

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

Odlišné fungovanie položiek medzi mužmi a ženami v Inventári Veľkej Päťky (BFI)

M I C H A L K O H Ů T

Katedra psychológie, Filozofická fakulta, Trnavská univerzita v Trnave

Anotácia Odlišné fungovanie položky (DIF) vyjadruje stav, kedy je pravdepodobnosť kľúčovej odpovede na položku odlišná medzi členmi porovnávaných skupín, ktorí majú rovnakú, alebo približne rovnakú mieru meranej vlastnosti (Clauser a Mazor, 1989). Štúdia sa zameriava na identifikáciu odlišne fungujúcich položiek v Inventári Veľkej Päťky (Big Five Inventory; John, Srivastava, 1999) medzi mužmi a ženami. Výskumný súbor tvorilo 1208 respondentov, 552 (45,7%) mužov a 656 (54,3%) žien vo veku od 17 do 67 rokov ($M = 31,57$; $SD = 12,89$). Pre skúmanie DIF bola použitá Mantelova metóda v programe DIFAS, ordinálna logistická regresia a metóda Poly-SIBTEST. Celkovo bolo identifikovaných 12 položiek, ktoré boli aspoň jednou metódou označené ako DIF, z tohto počtu 5 položiek vykazovalo aspoň pri jednej z použitých metód stredne silné alebo silné uniformné DIF. Z hľadiska počtu a sily odlišne fungujúcich položiek nie je dôvodné podozrenie z prítomnosti závažného skreslenia alebo znevýhodňovania mužov či žien. Možné príčiny odlišného fungovania položiek sú diskutované.

Kľúčové slová: odlišné fungovanie položiek, rodové rozdiely, Inventár Veľkej Päťky, skreslenie

V psychológii je využívané veľké množstvo rôznych meracích nástrojov, ktoré nachádzajú uplatnenie v psychologickej praxi a výskume. Neodmysliteľnou súčasťou psychodiagnostiky sú dotazníkové metodiky. Pre zaručenie ich kvality je potrebné skúmať ich psychometrické vlastnosti, napríklad vnútornú konzistenciu, vzťahy s inými metodikami alebo konštruktmi no dôležité je venovať pozornosť tiež základným stavebným jednotkám týchto dotazníkov – ich položkám.

1 Odlišné fungovanie položky

Odlišné fungovanie položky, z anglického *differential item functioning* (skratka *DIF*) vyjadruje stav kedy je pravdepodobnosť kľúčovej (najmä pri osobnostných dotazníkoch) alebo správnej (najmä pri psychodiagnostických

testoch) odpovede významne odlišná u respondentov z dvoch oddelených skupín, ktorí však majú rovnakú, alebo približne rovnakú mieru meranej vlastnosti (Clauser, Mazor, 1998). V takomto prípade je odpoveď na položku ovplyvnená nie len meranou vlastnosťou, na ktorú je položka zameraná, ale aj inou vlastnosťou v miere ktorej sa respondenti z dvoch skúmaných skupín líšia. Táto vlastnosť môže byť čiastkovou vlastnosťou tej ktorú meriame. V takomto prípade síce nastáva odlišné fungovanie položky, no nemožno ho jednoducho považovať za skreslenie, keďže jeho príčina je v pláne merania. Inak povedané, hoci majú respondenti rovnakú mieru meranej vlastnosti, špecificky sa líšia v určitom prejave, ktorý táto položka meria. DIF však môže byť spôsobené vlastnosťou, ktorá nie je zahrnutá v pláne merania, kedy môžeme hovoriť o narušení validity, vytváraní skreslenia a nespravodlivosti psychodiagnostickej metódy (Roussous, Stout, 1996; Shealy, Stout, 1993).

