

Preverjene poligrafske tehnike - analiza

Polona Selič

Članek predstavi značilne razlike med psihološkimi in poligrafskimi testi ter opisuje merske značilnosti obojih (veljavnost, objektivnost, zanesljivost, občutljivost). Prvič v relevantni strokovni literaturi na Slovenskem je prikazana celovita taksonomija poligrafskih tehnik. Skupna veljavnost posameznih poligrafskih tehnik je predstavljena v sumarni analizi dostopnih (objavljenih) validacijskih študij minulih desetletij, ki so zadevale poznane/uveljavljene poligrafske tehnike z ocenami njihove zanesljivosti (veljavnosti), posebej za iskrene testirane osebe in one, ki so zavajale, ter z določenim zunanjim neodvisnim merilom »poznane resnice«. Vse poligrafske zapise so ovrednotili poligrafski preiskovalci po preverjenih pravilih vrednotenja, brez uporabe računalniških algoritmov vrednotenja.

Pregled rezultatov (tabele 4–8) pokaže, da dokaznim standardom ustreza samo Utah ZCT, vojaški ZCT pa praga ne dosega zgolj za eno odstotno točko. Za preiskovalna poligrafranja je dopustna raba vojaškega ZCT, Reidove tehnike (CQT), tehnike prikritih informacij (CIT ali GKT) ter testa R-I. Ob tem je nujno opozoriti, da lahko v prihodnje pričakujemo tudi drugačne rezultate validacijskih študij, saj so se pravila odločanja uveljavila kot preverjen dejavnik kakovostne izvedbe psihofiziološke preiskave verodostojnosti izjav s poligrafom pri enotematskih testih, ki vpliva na točnost poligrafskega rezultata. Skupna povprečna veljavnost (tabela 9) morda pove manj o metodah kot pa o do sedaj opravljenem validiranju teh metod. Ne zadeva tudi znanstveno preverjenih principov izvedbe psihofiziološke preiskave verodostojnosti izjav s poligrafom, na katere je ponovno opozorjeno.

Ključne besede: veljavnost poligrafskih testov, taksonomija poligrafskih pristopov k odkrivanju zavajanja, testi s primerjalnim vprašanjem, testi prikritih informacij, pravila odločanja, točnost poligrafske preiskave, poligraf, psihofiziologija

UDK: 343.985.3:343.144

1 Psihološki in poligrafski testi

Sodobna psihologija večinoma temelji na pozitivistični filozofiji znanstvenega raziskovanja in hipotezi empirično preverja. Pri tem (po zgledu naravoslovnih disciplin) uporablja kvantitativne in kvalitativne metode. Redki, predvsem humanistični psihologi, na svojem raziskovalnem področju povsem dvomijo v veljavnost (uporabnost) rezultatov kvantitativnih metod. Nekateri koncepti izhajajo iz fizioloških pojmov (delovanje centralnega in vegetativnega živčnega sistema), drugi temeljijo na fenomenologiji ali na teoriji procesiranja informacij. Psihologija se v obravnavi intrapsihičnih procesov, ki jih ni mogoče neposredno opazovati (o njih lahko le sklepamo), najbolj razlikuje od družboslovnih disciplin (sociologije, antropologije, ekonomije in politologije). Zato je zgodovina psihologije v veliki meri določena z razvojem metod sistematičnega eksperimentalnega in neeksperimentalnega (v ožjem, psihofiziološkem pomenu) preučevanja, ki izhajajo iz opazovanja ter skušajo preseči omejitve

eksperimentalnega (psihofiziološkega) pristopa (Selič, 2007).

Sistematično spremljanje psihičnih procesov in zakonitosti v psihologiji upošteva zahteve po ciljnosti, načrtnosti, sistematičnosti, kontroli in objektivnosti. Tako gre za opredeljeni cilj/objekt opazovanja/merjenja, vnaprej določeni postopek, spremljanje vseh spremenljivk, ki lahko vplivajo na objekt merjenja, nadzorovane pogoje in postopke, pri katerih na rezultat ne morejo vplivati želje, predsodki in druge osebnostne značilnosti izvajalcev. V psihologiji je test specifično psihološki eksperiment ali poskus – osebi postavimo eno ali več nalog, iz rešitev/odgovorov pa sklepamo na (merjeno) lastnost. Test je tako preskus posameznikovega stabilnega vedenja, ki temelji na standardiziranih normah, na podlagi katerih lahko primerjamo dosežek ene osebe z dosežkom kriterijske skupine. S testom identificiramo individualne razlike, zato je uporaba psiholoških testov prvenstveno diferencialna (in s tem korelacijska) metoda.

Poligrafski testi, ki imajo korenine v psiholoških testih, so sestavljeni iz skrbno in natančno izbranih vprašanj v skladu s predpisanimi testnimi formati v posameznih poligrafskih tehnikah. Značilnost (posebnost) poligrafskih testov so enopomenski odgovori, prisilna izbira med »da« oziroma »ne«. Tudi

* Polona Selič, dr. znanosti, Katedra za družinsko medicino MF UL, Inštitut PARES

vprašanja v poligrafskem testu imajo funkcijo »dražljajev/stimulusov«, ki naj bi (pri testirani osebi) izzvali spremembe v delovanju (fiziološke spremembe). Odzivi se primerjajo z drugimi odzivi posameznika (na vprašanja drugačne vrste – nevtralna, relevantna, irelevantna, primerjalna/kontrolna ipd.) v isti testni situaciji.

Poligrafski testi za razliko od običajnih psiholoških testov niso povsem standardizirani, ampak se oblikujejo (sestavljajo) za vsak primer in vsakega posameznika posebej. To pomeni, da so poligrafski testi individualizirani ob uporabi uveljavljenih in preverjenih testnih formatov – vrsta vprašanja na posamezni poziciji je predpisana, vsebina je prilagojena obravnavanemu primeru in testirani osebi (njenemu besedišču, zmožnosti razumevanja ipd.). Kljub temu pa se tudi poligrafski testi oziroma vprašanja, ki jih zastavljamo testirani osebi, podrejajo psihološki, logični in semantični analizi. Poligrafski testi tako omogočajo preverjanje splošnih dejstev in/ali postavljanje posebnih vprašanj, ki so vezana izključno na posameznika oziroma na dejanje (ali kaznivo dejanje) (Abrams in Abrams, 1993:110).

Psihofiziološke tehnike preverjanja verodostojnosti izjav s poligrafom in psihološki testi so sestavljeni iz zaporedja procesov/postopkov – izvajanje (administracija) testa – predstavitev z navodilom in zapisovanje/beleženje odgovorov, vrednotenje – pretvorba surovih podatkov v (standardne ali točkovne) vrednosti ter razlaga (interpretacija) z diagnostičnimi ugotovitvami. Pri psiholoških testih je možna še primerjava individualnega rezultata z reprezentativnimi poprečji normirane populacije. Poligrafski testi imajo specifičen namen, zato so vsebinsko zasnovani posebej za testiranca. Pri tem so upoštewane njegove osebne lastnosti in značilnosti (obravnavanega) dejanja.

Druge razlike med psihološkimi in poligrafskimi testi so v merskih karakteristikah:

Veljavnost testa je natančnost, s katero test meri prav tisto lastnost, ki naj bi jo meril ali ki jo nameravamo meriti. Test je docela veljaven, če točkovna vrednost, ki smo jo s testom določenih lastnosti dobili, testirano osebo na normirani lestvici teh lastnosti enoumno postavi na določeno mesto. Razlikujemo vsebinsko, sestavno, sočasno in napovedno veljavnost (Selič, 2007).

Pri poligrafskih testih je pomembno odkrivanje zavajanja ali poznavanja okoliščin. Torej je veljaven (natančen) tisti test oziroma testni format, s katerim lahko razlikujemo osebe, ki zavajajo (ali poznajo okoliščine obravnavanega dogodka), od onih, ki govorijo resnico (»naivne« testirane osebe, neplete ne v obravnavani dogodek). Relevantni dražljaj je korektno formulirano vprašanje, izrečeno kar se da nevtralno, enako-

merno in brez posebnih poudarkov. Če vsa vprašanja v poligrafskem testu niso enako zastavljena, je lahko odziv testirane osebe vezan na način spraševanja in ne na vsebino. V takšnih primerih lahko gre za učinke presenečenja ipd. (Abrams, 1989:188–190; Kircher in Raskin, 2002:287–288).

