

Ali Bujasov "problemski test" sodi med staro šaro?

Pregled nekaterih metrijskih lastnosti testa za določevanje stopnje občutljivosti za probleme

VALENTIN BUCIK, KLAS BRENK*

POVZETEK

Pred 24 leti je bil v Zagrebu razvit Problemski test - test za preizkušanje sposobnosti lahkega in hitrega odkrivanja neuravnoteženosti v določenih situacijah, na kratko test občutljivosti za probleme. Test je še vedno pogosto uporabljan, v zadnjih letih pa smo pogrešali preverjanje njegovih metrijskih lastnosti. Poskušali smo analizirati testne naloge, določiti diskriminativnost in zanesljivost testa. Ugotovili smo, da je Pb-test še vedno zelo pripraven instrument za merjenje občutljivosti za probleme, ki je ena temeljnih lastnosti mišljenja v problemskih situacijah in ki je tesno povezana z G-faktorjem inteligentnosti. Test je ustrezno težak, težavnost je enakomerno porazdeljena in narašča linearno proti koncu testa. V povprečju so preizkušanci dosegali polovični uspeh, distribucija individualnih dosežkov pa je blizu normalni. Več vrst koeficientov zanesljivosti kaže dovolj visoko stopnjo zanesljivosti in nudi garancijo za uporabnost Pb-testa tudi v prihodnje pri testiranju ljudi različnih starosti, spolov in izobrazbenih ravni ter struktur.

ABSTRACT

SHOULD WE THROW THE BUJAS' "PS-TEST" INTO A GARBAGE CAN?

(The psychometric analysis of the "Problem-solving" test)

24 years ago, a "Problem-solving" test was developed at Department of psychology in Zagreb. This is the test with which the ability of easy and quick detection of unbalanced situations in problem solving can be measured. The application of the test is still frequent, but there is lack of reliability and validity analyses in the last few years. The item-analysis is reported in this article, and also the discriminant validity, split-half reliability and internal consistency of the test are checked. It is concluded that PS-test is still a very applicable instrument for measuring the sensitivity of problem detection, which is one of the essential cognitive characteristics, and which is highly correlated with G-factor of intelligence. The test is difficult to an appropriate level, and the difficulty of items arises from the beginning to the end of test and the distribution

* - asist. mag. Valentin BUCIK, izr. prof. dr. Klas BRENK,
Univerza v Ljubljani, Filozofska Fakulteta Oddelek za psihologijo Katedra za psihološko
metodologijo Aškerčeva 12 61000 LJUBLJANA

of the test scores tends to be normal. The reliability coefficients show that PS-test is a good measure of intelligence for different groups of subjects regarding gender, age, education level and occupation.

V procesu raziskovanja teorij inteligentnosti in konstrukcije novih merskih instrumentov za merjenje različnih vidikov inteligentnosti je **Z. Bujas** pred več kot dvajsetimi leti v Zagrebu s sodelavci razvil prvo obliko t.i. "Problemskega testa". To naj bi bil nov instrument za preučevanje ene od temeljnih lastnosti razmišljanja v problemskih situacijah - **občutljivosti za probleme**, ki jo klasični testi inteligentnosti nehoti zanemarjajo, bodisi zaradi možnosti, da preizkušanec rešuje eno nalogo s pomočjo pri drugi nalogi osvojenega načina reševanja, bodisi zaradi preveč natančnih predhodnih navodil o načinu reševanja posameznih nalog oziroma skupin nalog.

Avtor je menil (**Bujas, 1966; Krković in Kolesarić, 1970**), da je uspešno reagiranje človeka v novih problemskih situacijah odvisno predsem od njegove sposobnosti vpogleda v neuravnoteženost med strukturami v okolju oziroma v neravnotežje med situacijo in cilji, h katerim teži. Iz tega naj bi izhajala osebnostna poteza občutljivosti za probleme.

Opredelitev osnovne ideje za konstrukcijo takšne vrste testa je razvidna iz avtorjevih besed ob predstavitvi testa:

"Ljudje z različno razvito inteligentnostjo se razlikujejo po tem, da inteligentnejši individuum odkriva neuravnoteženosti tudi v tistih situacijah, ki se manj inteligentnim zdijo uravnotežene in jih zato ne pripravijo do miselne aktivnosti. To sposobnost lahkega in hitrega odkrivanja neuravnoteženosti v določenih situacijah lahko imenujemo občutljivost za probleme" (Bujas, 1966; Krković in Kolesarić, 1970).

