

# Acta Psychologica Tyrnaviensia

# 21

25. výročie Katedry psychológie FF TU  
na obnovenej Trnavskej univerzite (1992-2017)

TRNAVA 2017

ACTA PSYCHOLOGICA TYRNAVIENSIA 21

EDITOR

Doc. PhDr. Marián Špajdel, PhD.

VYDAVATEĽ A TLAČ

Spolok Slovákov v Poľsku  
v spolupráci s Filozofickou fakultou  
Trnavskej univerzity v Trnave

Towarzystwo Słowaków w Polsce  
ul. św. Filipa 7, 31-150 Kraków  
zg@tsp.org.pl, www.tsp.org.pl

Filozofická fakulta Trnavskej univerzity v Trnave  
Hornopotočná 23, 918 43 Trnava  
+421 33 5939303  
katpsych@truni.sk, fff.truni.sk

© Towarzystwo Słowaków w Polsce, 2017

© Filozofická fakulta Trnavskej univerzity v Trnave, 2017

ISBN 978-83-8111-026-6

# Aplikácia moderných psychometrických postupov v adaptácii a analýze psychologických testov

PETER HALAMA · PETER ŽITNÝ · MICHAL KOHÚT

Katedra psychológie, Filozofická fakulta Trnavskej univerzity v Trnave

**Anotácia** · Štúdia oboznamuje čitateľov s psychometrickým výskumom, ktorý bol realizovaný na Katedre psychológie Filozofickej fakulty Trnavskej univerzity od čias jej vzniku až po súčasnosť. Počiatky psychometricky orientovaného výskumu boli v znamení štúdií zameraných na psychometrickú analýzu a adaptáciu významných diagnostických metód. Prvý projekt s psychometrickým cieľom sa venoval výskumu psychoterapie, pričom v rámci neho bolo adaptovaných niekoľko metód na overenie efektu terapie (napr. CORE-OM, OQ-45). Neskoršie projekty sa zameriavali viac na moderné psychometrické prístupy umožňujúce pokročilú analýzu meracích nástrojov či testovacie postupy (teória odpovede na položku, počítačové adaptívne testovanie, odlišné fungovanie položiek). Štúdia prezentuje výsledky získané v rámci psychometricky orientovaného výskumu na katedre, ako aj jeho významu pre psychologickú komunitu.

**Kľúčové slová** · psychometria, teória odpovede na položku, odlišné fungovanie položiek

Psychometria je vo všeobecnosti definovaná ako pomocná psychologická veda, ktorá sa zaoberá teoretickými problémami merania v psychológii (Urbánek, 2002). Jej úlohou je formulovať princípy konštrukcie a hodnotenia meracích nástrojov v psychológii, ako aj spôsoby a postupy, ktoré umožňujú interpretáciu ich výsledkov. Psychometria slúži najmä pri vývoji nových testových metód, pri ich štandardizácii či revízii, a jej poznatky sú nevyhnutné pri pochopení teoretického základu psychodiagnostiky (Halama, 2011).

Na katedre psychológie sa psychometrický výskum realizuje dlhodobo, pričom pracovníci katedry získali niekoľko grantových projektov, ktorými bol tento výskum financovaný. Začiatky a inšpirácie psychometrického výskumu na katedre sú spojené s menom nášho, bohužiaľ už zosnulého, kolegu Mariána Košča. V rámci svojho klinického zamerania, dával veľký dôraz na psychologickú diagnostiku, pričom zdôrazňoval, aká potrebná je kvalitná

psychometrická analýza nástrojov, aby mohli dobre slúžiť svojmu účelu. Sám sa venoval analýze vybraných diagnostických nástrojov, a povzbudzoval kolegov k realizácii psychometrických štúdií. Aj jeho zásluhou sa súčasťou študijného programu psychológie stal predmet Teória psychodiagnostiky, kde sa študenti oboznamovali s teoretickými základmi psychodiagnostiky vrátane psychometrie. Aj keď sa toho nedožil, inšpiroval vznik pôvodnej slovenskej učebnice s názvom Princípy psychologickéj diagnostiky (Halama, 2005), ktorá sa stala povinnou literatúrou pre študentov na tomto predmete a dočkala sa aj druhého vydania (Halama, 2011). Tým bol nastavený trend, že pracovníci katedry psychológie nie sú len pasívnymi užívateľmi psychodiagnostických metód vo svojej praxi a výučbe, ale svojou aktivitou prispievajú k zvyšovaniu kvality a dostupnosti psychodiagnostických nástrojov na Slovensku.