Objektivnost testa je stopnja neodvisnosti rezultatov in njihove interpretacije od ocenjevalca. Test je docela objektiv, kadar različni ocenjevalci dobijo povsem enake rezultate (t. i. medosebna skladnost ocenjevalcev). Objektivnost zjema in zadeva izvajanje (standardizacija pogojev), vrednotenje in interpretacijo. Slednja je največja pri normiranih testih storilnosti (Selič, 2007).

Poligrafski test se objektivnemu približa le, ko se dobljeni rezultati (reakcije) ocenijo na način, ki izključuje subjektivno oceno (presojo) izpraševalca in drugih spremenljivih dejstev. Zato morajo biti vprašanja v poligrafskem testu pravilno strukturirana in postavljena na način, ki izključuje dvoumne reakcije. Pri poligrafskih testih je pomemben način izvedbe (monoton, enakomeren način spraševanja, upoštevanje predpisanih presledkov) in način vrednotenja poligrafskih zapisov, ki je predpisan. Uporaba sedemstopenjske ocenjevalne lestvice in preverjenih pravil odločanja pri testih s primerjalnimi vprašanji sta del standardnega testnega postopka (Abrams, 1989:47–48).

Zanesljivost je stopnja stabilnosti, s katero test meri vedenje ali lastnosti testiranca. Povsem zanesljiv test daje v različnih okoliščinah iste rezultate. Zanesljivost lahko povečamo s podaljševanjem testa (povečamo obseg in število nalog), z zvečanjem občutljivosti in izboljšanjem objektivnosti (Selič, 2007).

Zagotavljanje zanesljivosti s podaljševanjem je pri poligrafskih testih omejeno, saj je dolžina celotnega testiranja (aplikacije vseh testov s ponovitvami) odvisna od količine razpoložljivih (prikritih) informacij ter drugih psiholoških in fizioloških dejstev (elementi dejanja/dogodka/KD, adrenalinska izčrpanost ipd.). Nedvomno pa je s kar najbolj standardiziranim načinom izvedbe poligrafskih testov možno doseči, da beležimo odzive na vprašanja o dogodku/dejanju (= predmet merjenja) in ne na delovanje poligrafskega preiskovalca ali kakršnihkoli drugih neidentificiranih intervenirajočih dejavnikov (Abrams, 1989:199–201, Kircher in Raskin, 2002:309–316).

Občutljivost testa se kaže v razpršitvi rezultatov. Čim večja je razpršitev (čim večje so individualne razlike med ocenjevanimi osebami), tem večja je občutljivost testa. Razpršitev rezultatov je odvisna predvsem od dolžine testa in težavnosti nalog (pojmem težavnosti ni psihološki, marveč statistični: čim večje je število rešitev, tem »lažja« je naloga) (Selič, 2007).

Občutljivost poligrafskega testa ni povezana z dolžino vprašalnika oziroma z dolžino vprašanj. Da s poligrafskim testom identificiramo osebe, ki dajejo lažne odgovore na postavljena vprašanja, so pomembna tista relevantna vprašanja, s katerimi lahko zaznamo specifične reakcije posameznika, ter njihova primerjava s primerjalnimi vprašanji, kadar gre za teste s primerjalnimi vprašanji (Selič, 2004:368). V postopku vrednotenja v glavnem opisujemo štiri kategorije reagiranja po posameznih parametrih – brez reakcije (točkovna vrednost 0), šibka reakcija (točkovna vrednost +/- 1), srednje intenzivna reakcija (točkovna vrednost +/- 2) in zelo močna reakcija, ki jo ovrednotimo s +/- 3 – takšna ocena je dopustna samo enkrat v testu po posameznih parametrih (dihanje, elektrodermalna odzivnost, kardiovaskularne spremembe). Način govorjenja poligrafskega preiskovalca lahko okrepi ali oslabi odziv testirane osebe na posamezne kategorije vprašanj (Kircher in Raskin, 2002:287–288).

Pri poligrafskih testih nismo zainteresirani za zvezno razporeditev izmerjenih/ocenjenih vrednosti. Prizadevamo si za čim bolj nedvoumno razlikovanje med iskrenimi (nevpletenimi) in neiskrenimi (vpletenimi) testiranimi osebami, brez nedoločnih ocen.

Merske značilnosti poligrafskih testov so ustrezne in zadovoljujoče, kadar je njihova izvedba kar najbolj standardizirana, poligrafski preiskovalec pa s svojim vedenjem (tudi z govorjenjem) ne izzove dodatnih/neželenih/nekontroliranih odzivov pri testirani osebi. Vsakršna hotena ali nenamerna ali nenadzorovana intervencija poligrafskega preiskovalca lahko povzroči napačno oceno – bodisi lažno pozitivno (poligrafski preiskovalec meni, da testirana oseba zavaja, čeprav ne) bodisi lažno negativno (zavajanje testirane osebe oceni kot iskrenost). Z merskimi značilnostmi poligrafskih testov so tesno povezana pravila odločanja. Pravilo odločanja narekuje, pri katerih (točkovnih) vrednostih oziroma najmanjših/največjih vrednostih točk pri posameznih primerjavah (odzivov na relevantno in primerjalno vprašanje; R-P), gre za zavajanje ali iskrene odgovore.

2 Taksonomija najpomembnejših pristopov k odkrivanju zavajanja

Sodobna poligrafija razlikuje med testi prepoznavanja in testi za odkrivanje zavajanja. Leta 2003 je Nacionalni svet za raziskovanje (raziskovalno dejavnost) Združenih držav Amerike (*National Research Council*, 2003) namenil posebno poročilo pregledu razpoložljivih znanstvenih podlag, prednostim in omejitvam različnih metod odkrivanja zavajanja ter dal poseben poudarek poligrafski metodi. Poročilo upošteva oba temeljna pristopa psihofiziološkega preiskovanja verodostojnosti izjav s poligrafom, in sicer prepoznavanje in odkrivanje zavajanja. Koncept prepoznavanja izhaja iz predpostavke,

da bo izpostavljenost znanemu dražljaju (podatki, slike ipd.), umeščnemu v set smiselno podobnih, vendar ne povezanih dražljajev, izzvala jasen psihofiziološki odziv. Tehnike, s katerimi odkrivamo zavajanje, ne temeljijo na prepoznavi znanih (skritih) posebnosti/podrobnosti, marveč primerjajo odziv na neposredno vprašanje o obravnavanem dogodku z odzivi na druga vprašanja, uvedena po predpisanem postopku.

1 TESTI PREPOZNAVANJA

1.1 PRILAGODITVENI TESTI

V Sloveniji je uveljavljen izraz stimulativni, oziroma stimulatивно-prilagoditveni testi, kar manj kaže namen in vlogo teh testov kot izraz »prilagoditveni« – izvajajo se ob koncu predtestnega razgovora z namenom preverjanja psihofizioloških odzivov testirane osebe, spoznavanja testnega postopka in poznejšega ozaveščenega prostovoljnega pristanka testirane osebe na izvedbo psihofiziološke preiskave verodostojnosti izjav s poligrafom.

1.2 TESTI VRHUNCA NAPETOSTI (POT, *Peak of Tension*)

Test vrhunca napetosti je zasnovan enako kakor psihološki testi z različnimi možnimi izbirami/odgovori (*multiple choice*). Ena od alternativ (relevantni ali kritični odgovor) je povezana s posebnim (obravnavanim) dogodkom, za katerega predpostavljamo, da je testirani osebi poznan – bodisi je v dogodku sodelovala ali nekaj o njem ve. Druge alternative (odgovori) so z dogodkom nepovezane, vendar so smiselne, in oseba, ki obravnavanega dogodka ne pozna, ne more razlikovati med relevantnimi in kontrolnimi odgovori oziroma lahko o »pravilnem« odgovoru zgolj ugiba po verjetnosti (Ben-Shakhar, Bar-Hillel in Liebllich, 1986). Običajno se vrednoti zgolj elektrodermalna komponenta.

