

KONSTRUKCIJA TESTA DEKLARATIVNOG I PROCEDURALNOG ZNANJA¹

Gordana Mišćević-Kadijević²

Učiteljski fakultet u Beogradu

U radu se prikazuje nastanak testa znanja, namenjenog desetogodišnjim učenicima, kojim se utvrđuje usvojenost deklarativnog i proceduralnog znanja o građi ljudskog tela u okviru nastave prirode i društva u četvrtom razredu osnovne škole. Finalna verzija testa imala je šesnaest zadataka, jasno izraženu faktorsku strukturu o čemu svedoči urađena konfirmatorna faktorska analiza, utvrđenu validnost, kao i dobre metrijske karakteristike u domenu pouzdanosti, diskriminativnosti i objektivnosti. Pored prikaza karakteristika celog testa i dva podtesta u domenu pouzdanosti, analizirani su i svi pojedinačni zadaci po pitanju diskriminativnosti. Prikazane metrijske karakteristike ukazuju da se radi o solidnom mernom instrumentu.

Ključne reči: deklarativno znanje, proceduralno znanje, pouzdanost, diskriminativnost, objektivnost.

PRISTUP PROBLEMU

Učenje zahteva sticanje različitih tipova znanja (Anderson et al. 2001; Kadijević, 2004) u čijoj se osnovi nalaze usvajanje deklarativnog i proceduralnog znanja i njihova koordinacija (Anderson, 1983). Merenje usvojenosti ova dva tipa znanja zahteva izradu odgovarajućih testova znanja. Veoma zahtevna metodologija izrade mernih instrumenata (npr. Asher & Deveraj, 2001; cf. Momirović, 1997) u praksi se najčešće svodi na ispitivanje validnosti, faktorske strukture (u ovom slučaju dvofak-

¹ Članak predstavlja rezultat rada na projektu »Promene u osnovnoškolskom obrazovanju: problemi, ciljevi i strategije« čiju realizaciju finansira Ministarstvo nauke R Srbije (2006-2010).

² ✉: gordana.miscevic@uf.bg.ac.rs

torske), pouzdanosti i diskriminativnosti (videti, recimo, članke referisane u bibliografskoj bazi podataka Ebsco dostupnoj posredstvom servisa Kobson). Stoga u ovom radu, koji se bavi merenjem deklarativnog i proceduralnog znanja o građi ljudskog tela u okviru nastave prirode i društva u četvrtom razredu osnovne škole, proučavamo upravo ove četiri karakteristike razvijenog mernog instrumenta.

Jedan deo sadržaja, u nastavi prirode i društva, odnosi se na građu ljudskog tela. Nastavna tema *Čovek* obuhvata osam celina i to: čulne organe, nervni sistem, sistem organa za kretanje, sistem organa za varenje, sistem organa za izlučivanje, sistem organa za disanje, sistem organa za krvotok i reproduktivne organe. Za svaku od navedenih celina formulisan je po jedan deklarativni i jedan proceduralni zadatak, od toga su dva zadatka integrativna, pitanjem u njima se zahteva znanje iz nekoliko od navedenih osam celina.

Da bi se mogla utvrditi usvojenost deklarativnog i proceduralnog znanja učenika četvrtog razreda, potrebno je konstruisati kvalitetan test znanja. U narednim odeljcima nastojaćemo da prikazemo nastanak jednog takvog instrumenta, kao i podatke dobijene njegovim korišćenjem inicijalno, pre obrade nastavne teme *čovek*, finalno — neposredno nakon obrade poslednje nastavne jedinice navedene teme, i naknadno — pet meseci nakon finalnog testiranja. Prilikom testiranja rukovodili smo se postojećim sugestijama o uvođenju ispitanika u test situaciju i uputstvima o popunjavanju testa (Bukvić, 1996). Uzorak ispitanika se, inicijalno i finalno, sastojao od 259 učenika iz 11 odeljenja četvrtog razreda dve beogradske osnovne škole, dok se prilikom naknadnog testiranja sastojao od 148 ispitanika. Smanjeni broj ispitanika tokom naknadnog testiranja bio je posledica prelaska učenika iz četvrtog u peti razred, promene boravišta, kao i izostanaka sa časa u kome je vršeno testiranje.