V Problemskem testu¹ preizkušanec torej sam, brez posebnega navodila za reševanje, odkrije "problem", neuravnoteženost znotraj posamezne naloge, in ko v tem uspe, če uspe, zaradi enostavnosti problema rešitev naloge ni več vprašljiva. Toda odkritje neuravnoteženosti znotraj ene naloge subjektu ne pomaga pri reševanju naslednjih nalog; tu mora problem na novo odkriti (**Radošević in Krizmanić, 1973**). Test torej zahteva od preizkušanca pri reševanju posameznih nalog različne miselne pristope.

Raziskovalci so na različnih vzorcih preizkušancev ugotovili dobre merske karakteristike Pb-testa (**Krković in Kolesarić, 1970; Matešić, 1983**). Primerjalne študije kažejo na visoko napovedno vrednost rezultatov Pb-testa glede na Bujas-Petzovo M-serijo za merjenje inteligentnosti² (**Krković in Kolesarić, 1973**), glede na Domino test ($r = 0.62$), glede na Test poznavanja tujk ($r = 0.49$) ter glede na teste znanja na sprejemnem preizkusu za vpis na študij psihologije v Ljubljani (Bucik, 1987) ter v Zagrebu³ (**Krković in**

1 - Nadalje označevan kot Pb-test.

2 - $r = 0.81$, kar je največja doslej objavljena korelacija Pb-testa s kakim drugim testom inteligentnosti. M-serija je sicer baterija testov, zelo visoko saturirana z G-faktorjem.

3 - Dobljene so bile tudi relativno visoke stopnje povezanosti s Testom večkratne klasifikacije ($r = 0.54$), Testom numeričnega rezoniranja ($r = 0.70$) ter Testom poznavanja tujk (v Zagrebu razvit objektivni test izbirnega tipa s 100 testnimi vprašanji, kjer mora subjekt med petimi alternativami izbrati pravilni slovenski sinonim za določeno tujko ($r = 0.56$)).

Kolesarić, 1970, str.77) in glede na uspešnost v osnovni šoli (Gregorač, Toš in Žagar, 1988) itd.

PROBLEM

Kljub temu, da je bil v zadnjih letih Pb-test na podlagi dobrih rezultatov predhodnih analiz merskih lastnosti često uporabljan v baterijah instrumentov za različne selekcijske preizkuse, v strokovni literaturi ne najdemo svežih preverjanj njegovih merskih karakteristik. Zaradi časovne oddaljenosti opravljenih analiz so se pričeli pojavljati dvomi v nekatere, od uporabe odvisne metrijske karakteristike testa, npr. težavnost, zanesljivost, diskriminativnost itd.

METODA

Analizo smo izvedli na vzorcu 255 preizkušancev z različnimi stopnjami in smermi izobrazbe. 102 preizkušanca sta bila moškega, 153 pa ženskega spola. Starost preizkušancev v času testiranja se je gibala od 14 do 46 let s povprečno vrednostjo (M) 26.4 let in standardnim odklonom (σ) 6.43 let. Testiranja so bila izvedena v letih 1989 in 1990.

V analizo smo vzeli slovenski prevod Pb-testa s 66 testnimi nalogami. Iz skupine 70 nalog v prvotnem testu so bile namreč 4 naloge zaradi neustreznosti izločene. Preizkušanci imajo za reševanje na voljo 40 minut.