### **1 Psychometrická analýza vybraných diagnostických nástrojov (grant VEGA č. 1/4520/07)**

Samotný psychometrický výskum na katedre psychológie sa dá rozdeliť do niekoľkých etáp, ktoré sa do určitej miery prekrývajú s riešenými grantovými projektmi. V rámci prvotnej fázy, ktorej začiatok sa nedá úplne presne identifikovať, bol kladený dôraz najmä na analýzu psychometrických vlastností vybraných meracích či psychodiagnostických nástrojov a prípadnú adaptáciu významných zahraničných metód na slovenské podmienky. Už v roku 1998 publikoval M. Košč psychometrickú štúdiu zameranú na analýzu Rey-Osterriethovej komplexnej figúry (Košč, Jariabková, 1998), ktorá je významným nástrojom v neuropsychologickej psychodiagnostike. Jedna z najvýznamnejších adaptácií sa uskutočnila v spolupráci s Ústavom experimentálnej psychológie SAV. Jej predmetom bola adaptácia svetovo široko používaného nástroja na merania piatich osobnostných črt s názvom NEO-FFI (Ruisel, Halama, 2007). Tento 60 položkový dotazník meria päť osobnostných črt tak, ako sú definované v teórii piatich veľkých faktorov, a to neurotizmus, extravertzia, otvorenosť voči skúsenosti, privetivosť a svedomitosť. Jeho adaptácia a štandardizácia bola realizovaná na výskumnom súbore viac než 1000 osôb, pričom bola realizovaná analýza reliability, analýzy položiek ako aj overená jej faktorová štruktúra. Slovenský manuál, ktorý vyšiel vo vydavateľstve Testcentrum, sprístupnil dotazník slovenským psychológom v praxi aj vo výskume.

Výrazným posunom v psychometrickom výskume na katedre psychológie FF TU bolo získanie grantového projektu VEGA č. 1/4520/07, ktorého vedúcim riešiteľom bol vtedajší pracovník katedry a významný odborník na

výskum psychoterapie L. Timulák. Grantový projekt mal názov Psychologická zmena a poskytovanie psychoterapie v praxi, a jeho primárnym cieľom bolo riešenie teoretických a metodologických súvislostí psychoterapeutickej zmeny a jej zhodnocovania. Podstatný priestor v riešení projektu bol venovaný psychometrickým analýzám a zberu normatívnych dát vybraných dotazníkov na meranie psychoterapeutickej zmeny, ktorej sa venoval vtedajší doktorand a neskôr aj pracovník katedry M. Bieščad. V rámci riešenia projektu sa uskutočnila slovenská adaptácia dotazníka CORE-OM. Ide o 34-položkovú sebakýpoveďovú škálu zostrojenú za účelom sledovania zmien u pacientov/klientov v priebehu terapie. Je určená pre široký okruh pacientov a klientov, takisto pre celú škálu zariadení zabezpečujúcich starostlivosť o duševné zdravie. Škála má štyri základné dimenzie: well-being, problémy/symptómy, interpersonálne fungovanie a rizikové správanie. Pri adaptácii a psychometrickom overovaní škály bola nadviazaná intenzívna spolupráca s členom autorskej skupiny originálnej britskej verzie škály C. Evansom, ktorý podporil adaptáciu dotazníka a výrazným spôsobom pri nej asistoval. Overovanie škály na prvotných výberoch klinického aj neklinického charakteru vykázalo dobré psychometrické vlastnosti a dostatočne diskriminovalo medzi klinickou a neklinickou populáciou (Gampe et al. 2007). CORE-OM sa osvedčilo aj na meranie zmeny v prostredí konkrétnej kliniky, pričom preukázalo dostatočnú citlivosť na zmenu (Bieščad, Hašto, 2008). Najvýznamnejším výstupom tohto výskumu bolo vypracovanie slovenskej verzie manuálu (Bieščad, 2009), do ktorého boli zapracované poznatky o psychometrických vlastnostiach a normatívne údaje vypočítané na základe zberu dát na klinickej (n=173) a neklinickej populácii (n=237). Dotazník je tak voľne dostupný pre slovenských výskumníkov ako aj psychológov v praxi, ktorí majú potrebu merať účinnosť vlastných terapeutických intervencií.

Okrem dotazníka CORE-OM bol v rámci riešenia projektu realizovaný projekt a prvotná adaptácia Dotazníka výsledku-45 (angl. Outcome Questionnaire, OQ-45). Ide takisto o dotazník na meranie efektu psychoterapie, pričom tento dotazník obsahuje tri dimenzie: symptómy distresu (depresia a úzkosť), interpersonálne vzťahy (osamelosť, konflikty) a sociálna rola (ťažkosti v práci, škole atd.). Analýza dotazníka potvrdila jeho validitu vo vzťahu k iným dotazníkom, ako Beckov inventár depresie, SCL-90 a Škály hodnotenia výsledku (Zvarová, Bieščad, 2007). V ďalšej fáze bola na väčšom klinickom aj bežnom výbere realizovaná analýza viacerých nástrojov, pričom boli porovnávané štyri svetovo najpoužívanejšie dotazníky: CORE-OM a OQ-45, Symptom Checklist-10-R, a Škála hodnotenia výsledku (ORS). Všetky dotazníky ukázali dostatočné psychometrické vlastnosti, boli tiež odvodené