Testi vrhunca napetosti so posebna vrsta testov vedenja/poznavanja okoliščin (GKT, *Guilty Knowledge Test*), pri katerih je kritično vprašanje pozicionirano v sredini liste, testirana oseba vnaprej pozna vrstni red vprašanj in ob prepoznavanju pride do značilnega »vrhunca napetosti«. Pričakovanje tega (kritičnega vprašanja) zato izzove močnejši odziv na samo kritično vprašanje kakor pri testih vedenja (Furedey in Ben-Shakhar, 1991).

A/ TESTI Z ZNANO REŠITVIJO

Test z znano rešitvijo je POT s praviloma sedmimi odgovori (alternativami), od katerih samo eden ustreza obravnavanemu dogodku, preostali pa so izmišljeni, vendar smiselni. Arther (1970) je modificiral format z uvedbo »lažnega ključa«, vedno postavljenega na drugo pozicijo v nizu odgovorov. S tem naj bi odpravil nevarnost lažno pozitivnih ocen pri nevpleteni osebi in nevtraliziral model odzivanja, poznan kot »spot responding«, ko pri testirani osebi pride do spremembe psihofiziološkega dogajanja vedno na isti poziciji (v določenem časovnem intervalu oziroma ciklično), ne glede na vrsto

vprašanja na tej poziciji. »Pravi ključ« (relevanten odgovor) je na predpisani četrti poziciji.

V strokovni literaturi je za teste z znano rešitvijo mogoče najti tudi izraz test Keeler tip A.

B/ TESTI PO PREDPOSTAVKAH

Teste po predpostavkah nekateri avtorji imenujejo tudi »iskalne« POT (SPOT, *Searching Peak of Tension*) ali test Keeler tip B. Običajno zajemajo šest smiselnih alternativ, na sedmi poziciji pa je vprašanje »kaj drugega«. V praksi so namenjeni odkrivanju dokazov, predmetov in drugih okoliščin, povezanih z obravnavanim dogodkom. Zvezni preiskovalni urad (FBI, 1985) je izdal navodilo za izvedbo poligrafskim preiskovalcem ter zahteval pri SPOT postavljanje manj verjetnih (ali celo neverjetnih) alternativ na prvi dve poziciji ter na zadnji dve poziciji, dovolil pa je uporabo vizualnih pripomočkov (dražljajev), kot so zemljevidi, slike ipd.

1.3 TESTI PRIKRITIH INFORMACIJ (tudi TESTI VEDENJA/ POZNAVANJA OKOLIŠČIN (CIT, *Concealed Information Technique*; GKT)

Pri testnem formatu CIT testirana oseba ne pozna zaporedja vprašanj, kritično vprašanje (ključ) je pozicionirano po slučaju (kar odpravi problem *spot-responderjev*), nikdar pa na prvi poziciji. V nizu odgovorov morajo biti vsaj štiri alternative, prva pozicija vedno deluje kot mašilo in se ne vrednoti. Zaželeno je serija štirih do desetih testov, včasih se analizirajo (vrednotijo) vsi parametri, ne pa vedno.

Abrams (1989,120–127) je izdelal posebna priporočila za formulacijo alternativ (odgovorov) in poudaril njihovo uravnoveženost, izenačenost po številu besed ter enopomenskost. Če testirana oseba ponovi odgovor in nato odgovori »ne« (npr.: »dežnik, ne«), se veljavnost poveča, je ugotovil.

V začetku devetdesetih let prejšnjega stoletja je Poligrafski inštitut obrambnega ministrstva revidiral protokol za izvedbo GKT (Ansley, 2008) in prepovedal rabo vprašanja »Ali veste ...?«, ker naj bi takšna formulacija izzvala laganje pri vpleteni osebi na vse alternative (odgovore). Veljavna (predpisana) formulacija tudi po dvajsetih letih ostaja »Ali je bila/o ...?« (Yankee, 1991).

Vrednotenje psihofizioloških odzivov v objavljenih študijah CIT dosledno sledi Lykkenovim navodilom (Lykken, 1959) – elektrodermalni (EDR) odziv se vrednoti z vrednostmi od 0 do 2. Če je najmočnejši EDR-odziv na ključnem vprašanju, je ovrednoten z 2 točkama, če je na ključnem vprašanju drugi najmočnejši odziv, dobi točkovno vrednost 1, vsi preostali odzivi dobijo vrednost 0, prva pozicija se ne vrednoti.

Namesto pravil odločanja (o zavajanju oziroma prepoznavanju prikritih informacij) pri CIT govorimo o verjetnostih (Krapohl, McCloughan in Senter, 2009; tabela 1).

V skrajnem levem stolpcu je število izvedenih CIT, v zgornji vrstici tabele pa točkovne vrednosti, pripisane EDR-odzivom.

Če na primer po izvedenih šestih CIT dobimo skupno točkovno vrednost 9, je verjetnost, da t. i. »naivna« (nevpletena) testirana oseba pozna ključni odgovor, 1 %, če bi pri enaki seriji izvedenih CIT dobila 12 točk, bi bila takšna verjetnost manjša od desetinke odstotka, če bi bila točkovna vrednost 3, pa bi to pomenilo 69 % možnosti, da »naivni« subjekt pozna okoliščine obravnavanega dejanja.

2 TESTI ZA ODKRIVANJE ZAVAJANJA

2.1 TESTI S PRIMERJALNIM (KONTROLNIM) VPRAŠANJEM

V štiridesetih letih prejšnjega stoletja sta Inbau in Reid predstavila testni format, ki sta ga poimenovala test s kontrolnim

Tabela 1: Verjetnosti, da »naivna« (nevpletena) testirana oseba pozna prikrite informacije

| Število CIT | TOČKOVNE VREDNOSTI (zbir točkovnih vrednosti EDR) | | | | | | | | | | | | | | |
|-------------|---|------|------|------|------|------|------|------|------|------|------|------|------|------|--|
| | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | |
| 2 | 0,12 | 0,04 | | | | | | | | | | | | | |
| 3 | 0,28 | 0,13 | 0,03 | 0,01 | | | | | | | | | | | |
| 4 | 0,44 | 0,25 | 0,10 | 0,04 | 0,01 | 0,00 | | | | | | | | | |
| 5 | 0,58 | 0,38 | 0,20 | 0,09 | 0,03 | 0,01 | 0,00 | 0,00 | | | | | | | |
| 6 | 0,69 | 0,50 | 0,31 | 0,17 | 0,08 | 0,03 | 0,01 | 0,00 | 0,00 | 0,00 | | | | | |
| 7 | 0,78 | 0,61 | 0,42 | 0,26 | 0,14 | 0,07 | 0,03 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | | | |
| 8 | 0,84 | 0,70 | 0,53 | 0,36 | 0,22 | 0,12 | 0,06 | 0,03 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | |

vprašanjem (CQT; *Control Question Test*), ki je vključeval vprašanje, ki izzove verjetno laž (pri testirani osebi) in t. i. »kompleks krivde«. Pozneje sta namesto tega uvedla dve kontrolni vprašanji, ki naj bi izzvali verjetno laž. Vsebinsko spremembo kontrolnih/primerjalnih vprašanj je uvedel Backster (1969) – po njegovem konceptu se morata relevantno in primerjalno vprašanje razlikovati po vsebini in prostorsko-časovnih determinantah.

V zadnjih letih praviloma govorimo o različnih tehnikah primerjave polj (pas odziva na relevantno in primerjalno vprašanje, *zone comparison*), med katerimi izstopajo vojaški ZCT (*Army Zone Comparison Test*), zvezni ZCT (*Federal Zone Comparison Test*) in Utah ZCT (razvit in preverjen v minulih 40 letih na Univerzi Utah).

Testni formati vključujejo nevtralna, relevantna in primerjalna, nekateri tudi simptomatska vprašanja. Primerjalna vprašanja naj bi pritegnila pozornost nevpletene osebe. Lahko gre za t. i. »dogovorjeno laž« ali »verjetno laž«, ki je časovno razmejena od obravnavanega dogodka, ali pa ne. Dogovorjena laž je nikalni odgovor na vprašanje, na katero bi vsakdo odgovoril pritrnilno – testirana oseba dobi navodilo, da se na vprašanje zlaže. Testi s primerjalnim vprašanjem, kjer je odgovor trivialna dogovorjena laž, so manj točni (veljavni) od onih, kjer je dogovorjena laž osebne narave.