FAKTORSKA STRUKTURA TESTA

Od šesnaest zadataka, u okviru testa, osam se odnosi na deklarativno, a osam na proceduralno znanje. Konfirmatorna faktorska analiza realizovana je u programu EQS³. Testiran je dvo-faktorski model: zadaci deklarativnog znanja nasuprot zadacima proceduralnog znanja. Osnovne indekse fitovanja modela koji ukazuju na stepen slaganja analiziranih standardizovanih skorova na pojedinačnim zadacima sa testiranim dvo-faktorskim modelom daje tabela 1. Vrednosti ovih indeksa ukazuju na dobro fitovanje modela i opravdavaju podelu testa od 16 zadataka na dve grupe od po 8 zadataka koji se odnose na deklarativno, odnosno proceduralno znanje.

³ Program EQS je dostupan na Internetu na www.mvsoft.com (o korišćenju ovog programa videti npr. Schumacker & Lomah, 2004).

Tabela 1: Osnovni indeksi fitovanja modela za tri merenja znanja

Indeksi fitovanja modela	Inicijalno merenje	Finalno merenje	Naknadno merenje
$\chi^2 (df, p)$	120,597 (99; 0,069 ¹)	106,094 (99; 0,295 ¹)	113,149 (99; 0,157 ¹)
TLI ²	0,943	0,986	0,937
RMSEA ³	0,029	0,017	0,031

¹ Nije značajno; ² Od 0 (nema fitovanja) do 1 (perfektno fitovanje); vrednosti oko 0,95 ukazuju na dobro fitovanje modela; ³ Vrednosti manje od 0,05 ukazuju na dobro fitovanje modela.

Pitanja koja se odnose na deklarativna znanja

U ovom podtestu, u skladu sa teorijom Vigotskog, ispituje se prvi aspekt/nivo, a to su, kako ih on naziva manifestni sadržaji, tj. činjenice, podaci, informacije specifične za jedan domen znanja. Taj aspekt je blizak onim znanjima koja se u kognitivnoj psihologiji nazivaju deklarativna i/ili konceptualna znanja (Ivić, 1992). Navedeno mišljenje podržavaju i drugi istraživači (Anderson, 1982). Ovaj podtest se odnosi na činjenično i pojmovno znanje, u koje se ubrajaju osnovni elementi koje učenici moraju znati da bi bili upoznati sa disciplinom, kao i organizovanje celine znanja, pojmova, principa itd. (Anderson et al., 2001).

Podtest sadrži 8 pitanja u kojima se traži deklarativno znanje (odgovara se biranjem tačnog odgovora od ponuđenih, dopunom, zaokruživanjem onoga što pripada ili ne pripada grupi, obeležavanjem pravilnog redosleda kojim se odvija neki proces u telu, povezivanjem pojmova sa leve i desne strane, odnosno uklapanjem podređenih u nadređene pojmove). Navodimo primere nekih pitanja u kojima se proverava deklarativno znanje.

- Pitanja u kojima se traži izdvajanje, ili grupisanje podređenih u okviru nadređenog pojma:

1. Zaokruži ono što pripada organima za disanje :

dušnik, pretkomora, jednjak, bubna opna, dušnjačke cevčice, plućni mehurići.

15. Poveži na odgovarajući način reči koje se nalaze na levoj strani sa rečima koje se nalaze na desnoj strani:

kičmena moždina

jednjak

bubrezi

mozak

bešika

organi za varenje

nervni sistem

organi za izlučivanje

- Pitanja u kojima se traži obeležavanje pravilnog redosleda kojim se odvija neki proces u telu:

7. Obeleži brojevima od 1 do 4 kako nastaje novi život.

(Broj 1 je ono što se prvo dešava, a broj 4 ono što se poslednje dešava)

1

Spajaju se spermatozoid muškarca i jajna ćelija žene.

Beba počinje da raste u majci.

--

Oplođena jajna ćelija ulazi u matericu.

Žena se porađa i beba izlazi napolje.

Deklarativna znanja učenika merena su pomoću zadataka 1, 3, 5, 7, 8, 11, 14 i 15.