Vo všeobecnosti rozoznávame dva typy DIF. Odlišné fungovanie položiek môže byť uniformné, to znamená, že jedna skupina má vyššiu pravdepodobnosť kľúčovej odpovede v celom rozsahu meranej vlastnosti, alebo neuniformné, kedy sa tento smer mení – jedna skupina má vyššiu pravdepodobnosť kľúčovej odpovede pri nižšej miere meranej vlastnosti, druhá skupina pri vyššej. Pre skúmanie oboch typov DIF vzniklo viacero štatistických metód pri dichotómnych položkách ako aj polytómnych položkách. V tejto štúdiu využívame celkovo tri metódy, dve zamerané na uniformné DIF a jednu, ktorá dokáže odhaliť uniformné ako aj neuniformné DIF. Prvou z nich je Mantelova chí-kvadrát metóda (Mantel, 1963), ktorá je rozšírením Mantel-Haenszelovej chí-kvadrát metódy určenej pre skúmanie uniformného DIF v dichotómnych položkách. Analyzovať položky touto metódou je možné v štatistických softvéroch, no existujú aj špeciálne programy, napr. DIFAS verzia 5.0 (Penfield, 2013, 2005) v ktorých je táto metóda adaptovaná a doplnená o ďalšie automatické výpočty. Táto metóda je určená len pre skúmanie uniformného DIF. Druhou, rovnako zameranou metódou je Poly-SIBTEST (Chang, Mazeo, Roussos, 1996). Pri oboch metódach ide o porovnanie pravdepodobnosti určitej odpovede pri určitej miere vlastnosti medzi skúmanými skupinami. Obe metódy sú neparametrické, no Poly-SIBTEST sa mierne líši v tom, že nevyužíva priamo hrubé skóre, ale odhad pravého skóre. Pre skúmanie oboch typov DIF možno využiť napríklad parametrickú metódu ordinálnej logistickej regresie. Pri tejto metóde porovnáваме tri modely. Prvý obsahuje odpoveď na položku ako závislú premennú a hrubé skóre ako nezávislú. V druhom modeli pridávame k nezávislým premenným premennú vyjadrujúcu členstvo v skupine a v treťom modeli pridávame interakciu týchto premenných. Porovnaním vhodnosti modelov 1 a 2 môžeme identifikovať uniformné DIF, alebo neuniformné porovnaním

modelu 2 a 3. Závažnosť DIF je možné posúdiť na základe rozdielu v hodnotách pseudo- R^2 (Zumbo, 1999).

2 Metóda

2.1 Súbor

Výskumný súbor tvorilo 1208 respondentov, 552 mužov (45,7%) a 656 žien (54,3%) vo veku od 17 do 67 rokov. Priemerný vek respondentov bol 31,57 (SD= 12,89), priemerný vek u mužov bol 32,59 (SD= 12,22) a u žien bol 30,71 (SD= 12,29). Dáta boli vyzbierané prostredníctvom online dotazníka. Zber dát bol plne anonymný a dobrovoľný. Všetci respondenti vyjadrili súhlas s účasťou vo výskume ako aj s použitím dát pre výskumné účely.

2.2 Metodika

Inventár Veľkej Päťky [Big Five Inventory – BFI, John, Srivastava, 1999] je osobnostný inventár zameraný na meranie 5 faktorov osobnosti. Obsahuje 44 položiek, 8 pre Extraverziu, 9 pre Prívetivosť, 9 pre Svedomitosť, 8 pre Neuroticizmus, 10 pre Otvorenosť, a. Všetky položky majú spoločný úvod: „Vnímam sa ako niekto, kto...“, po ktorom nasledujú rôzne charakteristiky (napr. robí svoju prácu svedomito, je vynaliezavý). Následne respondenti vyjadrujú svoj súhlas alebo nesúhlas s týmto výrokom prostredníctvom 5 bodovej Likertovej škály: 1 – rozhodne nesúhlasím; 2 – trochu nesúhlasím; 3 – ani súhlasím ani nesúhlasím; 4 – trochu súhlasím; 5 – rozhodne súhlasím. Nami zistená vnútorná konzistencia jednotlivých faktorov je dostačujúca, 0,801 pre Extraverziu, 0,701 pre Prívetivosť, 0,771 pre Svedomitosť, 0,830 pre Neuroticizmus a 0,772 pre Otvorenosť.

2.3 Analýza

Pre prípravu a spracovanie dát sme použili MS Excel a štatistický program SPSS. Pre skúmanie DIF sme použili program DIFAS 5.0 (Penfield, 2013), SPSS a program DIFPACK/SIBTEST (Chang, Mazeo & Roussos, 1996; Shealy & Stout, 1993). Grafy položiek boli vypočítané v programe SPSS a vytvorené v MS Excel.

Program DIFAS 5.0 (*Differential item functioning analysis system 5.0*; Penfield, 2013) je program pre analýzu DIF použitím Mantelovej metódy (Mantel, 1963) pre polytómne položky. Výsledkom analýzy polytómnych položiek je niekoľko ukazateľov DIF (Penfield, 2013, 2005): Mantelov chí-kvadrát, Liu-Agrestiho kumulatívny pomer pravdepodobnosti (L-A LOR),

štandardná chyba L-A LOR (LOR SE), štandardizované L-A LOR (LOR Z) a ďalšie. Aby bola položka chápaná ako DIF musela vykazovať signifikantný Mantelov chí-kvadrát pri $p < 0,01$ a zároveň mať LOR Z väčšie ako 1,96 (Penfield, 2013). Ak splnila tieto podmienky, bola ďalej kategorizovaná podľa noriem od Penfielda (2007): nevýznamné alebo slabé DIF (kategória A) ak $|L-A LOR| < 0,43$; stredne silné DIF (kategória B) ak $|L-A LOR|$ je v rozmedzí 0,43 až 0,64, alebo ak je väčšie ako 0,64 no nie signifikantne väčšie ako 0,43; silné DIF (kategória C) ak je $|L-A LOR| > 0,64$ a zároveň signifikantne väčšie ako 0,43. Po uskutočnení prvej analýzy nasledoval druhý krok v ktorom sme overovali DIF u položiek ktoré boli na základe spomenutých pravidiel zaradené do kategórie B alebo C, no iba v prípade, že takéto položky boli v škále aspoň dve. V takom prípade sme uskutočnili analýzu s použitím upraveného skóre, ktoré obsahovalo overovanú položku, no neobsahovalo ostatné DIF položky zaradené do kategórie B alebo C.