Horowitz, Kircher, Raskin in Honts so leta 1997 izvedli laboratorijsko študijo. Na osmih skupinah testiranih oseb, ki so jih razdelili na vpletene in nevpletene ter temu sledili v izvedbi eksperimenta, so aplicirali test R-I, test s primerjalnim vprašanjem, kjer je odgovor trivialna dogovorjena laž, test s primerjalnim vprašanjem, kjer je odgovor dogovorjena laž osebne narave in test s primerjalnim vprašanjem, kjer je od-

govor verjetna laž. Pri obeh skupinah je bila največja točnost dosežena pri testu s primerjalnim vprašanjem z dogovorjeno lažjo osebne narave. Študija še ni bila ponovljena, kar ob majhnem številu testiranih oseb, vključenih v eksperimentalne skupine, ne omogoča trdnejših ugotovitev (Blalock, 2009).

Pri testih s primerjalnimi vprašanji je par (pas, polje) relevantno-primerjalno vprašanje običajno uporabljen trikrat v vsakem testu. Razlike v psihofizioloških odzivih se vrednotijo na sedem-stopenjski lestvici (+3 do -3), gre za vrednotenje amplitude, frekvence, trajanja ali kombinacije teh parametrov pri psihofizioloških odzivih.

2.2 TESTI S PRIMERJALNIM (KONTROLNIM) VPRAŠANJEM

2.2.1 DOGOVORJENA LAŽ

A/ TRIVIALNA
B/ OSEBNA

2.2.2 VERJETNA LAŽ

A/ IZKLJUČUJOČ ČASOVNI INTERVAL
B/ VKLJUČUJOČ ČASOVNI INTERVAL

2.3 TESTIBREZPRIMERJALNEGA (KONTROLNEGA) VPRAŠANJA

TESTI RELEVANT-IRELEVANT

Testi R-I se običajno uporabljajo kot presejalni testi. Gre za večtematske teste (v enem testu zastavimo običajno tri relevantna vprašanja), format v presejalne namene je predpisan (Blalock, 2009), postopek pa preverjen. Morebitni odziv testirane osebe na katerokoli relevantno vprašanje se v nadaljevanju preveri z enotematskim testom (ZCT).

Tabela 2: Primerjava točnosti ocen pri različnih testih za odkrivanje zavajanja – laboratorijska študija (Horowitz, Kircher, Honts in Raskin, 1997)

| Eksperimentalne skupine (N=15) | Rezultati (%) | | | |
|--------------------------------|----------------|---------------|------------|------------------------|
| | Pravilna ocena | Napačna ocena | Nedoločeno | Delež (%) pravil. ocen |
| VPLETENI | | | | |
| Test relevant-irelevant | 100 | 0 | 0 | 100 |
| Trivialna dogovorjena laž | 53 | 20 | 27 | 73 |
| Dogovorjena laž osebne narave | 73 | 14 | 13 | 84 |
| Verjetna laž | 53 | 20 | 27 | 73 |
| NEVPLETENI | | | | |
| Test relevant-irelevant | 20 | 73 | 7 | 22 |
| Trivialna dogovorjena laž | 67 | 13 | 20 | 84 |
| Dogovorjena laž osebne narave | 87 | 13 | 0 | 87 |
| Verjetna laž | 80 | 12 | 7 | 86 |

Odkrivanje zavajanja je uporabnejše, saj ne zahteva »prikritih« informacij, vendar je izvedba testov poznavanja manj zahtevna. Izvedba tehnik za odkrivanje zavajanja zahteva ustrezno strokovno usposobljenost poligrafskega preiskovalca in nadzor kakovosti, sicer prihaja do napak, ki pa ne izražajo vrednosti same tehnike. Za tehnike/teste odkrivanja zavajanja *National Research Council* (NCR, 2003) opisuje veljavnost v intervalu 81 do 91 %, z vrednostjo mediane pri 86 %. Pri tehnikah prepoznavanja je po oceni NCR interval veljavnosti med 85 in 96 %, vrednost mediane pa 88 %. Raziskave, ki temeljijo na resničnih primerih, bolj ali manj dosledno ugotavljajo, da je psihofiziološka preiskava verodostojnosti izjav z uporabo poligrafa natančnejša (veljavnejša) pri odkrivanju zavajanja kakor pri identifikaciji iskrenih (izločanju nevpoletenih v obravnavani dogodek/dogajanje). To je razumljivo, saj je večina avtorjev (tehnik) izšla iz polja preiskovanja kaznivih dejanj, kar velja tudi za pravila odločanja (na primer Light, 1999).

V sodnih postopkih v ZDA morajo uporabljene metode zadostiti merilom Ameriškega združenja za testiranje in materiale (*American Society for Testing and Materials*, ASTM, 2005), uporabljena poligrafska metoda pa upošteva standard »Marinovega protokola« (Marin, 2000, 2001) tudi v delu, ki se nanaša na uravnoteženo tveganje (možnost) napačno pozitivnih in napačno negativnih ocen. Zato je pozornost usmerjena v (pre)oblikovanje pravil odločanja tako, da dosežemo kar največji delež pravilnih odločitev ob enaki verjetnosti napake ene ali druge vrste (Krapohl, 2006). Pravila odločanja oblikujejo zadnjo in ključno fazo numeričnega vrednotenja psihofizioloških odzivov po uporabi poligrafskih testov. Pravila odločanja so mejne vrednosti, ki vplivajo tudi na delež napačno negativnih/pozitivnih in nedoločenih ocen. Pravilo, ki zmanjša delež napačno negativnih ocen, avtomatično poveča delež napačno pozitivnih ocen in nasprotno. V predkazenskem postopku, ko gre za zbiranje dokazov, je škodljivost napačne pozitivne ocene povsem drugačna kot v dokaznem postopku. Zato so pravila odločanja takšna, da minimalizirajo možnost napačno negativne ocene. Delež pravilnih ocen je rezultat sovplivanja kompetentnosti poligrafskega preiskovalca, uporabljene tehnike in vrednotenja psihofiziološkega zapisa. Veljavna merila, ki izhajajo iz »Marinovega protokola«, so delež pravilnih ocen nad 86 % in delež nedoločenih ocen pod 20 %. Tako izboljšane merske karakteristike zavračajo pavšalne in nepoučene kritike, nadzor kakovosti dela poligrafskih preiskovalcev, vključenih v sodne postopke, pa povečuje preglednost poligrafskega postopka in omogoča tehtanje ocene.

Objavljene in ponovljene raziskave so potrdile točnost samo pri nekaterih poligrafskih tehnikah, kar šteje kot merilo znanstvene ustreznosti (ASTM, 2005). Pri tem je prag točnosti (natančnosti, veljavnosti) po standardih ASTM postavljen nad 90 % pri poligrafriranih in nad 80 %

pri drugih aplikacijah poligrafske metode. V obeh primerih je dovoljen manj kakor 20-odstotni delež nedoločenih ocen.

3 Vpliv pravil odločanja na veljavnost testov s primerjalnimi vprašanji

Mnenja NCR (*National Research Council*, 2003) in Ameriškega združenja za testiranje in materiale (*American Society for Testing and Materials*, ASTM, 2005) so potrdila smiselnost dotakratnega intenzivnega raziskovalnega angažmaja na Poligrafskem inštitutu obrambnega ministrstva, usmerjenega v preverjanje in redefiniranje pravil odločanja pri posameznih tehnikah (Senter, 2003; Senter, Dollins in Krapohl, 2004; Senter in Dollins, 2004; Senter in Ryan, 2005). Do Senterjevih obsežnih analiz v začetku tega desetletja so veljavna pravila odločanja temeljila bolj na dogovoru kakor na neoporečnih raziskovanih podatkih. Sedanja pravila določajo različna merila za enotemske in večtematske teste. Pri enotematskih testih (na primer ZCT, *Zone Comparison Test*) je merilo zavajanja skupna točkovna vrednost (-6) ali manj ali katerakoli vrednost -3 ali manj, odsotnost zavajanja pa določa meja skupna vrednost (+6) ali več in vsak odgovor (+1) ali več. Vse druge vrednosti terjajo nedoločeno oceno. Pri večtematskih testih, ki se opuščajo (na primer MGQT, *Modified General Question Test*) je kriterij zavajanja katerikoli odgovor (-3) ali manj, iskrenosti pa vsak odgovor (+3) ali več.