ich percentilové normy a indexy spoľahlivej zmeny (Bieščad, 2007). Bola tiež skúmaná senzitivita jednotlivých dotazníkov na psychoterapeutickú zmenu, pričom ako najcitlivejší sa ukázal CORE-OM, pričom zmeny reflektovali teoretické formulácie o typickom priebehu terapeutickje intervencie (Bieščad, 2007, Bieščad, Timulák, 2008). Tento fakt bol bližšie analyzovaný v štúdiu Bieščada (2008), ktorá potvrdila, že klient s najkratšou dobou liečenia dosiahli najväčšie zmeny v oblasti well-beingu, klienti so strednou dobou liečby v oblasti symptómov, a klienti s najdlhšou liečbou vo všetkých oblastiach vrátane sociálneho fungovania.

Časť psychometrických analýz v rámci projektu sa zameriavala aj na nástroje na meranie životnej zmyslupnosti a to vzhľadom na rozvoj logoterapie na Slovensku, ku ktorému kľúčovým spôsobom prispeli aj pracovníci katedry. Bol podaný prehľad nástrojov na meranie životnej zmyslupnosti a príbuzných konštruktov vychádzajúcich najmä z logoterapie (Halama, 2009a). V dvoch psychometrických štúdiách bola realizovaná analýza dvoch nástrojov na meranie životnej zmyslupnosti. Pôvodná slovenská Škála životnej zmyslupnosti preukázala dostatočnú vnútornú konzistenciu avšak pôvodná trojkomponentová štruktúra sa ukázala ako menej stabilná. Boli vypočítané aj stenové normy tejto škály (Halama, 2008). Svetovo najpoužívanejší dotazník na meranie životnej zmyslupnosti Purpose in life Questionnaire bol analyzovaný na bežnej slovenskej populácii. Potvrdila sa jeho dostatočná vnútorná konzistencia, pričom bola identifikovaná trojfaktorová štruktúra (Halama, 2009b).

## **2 Aplikácie teórie odpovede na položky (grant VEGA č. č. 1/0228/10)**

Ďalší výskumný projekt z oblasti psychometrie bol poskytnutý grantovou agentúrou VEGA na tému teórie odpovede na položku, čo je špecifický alternatívny prístup k psychometrickej analýze a hodnoteniu testov. Projekt bol riešený v rokoch 2010 až 2011 a jeho riešiteľmi boli Peter Halama a Peter Žitný. V publikáciách Halamu (2010, 2011) bol predstavený koncept teórie odpovede na položku (IRT) s dôrazom na základné princípy a modely, s ktorými tento koncept pracuje. Podrobne boli popísané dvoj a troj parametrové modely pre dichotómne položky ako aj model stupňovanej odpovede (GRM) pre položky typu Likertovej škály, ktoré boli neskôr použité vo výskumných analýzach. Bola uskutočnená komparácia teórie odpovede na položky s bežne používanou klasickou teóriou testov (CTT), a to v rovine teoreticko-metodologickej ako aj v rovine empirickej. V rovine teoreticko-metodologickej boli identifikované matematicko-štatistické rozdiely najmä v oblasti štandardnej

