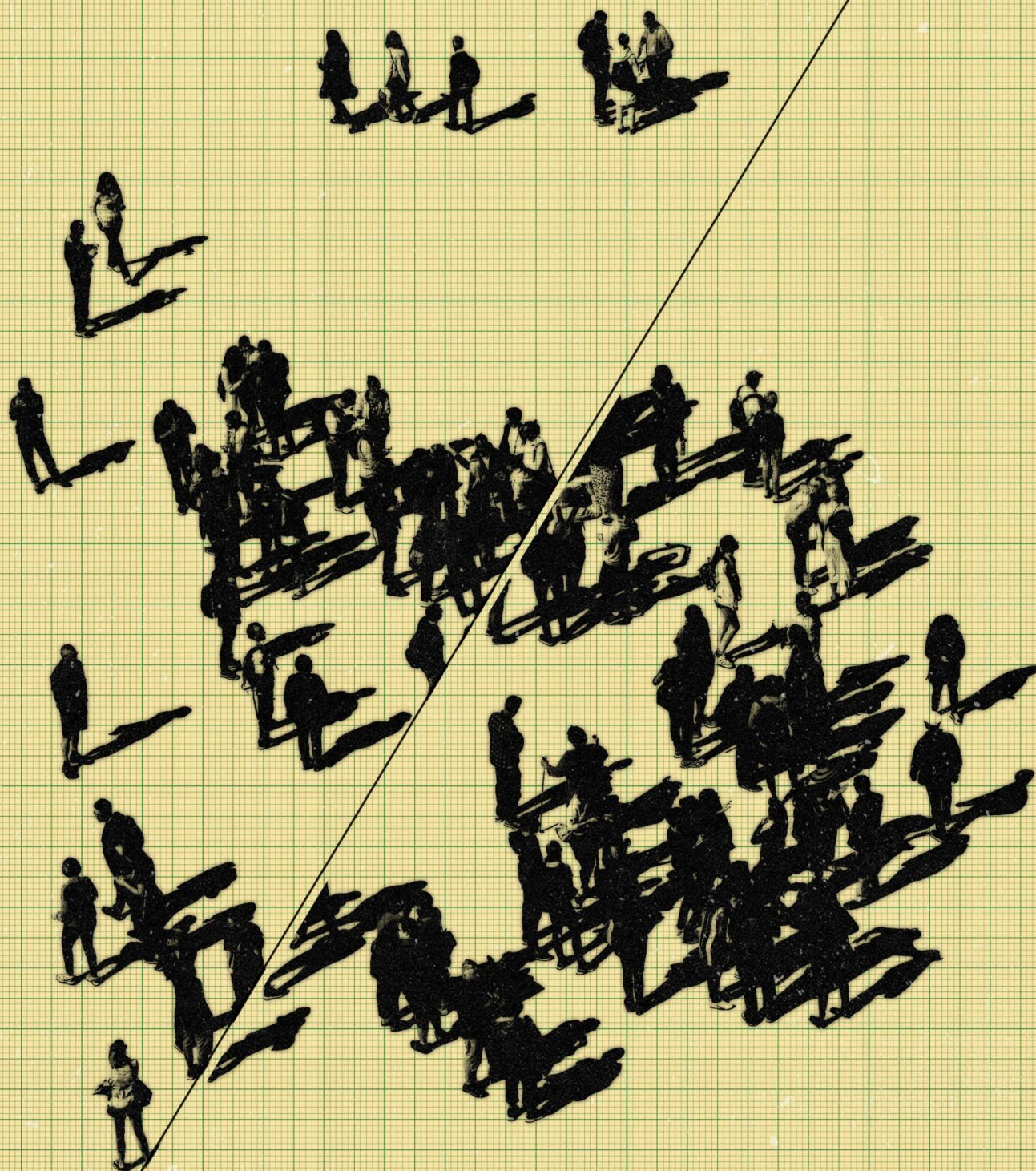


Michal Kohút

Meriame to isté?

Invariantnosť merania piatich faktorov osobnosti
v súvislosti s rodom a vekom



/ Filozofická fakulta TU v Trnave



Michal Kohút

Meriame to isté?

