

PREGLED EVROPSKIH IN MEDNARODNIH POBUD ZA MEDNARODNE STANDARDE UPORABE TESTOV

Dave BARTRAM

SHL Group PLC, Surrey, England

Prevedla in priredila

Duška BOBEN

Center za psihodiagnostična sredstva, Ljubljana

Povzetek: Predstavljeni so vsi glavni mednarodni, evropski in nacionalni projekti, povezani s standardi za uporabo testov: ameriški tehnični standardi, britanski certifikat za kompetentne uporabnike testov, delo in raziskava delovne skupine za teste in testiranje v evropskem združenju psihologov, evropske smernice za proces ocenjevanja*, mednarodne smernice za priredbe testov ter mednarodne smernice za uporabo testov.

KLJUČNE BESEDE: psihološki testi, testiranje, standardi, smernice, uporaba testov, priredbe testov, kompetentnost, ocenjevanje

* Za assessment (še) nimamo primerne slovenskega prevoda. Še največkrat uporabljamo ocenjevanje, ali psihološko ocenjevanje. Opredelelitev "assessmenta" je zapisana v poglavju GAP projekt EAPA.

Naslova/addresses:

- Prof. Dave Bartram, SHL Group PLC, 3 AC Court, Thames Ditton, Surrey KT7 0SR, England;
- Duška Boben, univ. dipl. psih, Center za psihodiagnostična sredstva, d.o.o., Dunajska 106, 1000 Ljubljana, Slovenija

THE NEED FOR INTERNATIONAL GUIDELINES ON STANDARDS FOR TEST USE: A REVIEW AND INTERNATIONAL INITIATIVES.

Dave BARTRAM

SHL Group PLC, Surrey, Anglija

Translated and adapted by

Dušica BOBEN

Center of Psychodiagnostics Ltd., Ljubljana, Slovenia

Abstract: *The international, european and national initiatives for international guidelines for psychological testing are presented and the need for these guidelines is discussed. Topics discussed include the development of test standards in the United States and United Kingdom, as well as projects initiated by the International Test Commission and the Task Force for Test and Testing at EFPPA.*

KEYWORDS: international guidelines, psychological testing, test standardization, professional standards, test adaptation, test use, competence, assessment

CC=2220 3450

Dave Bartram je direktor mednarodnih raziskav in razvoja pri založbi SHL, pred tem je bil profesor na Univerzi v Hullu, Velika Britanija. Ima licenco psihologa za področje zaposlitev in je član Britanskega psihološkega društva. V društvu je bil odgovoren za razvoj posebnega pristopa do standardov za teste, tj. kompetentnosti uporabnikov testov. Je član delovne skupine za teste in testiranja pri EFPPA. Trenutno vodi tudi projekt o mednarodnih smernicah za uporabo testov pri Mednarodni komisiji za teste (International Test Commission – ITC). Dave Bartram nam je za namene društva prijazno dovolil prevesti svoj prispevek, objavljen v reviji *European Psychologist*, 3 (2), ki je na straneh 155-163 izšel leta 1998. Ker se je medtem zgodilo že veliko novega, dodajamo informacije o novostih in seveda komentar, kakšno je stanje v Sloveniji.

UVOD

V večini razvitih držav pogosto uporabljajo psihološke teste, nekoliko manj v nerazvitih državah. Od vseh kategorij testov je pedagoško testiranje najbolj razširjeno (Bartram in Coyne, 1998). Večina narodov je spoznala pomembnost vlaganja v izobraževanje, saj je testiranje na področju izobraževanja ključ do merjenja potencialov in učinkovitosti. To pa je večinoma povezano z narodovimi omejenimi viri.

Prav tako je zelo razširjeno testiranje na področju zaposlovanja, posebno za izbor pri zaposlitvi in pri svetovanju za kariero. Uporabo utemeljujejo s tem, da mora biti proces racionalen in objektivni ter da je pri povezovanju potencialnega ujemanja kvalitete posameznika z zahtevami dela nepristranskost ključna pri izbiri za delo. Države, kjer tradicija selekcije za delo ne temelji na nepristranskosti (npr. sorodstvene zveze, status v družbi itn.), so manj navdušene nad dobrimi stranmi psihološkega testiranja.

Skoraj vse večje industrijske države sveta uporabljajo ali začenjajo uporabljati psihološko testiranje pri poklicnem ocenjevanju. Njegova uporaba je močnejša v tradicionalnih "zahodnih" kulturah (tj. Evropa, Severna Amerika, Avstralija itn.), vendar pa interes za testiranje in uporabo testov narašča tudi v kulturah, ki nimajo močnih korenin v psihometrični tradiciji in zahodnega stališča o kriteriju ujemanja posameznika z zahtevami dela. Naraščajoče zanimanje se je pokazalo na Srednjem vzhodu in v državah Pacifika. Globalizacija industrije je povzročila širitev testiranja in selekcije od velikih mednarodnih podjetij na male nacionalne organizacije, kjerkoli te tekmujejo na lokalnem trgu.

Med založniki testov so mnoge mednarodne organizacije, ki prodajajo svoje teste v veliko število držav. V vsaki državi verjetno različno klasificirajo uporabnike, upoštevajo zakonske omejitve uporabe testov itn. Ne le organizacije, tudi posamezniki postajajo bolj globalni. Znotraj Evropske skupnosti je uveljavljen princip svobodnega gibanja delovne sile med državami članicami. Ali tudi naša trenutna praksa ocenjevanja podpira ta princip? Kako nepristransko uporabiti npr. angleški postopek ocenjevanja pri skupini kandidatov, ki vključuje tudi posameznike iz npr. Nemčije ali Grčije? Kako oblikujemo in uporabljamo postopek ocenjevanja, da zagotovimo nepristranskost? Ko so se mnogi raziskovalci (Arvey in Faley, 1988; Eyde, Robertson in Krug, 1993; Foxcroft, 1997), založniki (Saville in Holdsworth, 1991) in državni organi (EOC, 1988; CRE, 1992, 1996 v Veliki Britaniji) ukvarjali z nepristranskostjo in dobro prakso pri selekcijah znotraj države, so

posvečali manj pozornosti mednarodnemu področju. Vendar pa Mednarodna komisija za teste (Hambleton, 1994; Van de Vijver in Hambleton, 1996) trenutno pripravlja mednarodna priporočila – smernice za priredbe testov, kar je eden od vidikov mednarodne uporabe testov.