chyby merania (v IRT je rozdielna na rôznych úrovniach meranej vlastnosti), v oblasti reliability a dĺžky testu (v CTT je reliabilita závislá na počte položiek, v IRT nie), a v oblasti odvodu skóre testu (v CTT normatívne, v IRT probabilisticky). V empirickej rovine boli analyzované výsledky iných štúdií porovnávajúcich oba prístupy pri praktickom meraní. IRT aj CTT prinášali podobné výsledky v položkovej analýze, avšak koeficienty odhadnuté prostredníctvom IRT boli stabilnejšie v rôznych podmienkach a testových súboroch. Určité rozdiely medzi oboma prístupmi boli identifikované pri zaradovaní do skupín podľa typov a pri meraní psychoterapeutickej zmeny, kde sa tieto dva prístupy nie vždy zhodovali. Výskum Halamu a Bieščada (2011a; 2011b) sa zamerával na porovnanie IRT a CTT v oblasti merania psychoterapeutickej zmeny. Vo výskume bolo analyzované, ako sa aplikácia každého z prístupov prejaví pri identifikácii signifikantnej therapeutickej zmeny pričom ako nástroj na identifikáciu signifikantnej zmeny bol v oboch prípadoch použitý 95-percentný konfidenčný interval štandardnej chyby merania, ktorý bol aplikovaný na rozdielové skóre merania pred a po intervencii. Ako merací nástroj bol použitý sebaučtový symptomatický dotazník CORE-OM, ktorý bol v dvoch fázach (pred a po intervencii) administrovaný 166 klientom v psychologickú a psychiatrickú starostlivosť. Ukázalo sa, že vyhodnotenie zmeny prostredníctvom IRT vedie k odhadnutiu väčšieho počtu osôb ako tých, u ktorých nastala signifikantná zmena. Pri ďalšej analýze sa ukázalo, že tento rozdiel súvisí s povahou štandardnej chyby merania u oboch prístupov, pričom v CTT je stabilná na celej úrovni meranej črty, avšak v IRT je flexibilná a závisí na presnosti merania na konkrétnej úrovni meranej vlastnosti. Pri meracích nástrojoch s vhodnými položkami pre požadovanú úroveň tak môže byť IRT citlivejším pri odhaľovaní signifikantnej zmeny než klasická teória testov. Vo výskumnej štúdií (Žitný, 2011b) bol Dotazník senzitivity voči nespravodlivosti (kognitívna úroveň; Lovaš, 1995) analyzovaný metódami v rámci klasickej teórie testov ako aj prostredníctvom metód teórie odpovede na položku (Samejmov model GRM). V súlade s niektorými predchádzajúcimi zisteniami (napr. Fan, 1998; Halama, Bieščad, 2006; Halama, Bieščad, 2009; Ndalichako, Rogers, 1997) sa potvrdilo, že informácie získané CTT a IRT prístupom sú do značnej miery vzájomne komplementárne, čím sú obidve psychometrické koncepcie využiteľné pri výbere vhodných položiek meracieho nástroja a zlepšenie jeho psychometrických vlastností. V prehľadovej štúdií (Žitný, 2011a) bola realizovaná komparácia efektívnosti, presnosti a validity počítačového adaptívneho testovania založeného na teórii odpovede na položku s tradičnou formou papier-ceruzka a počítačovou formou administrácie v teoretickej rovine a na základe doterajších výskumov. Štúdia

pojednáva o zisteniach 15-tich výskumných štúdií z oblasti testovania schopností, klinickej psychológie, testovania osobnosti a zdravotníckej starostlivosti zameraných na skúmanie reliability, užitočnosti (v zmysle úspory položiek) a validity (v zmysle korelácií s existujúcimi nástrojmi) CAT. Celkovo boli tieto zistenia povzbudivé – CAT poskytuje efektívny prostriedok pre získanie optimálneho množstva informácie potrebnej pre zodpovedanie posudzovanej otázky, a to využitím minimálneho množstva času a/alebo počtu položiek pre získanie danej informácie. CAT skóre silno korelovalo so skóre z celej položkovej banky (rozpätie  $r = 0,83 - 0,99$ ) a stredne silno so zaužívanými nástrojmi (rozpätie  $r = 0,58 - 0,83$ ) poskytujúc dôkazy pre reliabilitu, validitu a porovnateľnosť adaptívnych nástrojov. Vo výskumnej štúdii (Žitný et al., 2012) bola v spolupráci s odborníkmi z Psychologického ústavu AV ČR analyzovaná prediktívna a konštruktová validita CAT využitím real-data simulácie za účelom porovnania počítačového adaptívneho testovania s administráciou testov kognitívnych schopností vo forme papier-ceruzka a cez počítač. Výskum bol vykonaný na súbore 801 slovenských študentov stredných škôl (363 mužov, 438 žien). Testová batéria pozostávala z testov kognitívnych schopností: Test intelektového potenciálu (Říčan, 1971) a Viedenský maticový test (Vonkomer, 1992). Základný dizajn štúdie zahŕňal tri typy údajov pre porovnanie, reálnu administráciu papier-ceruzka (567 účastníkov: 279 mužov, 288 žien) a počítačovú administráciu (236 účastníkov: 84 mužov, 152 žien), a post-hoc simulačnú počítačovú adaptívnu administráciu založenú na „re-administrácii“ údajov z počítačovej verzie (236 účastníkov). Celkovo sa z výsledkov ukázalo, že prediktívna a konštruktová validita CAT bola adekvátne porovnateľná cez módy administrácie. V súlade s predchádzajúcim výskumom CAT používa len malé množstvo položiek dávajúc výsledky, ktoré, pokiaľ ide o validitu, sú len nepatrne odlišné od výsledkov tradičnej verzie papier-ceruzka alebo administrácie cez počítač. Simulovaná CAT administrácia TIP bola zhruba o 56 % a VMT o 54 % úspornejšia ako tradičné verzie. Tieto úspory položiek sú väčšie než je bežné u non-IRT CAT aplikácií. Kontrola údajov o administrácii položiek ukázala, že celá položková banka bola adaptívne administrovaná len 4.7 % (TIP) a 5.1 % (VMT) účastníkom.

