

KSENIJA BOSNAR
Filozofski fakultet
Zagreb

POSLEDICE PROMJENE U NAČINU VREDNOVANJA URATKA U TESTU NA NEKE METRIJSKE KARAKTERISTIKE TESTA HIJERARHIJSKOG SREĐIVANJA

SAŽETAK

U radu se prikazuju posljedice promjene načina vrednovanja uratka u zadatku na osjetljivost mjerenja testom i neke druge metrijske karakteristike testa, kada se uradak u testu mjerio skalom finijom od uobičajene (vrijednosti 0 – 1). Upotrebljena je skala 0 – 8 bodova, i to je dovelo do veće osjetljivosti i pouzdanosti u primijenjenom testu Hijerarhijskog sređivanja (Z. Bujas).

UVOD

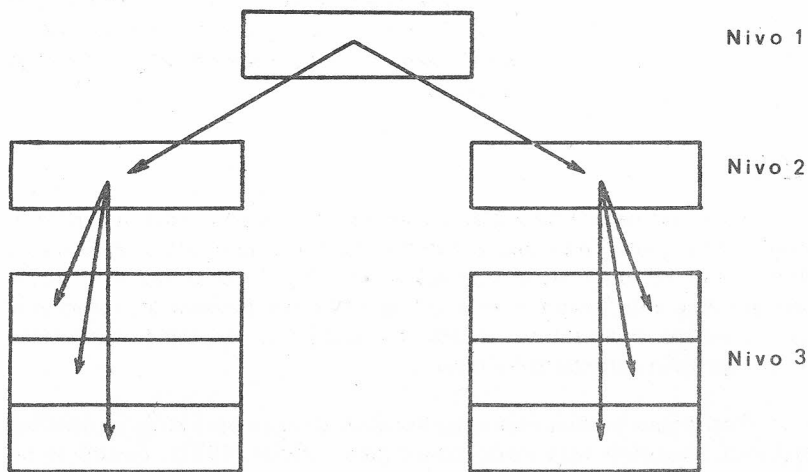
Jedan od načina poboljšanja osjetljivosti mjerenja testa, jest poboljšanje u osjetljivosti mjerenja zadatkom, tj. mjerenje uratka u zadatku na finijoj skali od uobičajene s vrijednostima 0 i 1. Ovaj rad bavi se, na primjeru mjerenja Testom hijerarhijskog sređivanja, posljedicama promjene načina vrednovanja uratka u zadatku na osjetljivost mjerenja testom i neke druge metrijske karakteristike testa.

Test hijerarhijskog sređivanja konstruiran je za ispitivanje „...sadržaja pojmova i vanjskih veza među pojmovima“ (Bujas, 1977). Sastoji se od dvadeset zadataka iste forme. U svakom zadatku ispitanik treba odabrati devet od dvanaest predloženih riječi—pojmova i ispravno ih smjestiti u grafičku shemu, što mu donosi jedan bod.

Test je primijenjen u veljači 1979. godine na učenicima osmih razreda osnovne škole u Kozarcu (SR B i H). Primijenjen je grupno na 325 ispitanika, s vremenom za rješavanje od 55 minuta. Test se pokazao neprimjeren, tj. suviše težak ispitanicima. Stupanj težinske primjerenosti testa, određen uz pomoć aritmetičke sredine indeksa lakoće svih zadataka u testu, jest $\bar{p} = 0.128$. Test slabo diferencira ispitanike u području brojčano niskih rezultata, tako da u rasponu od 0–4 boda nalazimo čak 265 rezultata, tj. preko 80% rezultata.

Međutim, ako pogledamo na koje su sve načine ispitanici odgovarali na zadatke u testu, vidjet ćemo da i unutar grupe ispitanika s istim brojem bodova, postoje razlike u uratku. Naime, 0 bodova u zadatku dobio je i ispitanik koji je osam (od devet) riječi–pojmova odabrao i smjestio ispravno, kao i ispitanik koji je cijeli zadatak postavio pogrešno. Tako su 0 bodova u testu dobili i oni ispitanici koji nisu uradili gotovo ništa, i oni koji su se u više navrata približili potpunom i ispravnom rješenju.

Da bismo ispravno riješili zadatak, potrebno je uočiti odnose među pojmovima, te položaje pojmova unutar tih odnosa. Među pojmovima nalazimo osam neposrednih relacija (između nivoa 1–2 i nivoa 2–3) i šest posrednih relacija (između nivoa 1–3), slika 18.