**Invariantnosť merania piatich faktorov osobnosti
v súvislosti s rodom a vekom**

Michal Kohút

**Meriame to isté? Invariantnosť merania piatich faktorov osobnosti v súvislosti s rodom
a vekom**

(Vedecká monografia)

Publikácia vznikla v rámci riešenia grantového projektu VEGA č. 1/0223/22.

Recenzenti:

Prof. PhDr. Tomáš Sollár, PhD.

PhDr. Eva Ballová Mikušková, PhD.

© Michal Kohút, 2022

© Filozofická fakulta TU v Trnave, 2022

Ilustrácia na obálke: © Ján Kurinec, 2022

ISBN 978-80-568-0507-7

Publikácia neprešla jazykovou korektúrou.

Obsah

Úvod	7
1. Validita psychodiagnostických metodík	8
2. Odlišné fungovanie položky.....	12
2.1. Uniformné a neuniformné DIF	14
2.2. Dôležitosť skúmania odlišného fungovania položiek	15
2.3. Metódy skúmania odlišného fungovania položiek.....	19
3. Invariantnosť merania	23
3.1. Postup testovania invariantnosti modelu.....	24
4. Päťfaktorový model osobnosti	27
4.1. Inventár Veľkej Päťky 2.....	28
4.2. Päť veľkých faktorov osobnosti v súvislosti s pohlavím – odlišné fungovanie položiek a invariantnosť merania.....	29
4.3. Päť veľkých faktorov osobnosti v súvislosti s vekom – odlišné fungovanie položiek a invariantnosť merania.....	33
5. Výskumné súbory a analýza dát.....	35
5.1. Analýza odlišného fungovania položiek BFI-2.....	36
5.2. Analýza invariantnosti merania BFI-2	37
6. Invariantnosť merania a odlišné fungovanie položiek BFI-2 medzi mužmi a ženami	39
6.1. Extraverzia	39
6.2. Prívetivosť	46
6.3. Svedomitosť	53
6.4. Negatívna emocionálnosť.....	60
6.5. Otvorenosť.....	69
7. Invariantnosť merania a odlišné fungovanie položiek BFI-2 podľa veku	75
7.1. Extraverzia	76
7.2. Prívetivosť	79

7.3. Svedomitost'	81
7.4. Negatívna emocionalita.....	84
7.5. Otvorenost'.....	86
8. Zhrnutie a diskusia výsledkov	90
8.1. Problémy pri skúmaní rovnosti merania	90
8.2. Konzistentnosť identifikovaných položiek	91
8.3. Meriame to isté vzhľadom na pohlavie?	92
8.4. Meriame to isté vzhľadom na vek?	95
8.5. Limitácie výskumu a odporúčania pre ďalšie výskumy a prax	96
Záver.....	98
Použitá literatúra	99

Úvod

Táto monografia je inšpirovaná našou dizertačnou prácou (Kohút, 2018) zameranou na odlišné fungovanie položiek medzi mužmi a ženami vo vybraných osobnostných dotazníkoch. V rámci tejto monografie sme sa zamerali konkrétne len na Inventár Veľkej Päťky 2, pričom sa zaoberáme jeho validitou vzhľadom na pohlavie a vek. Monografia prináša dôkazy o validite piatich veľkých faktorov osobnosti na slovenskej populácii. Postupne predstavíme koncept odlišného fungovania položiek a invariantnosti merania. Ďalej v krátkosti priblížime teóriu piatich osobnostných faktorov so zameraním najmä na doterajšie zistenia ohľadom ich súvislostí s pohlavím a vekom. Vo výskumnej časti predstavíme výsledky originálneho výskumu, v ktorom päť faktorov osobnosti meriame s využitím slovenskej verzie Inventára Veľkej Päťky 2 (Halama et al., 2020). Výskum je zameraný na identifikovanie odlišne fungujúcich položiek a zhodnotenie rovnosti merania vzhľadom na pohlavie a vek respondentov. Na vek nahliadame skrz dve kategórie, konkrétne mladá, respektíve vynárajúca sa dospelosť a starší. Jedným z cieľov je zhodnotenie, či je možné slovenskú verziu Inventára Veľkej Päťky 2 spoľahlivo používať pri výskume mladej dospelosti. Touto témou sa zaujímame v rámci riešenia grantového projektu VEGA 1/0223/22 „Prediktory optimálneho psychologického fungovania v čase tranzícií v období vynárajúcej sa dospelosti“.