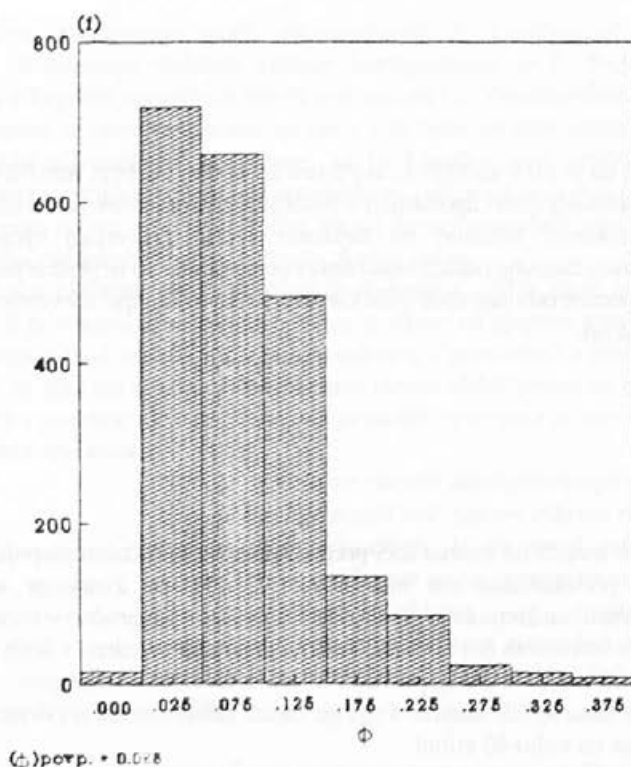
REZULTATI

Analiza testnih nalog:

O tem, da poznavanje principa reševanja ene naloge preizkušancu ne pomaga mnogo pri reševanju drugih nalog, govori povprečni interkorelacijski koeficient med vsemi nalogami. Glede na to, da gre pri odgovorih na posamezno testno nalogo za dihotomizirane vrednosti (pravilno - nepravilno), je prikladna mera interkorelacije ϕ - koeficient za 2x2 tabele. Njegova povprečna vrednost za 2145 interkorelacij znaša 0.088, razpršenost vrednosti posameznih ϕ -koeficientov pa je prikazana na Sliki 1.

Indeks "težavnosti" posameznih nalog (opredeljen kot propore števila preizkušancev, ki so nalogo pravilno rešili) variira od 0.992 pri 1. nalogi, ki je nista pravilno rešila le dva preizkušanca od 255-ih, do 0.012, pri 59.nalogi. Pogostost posameznih vrednosti Indeksov je prikazana na sliki 2. Videti je, da je težavnost nalog v testu relativno enakomerno razporejena od najlažjih do najtežjih.

Slika 1: Distribucija koeficientov interkorelacij nalog v Pb-testu (N = 2145)

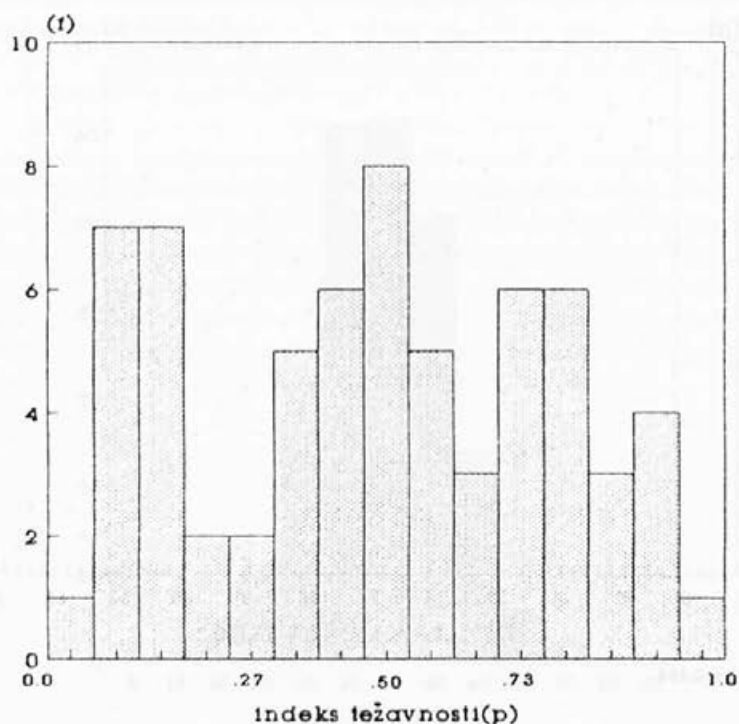


Povprečni indeks znaša 0.484^4 , kar pomeni, da je za ta vzorec preizkušancev test "ravno prav" težak, saj so najprimernejši testi z indeksom težavnosti okrog 0.5 - ko ni test ne prelahek in ne pretežak (Krkovič, 1978). Iz Slike 3 pa je razvidno, kako so naloge po težavnosti distribuirane glede na svoje "delovno mesto" v testu.