Veliko je razlogov, zakaj so potrebne smernice za uporabo testov na mednarodni ravni:

Države se zelo razlikujejo v stopnji zakonskega nadzora (če jo sploh imajo) glede testiranja in njegovih posledic za testirane. Nekatera nacionalna psihološka društva imajo pravila o registraciji, druga ne; nekatera imajo mehanizme kontrole standardov za uporabo testov pri nepsihologih, druga ne. Obstoječe mednarodno sprejete smernice bi nacionalnim psihološkim združenjem in drugim relevantnim profesionalnim telesom ter organizacijam pomagal pri njihovih prizadevanjih za razvoj standardov, predvsem v tistih državah, kjer teh standardov še nimajo ali pa so nepopolni.

Dostop do testnih materialov (tj. pravica nakupa ali uporabe) se zelo razlikuje od države do države. V nekaterih državah je dostop dovoljen samo psihologom, v drugih so uporabniki registrirani pri formalno odobrenih nacionalnih distributerjih testov, spet drugod lahko uporabniki testov dobijo materiale od dobaviteljev v državi ali neposredno od dobaviteljev iz tujini brez omejitev.

Pred kratkim narejena raziskava (Bartram in Coyne, 1998) je pokazala, da tako na področju izobraževanja kot zaposlitve nepsihologi številčno daleč presegajo psihologe. Samo na področju kliničnega testiranja (ki pa ima glede na celotno področje testiranja relativno majhen delež) so uporabniki testov večinoma psihologi. Še več, zelo veliko uporabnikov testov je slabo usposobljenih, ali pa sploh niso (specializacije, izobraževanja) usposobljeni za to področje. To velja ne le za nepsihologe, ampak tudi za psihologe.

Kar nekaj dobro znanih instrumentov se je pojavilo na internetu brez vednosti avtorjev ali izdajateljev (kršenje avtorskih pravic) in brez zaščite testa.

Znotraj testiranja za zaposlitev je velika mednarodna mobilnost dela povečala zahteve po testih. Uporabljajo jih pri kandidatih za zaposlitev iz različnih držav in pogosto testirajo v eni državi v korist zaposlovalca iz druge države.

V ZDA in Evropi razvijajo ocenjevanje na daljavo s pomočjo interneta, tako na področju izobraževanja kot zaposlovanja. To povzroča celo vrsto težav, povezanih s standardi za izvajanje in nadzorom nad procesom testiranja, vključno z zaščito testa.

Naraščajoče število uporabnikov testov z relativno slabim usposabljanjem in vedno pogostejše mednarodno testiranje pomenijo potencialno zelo resen problem za prihodnost. To vodi k vedno slabšemu delu in s tem k diskreditiranju psihološkega in pedagoškega testiranja. Obstaja resnična nevarnost, da bo brez primerne nadzora kakovosti celotna metodologija objektivnega testiranja zavržena – ne zaradi njene neučinkovitosti ali neprimernosti, ampak zaradi napačne uporabe in zlorab.

Nacionalna psihološka društva so v svojih državah odgovorna za opredelitve standardov psihološkega testiranja. Številni članki o nacionalnih težavah s testiranjem in testi kažejo različnosti med državami in skupinami držav (Bartram, 1996; Evers, 1996; Mardberg, 1996; Camera, 1997; Davidson, 1997; Foxcroft, 1997; Gregoire, 1997; Koene, 1997; Simmer, 1997). Do nedavnega ni bilo možno sistematično primerjati standardov med državami, niti ni bilo načina, kako usposobiti države, da bi se druga od druge učile na izkušnjah.

Zadnja leta pa je postalo očitno, da morajo posamezna psihološka društva izmenjati izkušnje o vzajemnih težavah. Ni več mogoče graditi na domnevi, da je ena država zaprt sistem. Države postajajo odprti sistemi, ki je pod močnimi zunanjimi vplivi in pritiski. Kot rezultat so vzpostavljeni različni mehanizmi, ki omogočajo boljšo koordinacijo prizadevanj in razširjanje informacij na mednarodni ravni. Te mednarodne mreže lahko vidite na evropski ravni (npr. Evropsko združenje za psihološko ocenjevanje – EAPA – European Association of Psychological Assessment; Evropska federacija profesionalnih psiholoških združenj – EFPPA – European Federation of Professional psychologists' Associations) in na širši mednarodni ravni, npr. Mednarodna komisija za teste – ITC – International Test Commission.

Prispevek najprej na kratko predstavi pristop Ameriškega psihološkega združenja (APA – American Psychological Association) in Britanskega psihološkega društva (BPS – British Psychological Society) k standardom za teste. Nato govori o skupini povezanih projektov, ki se nanašajo na testiranje: o delu delovne skupine za teste in testiranje pri EFPPA, o projektu GAP pri EAPA, o ITC projektu o mednarodnih smernicah za priredbo testov. In na koncu še o projektu ITC-ja o mednarodnih smernicah za standarde pri uporabi testov.

Za namen tega prispevka je "uporabnik testa" opredeljen kot nekdo, ki

- je kompetenten, da izbere med testi na podlagi informacij v priročnikih in drugih virih in z upoštevanjem konteksta, v katerem bo test uporabljen;
- zna test izvesti;
- zna nepristransko in pravilno uporabiti rezultate testa;
- zna znotraj organizacije dati usmeritve o primernosti testiranja za različne namene.