Pitanja koja se odnose na proceduralna znanja

U ovom podtestu, u skladu sa teorijom Vigotskog, ispituje se drugi aspekt/nivo, a to su, kako ih on naziva instrumentalni sadržaji, tj. komponente znanja koje se odnose na metode, tehnike, procedure, veštine. Taj aspekt je blizak znanjima koja se u kognitivnoj psihologiji nazivaju proceduralnim znanjima (Anderson, 1982; Ivić, 1992). Proceduralna znanja se definišu kao znanja *kako* nešto uraditi, koja u sebi uključuju znanja kriterijuma koji su potrebni da bi se različite procedure završile do kraja (Anderson et al., 2001). Ta znanja razmatraju procese. Ponekad je to fiksiran redosled koraka, a ponekad treba doneti odluku koji korak treba sledeći načiniti.

Podtest sadrži 8 pitanja u kojima se traži proceduralno znanje (znanje praktične primene određene procedure; praktična primena znanja). Sva pitanja u kojima se traži proceduralno znanje su otvorenog tipa, da bi se kod svakog učenika utvrdilo da li postoji procedura u opisu praktične primene znanja i da li je redosled koraka pravilno postavljen. Navodimo primere nekih pitanja u kojima se proverava proceduralno znanje.

6. Kako ćeš sebi pomoći ako su ti trun, ili trepavica upali u oko?

9. Desilo ti se da ti je potekla krv iz nosa. Šta sve treba da radiš?

Proceduralna znanja učenika su se merila pomoću zadataka 2, 4, 6, 9, 10, 12, 13 i 16.

Sva pitanja su se odnosila na sadržaje nastavnih celina u okviru nastavne teme *čovjek*.

METRIJSKE KARAKTERISTIKE

Validnost

Test znanja smo konstruisali u saradnji sa učiteljima odeljenja obuhvaćenim preliminarnim eksperimentalnim istraživanjem i školskim pedagogom. Instrument od 16 pitanja nastao je od prethodnog seta koji je sadržao 80 pitanja, iz kog su učitelji birali i modifikovali zahteve za koje su procenili da odgovaraju uzrastu učenika i da pokrivaju predviđene sadržaje. Dva asistenta (metodičari nastave prirode i društva) i autor kao treći, nezavisno su klasifikovali pitanja za učenike na ona u kojima se traži deklarativno znanje i na ona koja sadrže znanja o primeni procedure, proceduralna znanja. Izdvojeni su zadaci u kojima su se sva tri procenjivača složila oko klasifikacije. Nakon zadavanja nekoliko verzija testova, došli smo do konačne. Zadržan je po jedan zadatak koji se odnosi na deklarativno i po jedan za proceduralno znanje za svaku od proučavanih nastavnih celina. Kao i kod većine testova znanja, prilikom konstrukcije ovog testa primenjene su logička i sadržajna validacija. To znači da smo utvrđivali da li je test u skladu sa zahtevima nastavnog programa i sadržajima na koje se odnosi. Smatra se da je valjan ako meri znanje iz određene oblasti, a ne inteligenciju ili karakterne osobenosti crte ličnosti učenika, što je u slučaju ovog istraživanja postignuto kroz prethodno pomenutu višestruku validaciju.

Pouzdanost

Test je pouzdan ako bi nakon ponovne primene registrovao iste, ili veoma slične rezultate kod istih ispitanika. U našoj pedagoškoj literaturi pouzdanost testa se obično računa pomoću Spirmanove split-half tehnike, koja pouzdanost procenjuje na bazi pouzdanosti podtestova sa neparnim, odnosno parnim zadacima. Pri tome se zaboravlja da test možemo podeliti na dva podtesta na mnogo načina (ako test ima 10 zadataka takvih podela ima 252 zbog $\frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$), koji obično dovode do različite vrednosti za pouzdanost što nameće pitanje koju od njih bi trebalo uzeti. Takođe se zaboravlja da je u svom radu iz 1951. godine Kronbah pokazao da njegov koeficijent alfa predstavlja prosek Spirmanovih split-half pouzdanosti za sve moguće podele testa (Cronbach, 2004: 396). Stoga smo u ovom istraživanju pouzdanost izražavali preko široko korišćenog Kronbah alfa:

$$\left(\frac{n}{n-1} \right) \cdot \left(1 - \frac{\sum_{i=1}^n S_i^2}{S^2} \right)$$

pri čemu je n broj zadataka na testu, suma se odnosi na sumu varijansi skorova pojedinačnih zadataka, a imenilac razlomka u drugoj zagradi tiče se varijanse ukupnog skora na testu. Pouzdanost korišćenog testa i dva podtesta računatog za standardizovane skorove na pojedinačnim zadacima sadrži tabela 2.