Trend prehajanja nalog od lažjih k težjim nalogam je jasen (in to je lastnost, ki jo pripisujemo kvalitetno konstruiranim testom), čeprav je treba opozoriti, da naraščanje težavnosti nalog proti koncu testa ni samo posledica težavnosti posameznih nalog, temveč tudi dejstva, da je Pb-test v prvi vrsti test hitrosti in šele nato test moči. To pomeni, da

4 - Torej so preizkušanci v povprečju rešili ca. 48 % vseh nalog.

Slika 2: Distribucija nalog glede na indeks težavnosti (p)- Pb-test

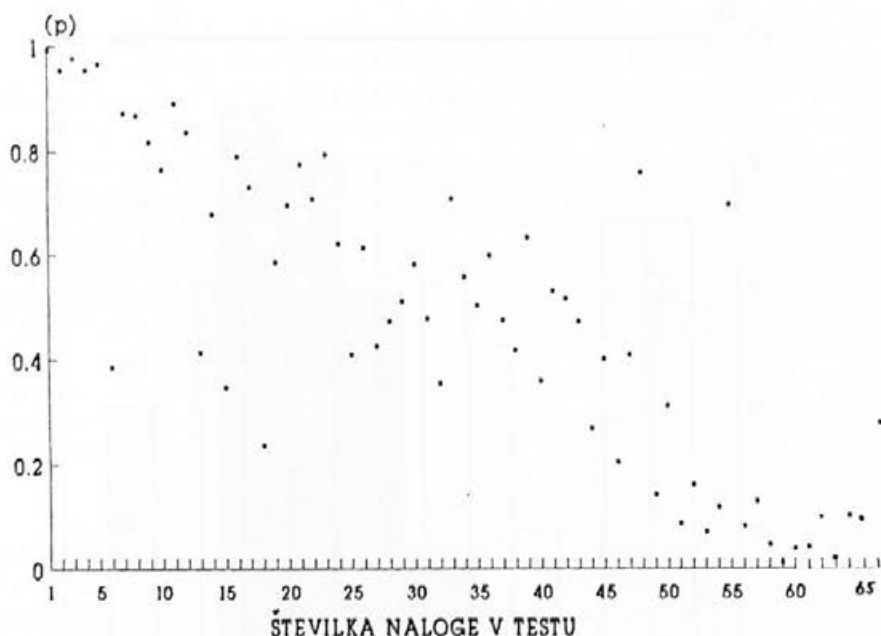


p (povp.) = 0.484

nekateri preizkušanci zaradi pomanjkanja časa verjetno ne uspejo niti začeti reševati nalog na koncu testa, kar vpliva na izrazitejši koeficient težavnosti.

Povprečni dosežek na Pb-testu je 31.95 točk (glej tudi op. 4), standardna deviacija pa 7.77 točke. Pri izračunu t-testa nismo med spoloma našli nobene statistično pomembne razlike v aritmetičnih sredinah dosežkov na testu, prav tako se mlajši preizkušanci niso pomembno razlikovali od starejših, ko smo jih razdelili v grupi pod in nad Mediano ($Me=26$ let). Distribucija bruto dosežkov na Pb-testu je prikazana na Sliki 4. Dosežki so grupirani v 10 razredov in kažejo relativno normalno porazdelitev okrog sredine. Najnižji dosežek je znašal 9 točk, najvišji pa 53, kar je okrog 80 % maksimalnega seštevka 66 točk. Zanimivo