Veríme, že táto monografia bude prínosom pre čitateľov, ktorí sa zaujímajú o teóriu piatich faktorov osobnosti, spomínaný inventár alebo majú záujem dozvedieť sa viac o invariantnosti merania a odlišnom fungovaní položiek.

1. Validita psychodiagnostických metodík

Validitu psychodiagnostických metodík možno najjednoduchšie zadefinovať ako schopnosť metódy merať to, na čo je určená. Štandardy pre pedagogické a psychologické testovanie definujú validitu ako mieru, v ktorej empirické dôkazy a teórie podporujú interpretáciu testového skóre pri odporúčanom spôsobe použitia testu (AERA, APA, NCME, 2001). Tieto dve používané definície poukazujú na rôzne možnosti chápania a používania validity v psychometrii. Túto problematiku riešili napr. (St-Onge et al., 2017), ktorí poukázali na tri možné spôsoby nahliadania na validitu diagnostických metodík, ktoré sa vyskytujú vo vedeckej literatúre (ich štúdia je zameraná prevažne na oblasť zdravotníctva, no dá sa vhodne použiť aj pre účely psychológie alebo edukácie). Prvý z nich sa týka metodiky samotnej. Podľa tohto chápania je validita ako vlastnosť patriaca metodike – metodika je validná, má dokázanú validitu. Tento spôsob nazerania na túto oblasť prirovnávajú k „známke kvality“ či „zlatej pečati osvedčenia“ (po anglicky „*gold seal of approval*“), to jest, cieľom snaženia je pomenovanie metodiky ako validnej. Takéto ponímanie validity je statické – limitáciou takéhoto prístupu je implicitný predpoklad, že dokázanie kvality metodiky v určitom kontexte znamená jej všeobecnú kvalitu. Od tejto limitácie sa odráža druhý spôsob chápania validity – validita ako proces dokazovania kvality. V tomto prípade sa *cesta stáva cieľom*, t.j. nejde o skonštatovanie platnosti metodiky, no o proces dokazovania možnosti interpretácie testového skóre prostredníctvom rôznych dôkazov, na rôznych skupinách, v rôznom čase či spôsobe použitia. Vďaka tomu sa môžeme vyjadrovať o väčšej či menšej miere platnosti a kvality určitej metodiky pri nejakom spôsobe použitia (napr. Downing, 2003) Možným problémom je využívanie dôkazov – príliš skeptický pohľad by nám nedovoľoval zovšeobecňovanie na iné výskumné súbory, t.j. zistenia by boli platné len pre skúmaný súbor respondentov, preto je v praxi potrebná určitá miera zovšeobecniteľnosti. Tretí spôsob nazerania na validitu sa týka spoločenských spojitostí psychodiagnostiky a jej dopadu na jednotlivca, ako aj širšiu spoločnosť. Do popredia sa kladie otázka užitočnosti a zmyslu psychodiagnostiky pre spoločnosť. Tento aspekt je nepochybne dôležitý, no dotýka sa nie len psychodiagnostickej metodiky, ale i metodológie, výberu metodík, kritérií, vytvorených záverov a pod. (St-Onge et al., 2017).

Ako sme na začiatku uviedli, validitu je potrebné dokazovať. Dôkazy o validite možno rozdeliť na základe zdrojov, z ktorých pochádzajú. Halama (2011) rozoznáva najčastejšie dôkazy založené na *obsahu testu*, *vzťahu výsledku testu s kritériom*, *vzťahu výsledku testu s inými konštruktmi*. Dôkazy na základe obsahu metodiky sa týkajú najmä posúdenia, či je