Metódu ordinálnej logistickej regresie sme uskutočnili prostredníctvom makra DIFLRT, ktoré je voľne dostupné na stránke podpory IBM, v štatistickom programe SPSS. V prvom kroku sme skúmali samostatne uniformné a neuniformné DIF s použitím hrubého skóre ako miery vlastnosti. Vypočítaní bol chí-kvadrát s jedným stupňom voľnosti pre porovnanie modelu uniformného a neuniformného DIF. Pokiaľ boli tieto modely signifikantne odlišné pre uniformné alebo neuniformné DIF pri $p < 0,01$ a zmena Nagelkerkeho R^2 bola aspoň 0,01 (podľa Crane et al., 2007), DIF tejto položky bolo považované za stredne silné alebo silné. V prípade identifikovania aspoň dvoch položiek, ktoré spĺňali túto podmienku v jednej škále, overovali sme tieto položky podobným spôsobom ako v prípade vyššie popísanej Mantelovej chí-kvadrát metódy.

Skúmanie DIF prostredníctvom Poly-SIBTESTu prebiehalo rovnako v dvoch krokoch. V prvom kroku sme skúmali DIF položiek s použitím hrubého skóre, ktoré však neobsahovalo analyzovanú položku. V prípade, že položka vykazovala signifikantné DIF pri $p < 0,01$ a hodnota Beta bola väčšia ako 0,2, bola nami položka označená ako odlišne fungujúca. Podmienku hodnoty Beta sme zvolili na základe návrhu kategorizácie DIF polytómnych položiek od Golia (2012), pričom položka vykazuje zanedbateľné alebo slabé DIF (A) ak je $|\beta/J| \leq 0,05$; stredne silné DIF (B) ak $0,05 < |\beta/J| \leq 0,1$ a silné DIF ak $|\beta/J| > 0,1$. J v tomto prípade vyjadruje počet odpovedových kategórií, v našom prípade 4 (škála 1 až 5, t.j. $5-1=4$). V prípade, že v jednej škále túto podmienku spĺňalo viacero položiek, bolo ich DIF overené v druhom kroku s použitím hrubého skóre, ktoré neobsahovalo ani jednu z týchto položiek.

3 Výsledky

Prostredníctvom Mantelovej chí-kvadrát metódy sme v prvom kroku identifikovali celkovo 12 položiek, ktoré vykazovali DIF pri $p < 0,01$; 12 položiek bolo identifikovaných prostredníctvom ordinálnej logistickej regresie a 10 prostredníctvom Poly-SIBTESTu. Celkovo ide o 12 položiek, ktoré boli aspoň jednou metódou označené ako DIF pri $p < 0,01$. Z tohto počtu 7 položiek vykazuje len zanedbateľné alebo slabé DIF (kategória A), 5 položiek bolo aspoň jednou metódou označené ako stredne silné alebo silné DIF (kategória B a C). V tabuľke 1 uvádzame výsledky Mantelovej chí-kvadrát metódy v programe DIFAS 5.0. Závažnejšie DIF bolo zistené v prípade položky 4 zo škály Neuroticizmus a pri položkách 25 a 30 zo škály Otvorenosť, pri ktorých bolo toto DIF potvrdené aj v druhom kroku analýzy, t.j. po purifikácii.

Tabuľka 1 Výsledky analýzy uniformného DIF v BFI (DIFAS)