Tabela 2: Pouzdanost (Kronbah alfa) testova i dela testova na tri merenja

Merenje	Svi zadaci	Zadaci za deklarativno znanje	Zadaci za proceduralno znanje
1. Inicijalno	0,74	0,70	0,66
2. Finalno	0,80	0,69	0,76
3. Naknadno	0,76	0,64	0,70

Imajući u vidu da se koeficijenti pouzdanosti ne manji od 0,7 obično smatraju zadovoljavajućim, korišćeni test i podtestovi imali su prihvatljivu pouzdanost (Nunnally, 1978; Fraenkel & Wallen, 1990). Niže pouzdanosti (0,64; 0,66 i 0,69) su posledica slabe odnosno visoke uspešnosti u rešavanju zadataka koju je **nemoguće** izbeći kod pedagoških eksperimenata poput našeg kada se koriste kriterijumski testovi znanja. Pre šezdeset godina Gutman je predložio šest mera donje granice pouzdanosti, od koje treća daje rezultat identičan Kronbahovoj alfi (Vehkalahti, 2000).

Diskriminativnost (osetljivost)

Osetljivost se postiže kada testom možemo utvrditi fine razlike kod ispitanika u onome što merimo. Naravno da je osetljiviji test koji ima više pitanja, ali bi ipak bilo nerealno učenicima četvrtog razreda osnovne škole zadati instrument sa više od 16 pitanja koliko je sadržao naš test, naročito kada se ima u vidu zahtevnost zadataka o proceduralnom znanju, kao i potreba balansiranog broja zadataka jedne i druge vrste znanja. Najpre ćemo izložiti analizu *zadataka* korišćenih u testu: konzistentnost (pouzdanost) pojedinačnih zadataka, njihov indeks težine, kao i indeks diskriminativnosti.

Konzistentnost (pouzdanost) pojedinačnih zadataka, merena korelacijom između standardizovanog skora određenog ajtema i prosečnog standardizovanog skora određenog podtesta znanja, prikazana je u tabeli 3. Za sve ajteme razmatrana korelacija nije bila manja od 0,3. Prihvatljivi su svi ajtemi za koje razmatrana korelacija nije manja od 0,2 (Ding et al., 2006). To znači da sve naše zadatke, nezavisno od vremena merenja znanja (inicijalno, finalno, ili naknadno), možemo smatrati konzistentnim (pouzdanim) test ajtemima.

Tabela 3: Konzistentnost (pouzdanost) zadataka u odnosu na ispitivano znanje i tri merenja znanja

Zadatak	Deklarativno znanje			Proceduralno znanje			Ukupno znanje		
	Ini	Fin	Nak	Ini	Fin	Nak	Ini	Fin	Nak
1.	0,56	0,47	0,34				0,49	0,38	0,30
2.				0,59	0,60	0,50	0,46	0,49	0,36
3.	0,56	0,59	0,60				0,46	0,50	0,50
4.				0,56	0,65	0,57	0,50	0,55	0,45
5.	0,52	0,60	0,63				0,41	0,51	0,60
6.				0,60	0,64	0,56	0,48	0,56	0,44
7.	0,70	0,66	0,64				0,57	0,51	0,60
8.	0,46	0,47	0,62				0,35	0,45	0,56
9.				0,56	0,61	0,55	0,44	0,52	0,39
10.				0,48	0,57	0,52	0,43	0,48	0,45
11.	0,63	0,61	0,50				0,48	0,50	0,46
12.				0,53	0,65	0,67	0,40	0,61	0,57
13.				0,55	0,63	0,55	0,42	0,56	0,47
14.	0,55	0,45	0,31				0,49	0,42	0,24
15.	0,57	0,62	0,53				0,49	0,50	0,49
16.				0,47	0,54	0,58	0,35	0,44	0,48

