

NOVIJE SPOZNAJE O VALJANOSTI SELEKCIJSKIH METODA:
OD DOKTRINE O SITUACIJSKOJ SPECIFIČNOSTI
DO POSTAVKE O GENERALIZACIJI VALJANOSTI

Branimir Šverko

Odsjek za psihologiju Filozofskog fakulteta u Zagrebu,
Ivana Lučića 3, 10000 Zagreb
e-mail: branimir.sverko@ffzg.hr

Sažetak

Usavršene tehnike meta-analize omogućile su tijekom posljednjih 20 godina plodonosnu spoznajnu integraciju tisuća izvornih validacijskih studija. Ona je bacila novo svjetlo na valjanost psihologijskih testova i drugih prediktora. Utvrđeni koeficijenti tzv. operativne valjanosti glavnih selekcijskih metoda nadilaze očekivanja i svojom visinom potvrđuju da je profesionalna selekcija djelotvoran postupak. Ustanovljeno je, također, da su dobiveni koeficijenti operativne valjanosti dobrim dijelom situacijski invarijantni. To podržava postavku o generalizaciji valjanosti, prema kojoj valjanost prediktora za neki posao utvrđena u jednom okruženju vrijedi i za slične poslove u drugom okruženju. U članku su prikazane recentne meta-analičke spoznaje o valjanosti testova inteligencije, testova ličnosti, intervjuja i drugih selekcijskih postupaka te raspravljene njihove implikacije.

Ključne riječi: profesionalna selekcija, selekcijske metode, generalizacija valjanosti, prediktori

S gotovo stogodišnjom poviješću istraživanja i praktičnog djelovanja, profesionalna selekcija jedno je od postojanih područja primijenjene psihologije. Pa ipak ima novosti koje su u posljednje vrijeme unijele živosti u ovo ustaljeno područje stručnog djelovanja psihologa. Jedna značajna novost jesu promjene i metodološke inovacije koje donosi razvoj novih tehnologija, posebice računalnih aplikacija i Interneta. Profesionalno usmjeravanje, regrutiranje i selekcija putem Interneta nisu više samo ekskluzivna budućnost, već sve prisutnija realnost (Bartram, 2000; Chartrand i Oliver, 2000; Šverko et al., 2002).

Ovaj rad je napisan u sklopu rada na projektu "Ljudski potencijali u promjenjivom svijetu rada" (No. 0130406), koji je odobrilo i djelomično financira Ministarstvo znanosti i tehnologije Republike Hrvatske

Drugi izvor povećane živosti jesu novije spoznaje o valjanosti selekcijskih postupaka. Njima se bavi i ovaj rad. Usavršene tehnike meta-analize omogućile su plodonosnu evaluaciju brojnih izvornih validacijskih studija. Ta je evaluacija kulminirala u posljednjih desetak godina i svojim je rezultatima bacila novo svjetlo na opću valjanost psihologijskih testova i drugih selekcijskih postupaka. U tome su suglasni autori recentnih preglednih radova u ovom području (Hough i Oswald, 2000; Murphey, 2000; Salgado, 1999; Robertson i Smith, 2001). Cilj je ovog rada prikazati najvažnije novije spoznaje o valjanosti selekcijskih metoda te prosuditi njihovo značenje za selekcijsku teoriju i praksu.

Tradicionalni selekcijski model i doktrina o situacijskoj specifičnosti prediktora

Da bismo shvatili značaj novijih spoznaja o valjanosti selekcijskih postupaka, treba se prvo podsjetiti tradicionalnog modela selekcijskog procesa. Prema ovom modelu, proces počinje analizom posla na temelju koje treba ustanoviti osobine važne za uspjeh u poslu, odnosno odabrati potencijalne prediktore. Na temelju analize posla utvrđuju se i relevantni kriteriji radne uspješnosti. Potom slijedi validacija – istraživanje na probnom uzorku kojem je cilj analizirati povezanost između prediktora i kriterija, dakle provjeriti valjanost odabranih prediktora. Tek nakon validacije pristupa se oblikovanju operativnog selekcijskog sustava i praktičnom odabiru osoblja.

Validacija odabranih prediktora je, dakle, nužan preduvjet projektiranja operativnog selekcijskog sustava. On proizlazi iz postavke tradicionalne selekcijske teorije da valjanost nije inherentno svojstvo testa, već specifične situacije u kojoj se test primjenjuje. Ta postavka utemeljena je na iskustvima prvih validacijskih studija, provedenih već u 20-im i 30-im godinama prošlog stoljeća, a osnažena je studijom o valjanosti psihologijskih selekcijskih testova koju objavio Ghiselli 1966. godine. U toj je studiji Ghiselli (1966) analizirao velik broj validacijskih studija koje su publicirane u razdoblju od 1919. do 1964. godine, a provjeravale su valjanost testova kognitivnih i psihomotoričkih sposobnosti i testova ličnosti, za raznovrsna zanimanja. Dva njegova ključna nalaza nisu bila ohrabrujuća. Prvo, ustanovio je da je opća razina valjanosti testova niska. Na primjer, prosječni koeficijent valjanosti za sve psihologijske testove iznosio je samo 0.22. Drugo, pokazao je da koeficijenti prognostičke valjanosti testova za istovrsne poslove dobiveni u različitim studijama imaju veliku varijabilnost. Na primjer, koeficijenti valjanosti testova inteligencije za uspjeh u administrativnom poslu varirali su u rasponu od -0,40 do +0,80, a koeficijenti valjanosti testova specijalnih sposobnosti za uspjeh u poslu rukovatelja stroja u rasponu od -0,55 do +0,65!

Ovako velike razlike u dobivenim koeficijentima valjanosti istih testova za istovrsne poslove, potkrijepile su tezu da je valjanost situacijski uvjetovana, od-

nosno ključnu postavku tradicionalne selekcijske teorije da *valjanost nije inherentno svojstvo testa, već specifične situacije u kojoj se test primjenjuje*. Test koji je dobar za odabir laboratorijskog tehničara u nekoj zdravstvenoj ustanovi, primjerice, ne mora biti dobar za odabir laboratorijskog tehničara u farmaceutskoj industriji. Valjanost testa se ne može generalizirati; nju treba iznova utvrditi za svaku pojedinu situaciju u kojoj se provodi selekcija. Neprijeporna preporuka praktičarima bila je stoga: izvršite validaciju svaki put iznova, jer svaki pojedini selekcijski program treba provjeriti! Smatralo se da se jedino validiranom programu selekcije može pridodati atribut "znanstveni" (Šverko, 1974).

Takvo je gledište vladalo sve do kraja 1980-ih godina. Promicali su ga teoretičari u području selekcijske teorije i vodeći pisci udžbenika iz psihologije rada. A praktičari su nijemo i nemoćno slušali. Nemoćno zato što najčešće nisu bili u stanju validirati svoje postupke, zbog raznih razloga: hitnosti zahtjeva za odabirom ljudi, nedovoljnog broja testiranih sudionika, nemogućnosti valjanog mjerenja radne uspješnosti, ili pak zbog vlastite preopterećenosti, pa i stručne nedoraslosti složenom zadatku validacije.

Moguće je da su pojedinci zbog toga doživljavali grižnju svoje profesionalne savjesti, jer nisu radili *lege artis*, tj. radili su suprotno proklamiranim profesionalnim načelima: prihvaćali su i odbijali ljude temeljem postupaka čiju valjanost nisu empirijski provjerili. No danas im savjest može biti mirnija, jer je tijekom posljednjih desetak i nešto godina doktrina o situacijskoj specifičnosti, nadomještena idejom o generalizaciji valjanosti.

Prve spoznaje o generalizaciji valjanosti

U osnovi postavke o mogućnosti generalizacije valjanosti nalazi se pretpostavka da opažena varijabilnost dobivenih koeficijenata valjanosti nije toliko izraz situacijske specifičnosti koliko rezultat nedostataka validacijskih studija. Ta ideja je odavno prisutna u psihometrijskoj teoriji, ali su je tek koncem 1970-ih oživotvorili Schmidt i Hunter (1977; Schmidt, Hunter, Pearlman i Shane, 1979). Oni su razvili jednu inačicu meta-analize, pomoću koje je bilo moguće testirati pretpostavku da se valjanost, izražena koeficijentima valjanosti, može generalizirati i na druge slične situacije. Postupak su kasnije detaljnije objasnili u svojoj knjizi (Hunter i Schmidt, 1990).