| Šk. | Pol. | Mantel | LA LOR | LOR SE | LOR Z | COX B | COX SE | COX Z | Kateg. | Zvýhod. |
|-----|-----------------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| E | 1 | 9,217 | -0,364 | 0,120 | -3,033 | -0,247 | 0,081 | -3,038 | A | Ž |
| E | 31 | 8,316 | 0,325 | 0,112 | 2,902 | 0,188 | 0,065 | 2,879 | A | M |
| P | 7 | 8,198 | -0,349 | 0,121 | -2,884 | -0,226 | 0,079 | -2,861 | A | Ž |
| S | 8 | 13,869 | -0,425 | 0,113 | -3,761 | -0,246 | 0,066 | -3,722 | A | Ž |
| S | 18 | 11,232 | 0,378 | 0,113 | 3,345 | 0,215 | 0,064 | 3,354 | A | M |
| S | 43 | 10,597 | 0,359 | 0,11 | 3,264 | 0,21 | 0,0646 | 3,251 | A | M |
| N | 4 | 56,052 | 0,891 | 0,119 | 7,487 | 0,495 | 0,0661 | 7,489 | C | M |
| N | 9 | 7,708 | -0,318 | 0,115 | -2,765 | -0,196 | 0,0705 | -2,78 | A | Ž |
| O | 5 | 12,366 | 0,421 | 0,119 | 3,538 | 0,275 | 0,0783 | 3,512 | A | M |
| O | 25 ¹ | 17,223 | 0,498 | 0,12 | 4,15 | 0,358 | 0,0861 | 4,158 | B | M |
| O | 30 ¹ | 32,251 | -0,696 | 0,123 | -5,659 | -0,406 | 0,0715 | -5,678 | C | Ž |
| O | 41 | 10,179 | -0,362 | 0,112 | -3,232 | -0,182 | 0,057 | -3,193 | A | Ž |

Poznámka. $p < 0,01$; A – zanedbateľné alebo slabé DIF, B – stredne silné DIF, C – silné DIF; M – muži, Ž – ženy; 1 – výsledok je z druhého kroku analýzy

V tabuľke 2 uvádzame výsledky ordinálnou logistickou regresiou. Položky identifikované ordinálnou logistickou regresiou sú v zhode s tými, ktoré sme identifikovali v programe DIFAS. Závažnejšie DIF sme zistili pri položke 4 zo škály Neuroticizmus a položkách 25 a 30 zo škály Otvorenosť. Celkovo nebola identifikovaná žiadna položka, ktorá by vykazovala neuniformné odlišné fungovanie pri $p < 0,01$.

Tabuľka 2 Výsledok analýzy uniformného a neuniformného DIF (ordinálna logistická regresia)

| Škála | Položka | χ^2_{13} | χ^2_{12} | χ^2_{23} | $\Delta NR^2 U$ | $\frac{\Delta NR^2}{NU}$ | Zvýhodňuje |
|-------|-----------------|---------------|---------------|---------------|-----------------|--------------------------|------------|
| E | 1 | - | * | - | 0,003 | 0,000 | Ž |
| E | 31 | * | * | - | 0,005 | 0,001 | M |
| P | 7 | - | * | - | 0,006 | 0,000 | Ž |
| S | 8 | * | * | - | 0,009 | 0,000 | Ž |
| S | 18 | * | * | - | 0,006 | 0,001 | M |
| S | 43 | * | * | - | 0,007 | 0,000 | M |
| N | 4 | * | * | - | 0,036 | 0,000 | M |
| N | 9 | - | * | - | 0,003 | 0,000 | Ž |
| O | 5 | * | * | - | 0,006 | 0,000 | M |
| O | 25 ¹ | * | * | - | 0,010 | 0,000 | M |
| O | 30 ¹ | * | * | - | 0,016 | 0,000 | Ž |
| O | 41 | * | * | - | 0,007 | 0,003 | Ž |

Note. * - $p < 0,01$; χ^2_{13} - celkové DIF; χ^2_{12} - uniformné DIF; χ^2_{23} - neuniformné DIF; $\Delta NR^2 U$ - zmena Nagelkerkeho R^2 pre uniformné DIF; $\Delta NR^2 NU$ - zmena Nagelkerkeho R^2 pre neuniformné DIF; Ž - ženy, M - muži; ¹ - výsledok je z druhého kroku analýzy

Analýzou prostredníctvom Poly-SIBTESTu sme identifikovali 10 položiek, ktoré vykazovali DIF pri $p < 0,01$. Závažnejšie DIF bolo zistené pri položke 8 a 18 zo škály Svedomitosť, položke 4 zo škály Neuroticizmus a položkách 30 a 41 zo škály Otvorenosť. Položky zo škály Svedomitosť a Otvorenosť boli overené v druhom kroku, pričom sa závažnejšie DIF potvrdilo pri položkách 8, 30 a 41. Pri položkách 18 a 25 bolo DIF signifikantné pri $p < 0,01$,

no z hľadiska kategórie bolo len zanedbateľné alebo slabé. Výsledky analýzy prostredníctvom Poly-SIBTESTu uvádzame v tabuľke 3.