Ta opredelitev pomeni, da ni take osebe, ki bi bila kompetentna uporabljati vse teste v vseh možnih situacijah. Namesto tega kompetentnost za uporabo testov potrebuje opredelitev, ki se nanaša na način uporabe in področje uporabnosti. Ker naj bi bili standardi in smernice napisani splošno in v vsemu prilegajočih se izrazih, je zelo pomembno, da ne pozabimo, da mora biti obravnavana njihova uporaba previdno, v izrazih kontekstualne specifičnosti. Oseba, ki je dobro usposobljena za uporabo osebnostnih testov pri selekciji in razvoju menedžmenta, je lahko povsem neprimerna za uporabo celo istih testov, npr. v klinični situaciji. Podobno oseba, ki je postala izvedenec v uporabi WISC-a v pedagoškem ocenjevanju, nima nujno kompetenc, potrebnih za uporabo podobnega inštrumenta npr. v svetovanju za kariero ali poklicno rehabilitacijo.

RAZVOJ STANDARDOV ZA TESTE V ZDA

Delo, ki so ga opravili v ZDA v zadnji polovici tega stoletja na področju standardov za teste, je imelo velik vpliv ne samo v ZDA, ampak tudi na standarde v drugih državah. Ameriško psihološko združenje (APA) je prvič objavilo *Tehnična priporočila za psihološke teste in diagnostične tehnike* (Technical Recommendations for Psychological Tests and Diagnostic Techniques) leta 1954. APA, AERA (American Educational Research Association – Ameriško združenje za raziskave v izobraževanju) in NCME (National Council on Measurement in Education – Nacionalni svet za merjenje v izobraževanju) so leta 1966 skupaj objavili *Standarde za pedagoške in psihološke teste in priročnike* (Standards for Educational and Psychological Tests and Manuals). Revizije teh so bile objavljene leta 1974 in 1985. Obliko standardov iz leta 1985 trenutno ponovno obnavlja Komite, v katerem sodelujejo predstavniki različnih ustanov, AERA, APA in NCME.

Standardi ZDA obsegajo širok spekter vsebin. Dosledno določajo tehnične standarde za konstrukcijo testov in ocenjevanja in zajemajo:

- veljavnost,
- zanesljivost in napake merjenja,
- razvoj in obnavljanja testa,
- lestvičenje, normiranje, primerljivost rezultatov in izenačevanje,
- testne izdaje, tj. tehnične pripomočke in vodnike za uporabnike.

Ta del standardov je oblikoval osnove za standarde, ki jih uporabljajo v mnogih drugih državah. Mnoga druga poglavja so lahko, z malo ali brez sprememb, uporabna za testiranje na splošno (npr. standardi za postopke izvajanja). Vendar pa je nekaj poglavij, ki se ukvarjajo z bolj specifično vsebino, povezano s situacijo ocenjevanja v ZDA (npr. poglavje o profesionalni in poklicni licenci in potrdilih). Ti standardi so trenutno najbolj uporabni vir za vsakogar, ki načrtuje razvoj standardov za uporabo v katerem koli kontekstu. Vendar pa je treba povedati, da niso razviti kot mednarodni standardi, ampak kot standardi za ZDA. Raven, do katere je vsak od standardov uporaben samo znotraj specifičnega ali omejenega kulturnega konteksta, potrebuje zelo previden razmislek, posebno kadar jih uporabimo kot osnovo za standarde v drugih državah.

Standardi ZDA so navodila za kakovostno uporabo testov, ne opredeljujejo in ne zagotavljajo pa mehanizma, kako dejansko doseči tako kakovost. Od leta 1985, ko so bili standardi objavljeni, razvijajo smernice o dobri praksi in raziskujejo načine nadziranja registracije testov. Omembe vredno delo s tega področja je delo skupine za kvalifikacijo uporabnikov testov (TUQWG – Test User Qualifications Working Group) pri JCTP (Joint Committee on Testing Practices – Združeni komite za testno prakso), ki je na podlagi podatkov promovirala dobro uporabo testov (Eyde, Moreland in Robertson, 1988; Eyde in sodelavci, 1993; Moreland, Eyde, Robertson, Primoff in Most, 1995) in delo JCTP pri razvoju in razširitvi kodeksa poštenega testiranja v izobraževanju (Code of Fair Testing Practices in Education, JCTP, 1988; Fremer, Diamond in Camara, 1989).

V Sloveniji so postali standardi širše znani preko hrvaškega prevoda iz leta 1992. Študenti psihologije poznajo standarde, saj jih obravnavajo pri predmetu Psihometrija II na dodiplomskem študiju na oddelku za psihologijo Univerze v Ljubljani. V Sloveniji smo jih želeli prevesti in uradno objaviti, vendar za to nismo dobili pravic. V ZDA pripravljajo novo verzijo standardov, ki jo bodo objavili leta 2000. V delovni skupini za teste in testiranja pri EFPPA smo se dogovorili, da bodo to verzijo poskušali prilagoditi evropskih razmeram in postaviti temelje, na katerih

bi vsako nacionalno društvo priredilo standarde zase z upoštevanjem specifičnosti držav³.

BRITANSKI PRISTOP: SPRIČEVALO O UPORABNIKOV KOMPETENTNOSTI

V Veliki Britaniji smo se odločili, da se usmerimo v način, kako doseči implementacijo standardov (npr. tistih, ki so opisani v standardih ZDA). Vodilni komite za standarde testov (SCTS – Steering Committee on Test Standards) pri Britanskem psihološkem društvu se je odločil, da razvije strategijo za reševanje obeh problemov: slabih testov in slabe uporabe testov. Ciljna skupina te strategije so bili bodoči uporabniki testov.

Pristop je temeljil na predpostavkah, da

- kompetenten uporabnik testov ne uporablja slabih testov;
- ve, kdaj testirati in kdaj ne;
- ve, kateri testi so primerni in kateri ne;
- ve, kako uporabiti teste in kako delati z ljudmi, ki so testirani.