Indeks težine pojedinačnih zadataka. Pošto se maksimalni broj poena koji učenik može da postigne na pojedinačnom zadatku razlikuje od zadatka do zadatka (to je učinjeno u cilju maksimizacije objektivnosti kriterijuma ocenjivanja), kod svakog zadatka je sirovi skor podeljen maksimalnim brojem poena, što je rezultiralo skorovima na skali od 0 do 1 (skor 0,60 kod nekog zadatka pokazuje da je učenik tačno odgovorio na 60% postavljenih zahteva kod tog zadatka). Tabela 4 sadrži podatke o indeksu težine pojedinačnih zadataka na inicijalnom, finalnom i naknadnom merenju. Proseci indeksa težine iznosili su: 0,47 za inicijalno; 0,65 za finalno i 0,63 za naknadno merenje. Ovi proseci su za ova tri merenja iznosili 0,62; 0,88 i 0,84 za deklarativno znanje, i 0,31; 0,43 i 0,42 za proceduralno znanje. Indeks težine najvećeg broja zadataka trebalo bi da bude između 0,3 i 0,9 a samim tim i njihov prosek (Ding et al., 2006). To znači da je korišćeni test sačinjen od zadataka prihvatljivog stepena težine. Zadaci su se na sva tri merenja grupisali prema tipu znanja koje proveravaju: prvih osam zadataka (lakši zadaci) odnose se na deklarativno znanje, dok se preostalih osam zadataka (teži zadaci) odnose na proceduralno znanje. Konfirmatornu faktorsku analizu realizovali smo na standardizovanim skorovima individualnih zadataka. Razlog nije onemogućavanje da se faktorsko grupisanje zadataka ostvaruje prema njihovoj težini, a ne vrsti znanja koju treba da mere. To je učinjeno zbog bolje pouzdanosti varijabli⁴ dek1, pro1, tot1, dek2, pro2, tot2, dek3, pro3 i tot3 kada se njihove vrednosti zasnivaju na standardizovanim, a ne reskaliranim skorovima, kao i bolje diskriminativnosti pojedinačnih zadataka koju prikazujemo u narednom odeljku.

⁴ „dek1” se odnosi na inicijalni nivo deklarativnog znanja, „dek2” na finalni, a „dek3” na naknadni. Slično se odnosi i na varijable pro-proceduralno znanje i tot-totalno, zbir deklarativnog i proceduralnog znanja.

Tabela 4: Težina pojedinačnih zadataka na tri merenja znanja

Inicijalno merenje (pre realizacije eksperimenta)		Finalno merenje (po realizaciji eksperimenta)		Naknadno merenje (pet meseci nakon finalnog)	
Zadatak	%	Zadatak	%	Zadatak	%
14.	87	14.	97	14.	97
7.	83	7.	94	7.	92
1.	66	15.	90	1.	85
3.	62	8.	86	3.	85
5.	53	11.	85	15.	85
15.	52	3.	84	11.	79
11.	51	5.	84	5.	75
8.	44	1.	80	8.	73
12.	42	12.	52	12.	49
9.	36	4.	47	9.	48
13.	35	16.	44	4.	47
4.	31	13.	44	16.	46
16.	29	9.	44	13.	41
2.	28	6.	40	6.	40
6.	25	10.	37	10.	32
10.	24	2.	34	2.	31