### 3 Odlišné fungovanie položiek (grant VEGA č. 1/0234/15)

Posledným a v súčasnosti ešte prebiehajúcim katedrovým projektom VEGA, ktorého obsahom je psychometrický výskum, je projekt VEGA č. 1/0234/15. Projekt sa venuje špecifickému konceptu s názvom odlišné fungovanie položky (differential item functioning, skr. DIF). Tento pojem vyjadruje stav, kedy



je odlišná pravdepodobnosť správnej (napr. pri intelligenčných testoch) alebo kľúčovej (napr. pri osobnostných dotazníkoch) odpovede na položku odlišná medzi oddelenými skupinami respondentov (rod, vek, miesto bydliska, viera a pod.), napriek tomu, že majú rovnakú alebo podobnú mieru meranej vlastnosti, ako napríklad verbálna inteligencia, extravertzia a podobne (Clauser, Mazor, 1998). V rámci riešenia grantovej úlohy bola prvotná snaha o informovanie odbornej verejnosti ohľadom konceptu odlišného fungovania položiek. Prehľadová štúdia Halamu a Kohúta (2016) predstavila túto tému najmä z hľadiska testového skreslenia a spravodlivosti v psychodiagnostike. Článok okrem teoretického ukotvenia a histórie konceptu informuje tiež o základných a využívaných postupoch štatistickej analýzy. Dôležitou súčasťou tejto štúdie je prehľad výskumných zistení ohľadom odlišného fungovania položiek psychodiagnostických testov a dotazníkov.

V ďalšej prehľadovej štúdii, Kohút (2016) informoval o druhom dôležitom využití skúmania DIF. Okrem pôvodného zamerania na spravodlivosť a odstránenie skreslenia, koncept DIF môže nájsť dôležité uplatnenie pri posudzovaní validity psychodiagnostických metodík. DIF v chápaní multidimenzionálnej teórie vzniká tým, že položka zachytáva okrem primárnej dimenzie, na meranie ktorej bola skonštruovaná, aj sekundárne dimenzie. V miere primárnej dimenzie sa respondenti zo skúmaných skupín nelíšia (rovnaká miera meranej vlastnosti), no odlišujú sa v miere sekundárnych dimenzií, resp. sekundárnej dimenzie, čím vzniká rozdielna pravdepodobnosť správnej alebo kľúčovej odpovede (Roussos, Stout, 1996) Je však dôležité podotknúť, že samotná prítomnosť DIF automaticky neznamená narušenie validity, keďže sekundárna dimenzia môže byť v spojení s meraným konštruktom, rozdiely sú teda reálne, no môže ísť aj o schopnosti alebo vlastnosti, ktoré nebolo v pláne merať – v takom prípade môžeme hovoriť o skreslení a narušenej validite.

Analýzu DIF v Teste Štruktúry Inteligencie 2000 R (I-S-T 2000 R) predstavili Kohút et al. (2016), ktorí sa zamerali na DIF medzi mužmi a ženami. V štúdii realizovanej v spolupráci s Výskumným ústavom detskej psychológie a patopsychológie, bolo identifikovaných viacero položiek, pri ktorých sa líšila pravdepodobnosť kľúčovej odpovede medzi mužmi a ženami. Zaujímavé zistenia sa týkajú najmä obsahu položiek vo verbálnych subtestoch tejto metodiky. Položky, pri ktorých mali muži vyššiu pravdepodobnosť kľúčovej odpovede sa týkali napríklad mechaniky alebo prírody, položky pri ktorých mali vyššiu pravdepodobnosť kľúčovej odpovede ženy sa týkali stravy a zdravia, kuchynského zariadenia, emócií alebo obrázkov a čítania. Odstránenie položiek, ktoré vykazovali DIF z hrubého skóre spôsobilo pomerne výrazný

rozdiel iba v škále Verbálne Analógie, kde bolo identifikovaných 6 DIF položiek a pri každej z nich majú ženy vyššiu pravdepodobnosť kľúčovej odpovede. Pre potvrdenie prítomnosti skreslenia sú potrebné ďalšie výskumy, no Kohút et al. (2016) otvorili priestor pre zamyslenie nad obsahom položiek a možným dopadom týchto obsahov v inteligenčnom testovaní.

Kohút a kol. (2017) skúmali DIF vo verbálnych subtestoch tohto I-S-T 2000 R, medzi skupinou obyvateľov miest (nad 10000 obyvateľov) a obcí (do 9999 obyvateľov). Celkovo identifikovali 4 položky pri ktorých mali vyššiu pravdepodobnosť kľúčovej odpovede obyvateľa miest a 2 pri ktorých to boli obyvateľia obcí. Všetky identifikované položky vykazovali len nízku mieru DIF a ich odstránenie nevyústilo do výraznejšej zmeny rozdielov medzi skupinami. Z hľadiska obsahu je zaujímavá položka, ktorá vykazovala najsilnejšie DIF zo všetkých identifikovaných. Táto položka priamo operuje s pojmom obec a vyššiu pravdepodobnosť kľúčovej odpovede majú obyvateľia obcí. Napriek prítomnosti DIF položiek, autori uvádzajú, že neexistuje dôvodné podozrenie z prítomnosti závažnejšieho skreslenia z hľadiska typu bydliska (Kohút a kol., 2017).