Slika 18 – Neposredne relacije između riječi–pojmova u zadatku THSa čije bi točno postavljanje donosilo po 1 bod (prikazane strelicama)

Ako za svaki ispravno postavljeni hijerarhijski odnos pretpostavimo da je korak bliže potpunom i ispravnom rješenju, broj ispravno uočenih pozicija i neposrednih relacija između riječi—pojмова mogli bismo smatrati stupnjem do kog je zadatak riješen. Ovim načinom ocjenjivanja zadatak se dijeli na osam dijelova potpunog rješenja. Za svaki ispravno postavljen neposredni hijerarhijski odnos, ispitaniku bi se mogao dati 1 bod, te bi u zadatku ispitanik mogao dobiti 0—8 bodova, što bi povećalo raspon mogućih rezultata u testu s 0—20 na 0—160 bodova u odnosu na klasični način ocjenjivanja testa.

Kako postoji mogućnost da se, uz novi način vrednovanja uratka, testovnim rezultatom ne mjeri ista olina kao i klasičnim, i kako novi način ocjenjivanja uratka u testu može promijeniti i veličinu pouzdanosti mjerenja testom, problem ove radnje možemo formulirati ovako:

1. Kako se promijenila osjetljivost mjerenja testom uz novi način ocjenjivanja uratka u odnosu na klasični način;

2. Kakva je korelacija između rezultata testa ocijenjenog na klasični i na novi način; na osnovi nje treba prosuditi da li će doći do promjene u valjanosti mjerenja testom s novim načinom ocjenjivanja uratka u testu;

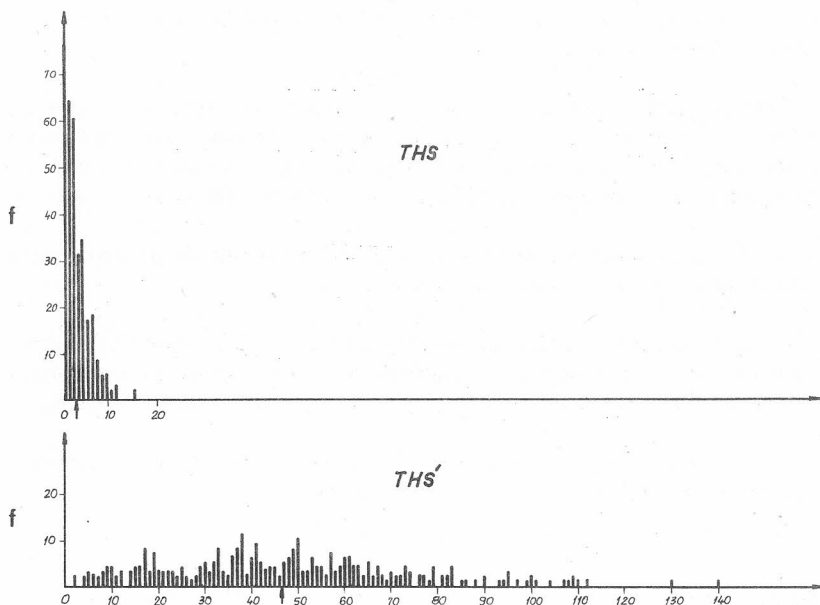
3. Da li je došlo do promjene u pouzdanosti mjerenja testom u funkciji novog načina ocjenjivanja uratka u testu?

REZULTATI

Na slici 19 prikazana je distribucija frekvencija rezultata Testa hijerarhijskog sređivanja vrednovanog na klasični način (THS), i distribucije rezultata testa ocijenjenog na novi način s 0—8 bodova u zadatku (THS'), obje s intervalom grupiranja od jednog boda.

Distribucija bruto rezultata testa ocijenjenog s 0—1 bodova u zadatku, zrcalnog je J oblika, s totalnim rasponom od 0—15 bodova. Nasuprot tome, distribucija rezultata testa ocijenjenog na novi način, je izrazito platičurtna, s više od osam puta većim totalnim rasponom rezultata od 2—140 bodova. Distribucija klasično vrednovanog rezultata pozitivno je asimetrična s aritmetičkom sredinom $M_{THS} = 2.6$. Distribucija rezultata testa ocijenjenog s 0—8 bodova u zadatku, također, je

pozitivno asimetrična, s aritmetičkom sredinom $M_{THS'} = 46.3$, što nam govori da je Test hijerarhijskog sređivanja, za ovu grupu isitanika, isuviše težak, bilo da je ocijenjen na klasični način, s 0–1 bodova u zadatku, ili ocijenjen na novi način, s 0–8 bodova u zadatku.



Slika 19 — Distribucija individualnih ukupnih rezultata (apscisa: ukupni broj bodova, ordinata: frekvencija ispitanika).
THS: distribucija rezultata THSa ocijenjenog na klasični način (1–0);
THS' distribucija rezultata THSa ocijenjenog od 0–8 bodova u zadatku.
Interval grupiranja = 1 bod. Strelice na apscisi označavaju položaje aritmetičkih sredina.