Indeks diskriminativnosti pojedinačnih zadataka. Diskriminativnost ajtema prikazana je u tabeli 5. Diskriminativnost je određena korišćenjem 25%-25% analize (uzeta je u obzir razlika proseka skorova pojedinačnih zadataka izraženih na skali od 0 do 1 za 25% najboljih i 25% najslabijih učenika na celom testu, odnosno podtestovima) jer ona, kako ističu Ding i sar. umanjuje verovatnoću potcenjivanja vrednosti indeksa diskriminativnosti usled nerazmatranja nestabilnih postignuća srednjih 50% učenika (Ding et al., 2006). Iako jedan broj zadataka ima slabiju diskriminativnost — obično se smatra da ajtem u testu ima dobru diskriminativnost ako vrednost indeksa nije manja od 0,3 (Ding et al., 2006) — prosečna diskriminativnost zadataka može se smatrati zadovoljavajućom, pri čemu je nešto bolja diskriminativnost bila prisutna kod zadataka kojima je mereno deklarativno znanje. S obzirom na to da smo mi u ovom istraživanju koristili standardizovane a ne reskalirane skorove od 0 do 1, odredili smo indekse diskriminativnosti standardizovanih ajtema pri čemu je, slično prethodnom, uzeta u obzir razlika proseka standardizovanih skorova pojedinačnih zadataka za 25% najboljih i 25% najslabijih učenika na celom testu, odnosno podtestovima. Ove indekse prikazuje tabela 6. Pošto su vrednosti indeksa bile veće od 1 kod 89% slučajeva (85 od 96 prikazanih brojeva), diskriminativnost zadataka čija je uspešnost rešavanja izražena standardizovanim skorovima bila je veoma dobra, što ide u prilog korišćenju standardizovanih, a ne sirovih, ili na 0-1 reskaliranih skorova. Granica dobre diskriminativnosti od 1 bi za podtest od 8 ajtema mogla da bude oko 0,9; a za test sa 16 ajtema oko 0,6. Ove brojeve smo dobili realizacijom sledeće simulacije: (1) generisani su standardizovani skorovi za 16 zadataka; (2) određeni su proseci ovako generisanih skorova za 8, odnosno 16 zadataka; (3) određene su razlike proseka ukupnog rezultata na testu sa 8, odnosno 16 ajtema za 25% najboljih i 25% najslabijih učenika. Simulacija je ponovljena 20 puta i prosečne vrednosti ovih razlika su iznosile 0,91 (za podtest sa

8 ajtema) i 0,64 (za test sa 16 ajtema). Kada je u novoj simulaciji izračunata srednja vrednost razlika standardizovanih skorova na 8 zadataka za 25% najboljih i 25% najslabijih učenika na podtestu od 8 zadataka dobili smo 0,86; dok smo za 16 zadataka dobili 0,59; što ide u prilog svrsishodnosti pomenutih granica od 0,9 i 0,6.

Tabela 5: Diskriminativnost zadataka (skorovi reskalirani na 0-1) u odnosu na ispitivano znanje i tri merenja znanja

Zadatak	Deklarativno znanje			Proceduralno znanje			Ukupno znanje		
	Ini	Fin	Nak	Ini	Fin	Nak	Ini	Fin	Nak
1.	0,47	0,45	0,25				0,46	0,35	0,17
2.				0,20	0,20	0,19	0,13	0,19	0,12
3.	0,61	0,45	0,46				0,58	0,45	0,39
4.				0,35	0,40	0,34	0,28	0,36	0,30
5.	0,63	0,44	0,66				0,53	0,40	0,60
6.				0,22	0,29	0,27	0,16	0,26	0,21
7.	0,54	0,23	0,30				0,51	0,21	0,30
8.	0,74	0,36	0,66				0,60	0,40	0,64
9.				0,38	0,43	0,35	0,24	0,35	0,23
10.				0,16	0,26	0,22	0,15	0,23	0,19
11.	0,57	0,43	0,38				0,51	0,37	0,33
12.				0,33	0,45	0,45	0,19	0,39	0,35
13.				0,19	0,30	0,23	0,13	0,30	0,24
14.	0,23	0,08	0,05				0,25	0,07	0,03
15.	0,44	0,26	0,20				0,42	0,20	0,19
16.				0,27	0,30	0,36	0,14	0,22	0,30
Prosek	0,53	0,34	0,37	0,26	0,33	0,30	0,33	0,30	0,29

Tabela 6: Diskriminativnost zadataka (standardizovani skorovi) u odnosu na ispitivano znanje i tri merenja znanja

Zadatak	Deklarativno znanje			Proceduralno znanje			Ukupno znanje		
	Ini	Fin	Nak	Ini	Fin	Nak	Ini	Fin	Nak
1.	1,40	1,43	1,02				1,33	1,03	0,68
2.				1,41	1,51	1,36	1,14	1,31	1,20
3.	1,34	1,37	1,32				1,24	1,25	1,05
4.				1,46	1,54	1,39	1,35	1,49	1,32
5.	1,50	1,27	1,43				1,12	1,11	1,32
6.				1,42	1,71	1,48	1,18	1,54	1,40
7.	1,64	1,22	1,29				1,38	1,07	1,22
8.	1,24	1,09	1,36				1,09	1,13	1,11
9.				1,31	1,60	1,36	1,10	1,44	1,06
10.				1,23	1,38	1,40	1,15	1,22	1,28
11.	1,72	1,32	1,43				1,28	1,15	0,94
12.				1,32	1,68	1,77	0,87	1,59	1,51
13.				1,29	1,59	1,48	0,95	1,48	1,43
14.	1,19	0,87	0,74				1,12	0,74	0,30
15.	1,43	1,33	0,94				1,18	1,02	0,90
16.				1,22	1,16	1,40	0,72	1,07	1,37
Prosek	1,43	1,24	1,19	1,33	1,52	1,46	1,14	1,23	1,13