Celotna strategija je bila usmerjena v to, kako doseči večjo kompetentnost uporabnikov pri testiranju in zanje zagotoviti boljše informacije o testih. To smo kasneje dosegli s postavitvijo Registra uporabnikov testov, z Revijo o selekciji in razvoju (Selection and Development Review), ki jo prejme vsak registriran uporabnik testov, ter s publikacijo rednih strokovnih pregledov testov.

Dosedanji certifikacijski postopek je povezan z uporabo psiholoških testov v situacijah ocenjevanja za zaposlitev. Ta vsebuje kvalifikacije na dveh ravneh:

- raven A, ki pokriva osnovna psihometrična načela in uporabo testov zmožnosti in sposobnosti in
- raven B, ki pokriva uporabo bolj kompleksnih instrumentov, posebno tistih za ocenjevanje osebnosti.

3 Medtem so novi standardi v ZDA že izšli: American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *Standards for educational nad psychological testing*. Washington, DC: American Educational Research Association.

Pristop, ki se osredotoča na kompetentnost, je radikalna sprememba. Poudarja kompetentnost uporabnikov in ne način, s katerim so kompetentnost pridobili. Opredelili smo standarde in mehanizme, ki povedo, kako oceniti kompetentnost in kako preveriti in spremljati kakovost ocenjevalnega procesa. Podrobnosti o razvoju BPS standardov in o pozitivnih spremembah pri njihovi implementaciji najdete v Bartram (1996).

Pri razvoju strategije se je britansko društvo najprej osredotočilo na psihološko testiranje v poklicnih situacijah. Razlog za ta izbor je bil pragmatičen: To je bilo področje, v katerem je uporaba testov rasla najhitreje in v katerem je bila najbolj vidna potreba po "nadzoru kakovosti". Delo se sedaj nadaljuje na razvoju osnovne ravni (A) certifikata za uporabnike testov v izobraževalnih situacijah. To bo potencialno vplivalo na veliko večje število uporabnikov testov. V Veliki Britaniji je večina uporabnikov testov v izobraževanju nepsihologov. S postavitvijo standardov in "lastništvom" nad procesom certificiranja uporabnikov je BPS postavilo jasno sliko o svoji vlogi kot telesu, ki je odgovorno za vzdrževanje dobrih standardov pri testiranju in uporabi testov. To sporočilo je seveda pomembno s praktičnega vidika, za stroko pa je pomemben tudi politični vidik.

V Sloveniji ni veliko nepsihologov, ki bi uporabljali psihološke teste. Tudi znanje diplomiranih psihologov o psihometriji in testih je dobro. Vendar se tudi pri nas pojavljajo tuja podjetja, ki naših strokovnih standardov ne poznajo, zakona o psihološki dejavnosti, ki bi bil tudi zanje obvezujoč, pa tudi še nimamo. Kljub temu bi se lahko veliko naučili iz izkušenj naših evropskih kolegov in začeli delovati preventivno: poskrbeti za standarde in ljudi (tudi nepsihologe) in jih seznaniti z njimi. Čeprav je večinsko mnenje psihologov, da je pomembnejša zakonska regulativa (raziskava DPS o testih in testiranjih, 1999), pa je načrtovanje mednarodnih kriterijev o kompetentnosti uporabnikov in kurikulumu za usposabljanje prav tako pomembno področje. To pa bo tudi ena od naslednjih nalog delovne skupine za teste in testiranja pri EFPPA, v kateri ima tudi DPS svojega člana.

DELOVNA SKUPINA ZA TESTE IN TESTIRANJE PRI EFPPA (TASK FORCE ON TEST AND TESTING – TFTT)

Evropska unija (EU) je lep primer za nastajajoče probleme v zvezi z razvojem testov in uporabniki testov. Direktive evropskega komiteja, povezane s svobodnim pretokom delovne sile in enakovrednosti kvalifikacij med evropskimi narodi, narekujejo domnevo o procesih "poštene selekcije", ki jo

bo težko zagotoviti v realnosti. Ne glede na razlike v praksi med različnimi evropskimi narodi (Shackleton in Newell, 1994) narašča uporaba testiranja sposobnosti in osebnosti za poklicno selekcijo. Na sestanku v Atenah (leta 1995) je EFPPA ustanovila delovno skupino - TFFT, ki je dobila nalogo, da pregleda vsebine, ki so povezane z razvojem evropske politike do testov in testiranja. Teme delovne skupine so:

- tehnični standardi,
- etična načela uporabe testov,
- prevodi in priredbe testov iz ene kulture (jezika) v drugo,
- usposabljanje in akreditacije uporabnikov testov – psihologov in nepsihologov,
- akreditacije testov – instrumentov – avtorske pravice in patenti,
- evropski zakoni, povezani s testi in testiranjem.

Na njenem prvem sestanku v Španiji (leta 1996) je bilo dogovorjeno, da bo glavna dejavnost delovne skupine zbrati podatke o testih in testiranju v Evropi. Da bi to dosegli, je Dave Bartram soglašal, da pripravi in razpošlje obsežen vprašalnik vsem državam članicam EFPPA. Ta vprašalnik se je nato razširil tudi na države, ki niso članice EFPPA, z namenom dobiti podatke tudi za druge naloge pri ITC.

Vprašalnik je bil namenjen profesionalnim združenjem, založnikom in drugim skupinam ali organizacijam, s ciljem zbrati nacionalne poglede na uporabo testov v posameznih državah. Razdeljen je bil na dva dela in se je nanašal na uporabo testov na štirih področjih dela: kliničnem, forenzičnem, pedagoškem področju in področju dela.