Polazna postavka je jednostavna ideja da se dobiveni koeficijent valjanosti u svakoj pojedinoj validacijskoj studiji sastoji od dvije komponente: *prave korelacije i pogreške*. Jedan niz koeficijenata valjanosti za isti posao, dobivenih u različitim organizacijama, koji imaju (pod)jednake prave korelacije, u realnosti će se dosta razlikovati - zbog različitog udjela komponente pogreške. Pogreške u svakom slučaju povećavaju varijabilitet dobivenih koeficijenata valjanosti. Zato

su i mogući onakvi rasponi među dobivenim koeficijentima valjanosti za isti posao, na koje je uputila Ghisellijeva (1966) studija.

Postoji nekoliko izvora pogrešne varijance u dobivenim koeficijentima valjanosti. Prvi izvor je *pogreška uzorkovanja*, koja je to veća što je uzorak manji. A zbog praktičnih ograničenja mnoge se validacije provode na malim uzorcima. Ghiselli je našao da su u mnogim studijama koje je on analizirao uzorci imali manje od 50 ispitanika. U to vrijeme prosječni uzorak imao je znatno manje od 100 ispitanika. Danas je taj prosjek porastao iznad 100, ali i dalje ima uzoraka s manje od 50 ispitanika (Salgado, 1998). Drugi izvor pogrešne varijance je *smanjenje varijabilneta rezultata u prediktorima*. Ona je prisutna u validacijskim studijama koje koriste "metodu prisutnih radnika", jer su tijekom vremena neuspješni radnici otpali, a najsposobniji napredovali. Ali i pri validaciji *metodom praćenja*, uključimo li u analizu samo radnike primljene nakon provedene selekcije, imat ćemo smanjenje varijabilneta. A ono uvijek utječe na smanjenje prave valjanost prediktora. Treći izvor varijabilnosti dobivenih koeficijenata valjanosti jesu *razlike u pouzdanosti i valjanosti kriterija*, koje mogu značajno varirati od jedne do druge validacijske studije. Konačno, na varijabilitet dobivenih koeficijenata valjanosti utječu i razlike u pouzdanosti prediktora te razlike u njihovoj faktorskoj strukturi.

Schmidt i Hunter su razvili meta-analitički postupak kojim se ispravljaju učinci ovih pogrešaka tako što se iz velikog broja studija, određenim postupcima integrativne analize, izračunaju korigirane vrijednosti. Te korigirane vrijednosti su procjene "pravih" koeficijenata valjanosti. Već koncem 1970-ih i početkom 1980-ih publicirano je nekoliko meta-analitičkih studija, koje su potakle promjenu tradicionalnog stava o situacijskoj ograničenosti koeficijenata valjanosti. Procijenjeno je da četiri izvora pogreške (pogreška uzorkovanja, restrikcija varijabilneta te razlike u pouzdanosti kriterija i prediktora) objašnjavaju oko 75% varijabilneta dobivenih koeficijenata valjanosti. Zbog toga su Schmidt i Hunter (1977; Hunter i Schmidt, 1982; 1990) odbacili doktrinu o situacijskoj specifičnosti.

Na temelju korigiranih koeficijenata valjanosti koji su dobiveni u prvim meta-analizama moglo se zaključiti da ima osnova za ideju o generalizaciji valjanosti, tj. postavku da valjanost testova za neki posao utvrđena u jednom okruženju, vrijedi i za slične poslove i na drugim mjestima (Guion, 1991). Evo nekih od zaključaka prvih meta-analiza koji na to upućuju:

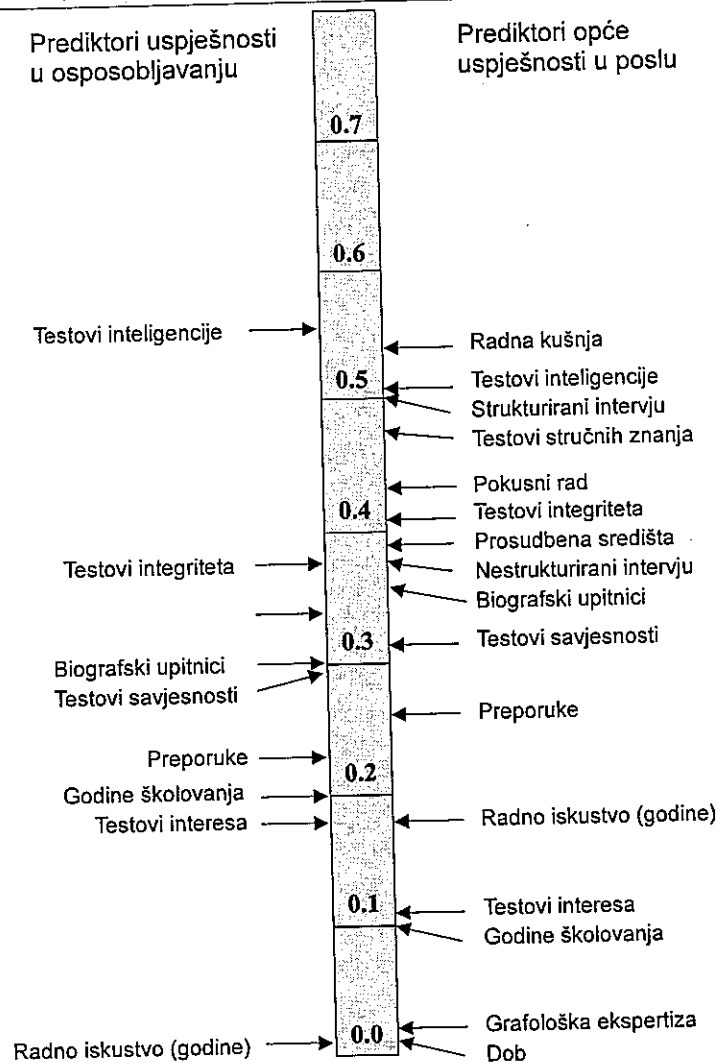
- Profesionalno konstruirani testovi kognitivnih sposobnosti valjani su prediktori uspješnosti u *svim* poslovima (Schmidt i Hunter, 1981).
- Testovi sposobnosti shvaćanja mehaničkih odnosa dobro predviđaju uspješnosti mehaničara; prosječni koeficijent valjanosti, korigiran za restrikciju varijabilneta i nepouzdanost kriterija, iznosi 0.78 (Schmidt i Hunter, 1977).

- Uspješnost poslovođa u različitim organizacijama može se valjano predvidjeti na temelju testova opće inteligencije, shvaćanja mehaničkih odnosa i spacijalnih sposobnosti (Schmidt, Hunter, Pearlman i Shane, 1979).
- Testovi opće inteligencije valjano predviđaju učinkovitost uredskih službenika; korigirani prosječni koeficijent valjanosti iznosi 0.67 (Schmidt i Hunter, 1977). I uže mentalne sposobnosti (perceptivna brzina, verbalna lakoća, numerička sposobnost i dr.) valjano predviđaju uspješnosti u različitim vrstama činovničkih poslova (Pearlman, Schmidt i Hunter, 1980).
- Testovi spretnosti prstiju valjani su prediktori radne uspješnosti radnika za montažnim stolovima (eng. *bench workers*); prosječni korigirani koeficijent valjanosti je 0.39 (Schmidt i Hunter 1977).
- *Programmer Aptitude Test* (McNamara i Hughes, 1961) valjano prognozira uspješnost programera; prosječni korigirani koeficijent valjanosti za uspjeh u osposobljavanju iznosi 0.91, a za uspjeh u radu 0.73 (Schmidt, Gast-Roseberg i Hunter, 1980).
- *Aptitude Index Battery* je valjan prediktor uspješnosti prodavača u svih 12 kompanija uključenih u studiju, ali visinu valjanosti moderira kvaliteta rukovođenja (Brown, 1981).