Standardna pravila pa ne zadostijo uporabi psihofiziološke preiskave verodostojnosti izjav s poligrafom v dokaznem (kazenskem) postopku (t. i. dokazno poligrafriranje). Napačna pozitivna ocena je v tem kontekstu veliko bolj problematična, saj so posledice neprimerno težje. V procesu kriminalističnega preiskovanja je mogoče napačno pozitivno oceno poligrafskega preiskovalca korigirati z drugimi kriminalističnimi opravili in forenzičnimi dokazi. Standardna pravila odločanja so se pokazala boljše za odkrivanje zavajanja in Senterjeve analize (2003) so opozorile na asimetrično distribucijo točkovnih vrednosti psihofizioloških odzivov - točkovne vrednosti tistih, ki zavajajo, so ob uporabi standardnih pravil odločanja pomembno manjše od nič (0), za razliko od iskrenih, katerih vrednosti se razporejajo bližje 0. Zato je Senter (2003) predlagal dvofazno pravilo odločanja, ki ga lahko imenujemo tudi pravilo odločanja v dokaznem postopku.

3.1 Pravilo odločanja v dokaznem postopku

Pravilo odločanja v dokaznem postopku zmanjša delež nedoločenih ocen pod 20 %, poveča poprečno zanesljivost ocen nad 86 % ter zagotavlja uravnoteženo vrednotenje odzivov tistih, ki zavajajo in onih, ki so iskreni, kar ustreza

»Marinovemu protokolu« (Selič, 2006). Pravilo odločanja v dokaznem postopku je dvofazno. V prvi fazi (test apliciran trikrat) se odločamo po kriterijih skupna točkovna vrednost (- 6) ali manj pomeni zavajanje, skupna točkovna vrednost (+ 4) ali več pomeni odsotnost zavajanja, vse druge vrednosti terjajo nedoločeno oceno – v tem primeru sledi druga faza. V drugi fazi test apliciramo še dvakrat, poligrama vrednotimo po ustaljenih pravilih, kriterij odločanja pa je, če je točkovna vrednost pri kateremkoli odgovoru manjša od (- 2), gre za zavajanje, vse drugo pomeni nedoločeno oceno,

Krapohl (2005) je primerjal rabo standardnih pravil odločanja (SPO) in pravil odločanja v dokaznem postopku (PDP) v štirih študijah. V primerih ugotovljenega zavajanja je bila ocena po SPO točna (veljavna) v 78,3 %, do napake je prišlo v 6,8 %, delež nedoločenih ocen pa je znašal 14,9 %. Ob uporabi PDP je bilo pravih ocen 80,8 %, do napake je prišlo v 10,4 %, delež nedoločenih ocen pa je znašal 8,9 %. Pri iskrenih so standardna pravila odločanja (SPO) zagotovila pravilno oceno v 61,5 % in nedoločeno oceno v 29,2 %, do napake pa je prišlo v 9,3 %, za razliko od ocen po pravih odločanja v dokaznem postopku (PDP), ki so bile statistično pomembno drugačne – pravih je bilo 80,3 %, nedoločenih 10,5 % in napačnih 9,3 % ocen. Skupni delež pravih odločitev je bil pri standardnih pravih odločanja (SPO) 89,6 % ob 22,1-odstotnem deležu nedoločenih ocen, pri pravih odločanja v dokaznem postopku (PDP) pa 89,1 % ob 9,6-odstotnem deležu nedoločenih ocen.

V nadaljevanju je Krapohl preveril rabo obeh pravil odločanja pri 100 primerih z znano resnico. V primerih zavajanja je bila veljavnost ocen po standardnih pravih odločanja (SPO) 82 % ob 4-odstotni napaki in 13-odstotnem deležu nedoločenih ocen. Ocenjevanje po pravih odločanja v dokaznem postopku (PDP) je bilo pravilno v 79 % ob 12-odstotni napaki in 8-odstotnem deležu nedoločenih ocen. PDP se je bolj obneslo pri iskrenih s 83 % pravih ocen, 12 % napak in 6 % nedoločenih ocen, v primerjavi s SPO, kjer je bil delež pravih ocen 57 %, napaka 8 % in delež nedoločenih ocen 17 %. V skupnem je bil delež pravih odločitev po standardnih pravih odločanja (SPO) 86 % ob 20 % nedoločenih ocen, po pravih odločanja v dokaznem postopku (PDP) pa je bilo pravih ocen 87 % ob 8 % nedoločenih ocen.

Krapohl (2005) ni odkril statistično pomembnih razlik med rezultati uporabe standardnih pravil odločanja in pravih odločanja v dokaznem postopku v skupnem deležu pravih ocen. Tudi natančnost ocen pri zavajanju se ne razlikuje statistično pomembno. Pravila odločanja v dokaznem postopku izboljšajo identifikacijo iskrenih in zmanjšajo delež nedoločenih ocen, standardna pravila odločanja pa so dosledno bolj učinkovita pri identifikaciji zavajanja. Opisano pa velja samo za enotemske teste (ZCT).

Leta 2006 je isti avtor študijo ponovil (Krapohl in Cushman, 2006) in k sodelovanju povabil deset neodvisnih ocenjevalcev (poligrafskih preiskovalcev), ki so vrednotili 100 primerov (300 poligramov) z znano resnico v procesu pridobivanja certifikata za »Marinov protokol« (da je njihova strokovna kompetentnost skladna z zahtevami »Marinovega protokola« in dosegajo delež pravih ocen nad 86 % (Veritas Center, 2005)), kar naj bi zagotovilo njihovo motiviranost za kar najboljši rezultat (pravilnost vrednotenja oziroma odkrivanja zavajanja/iskrenosti). Pri standardnih pravih odločanja je bil skupni delež pravih ocen 86,1 %, kar ustreza standardu ASTM v »Marinovem protokolu«. Tudi delež nedoločenih ocen (19,8 %) je bil pod predpisano mejo 20 %. Pri pravih odločanja v dokaznem postopku je bil skupni delež pravih ocen 87,2 %, delež nedoločenih pa 7,3 %. Pri standardnih pravih odločanja (SPO) je bil delež napačnih ocen pomembno nižji pri zavajanju, kar je potrdilo tezo, da uporaba SPO poveča občutljivost za zavajanje. Tudi rezultat pri PDP je potrdil Krapohlve ugotovitve (Krapohl, 2005), da so pravila odločanja v dokaznem postopku nepristranska, saj je bil delež pravih odločitev pri zavajanju in pri iskrenih primerih (brez zavajanja) statistično nepomembno različen.

Pravila odločanja v dokaznem postopku so se pokazala kot ustrežnejša pri primerih brez zavajanja (iskrenih), vendar je novejša študija opozorila na pomembno večji delež napačno negativnih ocen ($z = 4,47$, $p < 0,01$) z rabo teh pravil. Če povzamemo ugotovitve iz leta 2005 in 2006, lahko pomeni raba pravil odločanja v dokaznem postopku večje tveganje za napako pri primerih zavajanja, vendar prav raba tega pravila pomembno zniža delež nedoločenih ocen in poveča nepristranskost ocen, vse pa le pri enotematskem testu primerjave polj (ZCT).

Krapohl je tako dokazal, da so pravila odločanja v dokaznem postopku skladna z merili v »Marinovem protokolu« in primerna za izvedbo poligrafske metode v dokaznem pomenu. Opozoril pa je, da raba PDP ni primerna za preiskovanje verodostojnosti izjav s poligrafom v predkazenskem postopku (v kriminalističnem preiskovanju) zaradi hudih posledic, ki jih ima napačno negativna ocena. Opisano pomeni ključni premiki teorije in prakse psihofiziološkega preiskovanja verodostojnosti izjav s poligrafom od rabe arbitrarnih pravil odločanja k verificiranim pravilom odločanja. Izbor pravila odločanja je postal funkcija konteksta in namena in ne (več) osebnih preferenc ali časovnih omejitev pri izvedbi poligrafske metode. Pri preiskovanju kaznivih dejanj, kjer je osnovna funkcija poligrafske metode izločanje nevpletenih v obravnavano kaznivo dejanje oziroma identifikacija tistih, ki zavajajo (v zvezi z obravnavanim kaznivim dejanjem), je umestna raba standardnih pravil odločanja, ki povečajo občutljivost poligrafske metode na zavajanje in niso obremenjena z verjetnostjo napačno ne-

gativne ocene. Potrebam in zahtevam dokaznega postopka pa ustrezajo pravila odločanja v dokaznem postopku. Tako se je psihofiziološko preiskovanje zavajanja z uporabo poligrafa argumentirano odzvalo na standarde različnih kontekstov, kjer je raba poligrafske metode uveljavljena.