Tabuľka 3 Výsledky analýzy uniformného DIF v BFI (Poly-SIBTEST)

| Škála | Položka | Beta | Beta S.E. | Kategória | Zvýhodňuje |
|-------|-----------------|--------|-----------|-----------|------------|
| E | 1 | -0,136 | 0,050 | A | Ž |
| E | 31 | 0,165 | 0,062 | A | M |
| S | 8 ¹ | -0,202 | 0,058 | B | Ž |
| S | 18 ¹ | 0,183 | 0,059 | A | M |
| S | 43 | 0,195 | 0,060 | A | M |
| N | 4 | 0,457 | 0,057 | C | M |
| O | 5 | 0,178 | 0,047 | A | M |
| O | 25 ¹ | 0,176 | 0,043 | A | M |
| O | 30 ¹ | -0,357 | 0,056 | B | Ž |
| O | 41 ¹ | -0,278 | 0,069 | B | Ž |

Poznámka. $p < 0,01$; Beta S.E. – štandardná chyba hodnoty Beta; A – zanedbateľné alebo slabé DIF, B – stredne silné DIF, C – silné DIF; M – muži, Ž – ženy; ¹ – výsledok je z druhého kroku analýzy

Pre zhrnutie uvádzame v tabuľke 4 prehľad jednotlivých identifikovaných položiek, spolu s výsledkom analýzy v každej z použitých metód a informáciou o smere DIF ako o téme položky.

Pre zistenie dopadu identifikovaných DIF položiek sme ďalej prostredníctvom Poly-SIBTEST metódy vypočítali odlišné fungovanie skupiny položiek (DBF), vypočítali sme rozdiely medzi mužmi a ženami v pôvodných škálach a upravených škálach bez všetkých identifikovaných DIF položiek a bez položiek, ktoré v zhode aspoň dvoch použitých metód spĺňajú kritéria pre DIF. V pôvodnej škále Extraverzie sme zistili signifikantný malý rozdiel ($t_{1206} = -2,103$; $p = 0,036$) a po odstránení položiek 1 a 31 nenastala skoro žiadna zmena v tomto rozdiely ($t_{1206} = -2,147$; $p = 0,032$). Tieto položky taktiež nevykazovali spoločné odlišné fungovanie. Podobná situácia je aj v prípade škály Prívetivosť, kde bola ako DIF označená jedna položka, ktorá však nespĺňa kritéria pre závažné odlišné fungovanie. Jej vymazanie z hrubého skóre nemalo na malý rozdiel medzi mužmi a ženami ($t_{1206} = -4,768$; $p < 0,01$) závažnejší dopad ($t_{1206} = -4,237$; $p < 0,01$). V škále Svedomitosti sme zistili malý signifikantný rozdiel ($t_{1206} = -3,392$; $p < 0,01$) v smere žien dosahujúcich vyššie

Tabuľka 4 DIF položky identifikované v Big Five Inventory

| Pol. | Škála | DIFAS | OLR U | NU | SIBTEST | Zvýhodňuje | Téma |
|------|-------|----------------|----------------|----|----------------|------------|-----------------------------------|
| 1 | E | A | sig. | - | A | Ž | zhovorčivosť |
| 4 | N | C | * | - | C | M | skľúčenosť |
| 5 | O | A | sig. | - | A | M | originalita |
| 7 | P | A | sig. | - | - | Ž | nezištnosť |
| 8 | S | A | sig. | - | B ¹ | Ž | nedbanlivosť ² |
| 9 | N | A | sig. | - | - | Ž | uvoľnenosť ² |
| 18 | S | A | sig. | - | A ¹ | M | chaotickosť ² |
| 25 | O | B ¹ | * ¹ | - | B ¹ | M | vynaliezavosť |
| 30 | O | C ¹ | * ¹ | - | B ¹ | Ž | záujem o umenie |
| 31 | E | A | sig. | - | A | M | hanblivosť ² |
| 41 | O | A | sig. | - | B ¹ | Ž | nezáujem o umenie ² |
| 43 | S | A | sig. | - | A | M | rozptyliteľnosť ² |

Poznámka. E – Extraverzia, N – Neuroticizmus, O – Otvorenosť, P – Prívetivosť, S – Svedomitosť; A – zanedbateľné alebo slabé DIF, B – stredne silné DIF, C – silné DIF; sig. – je signifikantné, no nespĺňa podmienku pre DIF, * – spĺňa podmienku pre DIF; 1 – výsledok po purifikácii; 2 – položka je reverzne skórovaná