Svi ovi nalazi pridonijeli su tezi o mogućnost generalizacije valjanosti. Ohrabreni time, istraživači su nastavili s istraživanjima i u 1990-im godinama. Ones (2001) navodi da su se između 1990. i 1996. godišnje objavljivale u prosjeku 174 meta-analize, a između 1997. i 2000. godine čak 224 meta-analize godišnje. Osim za testove sposobnosti, analizirani su koeficijenti valjanosti za postupke provjere znanja, testove ličnosti, intervju i druge prediktore. To nam danas omogućuje da stvorimo pouzdaniju sliku o općoj i komparativnoj valjanosti pojedinih vrsta prediktora.

Novije spoznaje o valjanost različitih selekcijskih postupaka

Odličan pregled relevantnih meta-analiza o valjanosti različitih selekcijskih metoda pripremili su Schmidt i Hunter (1998). Njihov članak se temelji na meta-analizama različitih autora provedenih u posljednjih 20 godina, a obuhvaća rezultate validacijskih istraživanja provedenih tijekom 85 godina. Podaci se odnose na 19 različitih selekcijskih postupaka, čija je valjanost analizirana u odnosu na dvije vrste kriterijskih mjera – uspješnost u tijeku osposobljavanja za posao i uspješnost u samom poslu. Schmidt i Hunter su rezultate svoje integrativne analize prikazali u dvije tablice, popraćene opsežnim bilješkama o korištenim izvorima podataka.



Slika 1 Valjanost selekcijskih metoda i prediktora. Streljice pokazuju veličinu prosječnih koeficijenata valjanosti dobivenih u meta-analizama (nacrtano prema Schmidt i Hunter, 1998)

Za potrebe ovog članka, dio njihovih tabelarnih podataka¹ preoblikovan je u dijagram prikazan na Slici 1.

¹ Schmidt i Hunter (1998) su u svojim tablicama prikazali 19 selekcijskih metoda. Ovdje, na slici 1 izostavljene su tri: procjene koordiniranih procjenitelja (eng. *peer ratings*), koje se baš ne mogu koristiti pri odabiru novog osoblja i dvije specifične, u nas nepoznate metode procjene treninga i iskustva.

U sredini je grafički prikaz korelacijskog kontinuuma, a strelice uperene u njega pokazuju prosječni koeficijent valjanosti pojedinih selekcijskih metoda: lijeve strelice pokazuju valjanost u odnosu na uspješnost u osposobljavanju, a desne u odnosu na uspješnost u poslu. Vrijednosti na koje pokazuju strelice jesu prosječni koeficijenti kriterijske valjanosti *korigirani* za restrikciju varijabiliteta i nepouzdanost kriterija. Tako korigirane koeficijente valjanosti možemo smatrati "operativnim" valjanostima selekcijskih metoda pri njihovoj upotrebi za odabir osoblja iz neke skupine pristupnika, odnosno "pravim" valjanostima (Schmidt i Hunter, 1998). Uz nastavku su komentirane ove vrijednosti, uz navodene meta-analiza u kojima se one procijenjene.

Testovi inteligencije ili "opće mentalne sposobnosti" su, prema uvjerenju mnogih psihologa, najbolji i univerzalni prediktor uspješnosti u različitim poslovima i aktivnostima (Ree i Carretta, 1998). Podaci na Slici 1 potvrđuju ovu reputaciju: prosječni korigirani koeficijent valjanosti testova inteligencije iznosi 0.56 za uspjeh u osposobljavanju i 0.51 za uspjeh u poslu. Ova potonja vrijednost dobivena je u meta-analizi, provedenoj po narudžbi *U.S. Department of Labor* (Hunter, 1980; Hunter i Hunter, 1984), čija je baza podataka uključivala 32.000 radnika u 515 različitim poslovima. Ovako velik skup podataka omogućio je da se prosječni operativni koeficijenti valjanosti izračunaju i za kategorije poslova koji se razlikuju po stupnju složenosti. Tako je ustanovljeno da prognostička valjanost testova inteligencije raste sa složenošću poslova: 0,23 za nekvalificirane poslove, 0,40 za polukvalificirane, 0,51 za srednje složene, 0,56 za složene tehničke poslove i 0,58 za složene profesije i rukovodne poslove. Poslova srednje razine složenosti ima najviše (64% među svim poslovima), pa njihova prognostička valjanost (0.51) dobro reprezentira opću prognostičku valjanost testova inteligencije. Usput, ovaj moderacijski učinak složenosti posla, odnosno nalaz da su testovi inteligencije to valjaniji što su poslovi složeniji, značajan je prilog konstruktnoj valjanosti testova inteligencije.

Standardizirana radna kušnja je izravna provjera radnog umijeća na uzorku radnih operacija koje se i inače izvode u stvarnom poslu. Upravo zato što se izvodi uzorak radnih operacija, u engleskoj se literaturi za radnu kušnju koristi termin *work sample test*. Može se provesti u stvarnom radnom ambijentu, ili pak u izdvojenoj prostoriji i uz opremu koja omogućuje simulaciju radnog procesa. Bitno je da je uzorak radnih operacija reprezentativan za čitav posao te da je postupak provođenja radne kušnje i mjerenja učinka strogo standardiziran. To uključuje i izradu normi za vrednovanje individualnih rezultata. Radne kušnje se uspješno koriste pri odabiru uredskog osoblja i proizvodnih radnika (varilaca, mehaničara, električara, zidara i drugih). Koriste se i pri selekciji za prometna zanimanja; primjerice, sofisticirani simulatori leta omogućuju provođenje vrhunskih radnih kušnji. No radne kušnje je moguće ostvariti i bez skupe opreme. Uz poznavanje osnovnih principa standardizacije i malo domišljatosti, moguće ih je razviti

za gotovo svaki posao. Nedostatak im je što su skupe, jer za svaki posao treba razviti i standardizirati poseban postupak, a ispitivanje kandidata je individualno. Međutim ulaganje u razvoj i primjenu radne kušnje se u pravilu isplati, jer iskustva s njenom primjenom su odlična. Korigirani koeficijent valjanosti za "radnu kušnju" iznosi 0.54 (prema Hunter i Hunter, 1984). Kako se može vidjeti na Slici 1, to je najveći pojedinačni koeficijent valjanosti za prognozu radne uspješnosti, veći i od onog za testove inteligencije. Radna kušnja je odlična selekcijska metoda kad trebamo odabrati one koji poznaju svoj posao, ali je neprikladna za odabir onih koji ga tek trebaju učiti (zato je i nema na lijevoj strani Slike 1).

Intervju je najčešće korištena metoda odabira osoblja. Premda rane analize valjanosti tradicionalnog intervjuja nisu bile ohrabrujuće (Arvey i Campion, 1982), moderne meta-analize ga dobro ocjenjuju. Kako vidimo na Slici 1, *strukturirani intervju* je visoko rangirana selekcijska metoda. Njegova korigirana prognostička valjanost za prognozu uspjeha u poslu jednaka je onoj za testove inteligencije ($r_c = 0.51$; prema McDaniel, Whetzel, Schmidt, & Mauer 1994). To nadmašuje sve ranije stručne prosudbe o valjanosti intervjuja. Čini se da su dugogodišnja upozoravanja psihologa na nedostatke intervjuja i različite mjere koje su poduzimali za njegovo poboljšanje – urodili plodom. Treba naglasiti, međutim, da se dobra svojstva intervjuja ograničena samo na strukturirani intervju. Naime, neprijeporna i opetovana spoznaja svih meta-analiza jest da *prognostička valjanost intervjuja ovisi o stupnju njegove strukturiranosti*. Razgovor bez fiksnog formata, tj. *nestrukturirani intervju* – kod kojeg nema planirane pripreme za razgovor i osobine koje će se prosuđivati nisu unaprijed jasno određene, a sadržaj i način pitanja prepušteni su domišljatosti intervjuera, baš kao i način prosudbe dobivenih odgovora – ima znatno manju valjanost. Kako se iz Slike 1 može očitati, njegova korigirana prognostička valjanost iznosi 0.38. No ova vrijednost, također preuzeta iz analize McDaniela i suradnika (1994), možda je i precijenjena, jer, prema Salgadu (1999), nestrukturirani intervju ima vrlo skromnu valjanost od svega 0.20.