4 Vrednotenje posameznih poligrafskih tehnik

Iz raziskav, objavljenih v minulih štiridesetih letih, smo po merilih znanstvene validacije izločili tiste, ki so zadevale poznane/uveljavljene poligrafske tehnike z ocenami njihove točnosti (veljavnosti), posebej za iskrene testirane osebe in one, ki so zavajali, ter z določenim zunanjim neodvisnim kriterijem »poznane resnice«. V omenjenih raziskavah je vse poligrafske zapise ovrednotil poligrafski preiskovalec po preverjenih pravilih vrednotenja, brez uporabe računalniških algoritmov vrednotenja. Pri pregledu validacijskih študij ni bilo mogoče upoštevati vpliva različnih pravil odločanja, saj je bila večina študij izvedena v obdobju, ko ta še niso bila uvedena.

Predstavljene ugotovitve zadevajo samo dostopne raziskave in ne izražajo nujno kakovosti tehnik, ki so bile do zdaj manj ali manj ustrezno validirane. Prikazane so neponderirane srednje vrednosti (tabela 4–8).

4.1 Test prikritih informacij (CIT)

Prvo analizo validacijskih študij testov prikritih informacij (tudi testov vedenja/poznavanja okoliščin) je opravil Ansley leta 1992 (Ansley, 2008). Rezultate prikazujemo v tabeli 3.

Tabela 3: Zbirna tabela veljavnosti študij, ki jih je predstavil Ansley (2008)

| | Povprečna veljavnost | N | Odkrito zavajanje | N | Brez zavajanja | N |
|---------------------|----------------------|------|-------------------|------|----------------|-----|
| Vsi testi | 68 % | 4874 | 65 % | 4396 | 93 % | 478 |
| Testi vedenja (GKT) | 76 % | 1519 | 72 % | 1181 | 91 % | 338 |
| Drugi testi POT | 66 % | 3355 | 65 % | 3215 | 100 % | 140 |

Ansleyjeva analiza vključuje več kakor 80 študij na laboratorijskih in realnih primerih, zajetih je 4774 testiranih oseb, vendar je omejitev njegovih ugotovitev v premajhnem številu testiranih oseb v posamičnih študijah ter v neenakih ali celo nejasnih kriterijih vrednotenja poligramov. Zato je edini dopustni sklep (z upoštevanjem siceršnjih omejitev), da sta testna formata POT in GKT natančnejša pri izločanju nevpoletnih, ki ne zavajajo, kakor pa pri odkrivanju onih, ki zavajajo oziroma poznajo okoliščine obravnavanega dejanja.

Meta analizo 50 neoporečnih študij testov prikritih informacij je opravil MacLaren (2001), ponovil pa Kraphol (2006). Oba sta dobila skupno povprečno veljavnost 80 %. Pri osebah, ki so zavajale (N = 843) je bila točnost testa 76 %, pri iskrenih (N = 404) pa 83 %.

Tabela 4: Skupni rezultat validacijskih študij CIT

| Testirane osebe | Rezultati (%) | | |
|-----------------|---------------|--------------------|----------------------|
| | Število | Pravilna ocena (%) | Nedoločena ocena (%) |
| VPLETENI | 843 | 76 | 0 |
| NEVPLETENI | 404 | 83 | 0 |
| Skupaj | 1247 | 80 | 0 |

Kljub temu da ima CIT od vseh poligrafskih tehnik najboljšo teoretično podlago in več validacijskih raziskav kakor vse druge poligrafske tehnike skupaj, je v praksi najmanj uporabljan (Kraphol, McCloughan in Senter, 2009). Pred približno 50 leti ga je predstavil Lykken (1959, 1960) in ga opredelil nekoliko drugače, kakor je veljalo za do takrat uporabljeni format vrhunca napetosti (POT). Sodobni trendi dobrih praks v psihofiziološkem preiskovanju verodostojnosti izjav s poligrafom prepoznavajo vrednost CIT zlasti v primerih, ko lahko dajo testi s kontrolnim/primerjalnim vprašanjem manj kot optimalne rezultate (čustveno nasičene situacije/dogodki).

4.2 Vojaški ZCT (tudi zvezni ZCT)

Veljavnost tega testnega formata so raziskovali Blackwell (1998), Krapohl (2005) ter Yankee, Powell in Newland (1985).

Tabela 5: Skupni rezultat validacijskih študij vojaškega ZCT

| Testirane osebe | Rezultati (%) | | |
|-----------------|---------------|--------------------|----------------------|
| | Število | Pravilna ocena (%) | Nedoločena ocena (%) |
| VPLETENI | 141 | 97 | 9 |
| NEVPLETENI | 110 | 82 | 23 |
| Skupaj | 251 | 89 | 16 |

V vseh primerih so bila uporabljena preiskovalna pravila odločanja (ne dvostopenjska) ter vključeni izkušeni poligrafski preiskovalci, s stažem več kakor 5 let profesionalnega delovanja na področju psihofiziološkega preiskovanja verodostojnosti izjav s poligrafom.

4.3 Reidova tehnika (CQT)

Reidovo tehniko (CQT) sta validirala Jayne (1990) in Horvath (1977, 1988)

Tabela 6: Skupni rezultat validacijskih študij Reidovega CQT

| Testirane osebe | Rezultati (%) | | |
|-----------------|---------------|--------------------|----------------------|
| | Število | Pravilna ocena (%) | Nedoločena ocena (%) |
| VPLETENI | 88 | 88 | 7 |
| NEVPLETENI | 88 | 78 | 5 |
| Skupaj | 176 | 83 | 6 |

Uporabljeni so bili samo primeri, pri katerih so bila dosledno upoštevana pravila Reidove tehnike, zunanji kriterij veljavnosti je bil zagotovljen.

4.4 Test relevant-irelevant (presejalni)

Tabela 7: Skupni rezultat validacijskih študij testa R-I

| Testirane osebe | Rezultati (%) | | |
|-----------------|---------------|--------------------|----------------------|
| | Število | Pravilna ocena (%) | Nedoločena ocena (%) |
| VPLETENI | 79 | 90 | 0 |
| NEVPLETENI | 61 | 73 | 0 |
| Skupaj | 140 | 83 | 0 |

4.5 Utah ZCT

Utah ZCT je bil preverjen v sedmih validacijskih študijah, kar je dvakrat več kakor pri katerikoli drugi tehniki s primerjalnim vprašanjem (Honts, Hodes, Raskin, 1985; Honts, Raskin, Kircher, 1987 in 1994; Horowitz, Kircher, Honts in Raskin, 1997; Kircher, in Raskin, 1988; Raskin, in Hare 1978; Rovner, 1986).

Tabela 8: Skupni rezultat validacijskih študij Utah ZCT

| Testirane osebe | Rezultati (%) | | |
|-----------------|---------------|--------------------|----------------------|
| | Število | Pravilna ocena (%) | Nedoločena ocena (%) |
| VPLETENI | 143 | 91 | 13 |
| NEVPLETENI | 128 | 89 | 10 |
| Skupaj | 271 | 90 | 12 |

5 Znanstvena merila ustreznosti poligrafskih tehnik

ASTM (2005) zahteva za dokazne poligrafske rezultate točnost 90 % brez nedoločenih ocen, dopušča delež pod 20 % nedoločenih ocen, oziroma pri preiskovalnih poligrafskih rezultatih točnost (veljavnost) 80 % brez nedoločenih ocen in skupni delež nedoločenih ocen pod 20 %. Pregled tabel 4–8 pokaže, da dokaznim standardom ustreza samo Utah ZCT, vojaški ZCT pa praga ne dosega zgolj za eno odstotno točko. Za preiskovalna poligrafranja je dopustna raba vojaškega ZCT, Reidove tehnike (CQT), tehnike prikritih informacij (CIT ali GKT) ter testa R-I. Ob tem je nujno opozoriti, da lahko v prihodnje pričakujemo tudi drugačne rezultate validacijskih študij, saj so se pravila odločanja uveljavila kot preverjen dejavnik kakovostne izvedbe psihofiziološke preiskave verodostojnosti izjav s poligrafom pri enotematskih testih, ki vpliva na točnost poligrafskega rezultata.