Dva glavna dela sta vsebovala:

1. del

- število uporabnikov testov, področja in načine uporabe testov;
- uporabo psiholoških testov v raziskavah in ocenjevanju;
- usposabljanje uporabnikov testov in vlogo profesionalnega združenja (ali združenj), ki jo ima pri pripravi in/ali nadziranju kakovosti usposabljanja, povezanega z uporabo testov;
- sistem kategorizacije testov, nabave testov in testnega materiala, dostopnost priredb/prevodov, dostop do informacij o kakovosti testov;
- razširjenost lokalnih pripomočkov in ekspertiz o psihološkem testiranju in psihometriji;
- kodeks dobre prakse in smernice za uporabo testov;

- tehnične standarde testov, minimalne standarde za teste, in če so, kako so upoštevani;
- zakonski kontekst: statutarni (zakonski) nadzor, povezan s testi in testiranjem, druga kontrola;
- vpletenost psihologov in nepsihologov v proces testiranja.

2. del

- Indikatorje pomanjkanja kompetentnosti pri uporabi testov. Tistim, ki so odgovarjali, so bila zastavljena vprašanja o 86 različnih dogodkih – elementih, ki so jih zbrali Eyde in sodelavci (1993) v svoji analizi kritičnih dogodkov, povezanih z uporabo testov. Vsak je povezan s problemom glede uporabe testov.

Vprašalnik je bil pripravljen v angleščini, španščini, francoščini in nemščini. Skupaj 127 vprašalnikov je bilo poslanih v 48 držav. Od teh je bilo 30 evropskih držav (25 članic EFPPA), druge pa so predstavljale vzorec ITC. Še 41 vprašalnikov je bilo nato uporabljenih v projektu ALFA, ki pokriva špansko govoreče države v Evropi in Srednji ter Južni Ameriki. Devetinpetdeset (46,5 %) vprašalnikov je bilo iz 30 držav (62,5 % vzorca). Od teh je 20 evropskih držav. Kasneje je iz ALFA projekta prišlo še 11 (20, 8 %) vrnjenih vprašalnikov, kar predstavlja 9 držav. Od teh sta bila Peru in Portugalska že zastopana v ITC oz. EFPPA vzorcu.

Od vseh 168 razposlanih vprašalnikov se jih je pravočasno vrnilo 70 (41,7 %). Glede na dolžino in kompleksnost vprašalnika je to zelo dobro. Celotno število držav, ki so sodelovale v raziskavi, je 36. Čeprav se zdi število majhno, ne smemo pozabiti, da so bili tisti, ki so odgovarjali, izbrani previdno in da so to bili tisti, ki so zmožni posredovati zanesljive informacije ali pa lahko zberejo zanesljive informacije od drugih. Velikokrat so izpolnjevalci zapisali, da so odgovori sestavljeni iz skupinskih odgovorov.

Analiza se je osredotočila na evropske podatke (Bartram in Coyne, 1998). Poročilo o celotnih podatkih je bilo pripravljeno za člane delovne skupine EFPPA in člane sveta ITC. Informacije iz te raziskave se bodo uporabile za delo na ITC smernicah o uporabi testov.

Tudi slovensko društvo psihologov je sodelovalo pri tej raziskavi. Rezultati in še posebej primerjava s slovenskimi odgovori so bili predstavljeni na 2. kongresu slovenskih psihologov v Portorožu 1997. Delovna skupina je leta 1998–99 izvedla še eno raziskavo na isto temo. Tokrat je želela dobiti mnenje od članov psiholoških društev iz različnih evropskih držav. Vse države še niso izpeljale raziskave do konca. Slovenske rezultate, primerjane s španskimi, angleškimi in hrvaškimi, ste lahko

slišali na letošnjem 3. kongresu slovenskih psihologov v Portorožu oktobra 1999 in bodo objavljeni v eni naslednjih številke Psiholoških obzorij.

GAP PROJEKT EAPA

Delovna skupina za razvoj smernic za proces ocenjevanja (Guidelines for the assessment process – GAP) je imela prvi sestanek v Lizboni leta 1997. Usmerjenost te skupine je nekoliko drugačna od tistih iz drugih projektov, ki se osredotočajo bolj ozko na testiranje. Razliko med testiranjem in ocenjevanjem (assessment) je dobro razložila Fernandez Ballesteros (1997a, 1997b) in to lahko strnemo v naslednje:

- Testiranje je primarno usmerjeno na merjenje, medtem ko je ocenjevanje (assessment) usmerjeno na problem ali področje.
- Testiranje se sklicuje na standardizirane (ali formalizirane) instrumente za merjenje, medtem ko se ocenjevanje (assessment) sklicuje na širok spekter postopkov zbiranja podatkov in na celostno zbiranje podatkov in procesov odločanja, v katerih je testiranje lahko en del.
- Ocenjevanje (assessment) se osredotoča na opis in analize posameznika v določeni situaciji, medtem ko je težnja testiranja osredotočena na karakteristične skupine in podskupine posameznikov.
- Testiranje lahko zahteva temeljite (poštene, nepristranske) specifične veščine in kompetentnosti, medtem ko ocenjevanje (assessment) temelji na širokem obsegu psiholoških in tudi drugih znanj in ekspertiz.

Ta delovna skupina dela v tesni povezavi z EFPPA in ITC projekti, da bi zagotovila uspešen rezultat, ki bi vplival tudi na nekatere mednarodne dogovore.

PROJEKTI MEDNARODNE KOMISIJE ZA TESTE (ITC)

ITC je "združenje nacionalnih psiholoških društev, testnih komisij, založnikov testov in drugih organizacij, da bi promoviralo učinkovito testiranje in politiko ocenjevanja in pravičen razvoj, evalvacijo in uporabo pedagoških in psiholoških inštrumentov" (ITC Cilj, ITC Directory, 1995).

Med različnimi dejavnostmi je ITC odgovorno za številne mednarodne projekte. Pred nekaj leti je sprejelo politiko osredotočanja pozornosti na tista področja, kjer je mednarodna koordinacija najbolj pomembna. Kot posledica tega sta se razvila dva glavna projekta. Eden se nanaša na smernice pri priredbi testov, drugi na smernice za uporabo testov.