Testovi stručnih znanja su standardizirani testovi, u pravilu tipa papirolovka, koji ispituju znanja potrebna pri obavljanju određenog posla. To mogu biti znanja o radnim procesima, vrstama materijala, upotrebi različitih alata i sl. Najbolje je kad su testovi stručnih znanja konstruirani na temelju analize sadržaja konkretnog posla za koji se provodi selekcija. Kao i u slučaju radne kušnje, njihova upotreba je ograničena samo na one koji su prošli proces stručnog osposobljavanja. Testovi stručnih znanja imaju dobru pojavnu valjanost, visoko koreliraju s radnom kušnjom, a meta-analize svjedoče i o njihovoj dobroj prognostičkoj valjanosti. Hunter i Hunter (1984) procjenjuju da njihov prosječni korigirani koeficijent prognostičke valjanosti iznosi 0.48.

Pokusni rad je razdoblje sustavnog opažanja novoprimitljenih radnika (između 3 i 8 mjeseci), tijekom kojeg se procjenjuje zadovoljavaju li oni neke definirane standarde radnog ponašanja. Onima koji ne zadovoljavaju, privremeni radni

odnos prestaje. Prema tome, riječ je o procjeni radne uspješnosti i prilagodbe organizaciji tijekom početnog razdoblja rada. Ako su standardi izvrsnosti jasno definirani i dosljedno se primjenjuju, ovaj postupak može imati prihvatljivu prognostičku valjanost ($r_c = 0.44$; prema Hunter i Hunter, 1984). Međutim postupak je skup u provedbi, a procjenitelji – u pravilu nadređeni rukovoditelji – često su neskloni prosudbi kojom se radnik otpušta.

Preostala dva prediktora u tablicama 1 i 2 odnose se na testove ličnosti. Na njih tradicionalna, psihometrijski utemeljena psihologija rada nikad nije osobito računala. Ne zbog toga što ličnost ne bi bila važna, nego zato što je mjerenje osobina ličnosti bremenito metodološkim poteškoćama. Objektivni testovi ličnosti skupi su i nepraktični, projektivne tehnike imaju poteškoća s osnovnim metrijskim karakteristikama, a upitnici ličnosti omogućuju neiskreno odgovaranje (simuliranje i disimuliranje). Suzdržan stav prema upitnicima ličnosti potkrijepile su i rane pregledne studije (npr. Guion i Gottier, 1965) koje su pokazale da su koeficijenti valjanosti testova ličnosti preniski za upotrebu u selekcijske svrhe. No spoznaje o mogućnosti generalizacije valjanosti kognitivnih testova potakle su u 1990-im interes i za testove ličnosti. Meta-analize usmjerile su se prvenstveno na dvije vrste postupaka za ispitivanje ličnosti – tzv. testove integriteta i testove savjesnosti, za koje je vjerojatno bilo najviše izvornih studija. U oba slučaja riječ je o standardiziranim upitnicima ličnosti provjerenih metrijskih karakteristika, koje stoga ubrajamo u psihologijske testove.

Testovi integriteta su u nas slabo poznati. Integritet između ostalog znači "osobno poštenje, neovisnost, striktnu privrženost određenim vrijednostima i ponašanjima"². U sličnom značenju koristi se i kad je riječ o testovima integriteta. Uvijek je bilo zaposlenika koji su s mjesta rada otuđivali alat, skupocjen materijal i drugo. Znatno veći problem je gospodarski kriminal koji mogu počinuti oni na najodgovornijim položajima. Amerikanci su procijenili da privredni kriminal stoji njihove poslovne organizacije preko 67 milijardi dolara godišnje. Zato su poslodavci oduvijek pokazivali interes za neke postupke kojim bi se unaprijed, u trenutku zapošljavanja, mogli otkriti pojedinci skloni nečasnim radnjama. U SAD-u su u tu svrhu koristili čak poligrafska istraživanja, ali nakon što je upotreba "detektora laži" zakonski ograničena, porastao je interes za upitničke mjere, tj. za testove integriteta. Oni su utemeljeni na nalazima istraživanja da pojedinci koji su skloni krađi dijele neke stavove, koji mogu pomoći pri identifikaciji. Na primjer, oni drže za sebe da su pošteni ljudi u nepoštenom svijetu, često razmišljaju o krađi, pokazuju veću toleranciju prema onima koji krađu, vjeruju da većina ljudi krađe, vjeruju da se kradljivci međusobno ne odaju, skloni su nalaženju opravdanja za krađu (Smither, 1998). Testovi integriteta su, dakle, upitnici koji sadrže pitanja o stavovima prema nepoštenju, ali i o okolnostima kad su ispitanici sami

² Anić, V. i Goldstein, I. (1999), *Rječnik stranih riječi*. Zagreb: Novi Liber.

učinili nešto nepošteno. Prvi testovi integriteta bili usmjereni isključivo na identifikaciju pojedinaca koji su skloni krađi. Kasnije su razvijeni i općenitiji testovi, koji nastoje otkriti sklonost i prema drugim aspektima antisocijalnog ponašanja, kao što su neopravdano izostajanje s posla, bezobzirnost, namjerno destruktivno ponašanje, sklonost agresivnom ponašanju ili upotrebi droga.

Bez obzira na skepsu koju možemo imati prema ovakvim upitnicima, validacijske studije daju im potporu. Ones, Viswesvaran i Schmidt (1993) su objavili opsežnu meta-analizu testova integriteta na bazi podataka koja je raspolagala sa 665 koeficijenata valjanosti, dobivenih na različitim uzorcima ispitanika, s ukupno oko 600.000 podataka. Analiza je pokazala da testovi integriteta predviđaju detektirane krađe ($r_c = 0.13$), priznate krađe ($r_c = 0.42$) i opće "kontraproduktivno" ponašanje ($r_c = 0.39$). No najinteresantnije je da testovi integriteta predviđaju i radnu uspješnost! Na temelju 23 studije prognostičke valjanosti provedene na pristupnicima (čiji je ukupni broj bio 7.550), procijenili su da operativna valjanost testova integriteta za procjenu radne uspješnosti iznosi 0.41. Kao prediktori uspjeha u osposobljavanju ($r_c = 0.31$), testovi integriteta zaostaju samo za testovima inteligencije. To nas potiče na pitanje: zašto testovi integriteta uopće predviđaju uspješnost u poslu i osposobljavanju? Mogući odgovor na to pruža analiza njihove konstruktne valjanosti. U svojoj doktorskoj radnji Ones (1993; prema Ones i Viswesvaran, 2001) je istraživala povezanost testova integriteta s dimenzijama Petfaktorskog modela ličnosti i ustanovila da njihovi rezultati najviše odražavaju "savjesnost".

Savjesnost (eng. *conscientiousness*) je jedna od dimenzija Petfaktorskog modela ličnosti koji se nametnuo početkom 1990-ih (Digman, 1990; Goldberg, 1993). Ona uključuje sklonost sredenosti, planiranju, ostvarenju ciljeva i poštivanju rokova, a za njeno mjerenje postoje izravni testovi ili ljestvice ličnosti, poput one iz NEO upitnika ličnosti (Costa i McCrae, 1985). Prema proračunima Mounta i Barricka (1995; prema Schmidt i Hunter, 1998), ovakvi testovi savjesnosti podjednako predviđaju uspješnost u poslu ($r_c = 0.31$) i osposobljavanju ($r_c = 0.30$). To su skromne operativne valjanosti, ali ipak dovoljne za uzimanje u obzir testova savjesnosti kao potencijalnog prediktora u sastavu neke selekcijske baterije.

Međutim, zanimljivo je da su testovi integriteta, koji također mjere savjesnost, bolji prediktor radne uspješnosti od samih testova savjesnosti. Ones i Viswesvaran (2001) smatraju da je to zato što oni spadaju u sadržajno specifičnije, kriterijski orijentirane mjere ličnosti, namijenjene upravo prognozi određenih aspekata radnog ponašanja. Osim toga, uz savjesnost, testovi integriteta zahvaćaju još i dvije druge dimenzije Petfaktorskog modela – prijaznost (*agreeableness*) i emocionalnu otpornost. Faktorski složeniji, oni pokrivaju više komponenti radne uspješnosti, pa tako objašnjavaju i više kriterijske varijance.