skóre a táto situácia sa nezmenila ($t_{1206} = -3,764$; $p < 0,01$) ani po odstránení položiek 8, 18 a 43, ktoré nevykazovali spoločné signifikantné odlišné fungovanie. Odlišnú situáciu sme zistili v škále Neuroticizmus, kde ženy dosahujú vyššie skóre ($t_{1206} = -6,517$, $p < 0,01$). Po odstránení položky 4 a 9 sa tento rozdiel veľmi mierne prehĺbil ($t_{1206} = -7,318$; $p < 0,01$). Položka 4 funguje odlišne v smere mužov označujúcich vyššiu hodnotu odpovedi pri rovnakej miere neuroticizmu, zatiaľ čo položka 9 naopak. Odlišné fungovanie týchto položiek sa navzájom neruší, no prevláda v smere zvýhodňovania mužov. Ak by sme odstránili z hrubého skóre iba položku 4, rozdiel medzi rodmi ($t_{1206} = -7,709$; $p < 0,01$) by bol o niečo výraznejší (rozdiel v Cohenovom d o 0,06). V škále Otvorenosť sme zistili malý, signifikantný rozdiel ($t_{1206} = -2,600$; $p < 0,01$) s vyšším skóre u žien. Po odstránení identifikovaných položiek 5, 25, 30 a 41, ktoré vykazovali spoločné odlišné fungovanie v smere zvýhodňovania žien, sa zistený rozdiel v tejto škále stal nesignifikantným ($t_{1206} = -1,695$; $p > 0,05$). Položky 25 a 30, ktoré spĺňali kritéria pre uznanie za stredne závažne odlišne fungujúce spolu nevykazovali odlišné fungovanie a ich odstránenie

nepatrne znížilo veľkosť rozdielu v škále Otvorenosť medzi mužmi a ženami, ktorý zostal slabosignifikantný ($t_{1206} = -2,138$; $p = 0,033$). Dôležité je ale podotknúť, že zistený rozdiel v tejto škále je veľmi malý z hľadiska Cohenovho d. Výsledky týchto analýz zobrazujeme v tabuľke 5.

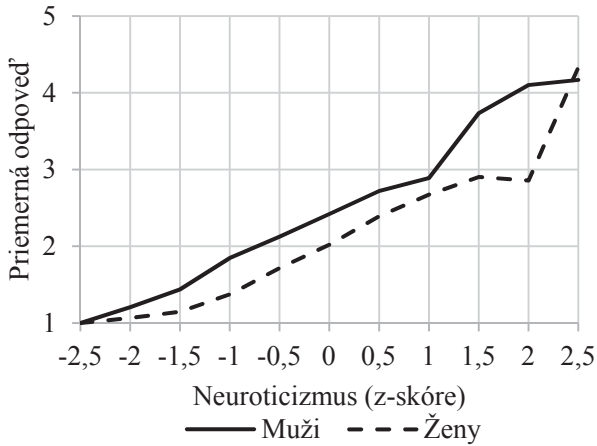
Tabuľka 5 Odlišné fungovanie skupiny položiek a dopad DIF na rozdiely medzi rodmi

| Šk. | DBF 1 | | DBF 2 | | Cohenovo d, 95% KI | | |
|-----|---------|-------|--------|------|---------------------------|---------------------------|---------------------------|
| | Beta | S.E. | Beta | S.E. | Originál | 1 | 2 |
| E | 0,059 | 0,088 | - | - | -0,12* [-0,23; -0,01] | -0,12* [-0,24; -0,01] | - |
| P | - | - | - | - | -0,28** [-0,39; -0,16] | -0,24** [-0,36; -0,13] | - |
| S | 0,199 | 0,124 | - | - | -0,20** [-0,31; -0,08] | -0,22** [-0,33; -0,10] | - |
| N | 0,332** | 0,085 | - | - | -0,38** [-0,49; -0,26] | -0,42** [-0,54; -0,31] | -0,44** [-0,56; -0,33] |
| O | -0,296* | 0,130 | -0,094 | 0,70 | -0,15* [-0,26; -0,04] | -0,10 [-0,20; 0,02] | -0,12* [-0,24; -0,01] |

Poznámka. Negatívne hodnoty pri Beta znamenajú odlišné fungovanie v smere zvýhodňovania žien. Negatívne hodnoty pri Cohenovom d znamenajú vyššie skóre u žien Originál – pôvodný set položiek, 1 – set položiek bez položiek podozrivých z DIF, 2 – set položiek bez položiek spĺňajúcich prísnejšie DIF kritéria

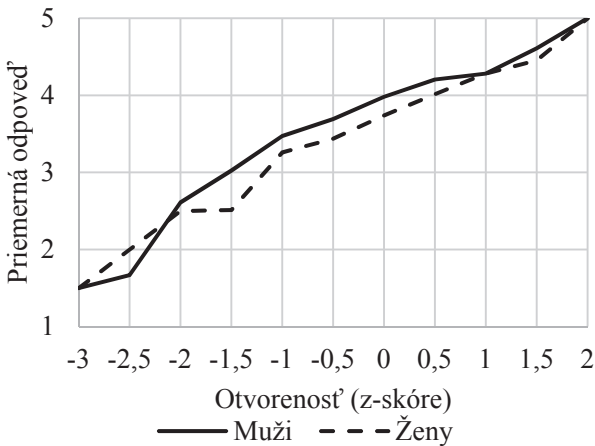
* - $p < 0,05$, ** - $p < 0,01$

V grafe 1 ilustrujeme rozdiel medzi rodmi v hodnotách priemerných odpovedí na položku 4 pri rôznej miere neuroticizmu. Táto položka má znenie „Vnímam sa ako niekto, kto je sklúčený, skleslý.“ Pri tejto položke muži udávajú vyššie skórované odpovede pri rovnakej (podobnej) úrovni neuroticizmu v porovnaní so ženami.