Skúmanie DIF nachádza uplatnenie aj pri osobnostných dotazníkoch. Kohút a Halama (2016) uskutočnili pilotnú analýzu odlišného fungovania položiek v slovenskej revidovanej verzii Eysenckovho osobnostného dotazníka (EPQ-R) celkovo identifikovali 8 položiek, ktoré podľa použitých kritérií vykazovali aspoň stredne silné DIF. Pri všetkých týchto položkách mali vyššiu pravdepodobnosť kľúčovej odpovede muži. Z hľadiska počtu DIF položiek bola najproblematickejšia škála Psychoticizmus, ktorá obsahovala 4 stredne silne alebo silne odlišne fungujúce položky. Neskôr, Kohút a Halama (v tlači) rozšírili tieto analýzy použitím iných metód a kritérií. Identifikovali 27 až 35 položiek (27-35%), v závislosti od použitej metódy skúmania, ktoré vykazovali rôzne závažné DIF medzi mužmi a ženami. Množstvo identifikovaných položiek však vykazovalo stredne závažné DIF len pri jednej z použitých metód, pričom v ostatných metódach sa toto DIF nepotvrdilo. Štúdia určitým spôsobom poukázala na nekonzistentnosť identifikovaných DIF položiek ako aj nekonzistenciu ich kategorizácie podľa závažnosti DIF. Napriek tomu, je vhodné použiť viacero metód pre skúmanie DIF, a to buď pre zachytenie čo najvyššieho počtu potenciálne odlišne fungujúcich položiek, alebo pre zníženie tohto počtu vybraním položiek, ktoré boli identifikované viacerými, popr. všetkými použitými metódami (Fidalgo, Ferreres, Muniz, 2004). Kohút a Halama (v tlači) uvádzajú 8 položiek, ktoré boli označené ako stredne silne alebo silne DIF všetkými tromi použitými metódami. Pri dvoch z týchto položiek majú vyššiu pravdepodobnosť kľúčovej odpovede ženy. Obe položky

sú zo škály Neuroticizmus a ich obsahom je téma zraniteľnosti citov a precitlivenosti. Najviac položiek, pri ktorých mali muži vyššiu pravdepodobnosť kľúčovej odpovede bolo v škále Psychoticizmus a ich obsahom boli témy ohľadom nízkej rozrušenosti z utrpenia dieťaťa alebo zvierťa, nízkej ľútosti nad zvieratom chytením do pasce alebo robením si žartov, ktoré môžu niekomu ublížiť. Okrem uvedenia výsledkov a porovnania jednotlivých metód skúmania DIF, autori v tejto štúdii diskutujú možné príčiny alebo faktory podieľajúce sa na vzniku DIF. Zhodnotenie možných faktorov vzniku odlišného fungovania položiek je veľmi dôležité pre posúdenie dopadu DIF na spravodlivosť a validitu psychodiagnostických metódik. Tejto téme sa venuje v poslednej prehľadovej štúdii Kohút (v tlači).

## Záver

Ak sa na psychometrický výskum realizovaný na katedre psychológie FF TU pozrieme z nadhľadu, dá sa vidieť tendencia, ktorá postupuje od zamerania sa na adaptáciu a štandardizáciu diagnostických metód, ktoré bolo typické pre počiatočné fázy výskumu, smerom k aplikácii a využívaniu novších pokročilých psychometrických metód, ako je teória odpovede na položku či koncept odlišného fungovania položiek. Tento posun reflektuje zmenu v prvotnom praktickom zameraní na sprístupnenie dôležitých nástrojov slovenskej odbornej verejnosti k oboznámeniu slovenských psychológov s novšími psychometrickými prístupmi, ktoré sú vo svete pomerne známe, avšak u nás sa využívajú zriedkavo. A hoci je ambíciou nášho psychometrického výskumu prispieť aj k medzinárodnému poznaniu, najvýznamnejším prínosom je sprostredkovanie moderných trendov v psychometrii slovenskému vedeckému publiku.