Sljedeći prediktor jesu *ekspertize prosudbenih središta*, koje se u nekim zemljama koriste u postupku odabira, osobito rukovoditelja. 'Prosudbeno središte' (eng. *assessment center*) nije neko mjesto u prostoru, nego selekcijski proces koji uključuje prosudbu pojedinaca pomoću različitih metoda. Naziv 'središte' dolazi otuda što se kandidati moraju okupiti na nekom izdvojenom mjestu radi temeljitih ispitivanja i opažanja. Proces uključuje intervju i različite psihološke testove, ali je naglasak na prosudbi kandidata na temelju opažanja njihova ponašanja u insceniranim situacijskim testovima. Na temelju svega toga se piše izvještaj o pojedincima i donose odluke. Korigirana prognostička valjanost ovog postupka je procijenjena na 0.37 (Gaugler, Rosenthal, Thornton i Bentson, 1987).

Između dvije vrste testova ličnosti svrstali su se *biografski upitnici*. Uz jednostavne i relativno objektivne podatke o pojedincu, ovi upitnici sadrže i različita pitanja o njegovoj životnoj povijesti, iskustvima iz djetinjstva i razdoblja školovanja, o aktivnostima kojima se bavio u životu, a mogu biti uključena i složenija pitanja, koja se odnose na pojedinčeve stavove, vrijednosti i preferencije te percepcije njegovih odnosa s okolinom. Ovi se upitnici boduju proporcionalno diskriminacijskoj valjanosti svakog pitanja, koja je se mora utvrditi prethodnom validacijskom studijom. Prosječna operativna valjanost biografskih upitnika za prognozu radnog uspjeha iznosi 0.35 (Rothstein i sur., 1990), a za predviđanje uspjeha osposobljavanju 0.30 (Hunter i Hunter, 1984).

Preporuke se temelje na pretpostavci da oni koji poznaju kandidata jesu dobar izvor informacija o njemu. Može ih priložiti sam pristupnik, ali su vjerojatno vjerodostojnije one koje novi poslodavac traži izravno od prethodnog poslodavca. Preporuke mogu biti u vidu narativnog iskaza, ali je bolje tražiti informacije pomoću strukturiranog upitnika usmjerenog na relevantne osobine. Općenito se drži da je vrijednost preporuka ograničena, zbog sklonosti spontanom uljepšavanju, ocjenama utemeljenim na sporadičnim dojmovima, površnom odgovaranju i sl. U skladu s time dobiveni su skromni koeficijenti operativne valjanosti preporuka: 0.26 za uspjeh u poslu i 0.23 za uspjeh u osposobljavanju (obje vrijednosti prema Hunter i Hunter, 1984).

Testovi interesa su važni pri profesionalnoj orijentaciji, jer su interesi ključna odrednica odabira zanimanja. No neki drže da oni mogu biti korisni i pri profesionalnoj selekciji, jer ako pojedinčevi interesi nisu usklađeni sa sadržajem posla, pojedinac će biti manje uspješan pri osposobljavanju i kasnijem radu. Rezultati meta-analize (Hunter i Hunter, 1984) pokazuju da je to u vrlo maloj mjeri točno: njihova valjanost iznosi 0.18 za uspjeh u osposobljavanju i svega 0.10 za uspjeh u radu.

Pri dnu našeg kontinuuma valjanosti na Slici 1 preostala su još tri demografska indikatora iskustva. *Godine radnog iskustva* imaju malu prognostičku valjanost za uspjeh u poslu ($r_c = 0.18$; prema Hunter i Hunter, 1984). Međutim, ta vrijednost je vjerojatno podcijenjena, jer istraživanja pokazuju (prema Schmidt i

Hunter, 1998) da je odnos između godina radnog iskustva i uspjeha u poslu zakrivljen: tijekom prvih 5 godina stjecanje iskustva uspješnost se povećava, a potom se stabilizira. *Godine školovanja* imaju još manju valjanost za uspjeh u poslu ($r_c = 0.10$), a neznatno veću za uspjeh u osposobljavanju ($r_c = 0.20$). *Kronološka dob* je na samom dnu našeg kontinuuma valjanosti. Nasuprot čestoj praksi otvorene ili prikrivene dobne diskriminacije pri zapošljavanju, meta-analize postojećih studija pokazuju da nema nikakve povezanosti između dobi i radne uspješnosti.

Na dnu našeg kontinuuma nalazi se još i *grafološka ekspertiza*. Spominjemo je ovdje jer u nekim zemljama ima poslodavaca koji pri odabiru osoblja koriste i usluge grafologa. Validacijske studije pokazuju da analiza rukopisa nema nikakve vrijednosti za predviđanje radne uspješnosti (Neter i Ben-Shakhar, 1989).

Dodatna valjanost pojedinih selekcijskih metoda

Izloženi podaci pokazuju da su testovi inteligencije najbolji pojedinačni prediktor: bez premca su u predviđanju uspjeha u osposobljavanju i među najboljim su prediktorima uspjeha u poslu. Osim toga njihova konstruktivna valjanost je dobro proučena, a cijena primjene u selekcijskom procesu manja nego za većinu drugih ozbiljnijih metoda. Zato testovi inteligencije zaslužuju osobit status među selekcijskim metodama – s pravom ih možemo smatrati *primarnom*, neizostavnom selekcijskom metodom, a sve ostale metode njihovim *dopunama* ili *dodacima*. Koliko dobivamo primjenom pojedinih dopuna?

Schmidt i Hunter (1998) su procijenili koliki se prirast valjanosti može teorijski ostvariti s pojedinim selekcijskim metodama kad se primijene u kombinaciji s testom inteligencije kao primarnom metodom. Dakle, upotrijebimo li neku selekcijsku metodu u optimalno ponderiranoj kombinaciji s testom inteligencije, za koliko ćemo povećati prognostičku valjanost iznad vrijednosti 0.51, odnosno 0.56, koju možemo postići upotrebom samog testa inteligencije? Taj prirast valjanosti naziva se *dopunska* ili *dodana valjanost* (eng. incremental validity). On ne ovisi samo o valjanosti pojedine dopune, već i o njezinoj interkorelaciji s inteligencijom. Smith i Hunter su procijenili dodanu valjanost za sve selekcijske metode, ali za potrebe ovog članka izdvojene su samo one koje ostvaruju najveći prirast.

U Tablici 1 prikazane su najbolje dopune testovima inteligencije za prognozu uspjeha u poslu. Najbolji su testovi integriteta. Premda je njihova osnovna valjanost umjerena (0.41), zbog toga što ne koreliraju s inteligencijom, oni daju najveću multiplu korelaciju ($R = 0.65$) i ostvaruju porast valjanosti od 27% iznad one koju možemo postići samo testom inteligencije. Slijede radna kušnja i strukturirani intervju s 24% povećanja valjanosti, pa testovi savjesnosti s 18% i konačno testovi profesionalnih interesa i probni rad, koji povećavaju valjanost za 14%.

Tablica 1 Najbolje dopune testovima inteligencije za prognozu uspjeha u poslu (prema Schmidt i Hunter, 1998)

Postupak	r_c	R	r_{ink}	Porast (%)	Regresijski koeficijenti	
					Intel.	Dopuna
Testovi inteligencije	.51					
Testovi integriteta	.41	.65	.14	27%	.51	.41
Standardizirana radna kušnja	.54	.63	.12	24%	.36	.41
Strukturirani intervju	.51	.63	.12	24%	.39	.39
Testovi savjesnosti	.31	.60	.09	18%	.51	.31
Testovi stručnih znanja	.48	.58	.07	14%	.36	.31
Pokusni rad	.44	.58	.07	14%	.40	.20

Objašnjenja: r_c = prosječni koeficijent valjanosti ispravljen za nepouzdanost kriterija i suženje varijabiliteta (tj. operativna ili "prava" valjanost); R = koeficijent multiple korelacije dobiven za kombinaciju s testom inteligencije kao osnovnim prediktorom; r_{ink} = porast valjanosti ostvaren dodavanjem metode testu inteligencije; % porasta = postotno povećanje valjanosti u odnosu na osnovnu valjanost testova inteligencije. U posljednja dva stupca su standardizirani regresijski koeficijenti ("beta-ponderi") za kombinaciju inteligencije i dopune.