Tabela 9: Razvrstitev poligrafskih tehnik glede na doseženo povprečno veljavnost v validacijskih študijah

| Tehnike | Rezultati (%) | |
|--------------------|----------------------------------|------------------------|
| | Veljavnost brez nedoločenih ocen | Delež nedoločenih ocen |
| Utah ZCT | 91 | 12 |
| Zvezni/vojaški ZCT | 89 | 16 |
| Test R-I | 83 | 0 |
| Reidov CQT | 83 | 6 |
| CIT (tudi GKT) | 80 | 0 |

Preseneča nižja povprečna veljavnost CIT, ki kaže razpoložljive raziskovalne podatke in ne nujno dejanskega stanja, saj ni bilo mogoče upoštevati števila uporabljenih CIT v vsaki študiji. Če upoštevamo, da je verjetnost reakcije na kritični dražljaj pri »naivni« testirani osebi funkcija točkovne vrednosti elektrodermalnega odziva in števila CIT v seriji (tabela 1), je razumljiva izračunana 80-odstotna skupna točnost (veljavnost) CIT, saj so po dostopnih podatkih v obravnavanih študijah uporabljali pretežno serije treh izvedenih CIT.

Slovenske izkušnje (Selič, 2002) kažejo pomembno večjo točnost CIT. V letih 1998–2002 smo izvedli več kakor 400 poligrafskih preiskav, skoraj izključno z rabo CIT. Če kot kriterij »resnice« vzamemo odkritje novih dokazov, okoliščin, storilca ipd. z uporabo rezultatov poligrafske preiskave, priznanje testirane osebe in/ali odkritje novih dokazov, okoliščin ipd. v razgovoru po aplikaciji poligrafskih testov ali odkritje novih dokazov, okoliščin, storilca ipd. z drugimi forenzičnimi preiskavami, o čemer lahko govorimo v 95 % primerov, lahko poročamo, da smo z aplikacijo povprečno petih CIT (največ 11 CIT, več o primeru Selič, 2000) pri testirani osebi dosegali 100-odstotno točnost poligrafskega rezultata. Podatek velja za »naivne« testirane osebe in tudi za tiste, vpletene v obravnavana (kazniva) dejanja. Podrobnejše analize ni mogoče predstaviti, saj dokumentacija avtorici ni več dostopna.

6 Sklep: Znanstveno utemeljeni in preverjeni principi izvedbe psihofiziološke preiskave verodostojnosti izjav s poligrafom

Objavljene validacijske študije, zajete v pregledu (tabela 4–8), kažejo razkorak med prakticističnim dojemanjem poligrafije, miti o poligrafiji in pristopi, ki zdržijo znanstveno verifikacijo. Analiza, ki jo predstavljamo (tabela 9), morda pove manj o metodah kakor o doslej opravljenem validiranju teh metod. Ne zadeva tudi znanstveno preverjenih principov izvedbe psihofiziološke preiskave verodostojnosti izjav s poligrafom, na katere smo že opozarjali (Selič, 2002, 2006).

Veljavni principi so temeljni elementi veljavnih tehnik (Krapohl, 2006). Seveda pa ni vsaka tehnika, ki vključuje veljavne principe, nujno tudi ustrezno validirana. Prav tako tehnika, ki se dolgo uporablja v nekem prostoru, ne upošteva nujno veljavnih znanstvenih principov, na kar smo opozorili v zvezi z Rosovim CQT (Selič, 2004). Rosov CQT-format zaradi generično RI-strukture otežuje numerično vrednotenje poligramov in povečuje možnost napake v ocenjevanju, in sicer v obeh smereh (lažno pozitivna ocena/lažno negativna ocena). Poleg tega je izvorna zamisel numeričnega vrednotenja primerjati dve sosednji vprašanji – kontrolno in relevantno, ki skupaj tvorita polje/pas/trak, kar v Rosovim CQT

ni mogoče. Nesporno je namreč ugotovljeno, da naslednje relevantno-irelevantno vprašanje izzove statistično pomembno nižje točkovne vrednosti kakor naslednje relevantno-primerjalno vprašanje pri osebah, ki zavajajo, kakor tudi pri tistih, ki odgovarjajo iskreno (v zvezi z obravnavanim dogodkom) (Krapohl, 2006). Kakršnekoli modifikacije in mešanje izvorno izključujočih se vprašanj oziroma formatov in/ali konceptov lahko povzroči velike težave in oteži interpretacijo (razvidno v Roso, 1996:149–159), kar je ugotovil tudi Horvath (1991), ko je v Reidovem CQT-formatu preverjal veljavnost Backsterjevih kontrolnih vprašanj. Takšni posegi v realnih primerih so nedopustni, običajno jih izvajajo praktiki s pomanjkljivo temeljno poligrafsko izobrazbo in pomanjkljivim razumevanjem filozofije testov s primerjalnimi vprašanji in vloge primerjalnih vprašanj sploh.

Dosledno upoštevanje veljavnih in znanstveno preverjenih principov lahko odgovori na večino izzivov pri aplikaciji poligrafskih tehnik. Ustrezen nadzor kakovosti izvedbe psihofiziološke preiskave verodostojnosti izjav s poligrafom pa mora zagotoviti, da se uporabljajo enotematski testi, ustrezna pravila odločanja, da se pri vrednotenju upošteva vsaj devet kriterijev, izvedba prilagoditvenega testa je nujna in poveča točnost odločitev pri »realnih« testih (v zvezi z obravnavanim dogodkom) (Krapohl, 2006). Odstopanje od teh pravil pomeni nestrokovno rabo poligrafskega instrumenta in zlorabo koncepta forenzične psihofiziologije.

Literatura

1. Abrams, S. (1989). **The complete polygraph handbook**. Lexington, MA/Toronto: Lexington Books.
2. Abrams, S., Abrams, J. (1993). **Polygraph testing of the pedophile**. Portland: Ryan Gwinner Press.
3. American society for testing and materials (2005). Standard guide for the conduct of Psychophysiological deception detection (PDD) screening examinations. ASTM E2386 – 04 Standard Guide for the Conduct of PDD Screening Examinations, dostopno na <http://www.astm.org/Standards/E2386.htm> (15. 8. 2009).
4. Ansley, N. (2008). The history and accuracy of guilty knowledge and peak of tension tests. **Polygraph**, 37 (1) 49–99.
5. Arther, R. (1970). Peak of tension: Question formulation. **Journal of polygraph studies**, 4 (5) 1–4.
6. Backster, C. (1969). **Standardized polygraph examination notebook** (5th ed. rev.). Backster Associates, San Diego: California.
7. Bell, B.G., Raskin, D.C., Honts, C.R., Kircher, J.C. (1999). The Utah numerical scoring system. **Polygraph**, 28 (1), 1–9.
8. Ben-Shakhar, G., Bar-Hillel, M., Lieblch, I. (1986). Trial by polygraph: Scientific and juridical issues in lie detection. **Behavioral Sciences and the Law**, 4 (4) 459–479.
9. Blackwell, N.J. (1998) PolyScore 3.3 and psychophysiological detection of deception examiner rates of accuracy when scoring examination from actual criminal investigations. **Polygraph**, 28 (2) 149–175.