obsah metodiky adekvátny. Pochádzať môžu napríklad od posudzovateľov, ktorí sú expertmi v oblasti, na ktorú je posudzovaná metodika zameraná. Experti môžu hodnotiť, či položky metodiky dostatočne zastupujú meraný konštrukt, či nie je niektorá oblasť v prevahe na úkor iných, alebo či nie sú iné, lepšie položky, ktoré by viac vystihovali meraný konštrukt. Vhodný príklad uvádzajú (John & Soto, 2007) – test pre študentov, ktorého otázky by mali reprezentovať témy a učivo (povinnú literatúru, skriptá), ktoré sa na predmete počas semestra prebrali. Pokiaľ tomu tak nie je, nemôžeme spoľahlivo tvrdiť, že test validne reprezentuje to, čo sa mali študenti naučiť. Okrem obsahu položiek môžu hodnotiť aj formuláciu, zrozumiteľnosť, možnosť odpovede a pod. Možným spôsobom je tiež kvalitatívny prístup, kedy požiadame respondentov aby nám priblížili to, akým spôsobom položku chápu, ako na ňu odpovedajú, vďaka čomu môžeme posúdiť, či položka vytvára u respondentov reakciu, ktorá je v súlade s plánom merania (napr. Willis, 2005). Príkladom môžu byť štúdie od Lund et al. (2014) alebo Berthelsen et al. (2016), ktorí pre získanie dôkazov o obsahovej validite využili rozhovory s respondentmi, pričom sa zaujímali o to, ako respondenti tieto položky chápu, či sa pri odpovedi na položky zamýšľajú nad témami, na ktoré sú metodiky zamerané a ako respondenti vnímajú odpoveďové škály. Vďaka tomu identifikovali vlastnosti, ktoré sú v rámci meranej oblasti dostatočne a nedostatočne reprezentované, čo prispieva k presnejším záverom vyvodením na základe výsledku metodiky.

Dôkazy založené na *vzťahu s kritériom*, alebo *kritériová validita*, hovorí o schopnosti metodiky predikovať inú premennú, nazývanú kritérium. Táto vlastnosť metodiky je veľmi dôležitá, pretože poukazuje na schopnosť predikovať aj niečo iné než samú seba (Říčan, 1977). Posúdenie prebieha prostredníctvom porovnania výsledku metodiky a inej premennej, ktorá sa v tomto prípade nazýva kritérium (John & Soto, 2007). Ak napríklad vytvoríme dotazník, ktorý by mal určovať vhodných a nevhodných kandidátov na určitú pracovnú pozíciu, môžeme skúmať, či kandidáti, ktorí boli týmto dotazníkom vybraní, sú úspešní v napĺňaní úloh danej pracovnej pozície, pozitívne hodnotení nadriadeným a pod. Ak by nami zostrojený dotazník nedokázal rozlíšiť medzi vhodnými a nevhodnými kandidátmi, jeho používanie by bolo neopodstatnené a zbytočné. Možné delenie kritériovej validity je spojené s časovým umiestnením kritéria, kedy rozoznávame, ako test rozlišuje kritérium v minulosti (postdiktívna validita) – napr. vzťah s rovesníkmi v detstve, kritérium, ktoré je v aktuálnom čase (konkurentná validita) – napr. osobnostné vlastnosti a kritérium, ktoré sa týka budúcnosti (prediktívna validita) – napr. už spomenutý dotazník pre výber zamestnancov (Halama, 2011). Príkladom skúmania kritériovej validity je napríklad štúdia od Bumberry et al. (1978), ktorí porovnávali výsledky Beckovho dotazníka depresie s hodnotením psychiatrov, ktorí mali

posúdiť hĺbku depresie u participantov. Zistili silnú koreláciu medzi skóre dotazníka a kritériom, na základe čoho konštatujú, že metodika je validná. Štúdie kritériovej validity môžu byť prínosné aj pri výbere vhodnej metodiky pre skúmanie určitej oblasti. Miller et al. (2014) porovnávali 3 metodiky pre diagnostiku narcistickej poruchy osobnosti. Výsledky týchto metodík porovnávali s hodnotením expertov. Výsledky poukázali na rozdiely medzi metodikami v schopnosti zachytiť rôzne aspekty narcizmu. Vďaka týmto zisteniam sa záujemcovia (napr. výskumníci, klinickí psychológovia) môžu lepšie rozhodnúť pri výbere metodiky.