Tablica 2 Najbolje dopune testovima inteligencije za prognozu uspjeha u osposobljavanju (prema Schmidt i Hunter, 1998)

Postupak	r_c	R	r_{ink}	Porast (%)	Regresijski koeficijenti	
					Intel.	Dopuna
Testovi inteligencije	.56					
Testovi integriteta	.38	.67	.11	20%	.56	.38
Testovi savjesnosti	.30	.65	.09	16%	.56	.30
Reference	.23	.61	.05	9%	.56	.23
Intervju (obje vrste)	.35	.59	.03	5%	.59	.19

Vidi objašnjenja ispod tablice 1.

U Tablici 2 prikazane su najbolje dopune za prognozu uspjeha u osposobljavanju. Ovdje odskakuju obje vrste testova ličnosti, a znatno manji prirast može se ostvariti preporukama i intervjuom.

Najnovija istraživanja

Istraživači su nastavili s meta-analizama i prvim godinama novog milenija. Tako su McDaniel, Morgeson, Finnegan, Campion i Braverman (2001) analizirali valjanost *situacijskih testova* tipa papir-olovka, s pomoću kojih se ispituje kako pojedinac prosuđuje i rješava različite situacije u poslu. Na temelju 102 koeficijenta valjanosti i ukupno 10.640 ispitanika, autori su ustanovili da situacijski testovi imaju "upotrebljivu" razinu valjanosti ($r_c = 0.34$). Istovremeno su Kuncel, Hezlett i Ones (2001) proveli opsežnu meta-analizu valjanosti GRE-a (baterija testova intelektualnih sposobnosti i znanja iz različitih područja) i prosječnih ocjena u dodiplomskom studiju za prognozu različitih kriterija uspješnosti u post-diplomskim studijima. Pokazali su da se valjanost oba prediktora može generalizirati na postdiplomske programe u različitim znanstvenim oblastima.

Osim toga, provedene su i značajne replikacije. Najnovija meta-analiza na europskim podacima (Salgado i Anderson, 2003) uvjerljivo je potvrdila univerzalnost spoznaja o generalizaciji valjanosti testova inteligencije. Salgado i Anderson su proveli opsežnu meta-analizu valjanosti testova opće mentalne sposobnosti i to na temelju izvornih validacijskih studija iz šest, kulturalno različitih zemalja Europske unije. Njihova baza podataka je uključivala 128 studija koje su kao kriterij koristile uspjeh u poslu (s ukupno 12.128 ispitanika) i 122 studije koje su kao kriterij koristile uspjeh u osposobljavanju (22.987 ispitanika). Za oba kriterija, uz primjenu korekcije za restrikciju varijabiliteta i nepouzdanost kriterija, nađene su izuzetno visoke operativne valjanosti (r_c je varirao između 0.60 i 0.70). Takve operativne valjanosti su nađene u svim zemljama, a mogle su se generalizirati preko različitih zanimanja u svakoj zemlji. Autori naglašavaju da su prosječne operativne valjanosti testova inteligencije, koje oni nalaze u europskim podacima, veće nego što pokazuju meta-analize američkih podataka, ali ne navode moguća obrazloženja za ovu razliku. U svakom slučaju, ovi europski podaci uvjerljivo potkrepljuju uvjerenje da su testovi inteligencije najbolji samostalni prediktor radne uspješnosti.

Proširenju spoznaja o valjanosti testova ličnosti pridonijeli su Hogan i Holland (2003). Oni su primijetili da dosadašnje meta-analize nisu bile teorijski utemeljene, a kao kriterij koristile su samo procjene opće radne uspješnosti. Hogan i Holland pretpostavljaju da se teorijskim "uparivanjem" pojedinih dimenzija ličnosti s njima primjerenim kriterijima ponašanja mogu potpunije profilirati podaci o valjanosti različitih dimenzija ličnosti. U svojoj meta-analizi, oni kao teorijski okvir koriste socio-analitičku teoriju R. Hogana, a dimenzije Petfaktorskog modela ličnosti analiziraju u odnosu na uže kriterijske konstrukte razvrstane u dvije šire skupine, ovisno o tome da li se odnose na "napredovanje" (*getting ahead*) ili "uklapanje" (*getting along*). Njihova analiza je pokazala da, u skladu s pretpostavkom, sve dimenzije Petfaktorskog modela najbolje predviđaju

sadržajno srodne kriterijske konstrukte, a najbolji koeficijenti operativne valjanosti za pojedine dimenzije ličnosti kretali su se u rasponu od 0.29 do 0.37.

Meta-analitičke spoznaje o dobroj valjanosti intervjua, potakle su zanimanje istraživača i za njegovu konstruktnu valjanost. Što zapravo mjere intervjui i razlikuju li se u tom pogledu različite vrste intervjua? Tim pitanjem bavile su se dvije recentne meta-analitičke studije.

Huffcutt, Conway, Roth i Stone (2001) su najprije sustavno razvili opsežnu taksonomiju konstrukata, koji bi mogli biti zastupljeni u intervjuiima, a potom su njihovu zastupljenost procjenjivali u realnim intervjuiima. U tom cilju analizirali su uzorak izvornih studija koje su provjeravale valjanost intervjua. Pronašli su ukupno 47 relevantnih studija, u kojima su intervjuieri procijenili 338 različitih osobina. Razvrstavanjem i kodiranjem ovih osobina, Huffcutt i suradnici su ustanovili da su to najčešće temeljne osobine ličnosti i socijalne vještine, a potom intelektualne sposobnosti i poznavanje posla. Međutim, ustanovili su i razlike u odnosu na vrstu intervjua. U *strukturiranim* intervjuiima više se procjenjuje poznavanje posla, praktične kognitivne vještine (rješavanje problema, donošenje odluka), socijalne vještine i usuglašenost kandidata s organizacijom, a u *nestrukturiranim* intervjuiima opća inteligencija, obrazovanje i iskustvo, interesi i tjelesna svojstva. Na temelju toga, a i na temelju meta-analitičkih procjena valjanosti pojedinih konstrukata za prognozu radne uspješnosti, autori su zaključili da su strukturirani intervjui u većoj mjeri usmjereni na konstrukte koji su bolje povezani s radnom uspješnošću. Prema tome, uz njihovu bolju pouzdanost, to može biti i sadržajno objašnjenje za empirijski ustanovljenu činjenicu (vidi Tablicu 1) da strukturirani intervjui svojom valjanošću nadmašuju nestrukturirane.

Salgado i Moscoso (2002) su pitanje konstruktne valjanosti intervjua pokušali riješiti na malo drugačiji način: analizom korelacija između prosudbi intervjua i rezultata intervjuiranih u testovima i upitnicima koji mjere različite konstrukte. U tom cilju su proveli niz meta-analiza u kojima su analizirali korelate dvije vrste intervjua: *konvencionalnog* (usredotočen na kvalifikacije, iskustvo i samoprosudbu kandidata) i *bihevioralnog* (usredotočen na pitanja o ponašanju na poslu i radnom iskustvu). Ustanovili su da konvencionalni intervjui prvenstveno odražava kandidatove socijalne vještine, opću inteligenciju, emocionalnu stabilnost i ekstraverziju, dok bihevioralni intervjui prvenstveno odražava radno iskustvo, poznavanje posla, situacijske prosudbe i socijalne vještine.

Istraživanje konstruktne valjanosti intervjua još je u začetku, ali izvjesno je da bi svojim spoznajama moglo doprinijeti boljem sadržajnom oblikovanju intervjua u cilju dodatnog povećanja njegove prognostičke valjanosti.