Mada se test, uz oba načina vrednovanja rezultata, u ovom ispitivanju pokazao preteškim, osjetljivost mjerenja, za dva način vrednovanja uratka u zadatku, bitno je različita.

Kako se raspon mogućih bruto rezultata za ova dva načina vrednovanja uratka vrlo razlikuje, te zbog asimetričnih distribucija rezultata, standardna devijacija i koeficijent relativnog varijabiliteta ne mogu nam poslužiti kao pokazatelji osjetljivosti mjerenja. O većoj osjetljivosti mjerenja testom vrednovanim s 0–8 bodova u zadatku, govori nam broj različitih opaženih bruto rezultata $BROR = 97$, koji je preko sedam puta veći od onog za klasično ocijenjen test $BROR = 13$. Indeks osjetljivosti mjerenja ($I.O. = (BROR - 1) / (N - 1)$, N = broj ispitanika) za klasični rezultat jest $I.O. = 0.037$. Za test ocijenjen s 0–8 bodova u zadatku, on je osam puta veći i iznosi $I.O. = 0.296$.

Da je test ocijenjen s 0–8 bodova u zadatku osjetljiviji, još je očiglednije iz relativnih frekvencija ispitanika po pojedinom opaženom bruto rezultatu ($N/BROR$). Za klasično ocijenjen test relativna frekvencija iznosi 25, dok je za novo ocijenjen test $N/BROR = 3.35$, okruglo tri ispitanika po jednom različitom opaženom bruto rezultatu.

Tabela 18

	M	σ	σ/M	BROR	I.O.	N/BROR
THS	2.554	2.612	1.023	13	0.037	25.
THS'	46.305	25.289	0.564	97	0.296	3.351

Aritmetičke sredine (M), standardne devijacije (σ), koeficijenti relativnog varijabiliteta (σ/M), broj različitih opaženih bruto rezultata (BROR), indeksi osjetljivosti mjerenja (I.O.) i relativne frekvencije ispitanika po pojedinom opaženom različitom rezultatu (N/BROR) za Test hijerarhijskog sređivanja ocijenjen s 0–1 bodova u zadatku (THS) i test ocijenjen s 0–8 bodova u zadatku (THS'). Vrijednosti su zaokružene na trećoj decimali.

Korelacija između rezultata testa vrednovanih s 0–1 u zadatku, i rezultata vrednovanih s 0–8 bodova u zadatku, iznosi $r_{xy} = 0.883$ (Pearsonov koeficijent produkt–momenta). Ta dosta visoka korelacija upućuje na to, da se predmeti mjerenja testom vrednovanim na dva načina, vjerojatno bitno ne razlikuju.

Uz novi način ocjenjivanja testa, s 0–8 bodova u zadatku, dobijen je veći koeficijent pouzdanosti mjerenja $r_{xx} = 0.780$ (računan po formuli za Cronbachov alfa–koeficijent nutarnje konzistencije) od onog za klasično vrednovane rezultate $r_{xx} = 0.714$ (računan po Kuder–Richardsonovoj formuli br. 20). Povećanu pouzdanost mjerenja testom, ocijenjenim na novi način, prati i smanjena relativna pogreška mjerenja, smanjio se nesistematski varijabilitet u rezultatima, te testom ocijenjenim s 0–8 bodova u zadatku, možemo razlikovati više rezultata, uz istu vjerojatnost da ta razlika nije uvjetovana slučajem, nego testom ocijenjenim na klasični način.

Tabela 19

	r_{xx}	Se/σ	Se
THS	0,714	0,535	1,397
THS'	0,780	0,469	11,860

Koeficijenti nutarnje konzistencije (r_{xx}), relativne pogreške mjerenja (Se/σ) i apsolutne standardne pogreške mjerenja (Se) za klasično ocijenjen uradak u testu s 0–1 bodova u zadatku (THS) i novo ocijenjen uradak s 0–8 bodova u zadatku (THS').

ZAKLJUČAK

Test hijerarhijskog sređivanja ocijenjen s 0–8 bodova u zadatku, je osjetljiviji i pouzdaniji od klasično ocijenjenog testa s 0–1 bodova u zadatku. Valjanosti testa, s rezultatima vrednovanim na klasični način i rezultatima ocijenjenim na novi način, vjerojatno se značajno ne razlikuju, no, da bi se novi način vrednovanja uratka u testu mogao bez rezerve prihvatiti kao zamjena za klasični način ocjenjivanja testa, tu bi pretpostavku trebalo provjeriti daljnjim istraživanjima.

LITERATURA

Bujas Zoran, (1977)

„Neki novi” testovi za ispitivanje intelektualne razvijenosti”, „Stručni skup psihologa „Dani Ramira Bujasa” 1976.”, Društvo psihologa Hrvatske, Zagreb 1977.