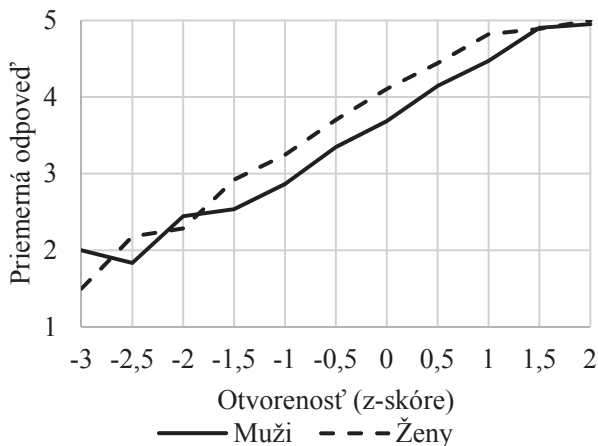
Graf 1: Znárodnenie položky 4

V škále Otvorenosť vykazujú dve položky DIF podľa určených kritérií. Prvá, položka 25 má znenie: „Vnímam sa ako niekto, kto je vynaliezavý.“ V tejto položke označujú muži vyššie skórované odpovede v porovnaní so ženami. Túto položku zobrazujeme v grafe 2.



Graf 2: Znárodnenie položky 25

Druhou identifikovanou položkou je položka 30, ktorá ma znenie: „Vnímam sa ako niekto, kto si cení umelecké a estetické zážitky.“ Pri tejto položke odpovedajú ženy vyššími hodnotami v porovnaní s mužmi. Túto položku znázorňujeme v grafe 3.



Graf 3: Znárodnenie položky 30

Výsledky, ktoré sme uviedli diskutujeme v ďalšej časti článku.

4 Diskusia

Štúdia sa zaoberá skúmaním odlišného fungovania položiek v Inventári Veľkej Päťky (BFI; John, Srivastava, 1999). Prostredníctvom metód Mantelovho chí-kvadrátu, ordinálnej logistickej regresie a Poly-SIBTESTu sme identifikovali celkovo 12 položiek, ktoré vykazovali rôzne závažné uniformné DIF. Päť z týchto položiek bolo aspoň jednou metódou označených za stredne silné a silné DIF a 3 položky boli takto označené v zhode všetkých troch metód. V žiadnej škále tohto dotazníka sme neidentifikovali položku, ktorá by podľa zvolených kritérií vykazovala neuniformné DIF. Posúdenie závažnosti DIF sa odvíja nie len od sily odlišného fungovania, no tiež možných príčin vzniku DIF. Samotná prítomnosť DIF je nevyhnutnou no nie dostačujúcou podmienkou pre označenie položky ako skresľujúcej (Roussos, Stout, 1996). Okrem reálneho rozdielu v miere položkou zisťovaných prejavov meranej vlastnosti môžu byť týmito príčinami napríklad sociálna žiadúcnosť odpovede, ktorá je

rozdielna medzi skúmanými skupinami, alebo napríklad iné chápanie obsahu položky (Mitchelson et al., 2009). Položky, ktoré boli označené aspoň ako stredne silne odlišne fungujúce v zhode troch metód ďalej diskutujeme.

Prvou z nich je položka 4 zo škály Neuroticizmus. Pri tejto položke muži priemerne označujú vyššiu odpoveď v porovnaní so ženami pri rovnakej miere neuroticizmu. Táto položka je zameraná na zisťovanie skleslosti a sklúčenosti, čo je spojené so smútkom a depresívnosťou (revidovaná forma položky je v druhej verzii dotazníka priamo zaradená do subškály Depresia, pozri Soto, John, 2017). Aj keď existujú určité rozdiely v prežívaní a vyjadrovaní emócií medzi mužmi a ženami (napr. Chaplin, 2006), nepodarilo sa nám nájsť výskumy, ktoré by DIF tejto položky spoľahlivo vysvetlili. Jedným z možných vysvetlení môže byť povaha škály, keďže neuroticizmus nezahŕňa len smútok, ale aj emocionálnu nestabilitu a úzkosť. Pre objasnenie sú však dôležité ďalšie výskumy, ktoré by sa zaoberali touto problematikou. Ďalšie dve DIF položky sú zo škály Otvorenosť. Položka pri ktorej označujú vyššie hodnoty muži sa týka vlastnej vynaliezavosti. Druhá DIF položka, pri ktorej vyššie skórujú ženy sa týka umeleckých a estetických hodnôt. DIF položiek zo škály Otvorenosť možno vysvetliť mierne problematickým, širokým záberom tejto škály (napr. Smith, 2002). Je možné, že DIF týchto položiek je spôsobené rodovými rozdielmi v špecifických a odlišných oblastiach, ktoré škála Otvorenosť zachytáva. Podľa zistení Costa, Terracciano a McCrae (2001), ženy dosahujú vyššiu úroveň v otvorenosti voči pocitom zatiaľ čo muži majú vyššiu mieru otvorenosti voči myšlienkam. Je možné, že tieto položky by nevykazovali DIF vzhľadom na užšie formulovanú subškálu, no nemožno vylúčiť vplyv sociálnej žiadúcnosti alebo rodových stereotypov.