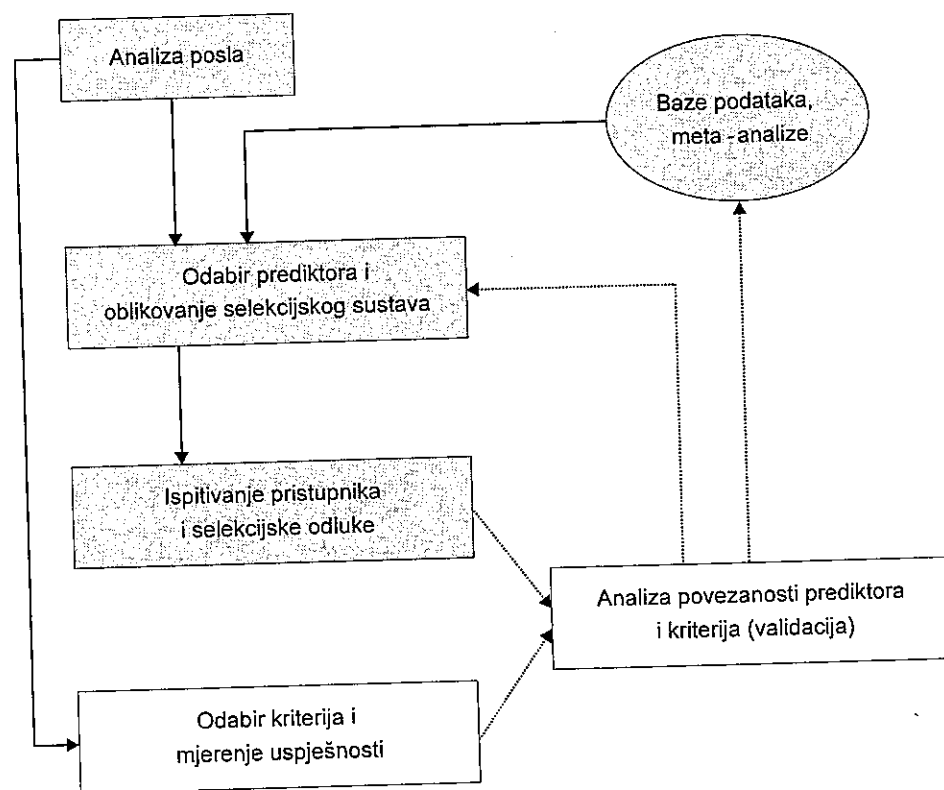
Zaključni osvrt: značenje spoznaja o generalizaciji valjanosti

Premda meta-analitičke studije nisu bez ograničenja, prikazane kumulativne spoznaje o valjanosti selekcijskih metoda izvanredno su znanstveno postig-

nuće. One su izravan dokaz sposobnosti znanstvenog istraživanja da proizvede zaista relevantne praktične rezultate. Značenje dobivenih rezultata je višestruko.

Prvo, dobiveni rezultati otklanjaju svaku sumnju u valjanost naših selekcijskih metoda i mogućnost praktičnog predviđanja radne uspješnosti. Utvrđeni koeficijenti operativne valjanosti za testove inteligencije i druge strukturirane dijagnostičke postupke svojom visinom potvrđuju da je profesionalna selekcija izvanredno djelotvoran postupak. Rijetko koja struka može se pohvaliti ovako impresivnom empirijskom verifikacijom svojeg profesionalnog djelovanja.

Drugo, dobiveni rezultati upućuju na neosnovanost doktrine o situacijskoj specifičnosti naših dijagnostičkih metoda. Meta-analize pokazuju da su operativne valjanosti naših selekcijskih postupaka u priličnoj mjeri situacijski invarijantne, jer su približno jednake valjanosti pojedinih prediktora dobivene u različitim organizacijskim okolnostima. Dakle, test koji je dobar za odabir laboratorijskog tehničara u nekoj zdravstvenoj ustanovi biti će dobar i za odabir laboratorijskog tehničara u farmaceutskoj industriji. Ova mogućnost generalizacije valida-



Slika 2. Revidirani model selekcijskog procesa

cijskih rezultata donosi praktičarima veliko ohrabrenje i olakšanje: ipak nije nužno za svaki novi selekcijski program provoditi skupu i složenu, ponekad i neizvedivu validaciju. Dovoljno će biti na temelju analize posla i poznavanja rezultata validacijskih meta-analiza odabrati najbolje prediktore i složiti selekcijski program. Sada je moguće revidirati tradicionalni selekcijski model. Moje grafičko viđenje takve revizije prikazuje Slika 2. Selekcijski program se, dakle, može utemeljiti samo na analizi posla i na podacima koji su akumulirani u bazama podataka i rezultatima meta-analiza, bez obveze da se prije operativne faze provede lokalna validacija³.

Treće, dobiveni rezultati dobar su putokaz praktičarima za unapređenje njihovih selekcijskih programa. Kako se vidi iz Slike 1, znatne su razlike u operativnim valjanostima pojedinih metoda: dok neke imaju visoku ili solidnu valjanost, druge su malo vrijedne ili bezvrijedne. Povremene snimke situacije u praksi, kako u svijetu (npr. Levy-Leboyer, 1994) tako i u nas (Šverko i Galešić, 2003), pokazuju da stvarna selekcijska praksa nije primjerena znanstvenim spoznajama. Dvije najbolje pojedinačne selekcijske metode (testovi inteligencije i radna kušnja) nedovoljno se koriste, a među intervjuima prevladavaju oni manje vrijedni – nesstrukturirani. Prema podacima u Tablici 1, najbolji se rezultati postižu udruživanjem testa inteligencije sa strukturiranim intervjuom ili testom ličnosti. Obje ove kombinacije daju visoku prognostičku valjanost, a kako su razmjerno jednostavne u provedbi, nude se kao dobro rješenje za većinu situacija u praksi. Dakako, uz pretpostavku da su testovi primjereni svojstvima pristupnika, a predlošci za intervju strukturirani na temelju analize konkretnih poslova.

Na kraju, kratak osvrt na pitanje koje se danas opravdano postavlja: Mogu li podaci validacijskih studija akumulirani tijekom proteklog stoljeća, poslužiti kao osnova za projektiranje selekcijskog procesa u današnje vrijeme velikih promjena? Naime, točno je da su na djelu krupni društveno-ekonomski procesi u čitavom svijetu, koji mijenjaju način rada, organizacijski ustroj i radne uvjete. Povećana globalna konkurencija, snažan utjecaj informacijske tehnologije, preoblikovanje poslovnog procesa, smanjivanje radnih organizacija, sve više organizacija koje nude usluge umjesto proizvoda te zamjena stabilnih i dobro određenih radnih mjesta s fleksibilnim i promjenjivim radnim ulogama – sve su to promjene koje utječu na prirodu ljudskog rada, a time potencijalno i na selekcijski

³ No premda nije obveza, validacija je i nadalje poželjna! U Međunarodnim smjernicama za upotrebu testova (ITC, 1999/2002) stoji da će kompetentni korisnici testova "kad je to moguće, validirati testove za njihovu upotrebu ili sudjelovati u formalnim validacijskim istraživanjima" (članak 2.9.5, str. 440). Čitav ovaj kolosalni skup spoznaja o valjanosti naših selekcijskih postupaka bio je i moguć zahvaljujući tome što su sposobni psiholozi, istraživači i praktičari, proveli velik broj validacija. Mislim da bi svaki ozbiljni praktičar, koji se intenzivno bavi selekcijom, svoju kompetenciju trebao potvrditi i nekim vlastitim prinosom korpusu validacijskih spoznaja.

proces (vidi npr. Cascio, 1995; Howard 1995; Herriot i Andersen, 1997). Mogu li nam u takvim okolnostima koristiti spoznaje validacijskih studija iz ranijih razdoblja?

Moj odgovor je *da*, jer dobivene spoznaje snažno podupiru postavku o generalizaciji valjanosti. Ako se valjanost može generalizirati preko različitih situacijskih okolnosti, onda je opravdano pretpostaviti da će se generalizirati i na okolnosti koje donose nove, aktualne promjene. Osim toga, kad je upravo promjena bitna odrednica stvarnosti, onda ponajviše ima smisla osloniti se na opću intelektualnu sposobnost i široke osobine ličnosti, prediktore s dokazanom generaliziranjem valjanosti.