Leta 1999 je član ITC postalo tudi Društvo psihologov Slovenije in pridružen član Center za psihodiagnostična sredstva, ki je tudi že prej sodeloval na ITC delavnicah. Tako bomo imeli v Sloveniji redne in tekoče informacije o svetovnih dogajanjih na tem področju in seveda tudi možnost, da sami kaj prispevamo.

SMERNICE ZA PRIREDBE TESTOV

Trinajstčlanski svet, ki predstavlja številne mednarodne organizacije, je v zadnjih nekaj letih pripravljal natančen sklop smernic za priredbe psiholoških in pedagoških testov za uporabo v različnih jezikovnih in kulturnih sredinah (Van de Vijver in Hambleton, 1996). To je zelo pomembno področje, saj teste uporabljajo v vedno več državah in velikokrat se test razvije v eni državi in potem prevede za uporabo v drugi državi.

Hambleton (1994) je opisal projekt v podrobnostih in kot rezultat predstavil 22 smernic. Te smernice lahko združimo v 4 glavne kategorije:

- tiste, ki govorijo o kulturnem kontekstu;
- tiste, ki zadevajo razvoj in priredbo instrumentov s tehničnega vidika;
- tiste, ki zadevajo testno izvedbo;
- tiste, ki zadevajo dokumentacijo in interpretacijo.

Vse posredno zadevajo uporabo testov in uporabnike.

MEDNARODNE SMERNICE – STANDARDI ZA UPORABNIKE TESTOV

Različni evropski projekti in tudi širše zbiranje informacij o tem so bili vir za začetek razvoja osnovnih standardov in smernic ter splošne politike, povezane z etično in strokovno uporabo testov. Glavna tema tega ITC projekta so standardi za dobro uporabo testov in pa vzpodbujanje k dobri praksi pri psihološkem in pedagoškem testiranju. Delo, ki ga je Mednarodna komisija za teste (ITC) do sedaj opravila s promocijo dobre prakse pri

priredbah testov, je pomemben korak naprej k zagotavljanju poenotenja kakovosti priredb testov med različnimi kulturami in jeziki. Kakor koli, pri psihološkem testiranju gre za ključni zadevi. Prvič, zagotoviti moramo, da dosega testi zahtevane minimalne tehnične standarde kakovosti. Drugič, vedeti moramo, da so ljudje, ki jih uporabljajo, kompetentni. V več pogledih je uporabnikova kompetentnost bolj pomembna. Kompetenten uporabnik ne ve samo, kako izvesti testiranje, vrednotiti in pravilno interpretirati, ampak mora biti tudi sposoben ločiti med primernimi in neprimernimi instrumenti.

Na sestanku v Atenah leta 1995 je Svet ITC sprejel predlog Dava Bartrama, da pripravi projekt, ki bi bil usmerjen v

- kontrolo in regulacijo dostopa do testnega materiala,
- smernice za pošteno in etično uporabo testov,
- standarde za usposabljanje in določitev kompetentnosti uporabnikov testov.

Dolgoročni cilj tega projekta je izdelava sklopa smernic, ki so povezane s kompetentnostjo (znanjem, veščinami, sposobnostmi in drugimi osebnostnimi lastnostmi) uporabnika testov. Te kompetence so podrobno opredeljene v sestavljenih praktičnih kriterijih skupaj s splošnim znanjem, ki jih človek potrebuje za kompetentno uporabo testov. Te kompetence pokrivajo:

- profesionalne in etične standarde testiranja,
- pravice kandidata (testiranca) in drugih udeležencev v procesu testiranja,
- ovrednotenje testov in njihovo izbiro,
- testno izvajanje, vrednotenje in interpretacijo,
- pisanje poročila in povratna sporočila.

Sedanji osnutek smernic predstavlja delo strokovnjakov za psihološko in pedagoško testiranje (psihologi, psihometri, založniki testov, konstruktorji testov) iz različnih držav. Namen tega dokumenta NI izumiti novih smernic, ampak zapisati splošne skupne vsebine, ki se pojavljajo v obstoječih smernicah, kodeksih, standardih in drugih primernih dokumentih, in ustvariti koherentno strukturo, ki bi jo vsakdo razumel in uporabljal.

Razvoj osnutka smernic: Smernice naj bi bile osnovne, skupne točke, s katerimi bi se lahko primerjali obstoječi lokalni standardi in iz katerih bi se lahko razvili bolj specifični standardi. Z uporabo smernic kot baze za razvoj lokalnih dokumentov (standardi, kodeksi, izjave o pravicah testirancev itn.) bi bila zagotovljena visoka raven konsistentnosti med narodi.

Priprava smernic se je začela z združevanjem velikega števila gradiv v zvezi s standardi, kodeksi, uporabo testov itn. iz številnih držav. Predvsem pa smo se opirali na:

- Standarde za pedagoško in psihološko testiranje [The American Educational Research Association (AERA), American Psychological Association (APA) in National Council on Measurement in Education (NCME) (1985)],
- Odgovornosti uporabnikov standardiziranih testov [American Association for Counseling and Development (AACD) (Schafer, 1992)],
- Dodatek smernicam za uporabo psiholoških testov [The Australian psychological Society (APS) (Kendall in sodelavci, 1997)],
- Raven A in raven B standardov za uporabo testov za zaposlitev [The British Psychological Society (BPS) (Bartram, 1995; 1996)],
- Smernice za pedagoško in psihološko testiranje [The CPA – Canadian Psychological Association, 1987].

Še posebno pomemben je bil APS-ov dokument, ki je združeval veliko tega, kar najdemo v britanski in ameriški publikaciji kot tudi publikaciji Južnoafriškega nacionalnega inštituta za psihološke raziskave (National Institute for Psychological Research – NIPR) in različnih založniških smernicah za uporabo testov. V APS je zapisano tudi veliko tega, kar so naredili v delovni skupini za kvalifikacijo uporabnikov testov (TUQWG – Test user Qualifications Working Group) pri JCTP.