## L I T E R A T Ú R A

- BIEŠČAD, M. (2007). *Aplikácia nástrojov merajúcich výsledky psychoterapie. Porovnanie citlivosti nástrojov merania v jednotlivých oblastiach terapeutickje zmeny*. Dizertačná práca. Trnava: Filozofická fakulta Trnavskej univerzity v Trnave.
- BIEŠČAD, M. (2008). Fázový priebeh psychoterapeutickej zmeny v skupinách pacientov s rôzne dlhou liečebnou intervenciou. In: Grác, J., Halama, P. (Eds.) *Acta Psychologica Tyrnaviensis* 11-12. Trnava, Filozofická fakulta TU, s. 45-55.
- BIEŠČAD, M. (2009). *Slovenská verzia dotazníka CORE-OM. Kompletný manuál pre použitie dotazníka v klinickej praxi*. Trnava: Filozofická fakulta Trnavskej univerzity v Trnave.

- BIEŠČAD, M., HAŠTO, J. A KOL. (2008). Monitorovanie výsledkov psychiatrickej a psychoterapeutickej starostlivosti v zariadení bežnej klinickej praxe. *Psychiatria-Psychoterapia-Psychosomatika*, 15(3), 166-173.
- BIEŠČAD, M., TIMULÁK, L. (2008). *Comparison of the CORE-OM, OQ-45, Outcome Rating Scale and Symptom Checklist-10R used in routine outcome monitoring in Slovakia*. Book of Abstracts. 39th Annual meeting of the Society for Psychotherapy Research, Barcelona, Spain, p. 74.
- CLAUSER, B. E., MAZOR, K. M. (1998). Using Statistical Procedures to Identify Differentially Functioning Test Items. *Educational Measurement: Issues and Practice*, 17, 1, 31-44.
- FAN, X. (1998). Item Response Theory and Classical Test Theory: An Empirical Comparison of their Item/Person Statistics. *Educational and Psychological Measurement*, 58, 3, 357-381.
- FIDALGO, A. M., FERRERES, D., & MUNIZ, J. (2004). Liberal and conservative differential item functioning detection using Mantel-Haenszel and SIBTEST: implications for Type I and Type II error rates. *The Journal of Experimental Education*, 7, 23-39.
- GAMPE, K., BIEŠČAD, M., BALÚNOVÁ-LABANIČOVÁ, E., TIMULÁK, L., EVANS, C. (2007). Slovenská adaptácia metódy CORE-OM (Clinical Outcomes in Routine Evaluation – Outcome Measure). *Česká a slovenská psychiatrie*, 103, 4-13.
- HALAMA P. (2009b) The PIL test in a Slovak sample: internal consistency and factor structure. *The International Forum for Logotherapy*, 2009, 32, 84-88.
- HALAMA, P. (2008). Škála životnej zmysluplnosti – psychometrické vlastnosti a predbežné normy. In: Grác, J., Halama, P. (Eds.) *Acta Psychologica Tyrnaviensis* 11-12. Trnava, Filozofická fakulta TU, s. 56-68.
- HALAMA, P. (2009a) Research instruments for investigating meaning of life and other logotherapeutic constructs. In: Batthyany, A., Levinson, J. (Eds.) *Existential psychotherapy of meaning*. Handbook of logotherapy and existential analysis. Phoenix, Zeig, Tucker & Theisen, Inc., s. 415-444.
- HALAMA, P. (2010) Teória odpovede na položku v psychologickom a pedagogickom testovaní. Porovnanie s klasickou teóriou testov. In: M. Špajdel (Ed.) *Acta Psychologica Tyrnaviensis* 13-14. *Špecifické psychologické prístupy k riešeniu aktuálnych výchovnovzdelávacích problémov*. Trnava, Filozofická fakulta TU, s. 43-60.
- HALAMA, P. (2011, prvé vydanie 2005) *Princípy psychologickej diagnostiky. 2. doplnené vydanie*. Trnava, Filozofická fakulta Trnavskej univerzity v Trnave.
- HALAMA, P., BIEŠČAD, M. (2006). Psychometrická analýza Rosenbergovej škály sebahodnotenia s použitím metód klasickej teórie testov (CTT) a teórie odpovede na položku (IRT). *Československá psychologie*, 50, 6, 588-603.
- HALAMA, P., BIEŠČAD, M. (2009). Item response theory analysis of the CORE-OM. In: *11th European Congress of Psychology ECP09*. Oslo, Norway 7 – 10 July 2009: Abstracts, CD-ROM, 63 pp.
- HALAMA, P., BIEŠČAD, M. (2011a). Comparison of Sensitivity of CTT and IRT Approaches in Measurement of Psychotherapy Change. In: *The 12th European Congress of Psychology*. Istanbul 2011, 4. - 8. July: Abstracts, CD-ROM, 138 pp.
- HALAMA, P., BIEŠČAD, M. (2011b). Meranie psychoterapeutickej zmeny: porovnanie klasického testového skóre a skóre odvodeného prostredníctvom teórie odpovede na položku. *Československá psychologie*, 55, 5, 400-411.
- HALAMA, P., KOHÚT, M. (2016). Odlišné fungovanie položiek (DIF) ako metóda analýzy testového skreslenia v psychodiagnostike. *Psychológia a patopsychológia dieťaťa*, 50, 1, 61-76.
- KOHÚT (v tlači). Odlišné fungovanie položiek (DIF): kvalitatívne posúdenie DIF položiek. *Studia doctoralia Tyrnaviensia* 2017.

