/Received: 23.11.2005
Sprejeto/Accepted: 10.02.2005