Slika 3: Koeficijentnosti (p) za vsako nalogo glede na njeno mesto - Pb-test



$p(\text{povp.}) = 0.484$

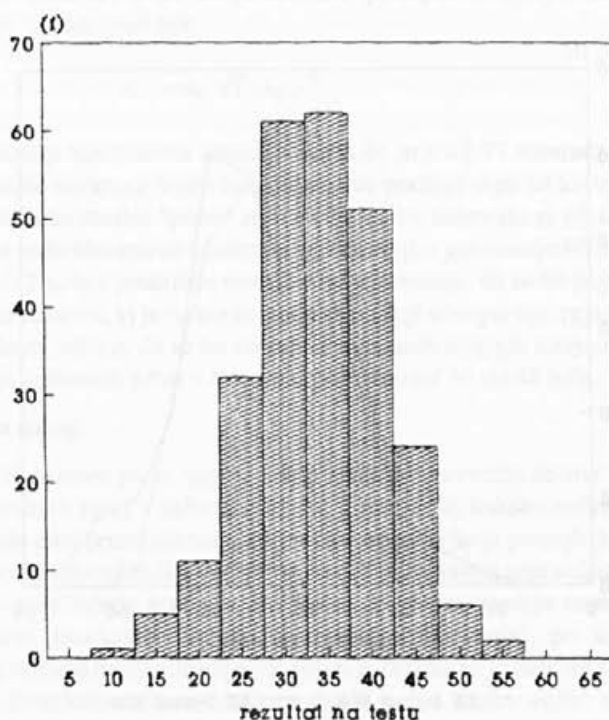
je, da nihče od 255 testirancev ni dosegel maksimalnega števila točk. Tudi to je ena od karakteristik "dobrega" testa.

Koeficijenti diskriminativne veljavnosti, tj. korelacije med uspehom na posamezni nalogi v testu in končnim testnim rezultatom, od naloge do naloge močno variirajo in znašajo od 0.003 do 0.719.

Zanesljivost:

Zanesljivost merjenja s Pb-testom smo odredili na dva načina: Indeks interne konsistentnosti kot odraz homogenosti testa smo določevali s pomočjo Kuder-Richardsove formule 20 (KR-20), katere rezultat je pri dihotomnih nalogah enak Cronbachovemu

Slika 4: Distribucija končnih dosežkov
za vseh 255 testirancev - Pb-test

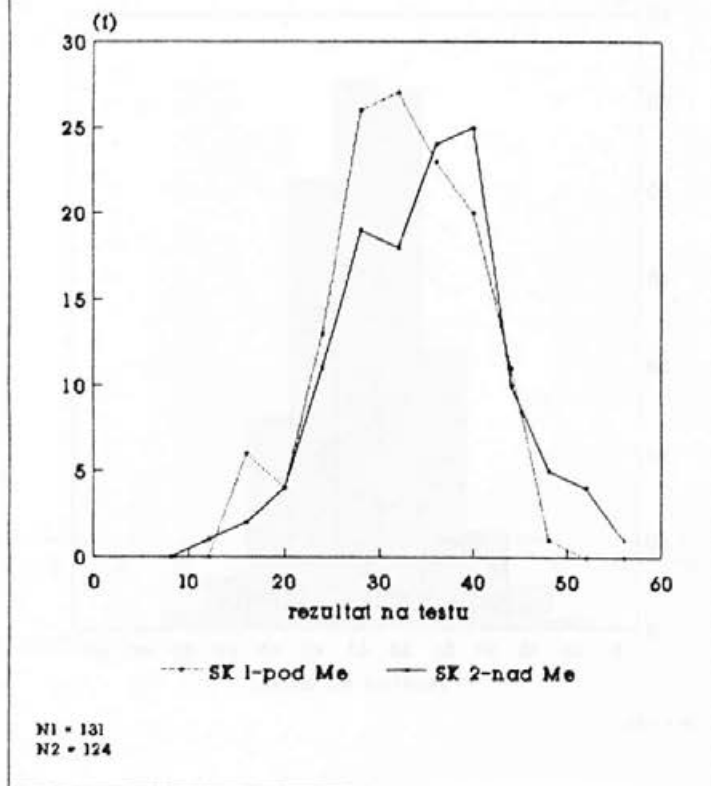


N = 255

koeficientu alfa (α) - Cohen, Montague, Nathanson in Swerdlik, 1988, str. 108-111). Zanesljivost testa, razdeljenega na dva enaka dela (t.i. "split-half"), smo preverjali s koreliranjem testnih rezultatov, seštetih iz parnih in neparnih nalog v testu⁵, ta korelacijski koeficient pa smo korigirali s Spearman-Brownovo formulo (ibid., str. 106-108; Krković, 1978).