10. Blalock, B. (2009). Directed Lie Screening Test (DLST). **Annual seminar of the American polygraph association**, Nashville, Tennessee (CD).
11. Correa, E.J., Adams, H.E. (1981). The validity of the pre-employment polygraph examination and the effects of motivation. **Polygraph**, 10 (3) 143–155.
12. Federal Bureau of Investigation (1985). Peak of tension (POT) test construction. **Annual seminar of the American polygraph association**, Reno: Nevada (CD).
13. Furedey, J.J., Ben-Shakhar, G. (1991). The roles of deception, intention to deceive, and motivation to avoid detection in the psychophysiological detection of guilty knowledge. **Psychophysiology**, 28 (2) 163–171.
14. Honts, C.R., Hodes, R.L., Raskin, D.C. (1985). Effects of physical countermeasures on the physiological detection of deception. **Journal of Applied Psychology**, 70 (1) 177–187.
15. Honts, C.R., Raskin, D.C., Kircher, J.C. (1987). Effects of physical countermeasures and their electromiographic detection during polygraph tests for deception. **Journal of Psychophysiology**, 1 (3) 241–247.
16. Honts, C.R., Raskin, D.C., Kircher, J.C. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. **Journal of Applied Psychology**, 79 (2) 252–259.
17. Horowitz, S.W., Kircher, J.C., Honts, C.R., Raskin, D.C. (1997). The role of comparison questions in physiological detection of deception. **Psychophysiology**, 34 (1) 108–115.
18. Horvath, F.S. (1977). The effect of selected variables on interpretation of polygraph records. **Journal of Applied Psychology**, 62 (2) 127–136.
19. Horvath, F.S. (1988). The utility of control questions and the effects of two control question types in field polygraph techniques. **Journal of Police Science and Administration**, 16 (3) 198–209.
20. Horvath, F.S. (1991). The utility of control questions and effects of two control questions tests on field polygraph techniques. **Polygraph**, 20 (1): 7–25.
21. Kircher, J.C., Raskin, D.C. (1988). Human versus computerized evaluations of polygraph data in a laboratory setting. **Journal of Applied Psychology**, 73 (2), 291–302.
22. Kircher, J.C., Raskin, D.C. (2002). Computer methods for the psychophysiological detection of deception. In: Kleiner, M. (ed.) **Handbook of polygraph testing**. San Diego/London: Academic Press, 287–326.
23. Krapohl, D.J. (2005). Polygraph Decision Rules for Evidentiary and Paired –Testing (Marin Protocol) Applications. **Polygraph**, 34 (3) 184–193.
24. Krapohl, D.J., Senter, S.M., Stern, B.A. (2005). An Exploration of Methods for the Analysis of Multiple-Issue Relevant/Irrelevant Screening Data. **Polygraph**, 34 (1) 47–61.
25. Krapohl, D.J., Cushman, B. (2006). Comparison of Evidentiary and Investigative Decision Rules: A Replication. **Polygraph**, 35 (1), 55–64.
26. Krapohl, D.J., McCloughan, J.B., Denter, S.M. (2009). How to use the concealed information test. **Polygraph**, 38 (1) 34–49.
27. Light, G.D. (1999). Numerical evaluation of the Army zone comparison test. **Polygraph**, 28 (1), 37–45.
28. Lykken, D.T. (1959). The GSR in the detection of guilt. **Journal of Applied Psychology**, 43 (6) 385–388.
29. Lykken, D.T. (1960). The validity of the guilty knowledge technique: The effects of faking. **Journal of Applied Psychology**, 44 (4) 258–262.
30. MacLaren, V.V. (2001). A qualitative review of the Guilty Knowledge Test. **Journal of Applied Psychology**, 86 (4) 674–683.
31. Marin, J. (2000). He said/She said: Polygraph evidence in court. **Polygraph**, 29 (4), 299–304.
32. Marin, J. (2001). The ASTM exclusionary standard and the APA »litigation certificate« program. **Polygraph**, 30 (4), 288–293.
33. National Research Council (2003). **The polygraph and lie detection**. National Academic Press: Washington, DC.
34. Raskin, D.C., Hare, R.D. (1978). Psychopathy and detection of deception in prison in a prison population. **Psychophysiology**, 15 (1) 126–136.
35. Roso, Z. (1996). **Poligraf u kriminalistici**. Zagreb: Ministarstvo unutrašnjih poslova Republike Hrvatske.
36. Rovner, L.I. (1986). The accuracy of physiological detection of deception for subjects with prior knowledge. **Polygraph**, 15 (1) 1–39.
37. Selič, P. (2000). Razgovor s osumljenim težjega kaznivega dejanja. **Revija za kriminalistiko in kriminologijo**, 51 (1) 42–49.
38. Selič, P. (2002). Možnosti in omejitve psihofiziološke preiskave verodostojnosti izjav s poligrafom v realnih pogojih slovenske kriminalistične prakse. **Varstvoslovje**, 4 (3) 235–244.
39. Selič, P. (2004). Ali je izvedba poligrfske metode ob razkritih podrobnostih obravnavanega dogodka res nesmiselna? : zanesljivost in veljavnost t. i. »direktne metode«. V: Lobnikar, B. (ur.). **Dnevi varstvoslovja** [Elektronski vir]. Zbornik prispevkov. Ljubljana: Fakulteta za policijsko-varnostne vede, 593–602.
40. Selič, P. (2004). Psihofiziološka preiskava verodostojnosti izjav s poligrafom. In: Maver, D. et al. **Kriminalistika: uvod, taktika, tehnika**. Ljubljana: Uradni list, 353–376.
41. Selič, P. (2006). Raziskovalni izzivi sodobnega psihofiziološkega preiskovanja verodostojnosti izjav s poligrafom. **Revija za kriminalistiko in kriminologijo**, 57 (4) 340–348.
42. Selič, P. (2007). Osnovni psihološki testi – psihološki instrumenti za merjenje in ocenjevanje osebnosti in sposobnosti. V: Kersnik, J. (ur.), **Diagnostične preiskave za vsakdanjo uporabo**. Ljubljana: Zavod za razvoj družinske medicine, 461–467.
43. Senter, S.M. (2003). Modified General Question Test Decision Rule Exploration. **Polygraph**, 32 (4) 251–264.
44. Senter, S.M., Dollins, A.B., Krapohl, D.J. (2004). A Comparison of Polygraph data Evaluation Conventions Used at the University of Utah and the Department of Defense Polygraph Institute. **Polygraph**, 33 (4) 214–223.
45. Senter, S.M., Dollins, A.B. (2004). Comparison of Question Series and decision Rules: A Replication. **Polygraph**, 33 (4) 223–234.
46. Senter, M.S., Ryan, A.H. (2005). The Impact of Averaging Assigned Scores on Polygraph Decision Accuracy. **Polygraph**, 34 (1) 10–24.
47. Veritas Center (2005). *Polygraph issues*. Dostopno na http://www.veritascenter.org/documents/Polygraph_FAQ.htm (25. 7. 2009)
48. Yankee, W.J. (1991). **Tehncial correction of curriculum**. Department of defense polygraph institute. Fort McClellan: Alabama.
49. Yankee, W.J., Powell, J.M., Newland, R. (1985). An investigation of the accuracy and consistency of polygraph chart interpretation by inexperienced and experienced examiners. **Polygraph** 14 (2) 108–117.

Analytical review of validated polygraph techniques

Polona Selič, Ph. D., Department of Family Medicine, Faculty of Medicine, Vrazov trg 2, 1000 Ljubljana, Slovenia;
PARES, Institute for Psychophysiological Studies, Koroška 26, 1000 Ljubljana, Slovenia

Introducing typical differences between testing in psychology and forensic psychophysiology, the main metric characteristics are described. A comprehensive taxonomy of polygraph techniques is presented for the first time in Slovenian scientific publishing. The non-weighted overall accuracy for different polygraph techniques was calculated, based on published validity studies in recent decades. The analytical overview includes studies referring to known and accepted polygraph techniques, which divided tested subjects into deceptive and non-deceptive subgroups, with a known "ground truth" (outside criterion) and using a numerical scoring system, but not with a computerized algorithm administered.

The results shown in Tables 4-8 indicate that the ASTM criterion is only met by Utah ZCT, with Federal ZCT missing the threshold by 1 per cent point. For investigative use of polygraph method, the administration of Federal ZCT, Reid Technique, Concealed Information Test and Relevant-Irrelevant Screening Test are admissible. The author expects that future validity studies may produce rather different outcomes, since using decision rules significantly increases the accuracy of the polygraph examiner's decisions and therefore the accuracy of the technique in question. Overall accuracy (Table 9) may be a function of validity research conducted so far, rather than the validity of polygraph techniques included per se. It does not expand on proven and accepted principles of conduct in the polygraph field, which should be known and obeyed.

Key words: validity of polygraph techniques, taxonomy of accurate approaches in deception detection, comparison question test, concealed information test, decision rules, accuracy of polygraph examination, polygraph, psychophysiology

UDC: 343.985.3:343.144