Diskriminativnost testa i dva podtesta prikazana je u tabeli 7. Diskriminativnost je određena korišćenjem 25%-25% analize, pri čemu je računata razlika proseka srednjih vrednosti standardizovanih skorova za sve zadatke za 25% najboljih, kao i za 25% najslabijih učenika na celom testu, odnosno podtestovima. Korišćeni instrument ima dobru diskriminativnost što se i očekivalo s obzirom na prethodno utvrđenu dobru diskriminativnost pojedinačnih zadataka. Napominjemo da istraživači diskriminativnost testa obično zasnivaju na ishodu testa Kolmogorova i Smirnova, kojim proveravaju da li raspodela određene varijable značajno ili ne odstupa od normalne raspodele, pa ukoliko nije nađeno značajno odstupanje smatraju da je diskriminativnost varijable zadovoljena. Pristup je ispravan, ali nije jasno da li je diskriminativnost ipak prisutna kada raspodela značajno odstupa od normalne. Kako je u našem slučaju bilo prisutno odstupanje od normalne raspodele (videti tabelu 8), morali smo koristiti prethodnu analizu u cilju utemeljenog razjašnjavanja da li je i pri odstupanju od normalne raspodele diskriminativnost ipak postignuta. Dakle, normalnost raspodele varijable može nekad biti narušena, a njena diskriminativnost ipak očuvana.

Tabela 7: Diskriminativnost testa i podtestova za tri merenja

Merenje	Svi zadaci	Zadaci za deklarativno znanje	Zadaci za proceduralno znanje
1. Inicijalno	1,14	1,43	1,33
2. Finalno	1,23	1,24	1,52
3. Naknadno	1,13	1,19	1,46

Tabela 8: Rezultati testa Kolmogorov-Smirnova za osnovne varijable u istraživanju

Varijabla	N	M	S	K-S Z	p
1. Deklarativno znanje (inicijalno merenje)	259	0,00	0,57	1,311	0,065
2. Proceduralno znanje (inicijalno merenje)	259	0,00	0,54	0,883	0,417
3. Ukupno znanje (inicijalno merenje)	259	0,00	0,45	0,619	0,838
4. Deklarativno znanje (finalno merenje)	259	0,00	0,56	3,396	0,000
5. Proceduralno znanje (finalno merenje)	259	0,00	0,61	0,917	0,369
6. Ukupno znanje (finalno merenje)	259	0,00	0,50	1,485	0,024
7. Deklarativno znanje (naknadno merenje)	148	0,00	0,53	1,970	0,001
8. Proceduralno znanje (naknadno merenje)	148	0,00	0,57	0,842	0,478
9. Ukupno znanje (naknadno merenje)	148	0,00	0,47	1,049	0,221

Objektivnost

Objektivnost pokazuje identitet, sličnost skorova, ili dijagnoza koje stručni ocenjivači dobijaju na osnovu istih podataka (Rajston, 1966). Smatra se da se ova metrijska karakteristika testa postiže ukoliko njegovi rezultati zavise samo od onoga šta je mereno, a ne od onoga ko te rezultate sređuje i ocenjuje. Ako bi istraživač

ocenjavao test prema svom nahodaenju, ukoliko ne postoje čvrsti i jasni kriterijumi za bodovanje odgovora, objektivnost merenja bila bi narušena mnogobrojnim pristranim i proizvoljnim odlukama. Objektivnost primenjenog testa znanja proveravali smo utvrđivanjem stepena slaganja rezultata koji su dobili dva nezavisna ocenjivača, ocenjujući isti test. Prvi ocenjivač je bio autor ovog rada, a drugi postdiplomac na specijalističkim studijama. Na osnovu broja poena kojim su ocenjivači bodovali test utvrđen je Pirsonov koeficijent linearne korelacije i on je iznosio $r_{xy}=0,99$ na inicijalnom, finalnom i naknadnom testiranju, što potvrđuje visoku objektivnost primenjenog testa znanja.