Napriek prítomnosti rodovo DIF položiek v dotazníku BFI sa nezdá, že by existovalo závažné podozrenie zo skreslenia medzi rodmi. K tomu dosvedčuje aj minimálna zmena rodových rozdielov vo väčšine škál po odstránení identifikovaných DIF položiek. Ďalšie výskumy by mali preveriť naše zistenia na iných súboroch, ktoré môžu byť špecifickejšie napríklad z hľadiska veku. Medzi hlavné limitácie nášho výskumu patrí najmä absencia skúmania príčin DIF a použitie len metód vychádzajúcich z klasickej teórie testov.

Záver

Slovenská verzia Inventára Piatich Faktorov obsahuje niekoľko položiek, ktoré vykazujú rôznu mieru uniformného odlišného fungovania medzi mužmi a ženami. Počet takýchto položiek, ktoré vykazujú stredne silné alebo silné,

resp. závažné odlišné fungovanie je nízky. Celkovo nebola zistená žiadna položka vykazujúca neuniformné odlišné fungovanie. Napriek prítomnosti DIF položiek, neexistuje dôkaz prítomnosti vážneho skreslenia z hľadiska rodu.

L I T E R A T Ú R A

- CHANG, H., MAZZEO, J., ROUSSOS, L. (1996). Detecting DIF for Polytomously Scored Items: An Adaptation of the SIBTEST Procedure. *Journal of Educational Measurement*, 33, 3, 333-353.
- CHAPLIN, T. M. (2006). Anger, Happiness, and Sadness: Associations with Depressive Symptoms in Late Adolescence. *Journal of Youth and Adolescence*, 35, 6, 977-986.
- 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.
- COSTA, P. J., TERRACCIANO, A., MCCRAE, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81, 2, 322-331.
- CRANE, P. K., GIBBONS, L. E., OCEPEK-WELIKSON, K., COOK, K., CELLA, D., NARASIMHALU, K., HAYS, R. D., TERESI, J. A. (2007). A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. *Quality of Life Research*, 16, 1, 69-84.
- GOLIA, S. (2012). Differential Item Functioning classification for polytomously scored items. *Electronic Journal of Applied Statistical Analysis*, 5, 3, 367-373.
- JOHN, O. P., SRIVASTAVA, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (vol. 2, pp. 102-138). New York: Guilford Press. 738 p.
- MANTEL, N. (1963). Chi-square tests with one degree of freedom: Extension of the Mantel-Haenszel procedure. *Journal of the American Statistical Association*, 58, 690-700.
- MITCHELSON, J. K., WICHER, E. W., LEBRETON, J. M., & CRAIG, S. B. (2009). Gender and Ethnicity Differences on the Abridged Big Five Circumplex (AB5C) of Personality Traits. *Educational and Psychological Measurement*, 69, 4, 613-635.
- PENFIELD, R. D. (2005). DIFAS: Differential Item Functioning Analysis System. *Applied Psychological Measurement*, 29, 2, 150-151.
- PENFIELD, R. D. (2007). An Approach for Categorizing DIF in Polytomous Items. *Applied Measurement in Education*, 20, 3, 335-355.
- PENFIELD, R. D. (2007). An Approach for Categorizing DIF in Polytomous Items. *Applied Measurement in Education*, 20, 3, 335-355.
- PENFIELD, R. D. (2013). DIFAS: 5.0 User's Manual. Online: http://soe.uncg.edu/wp-content/uploads/2015/12/DIFASManual_V5.pdf.
- ROUSSOS, L., STOUT, W. (1996). A Multidimensionality-Based DIF Analysis Paradigm. *Applied Psychological Measurement*, 20, 4, 355-371.
- SHEALY, R., STOUT, W. (1993). A model based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika*, 58, 159-194.
- SMITH, L. L. (2002). On the Usefulness of Item Bias Analysis to Personality Psychology. *Personality and Social Psychology Bulletin*, 28, 6, 754-763.

- SOTO, C. J., JOHN, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113, 117-143.
- ZUMBO, B. D. (1999). *Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores*. Ottawa: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Štúdia bola realizovaná s podporou agentúry VEGA, č. projektu 1/0234/15.