5 - Osnovna distinkcija med postopkoma je v "predmetu analize - metode razdeljevanja testa na polovici primerjajo eno polovico z drugo, mere notranje konsistentnosti pa primerjajo vsako od nalog s katerikoli drugo nalogo v testu posebej (Murphy in Davidshofer, 1988).

Slika 5: Distribuciji pravih odgovorov skupine nad in pod Mediano - Pb-test



Koeficient zanesljivosti Pb-testa, izračunan po formuli KR-20 znaša **0.85**, po KR-21 pa **0.81**. Ko smo test razdelili na dva enaka dela (v enem delu parne in v drugem delu neparne naloge) ter evaluirali parcialne rezultate na teh dveh "podtestih" za vsakega testiranca, smo izračunali koeficient korelacije med temi rezultati v višini **0.76**, po korekciji s SB formulo pa je koeficient zanesljivosti znašal **0.86** (prav toliko - **0.85** - znaša tudi Guttmanov "split-half" koeficient). V primerjavi z rezultati analiz Pb-testa, ki so nam na voljo (Krković in Kolesarić, 1970; 1973), so ti koeficienti v povprečju za okrog 0.06 enote nižji, kar pomeni, da se je test na podlagi naših rezultatov pokazal za malenkost manj zanesljivega. To bi lahko pripisali predvsem dejstvu, da gre v naši študiji za dokaj heterogen vzorec testirancev tako po višini in smeri izobrazbe, kot tudi po starosti, v nasprotju z vzorci, na

katerih je bila preverjana zanesljivost testa leta 1970. Kljub temu pa lahko na podlagi izračunanih koeficientov sklepamo na zadostno zanesljivost Pb-testa za uporabo v današnjem času. Standardna deviacija distribucije napak merjenja, oziroma standardna napaka merjenja, opredeljena kot

$$s_e = s_x \sqrt{1 - r_{tt}},^6$$

pri povprečnem koeficientu zanesljivosti 0.84 in $s = 7.77$ namreč znaša 3.08, kar je okroglo 40 % dane variance bruto rezultatov. Na podlagi tega lahko trdimo, da se bo z gotovostjo 68 % individualni "pravi" rezultat pojavil v intervalu $x_i \pm s_e$, kjer x_i pomeni dobljeni rezultat individuuma na konkretni preizkušnji, z gotovostjo 95 %, da se bo pojavil v intervalu $x_i \pm 2 s_e$ in s praktično stodontno gotovostjo, da se bo pojavil v intervalu $x_i \pm 3 s_e$. Za preizkušanca, ki je na konkretni preizkušnji dosegel npr. rezultat 42 točk, lahko torej s 5 % rizikom trdimo, da se bo ob nespremenjenih pogojih merjenja njegov rezultat tudi ob drugih priložnostih gibal v intervalu približno od 36 do 48 točk.

Način reševanja nalog:

V zvezi s Pb-testom pa se nam je zdelo zanimivo preveriti še eno hipotezo, ki jo na tem mestu omenjamo zgolj v informacijo. Ne glede na to, kakšen rezultat je kdo dosegel na testu, obstajajo med preizkušanci glede načina odgovarjanja precejšnje razlike. Nekateri na vsak način poskušajo rešiti čim več nalog, bodisi pravilno ali nepravilno, drugi pa naloge, pri katerih niso popolnoma gotovi v pravilnost odgovora, pustijo nerešene. Za vsakega preizkušanca smo izračunali "Indeks nepravilnega reševanja", pri katerem je število nerešenih nalog odštet od števila napačno rešenih, razlika pa je deljena z številom pravilno rešenih nalog. Preizkušance smo nato razdelili v dve enako veliki skupini (glede na Mediano). V prvo skupino (pod Me) so bili uvrščeni preizkušanci, ki so imeli več nerešenih kot napačno rešenih nalog, v drugo skupino pa tisti z večjim številom napačnih odgovorov. Distribuciji povprečnih dosežkov pravilnih odgovorov na testu za obe skupini sta razvidni iz slike 5.

Ne glede na težje vidne razlike na sliki je enosmerna analiza variance pokazala, da se skupini med seboj pomembno razlikujeta ($p=0.014$), in sicer preizkušanci v prvi skupini v povprečju za okrog tri točke zaostajajo za onimi v drugi skupini. Torej so tisti, ki poskušajo, se borijo in se pri težjih nalogah ne predajo, za kanček uspešnejši od onih, ki "si upajo" v odgovorne liste vpisati rezultat naloge le, še so prepričani v njegovo pravilnost.