- KOHÚT, M. (2016). Koncept odlišného fungovania položky ako spôsob zisťovania validity psychodiagnostických metód. In: E. Juríková, L. Tkáčik (Eds.) *Studia doctoralia Tyrnaviensia 2016* (pp. 54-63). Kraków : Towarzystwo Słowaków w Polsce, Filozofická fakulta Trnavskej univerzity v Trnave.
- KOHÚT, M., HALAMA, P. (2016). Odlišné fungovanie položiek podľa pohlavia v Eysenckovom osobnostnom dotazníku EPQ-R. In: M. Špajdel (Ed.) *Acta Psychologica Tyrnaviensia 19-20* (pp. 77-86). Kraków : Towarzystwo Słowaków w Polsce, Filozofická fakulta Trnavskej univerzity v Trnave.
- KOHÚT, M., HALAMA, P. (v tlači). Uniformné odlišné fungovanie položiek (DIF) u mužov a žien v Eysenckovom osobnostnom dotazníku: porovnanie troch metód analýzy. *Československá psychologie*.
- KOHÚT, M., HALAMA, P., DOČKAL, V., ŽITNÝ, P. (2016). Gender differential item functioning in Slovak version of Intelligence Structure Test 2000 – Revised. *Studia Psychologica*, 58, 3, 238-250.
- KOHÚT, M., HALAMA, P., DOČKAL, V., ŽITNÝ, P. (2017). Odlišné fungovanie položiek vo verbálnych subtestoch slovenskej verzie testu štruktúry inteligencie I-S-T 2000 R podľa miesta bydliska (pp. 243-251). In: L. Pitel (Ed.) *Sociálne procesy a osobnosť 2016*.
- KOŠČ, M., JARIABKOVÁ, K. (1998) Rey-Osterrieth Complex Figure: Ways of investigating cognitive processes. *Studia Psychologica*, 40, 266-270.
- LOVÁŠ, L. (1995). Nespravodlivosť v interpersonalných vzťahoch [Injustice in the interpersonal-relations]. *Československá Psychologie*, 39, 3, 203-212.
- NDALICHAKO, J.L., ROGERS, W.T. (1997). Comparison of finite state score theory, classical test theory, and item response theory in scoring multiple-choice items. *Educational and Psychological Measurement*, 57, 4, 580-589.
- ROUSSOS, L., STOUT, W. (1996). A Multidimensionality-Based DIF Analysis Paradigm. *Applied Psychological Measurement*, 20, 4, 355-371.
- RUISEL, I., HALAMA, P. (2007) *NEO päťfaktorový osobnostný inventár (podľa NEO Five-Factor Inventory P.T. Costu a R.R. McCraeho)*. Praha: Testcentrum - Hogrefe.
- ŘÍČAN, P. (1971). *Test Intelektového potenciálu - TIP*. 1. vyd. Bratislava: Psychodiagnostické a didaktické testy..
- URBÁNEK, T. (2002). *Základy psychometrie*. Brno: Masarykova univerzita.
- VONKOMER, J. (1992). *Viedenský maticový test - VMT*. 1. vyd. Bratislava: Psychodiagnostika.
- ZVAROVÁ, M., BIEŠČAD, M. (2007). *Overenie konštruktivej validity a vzájomných vzťahov Dotazníka výsledku (Outcome Questionnaire – 45) a Škály hodnotenia výsledku (Outcome Rating Scale)*. In Pilárik, L., Pohánka, M., Popelková, M. (Eds.), II. Medzinárodná konferencia doktorandov odborov psychológia a sociálna práca. Zborník z konferencie. [CD-ROM] Nitra: FSVaZ UKF.
- ŽITNÝ, P. (2011a). Presnosť, validita a efektívnosť počítačového adaptívneho testovania. *Československá psychologie*, 55, 2, 167-179.
- ŽITNÝ, P. (2011b). A psychometric analysis of the Sensitivity to Injustice Questionnaire using classical test theory and item response theory. *ECP 2011 - The 12th European Congress of Psychology*. [online], [citované 12.01.2012]. Dostupné na internete: <<http://tinyurl.com/zitny>>
- ŽITNÝ, P., HALAMA, P., JELÍNEK, M., KVĚTON, P. (2012). Validity of cognitive ability tests – comparison of computerized adaptive testing with paper-pencil and computer-based forms of administrations. *Studia Psychologica*, 54, 3, 181-194.