LITERATURA

- Arvey, R.D. & Campion, J.E. (1982). The employment interview: A summary and review of recent literature. *Personnel Psychology*, 35, 281-322.
- Bartram, D. (2000). Internet recruitment and selection: Kissing frogs to find princes. *International Journal of Selection and Assessment*, 8, 261-274.
- Brown, S.H. (1981). Validity generalization and situational moderation in the life insurance industry. *Journal of Applied Psychology*, 66, 664-670.
- Burke, M.J. (1984). Validity generalization: A review and critique of the correlation model. *Personnel Psychology*, 37, 93-115.
- Chartrand, J.M., & Oliver, L.W. (Eds.) (2000). Special issue: Career assessment and the Internet. *Journal of Career Assessment*, 8 (12), 1-104.
- Cornelius, E.T., Schmidt, F.L., & Carron, T.J. (1984). Job classification approaches and the implementation of validity generalization results. *Personnel Psychology*, 37, 247-260.
- Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417-440.
- Herriot, P. & Andersen, N. (1997). Selecting for Change: How will personnel and selection psychology survive? *International Handbook of Selection and Assessment* (str. 1-34), John Wiley & Sons Ltd
- Costa, P.T., & McCrae, R.R. (1985). *Manual for the NEO Personality Inventory*. Odessa, Fla: Psychological Assessment Resources Inc.
- Gaugler, B.B., Rosenthal, D.B., Thornton, G.C. III. & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, 72, 493-551.
- Ghiselli, E.E. (1966). *The validity of occupational aptitude tests*. New York: Wiley.
- Goldberg, R.L. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48, 26-34.
- Guion, R.M. Personnel assessment, selection, and placement. U: Dunnette, M.D. & Hough, L.M. (Eds.) *Handbook of industrial and organizational Psychology*, Vol. 2, (str. 327-397). Palo Alto, Ca: Consulting Psychologists Press. 1991.
- Guion, R.M. & Gottier, R.J. (1965). Validity of personality measures in personnel selection. *Personnel Psychology*, 18, 135-164.

- Hough, L.M. & Oswald, F.L. (2000). Personnel selection: Looking toward the future – remembering the past. *Annual Review of Psychology*, 51, 631-664.
- Howard, A. (1995). *The changing nature of work*. San Francisco: Jossey Bass.
- Huffcutt, A.I., Conway, J.M., Roth, P.L., & Stone, N.J. (2001). Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology*, 86, 897-913.
- Hunter, J.E. (1980). Validity generalization for 12,000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB). Washington, D.C.: U.S. Employment Service, 1980.
- Hunter, J.E. & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Hunter, J.E. & Schmidt, F.L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Beverly Hills, CA: Sage.
- ITC (1999). *International guidelines for test use. Version 2000*. Copyright: International Test Commission. (Hrvatski prijevod: Međunarodne smjernice za upotrebu testova, *Suvremena psihologija*, 2002, 5, 421-441).
- Kuncel, N.R., Hezlett, S.A. & Ones, D.S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for student selection and performance. *Psychological Bulletin*, 127, 162-181.
- Levy-Leboyer, C. (1994). Selection and assessment un Europe. U: H.C. Triandis, M.D. Dunnette i Hough, L.M. *Handbook of industrial and organizational psychology*, Vol. 4 (str. 173-190). Palo Alto, Ca.: Consulting Psychologists Press.
- McDaniel, M.A., Whetzel, D.L., Schmidt, F.L., & Mauer, S.D. (1994). The validity of employment interview: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599-616.
- McDaniel, M.A., Morgeson, F.P., Finnegan, E.B., Campion, M.A. & Braverman, E.P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, 86, 730-740.
- Neter, E. & Ben-Shakhar, G. (1989) The predictive validity of graphological inferences: A meta-analytic approach. *Personality and Individual Differences*, 10, 737-745.
- Ones, D.S. (2001). Contributions from psychological methodology to industrial, work, and organizational (IWO) psychology. Saopćenje na 10. Evropskom kongresu psihologija rada, Prag 16.-19. svibnja.
- Ones, D., & Viswesvaran, C. (2001). Integrity tests and other criterion focused occupational personality scales (COPS) used in personnel selection. *International Journal of Selection and Assessment*, 9, 1-9.
- Ones, D., Viswesvaran, C., & Schmidt, F.L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology Monograph*, 78, 679-703.
- Ree, M.J. & Carretta, T.R. (1998). General cognitive ability and occupational performance. In C.I. Cooper & I.T. Robertson (Eds.) *International Review of Industrial and Organizational Psychology*, Vol. 13 (str. 159-186). New York: Wiley.
- Robertson, I.T. & Smith, M. (2001). Personnel selection. *Journal of Occupational and Organizational Psychology*, 74, 441-472.

- Rothstein, H.R., Schmidt, F.L., Erwin, F.W. Owens, W.A., & Sparks, C.P. (1990). Biographical data in employment selection: Can validities be made generalizable? *Journal of Applied Psychology*, 75, 175-184.
- Salgado, J.F. (1998). Sample size in validity studies of personnel selection. *Journal of Occupational and Organizational Psychology*, 71, 161-164.
- Salgado, J.F. (1999). Personnel selection methods. In C.I. Cooper & I.T. Robertson (Eds.) *International Review of Industrial and Organizational Psychology*, Vol. 14 (str. 1-54). New York: Wiley.
- Salgado, J.F. i Anderson, N. (2003). Validity generalization of GMA tests across countries in the European Community. *European Journal of Work and Organizational Psychology*, 12, 1-17.
- Salgado, J.F. & Moscoso, S. (2002). Comprehensive meta-analysis of the construct validity of the employment interview. *European Journal of Work and Organizational Psychology*, 11, 299-324.
- Schmidt, F.L. & Hunter, J.E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, 62, 529-540.
- Schmidt, F.L. & Hunter, J.E. (1981). Employment testing: Old theories and new research findings. *American Psychologists*, 36, 1128-1137.
- Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124 (2), 262-274.
- Schmidt, F.L. & Hunter, J.E. & Caplan, J. (1981). Validity generalization results for two job groups in petroleum industry. *Journal of Applied Psychology*, 66, 261-273.
- Schmidt, F.L., Hunter, J.E., Pearlman, K., & Shane, G.S. (1979). Further tests of the Schmidt-Hunter Bayesian validity generalization procedure. *Personnel Psychology*, 32, 257-281.
- Schmidt, F.L., Gast-Rosemberg, I. & Hunter, J.E. (1980). Validity generalization results for computer programmers. *Journal of Applied Psychology*, 65, 643-661.
- Smither, R.D. (1998). *The psychology of work and human performance*. New York: Longman.
- Šverko, B. (1974). Teorijske i metodološke osnove izbora kadrova. U *Postupci kod izbora i prijema kadrova*, str. 7-27. Pula: Općinski zavod za zapošljavanje.
- Šverko, B., Akik, N., Babarović, T., Brčina, A., & Šverko, I. Validity of e-advice: The evaluation of an Internet-based system for career planning. *International Journal for Educational and Vocational Guidance*, 2002, 2, 193-215.
- Šverko, B. i Galešić, M. (2003). Seleksijska praksa u Hrvatskoj 2003. Rezultati intervjuiranja prigodnog uzorka sudionika selekcijskog procesa (neobjavljeno).
- Tett, R.P., Jackson, D.N. & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703-742.

NEW INSIGHTS INTO THE VALIDITY OF SELECTION METHODS: FROM THE DOCTRINE OF SITUATIONAL SPECIFICITY TO THE IDEA OF VALIDITY GENERALIZATION

Summary

Improved techniques of meta-analysis have made possible a fruitful integration of the findings of thousands of original validity studies. This has shed new light upon the validity of psychological tests and other predictors of occupational and educational success. The observed estimates of operational validities of the main selection methods surpass our expectations and confirm the view that personnel selection is an effective procedure. It has also been found that observed operational validities are mainly situationally invariant. This supports the notion of validity generalization, i.e. the view that predictive validity for a job established in one situation would also apply for similar jobs in different employment settings. The paper reviews recent meta-analytical findings about the validity of intelligence tests, personality tests, interview and other selection methods, and discusses their implications.

Key words: personnel selection, selection methods, validity generalization, predictors

Primljeno: 20. 9. 2003.