ZAKLJUČEK

Kljub temu, da je pričujoča analiza le delna in da je opravljena na relativno majhnem vzorcu, pa rezultati po našem mnenju kažejo na zadovoljivo stopnjo težavnosti, diskriminativne veljavnosti in zanesljivosti Pb-testa. Test dobro odkriva razlike v dosežkih

6 - Kjer je s_x standardna deviacija skupnih bruto rezultatov, r_{tt} pa koeficient kot pokazatelj odnosa med "pravo" in skupno variacijo rezultatov.

preizkušancev in že zgoraj omenjene korelacijske študije nakazujejo, da gre do te razlike v veliki meri na račun delovanja faktorja splošne sposobnosti (G). Za 253 preizkušancev, ki so bili testirani tako s Pb-testom, kot tudi s Testom poznavanja tujk, smo mogli izračunati koeficient korelacije, ki znaša 0.499, kar pomeni, da je okrog 25 odstotkov celotne variance lahko pripisemo delovanju istih faktorjev. Seveda pa bi bila analiza popolnejša, če bi lahko opravili pregled prognostične veljavnosti testa na naših podatkih. Vendar nas ta naloga še čaka in v tem trenutku si lahko pomagamo le z izsledki predhodnih, že opravljenih analiz, ki so omenjene v uvodu.

Vendar pa lahko zaključimo, da Bujasov Problemski test nikakor ne sodi v koš za odpadke in da je uporaben za preizkušance različnih starosti in izobrazbenega nivoja ter strukture. Dopadljiv je tudi zato, ker za razliko od nekaterih drugih testov vsebuje kratko ter enostavno navodilo za reševanje in zahteva relativno kratek čas za aplikacijo.

LITERATURA

- BUCIK, V. (1987). Metodološki problemi selekcijskega postopka ob vpisu na študij psihologije - pogled nazaj. *Anthropos*. 1-2, 250-259.
- BUJAS, Z. (1966). Testovi za ispitivanje inteligencije naših ljudi. Elaborat za Savjet za naučni rad SR Hrvatske. Zagreb.
- COHEN, R. J., MONTAGUE, P., NATHANSON, I. S. & SWERDLIK, M. E. (1988). *Psychological testing. An introduction to tests & measurement*. Mountain View: Mayfield Publishing Company.
- GREGORAČ, J., TOŠ, I. & ŽAGAR, D. (1988). Napovedna veljavnost Bujasovega Problemskega testa in Torrancovih testov ustvarjalnega mišljenja za učni uspeh učencev. Zbornik XVI. posvetovanja psihologov Slovenije, Radenci - 1987. Ljubljana: Društvo psihologov Slovenije. 66-72.
- KRKOVIĆ, A. (1978). *Elementi psihometrije I*. Zagreb: Sveučilište - Filozofski fakultet.
- KRKOVIĆ, A. & KOLESARIĆ, V. (1970). Prikaz novog testa za ispitivanje osjetljivosti za probleme. *Revija za psihologiju*. 1, 73-78.
- KRKOVIĆ, A. & KOLESARIĆ, V. (1973). Prilog odredjivanju simptomatske vrijednosti "Problemnog testa" i "M-serije". *Stručni skupovi "Dani R. Bujasa" 1970 i 1972*. Zagreb: Društvo psihologa SR Hrvatske. 175-187.
- MATEŠIĆ, K. (1983). Prikazi testova: Problemni test (A). *Primijenjena psihologija*. 4, 169-170.
- MURPHY, K. R. & DAVIDSHOFER, C. O. (1988). *Psychological testing. Principles and applications*. Englewood Cliffs: Prentice-Hall International, Inc.
- RADOŠEVIĆ, Z. & KRIZMANIĆ, M. (1973). Težina zadatka u Problemnom testu za ispitivanje inteligencije kod normalnih i shizofrenih. *Stručni skupovi "Dani R. Bujasa" 1970 i 1972*. Zagreb: Društvo psihologa SR Hrvatske. 195-201.