Analizirali smo vsebino prvega zapisa in izjave kategorizirali v 14 glavnih naslovov. Kjer je bilo potrebno, smo zato, da bi ujeli pravi pomen iz številnih različnih virov, zapisali kar posamezno izjavo. Izjave smo tudi spremenili v običajno obliko (npr.: kompetenten uporabnik testov *si bo prizadeval...* ali kompetenten uporabnik testov *lahko...*).

Ta začetna struktura 14 glavnih poglavij in njihovih vsebin je bila predstavljena v prvem osnutku dokumenta (Guidelines Framework Document). To so bila gradiva za mednarodno delavnico julija 1997 v Dublinu. Namen te ITC delavnice je bil pretehtati in kritično oceniti vse vidike okvirnega dokumenta, da bi naredili osnutek smernic, ki bi imele mednarodno vrednost in sprejemljivost. Udeleženci delavnice so dokument natančno pregledali in predlagali spremembe glede oblike, strukture in vsebine. Po delavnici smo dokument predelali (različica 2.0.) in razposlali vsem tistim, ki so bili pripravljeni dati pripombe. Osnutek dokumenta

pregledanih smernic (Guidelines Consultation Document - različica 3.1.) je vseboval že vse komentarje in pripombe na različico 2.0.

Kopije različice 3.1. pregledanih smernic in strukturiran list za odgovore smo poslali velikemu številu posameznikov in organizacij, da bi dobili nadaljnje komentarje. Vseh poslanih dokumentov je bilo 200. Dobili smo 28 podrobnih odgovorov, skupaj z združenimi odgovori iz APA, BPS in nekaterih drugih evropskih strokovnih združenj. Poleti 1988 smo smernice priredili glede na pripombe in ponovno poslali 200 kopij različice 4.1. v pregled. V tem drugem krogu smo dobili 18 formalnih odgovorov. Preko elektronske pošte ali osebno na sestankih pa smo prejeli tudi veliko neformalne podpore. Veliko truda smo vložili v pripravo sedanje različice smernic (5.0.), saj smo želeli upoštevati vse prejete odgovore. Brez izjeme so bili odgovori konstruktivni in v pomoč.

Namen smernic je pomoč in ne prisila. Zagotoviti moramo, da take smernice predstavljajo univerzalna ključna načela dobre uporabe testov brez namere, da bi prisilile k zakonski enotnosti razlik v funkcioniranju in praksi med državami ali med področji uporabe.

Predlagana struktura loči 3 glavne vidike kompetentnosti:

- Profesionalni in etični standardi dobre prakse, ki vplivajo na način, kako je proces testiranja izvajan, in na način, kako uporabnik testov ravna z drugimi, vključenimi v proces.
- Znanje, razumevanje in veščine, povezane s procesom testiranja: kaj mora biti uporabnik testov sposoben narediti.
- Znanje in razumevanje, potrebno za obveščanje in za podporo procesa testiranja.

Te tri komponente se v praksi prepletajo. Vendar je vseeno dobro, da razlikujemo med njimi.

Smernice se začenjajo z glavnim namenom, ki ga lahko označimo kot izjavo o poslanstvu uporabnikov testov. Ta zagotovi smer za vse, ki sledijo smernicam. O vsakem vidiku kompetentnosti uporabnikov testov je potrebno razmisliti, kako zagotavlja ta glavni namen.

Osnovni cilj smernic je izboljšanje testiranja z opisom kompetentnosti uporabnikov testov. Kompetenten uporabnik testov bo uporabljal teste pravilno, strokovno, etično, upošteval potrebe in pravice tistih, ki so vpleteni v proces testiranja, upošteval, zakaj testira in tudi širši okvir testiranja. To bomo dosegli ob zagotovitvi, da ima uporabnik testov vse potrebne

kompetentnosti, da izpelje proces testiranja, in znanje, razumevanje in informiranje o testih in testiranju. Izjava o obsegu opiše, skupaj s ključnim namenom, za koga so smernice, na katere oblike ocenjevanja se nanašajo in kakšen je kontekst ocenjevanja.

Podrobne smernice so predstavljene v dveh poglavjih:

- podroben opis kompetentnosti uporabnika testov v odnosu do etične uporabe testov,
- podroben opis kompetentnosti uporabnika testov v odnosu do dobre prakse pri uporabi testov.

Smernice so bile prečiščene skozi številne kroge konzultacij. Predstavljajo poskus združevanja vsebine zelo širokega spektra virov in predstavitev s skupno strukturo in standardno obliko.

Pripravljena je bila že zadnja različica smernic, imenovana verzija 2000. Opravljenih je bilo le nekaj manjših popravkov. ITC je smernice že potrdil, prav tako jih je sprejela tudi EFPPA. Sedaj pripravljajo pravila o prevodih in razširjanju. Uradno bodo smernice objavljene leta 2000. Tudi slovensko društvo (zanj Komisija za psihodiagnostična sredstva) končuje prevod smernic. Tako smo na 3. kongresu slovenskih psihologov v Portorožu 1999 prvič predstavili osnutek prevoda smernic. V načrtu pa imamo, da smernice kmalu po uradnem izidu originalnih izdamo tudi v slovenskem jeziku..

ZAKLJUČEK

Namen ITC projekta smernic standardov za uporabo testov je bil, da skupaj zapiše rezultate različnih nacionalnih in mednarodnih pobud. Odgovori, ki smo jih prejeli na ITC konzultacijsko listino, so bili zelo pozitivni. Obstaja zelo močna potreba po mednarodnem konsenzu in zagotavljanju konsistentnosti dobre prakse pri uporabi testov v vseh državah.

Delo ITC bo zagotovilo generične standarde (izvleček iz različnih virov: BPS, APA, EFPPA, EAPA in druga evropska združenja), ki uresničujejo univerzalne principe dobre prakse v psihološkem testiranju. Takšen sklop smernic lahko zagotovi skupno bazo, na kateri lahko posamezne države in posamezne specializacije znotraj države zgradijo svoje standarde, ki so prilagojene določenemu kulturnemu, socialnemu in političnemu kontekstu.