ZAKLJUČAK

Na osnovu prikazanih rezultata smatramo da je dobijen solidan merni instrument. Svedoci smo da se čak u nekim međunarodnim istraživanjima prikazuju rezultati i izvode zaključci na osnovu mernih instrumenata kod kojih nije dostignuta faktorska struktura (koja je, na primer, lakše dostižna u merenju dimenzija stavova nego tipova znanja) ili je njihova pouzdanost ispod granice od 0,7 (kritikovano u, recimo, Kadijevich, 2006; Eklöf, 2007). Stoga se nadamo da ovaj rad predstavlja mali prilog značajnijem obraćanju pažnje na ozbiljnost problema konstrukcije odgovarajućih mernih instrumenata, posebno u istraživanjima učenika mlađeg osnovnoškolskog uzrasta.

LITERATURA

- Anderson, J. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), 369-406.
- Anderson, J. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, C. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Asher, S., & Deveraj, S. (2001). An empirical comparison of statistical construct validation approaches. *IEEE transactions of engineering management*, 48(3), 319-329.
- Bukvić, A. (1996). *Načela izrade psiholoških testova*. Beograd: Zavod za udžbenike i nastavna sredstva.
- Cronbach, L. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 64(3), 391-418.

- Ding, L. Chabay, R. Sherwood, B., & Beichner, R. (2006). Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. *Physical Review Special Topics - Physics Education Research*, 2. Retrieved, December 20, 2008 from <http://prst-per.aps.org/pdf/PRSTPER/v2/i1/e010105>
- Eklöf, H. (2007). Self-concept and valuing mathematics in TIMSS 2003: Scale structure and relation to performance in a Swedish Setting. *Scandinavian Journal of Educational Research*, 51, 297-313.
- Fraenkel, J., & Wallen, N. (1990). *How to design and evaluate research in education*. New York: McGraw- Hill.
- Ivić, I. (1992). Teorije mentalnog razvoja i problem ishoda obrazovanja. *Psihologija*, 25 (3-4), 7-37.
- Kadijević, Đ. (2004). Matematičko mišljenje: Opšta ili posebna pravila zaključivanja. U S. Milanović-Nahod i N. Šaranović-Božanović (Ur.), *Znanje i postignuće* (189-202). Beograd: Institut za pedagoška istraživanja.
- Kadijevich, Dj. (2006). Developing trustworthy TIMSS background measures: A case study on mathematics attitude. *The Teaching of Mathematics*, 9(2), 41-51.
- Momirović, K. (1997). O realnoj egzistenciji psiholoških konstrukata. U K. Momirović (Ur.), *Provera egzistencije psiholoških konstrukata*. Beograd: Institut za psihologiju i Laboratorija za eksperimentalnu psihologiju.
- Nunnally, J. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw Hill.
- Rajston, Dž. (1966). *Vrednovanje u savremenom obrazovanju*. Beograd: Vuk Karadžić.
- Schumacker, R., & Lomah, R. (2004). *A beginner's guide to structural equation modeling*. New Jersey: Lawrence Erlbaum.
- Vehkalahti, K. (2000). *Reliability of measurement scales*. Finnish Statistical Society: Statistical Research Reports 17. Retrieved, December 20, 2008 from <http://ethesis.helsinki.fi/julkaisut/val/tilas/vk/vehkalahti/reliabil.pdf>

ABSTRACT

**DEVELOPMENT OF DECLARATIVE AND PROCEDURAL
KNOWLEDGE TEST**

Gordana Miscevic-Kadijevic

Teacher training faculty, University of Belgrade, Serbia

The paper presents the development of a knowledge test, designed for ten-year old pupils, that measures the attainment of declarative and procedural knowledge about human body acquired through a fourth-grade course on nature and society. The final version comprising sixteen tasks, had desired factor structure evidenced by the applied confirmatory factor analysis as well as appropriate metric characteristics of reliability, discriminativity and objectivity. Apart from the metric characteristics of the whole test and its two knowledge subtests as regards reliability, this paper also reports the discriminativity of particular tasks. The reported characteristics evidence that a solid instrument has been developed.

Key words: *declarative and procedural knowledge, validity, reliability, discriminativity, objectivity.*