Opredelitev smernic vedno vključuje iskanje ravnotežja med svobodo in odgovornostjo. Moramo biti previdni in ne preveč strogi, da ne zavremo eksperimentiranja pri uporabi novih metod ocenjevanja. Idealizem moramo obložiti tudi z realizmom: smernice naj bodo delovne (praktične) in naj opredelijo zdravo, strokovno delo znotraj sveta, kakršen je. Medtem ko je pravilno, da smernice za teste in testiranja vsebujejo močne etične imperativne (na primer pravico posameznika, da je poučen, preden pristane na testiranje), naj ne bi bile oblikovane le tako, da služijo strogo političnim namenom ali narekujejo družbeno akcijo. Naloga smernic in iz njih izhajajočih (nacionalnih ali mednarodnih) standardov je pomoč pri strokovni kakovosti ocenjevanja, ki se kaže v tem, da je primeren, natančen in pošten. Zagotavlja mehanizem, s katerim profesionalna psihološka združenja lahko opravijo svojo dolžnost, ki je skrb za ljudi.

LITERATURA

- American Educational Research Association, American Psychological Association in National Council on Measurement in Education (1985). *Standards for educational and psychological testing*. Washington DC: American Psychological Association.
- Arvey, R.D. in Faley, R.H. (1988). *Fairness in selecting employees*. Reading, Massachusetts: Addison Wesley.
- Bartram, D. (1996). Test qualifications and test use in the UK: The competence approach. *European Journal of Psychological Assessment*, 12, 62-71.
- Bartram, D. (Ur.) z Anderson, N., Kellett, D., Lindley, P.A., in Robertson, I. (uredniki svetovalci). (1995). *Review of personality assessment instrument (Level B) for use in occupational settings*. Leicester: BPS Books.
- Bartram, D. in Coyne, I. (1998). *The ITC/EFPPA survey of testing and test use within Europe*. Proceeding of the British Psychological Society's Occupational Psychology conference, Eastbourne, January.
- Bartram, D., Lindley, P.A. in Foster, J. (1992). *Review of psychometric tests for assessment in vocational training*. Leicester: BPS Books.
- Camara, W. (1997). Use and consequences of assessments in the USA: Professional, ethical and legal issues. *European Journal of Psychological Assessment*, 13, 140-152.
- Canadian Psychological Association (1987). *Guidelines for educational and psychological testing*. Ottawa: Canadian Psychological Association.
- Commission for Racial Equality (1992). *Psychometric tests and racial equality: A guide for employers*. London: Commission for Racial Equality.

- Commission for Racial Equality (1996). *Selecting train drivers at British rail: A fair test?* London: Commission for Racial Equality.
- Davidson, G. (1997). The ethical use of psychoogical tests: Australia. *European Journal of Psychological Assessment*, 13, 132-139.
- Equal Opportunities Commission (1988). *Avoiding sex bias in selection testing: Guidance for employers*. Manchester: Equal Opportunities Commission.
- Evers, A. (1996). Regulations Concerning Test Qualifications and Test use in the Netherlands. *European Journal of Psychological Assessment*, 12, 153-159.
- Eyde, L.D., Moreland, K.L. in Robertson, G.J. (1988). *Test user qualifications: A data-based approach to promoting good test use*. Report for the Test User Qualifications Working Group. Washington, DC: American Psychological Association.
- Eyde, L.D., Robertson, G.J. in Krug, S.E. (1993). *Responsible test use: Case studies for assessing human behaviour*. Washington, DC: American Psychological Association.
- Fernandez-Ballesteros, R. (1997a). Task force for the development of guidelines for the assessment process (GAP). *The International Test Commission Newsletter*, 7, 16-20.
- Fernandez-Ballesteros, R. (1997b). Guidelines for the assessment process. *European Psychologist*, 2, 352-355.
- Foxcroft, C.D. (1997). Psychological testing in South Africa: Perspectives regarding ethical and fair practices. *European Journal of Psychological Assessment*, 13, 229-235.
- Fremer, J., Diamond, E.E. in Camara, W.J. (1989). Developing a code of fair testing practices in education. *American Psychologist*, 44, 1062-1067.
- Gregoire, J. (1997). Regulation of testing practice in European French-speaking countries. *The International Test Commission Newsletter*, 7, 6-13.
- Hambleton, R. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10, 229-244.
- ITC Directory. (1995). International Test Commission.
- Joint Committee on Testing Practices (1988). *Code of fair testing practices in education*. Washington, DC: Joint Committee on Testing Practices.
- Kendall, I., Jenkinson, J., De Lemos, M. in Clancy, D. (1997). *Supplement to guidelines for the use of psychological tests*. Australian Psychological Society.
- Koene, C.J. (1997). Tests and professional ethics and values in European Psychologists. *European Journal of Psychological Assessment*, 13, 219-228.
- Mardberg, B. (1996). Improper use of tests in Sweden. *International journal of Selection and Assessment*, 4, 146-150.

- Moreland, K.L., Eyde, L.D., Robertson, G.J., Primoff, E.S. in Most, R.B. (1995). Assessment of Test user qualifications: A research-based measurement procedure. *American Psychologist*, 50, 14-23.
- Saville and Holdsworth Ltd. (1991). *Equal opportunities guidelines for best test practice in the use of personnel selection tests*. Surrey: Author.
- Schafer, W.D. (1992). *Responsibilities of users of standardized tests: RUST Statement Revised*. Alexandria, VA: American Association for Counseling and Development.
- Shackleton, V. in Newell, S. (1994). European management selection methods: A comparison of five countries. *International Journal of Selection and Assessment*, 2, 91-102.
- Simner, M.L. (1996). Recommendations by the Canadian Psychological Association for improving the North American Safeguards that help protect the public against test misuse. *European journal of Psychological Assessment*, 12, 72-82.
- Van de Vijver, F. in Hambleton, R. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1, 89-99.