Konštruktová validita je založená na dôkazoch vzťahu výsledku metodiky s inými konštruktmi. V praxi ide najčastejšie o skúmanie vzťahu medzi hrubým skóre validizovanej metodiky a iných metodík, ktoré sú zamerané na skúmanie rovnakej oblasti, pričom očakávame prítomnosť vzťahu – v takom prípade hovoríme aj o konvergentnej validite. Naopak, výsledok metodiky môžeme porovnávať s výsledkami metód (resp. konštruktov), pri ktorých vzťah neočakávame – nazývané aj divergentná validita. Posúdenie validity prebieha administráciou viacerých metodík jednej skupine respondentov a následným skúmaním vzťahov. Okrem iných metodík môžeme vziať v úvahu napríklad vek, rod alebo iné charakteristiky, ktoré by mali s našim konštruktom, a teda aj metodikou, súvisieť – napr. očakávané rozdiely medzi mužmi a ženami v prítomnosti, agresivite a pod. (Halama, 2011; John & Soto, 2007). Ako príklad uvádzame štúdiu od Crawford a Henry (2004), ktorí skúmali konštruktovú validitu škály pozitívneho a negatívneho afektu (PANAS). Po overení dvojfaktorovej štruktúry metodiky korelovali tieto dva faktory s dotazníkmi depresie a úzkosti. Výsledné vzťahy sú podľa očakávaní, čo je dôkazom konvergentnej validity. Autori sa zároveň zamerali na vzťah tejto metodiky s demografickými premennými, pri ktorých neočakávali výrazné vzťahy. Zistené slabé vzťahy slúžia ako dôkaz divergentnej validity. Spomenuté dôkazy o validite patria k najčastejšie uplatňovaným v procese validizácie psychodiagnostických metodík. Halama (2011) spomína tiež dôkazy založené na *dôsledkoch testovania*. Z tohto hľadiska sa pri validizácii metodiky zohľadňujú aj spoločenské dôsledky testovania, ktoré napovedajú o jeho užitočnosti a možných vedľajších dopadoch testovania – znevýhodňovanie, segregácia určitých skupín a pod. (St-Onge et al., 2017).

Proces validizácie psychodiagnostických metodík by sa mal ideálne opierať o čo najviac zdrojov resp. dôkazov validity. Štandardne sa výskumné štúdie zameriavajú na zisťovanie vzťahov medzi konštruktmi či výsledku metodiky a relevantným kritériom. Zaužívané je tiež skúmanie faktorovej štruktúry, prostredníctvom exploratívnej alebo konfirmatórnej faktorovej analýzy, ktorá napomáha skúmaniu, či je očakávaná štruktúra prítomná aj v reálnych dátach.

Menej pozornosti je však venovanej jednotlivým položkám. Z hľadiska kvantitatívneho je možné skúmať, ako metodiky či ich jednotlivé položky fungujú u rôznych skupín respondentov. Takýto prístup môže byť radený k dôkazom o konštruktivej validite no slúži tiež ako indikátor obsahovej validity.

2. Odlišné fungovanie položky

Ako sme v predchádzajúcej kapitole uviedli, existuje viacero možných prístupov ku skúmaniu psychometrických kvalít psychodiagnostických testov a dotazníkov. Pozornosť je mnohokrát v prípade psychometrických štúdií zameraná na fungovanie metodiky ako celku – skúmaný je vzťah výsledného hrubého skóre s inými metodikami, s úspešnosťou v určitej činnosti a pod., no menej často sa stretávame s vedeckými štúdiami, ktoré v rámci posudzovania metodík obracajú pozornosť na jednotlivé položky, z ktorých sú tieto metodiky zložené. V tejto kapitole približujeme prístup ku skúmaniu jednotlivých položiek, ktorý môže napomôcť odhaliť problémy v spravodlivosti psychodiagnostiky a validite psychodiagnostickej metodiky.

Odlišné fungovanie položky, z anglického *differential item functioning*, skratka DIF, označuje stav, kedy položka funguje odlišne medzi oddelenými skupinami respondentov (napr. muži a ženy, vekové kategórie, a pod.). K odlišnému fungovaniu položky dochádza vtedy, keď je pravdepodobnosť kľúčovej (najmä pre osobnostné dotazníky) alebo správnej (najmä pre výkonové testy) odpovede rozdielna medzi oddelenými skupinami participantov, pričom ich miera meranej vlastnosti je približne rovnaká (Clauser & Mazor, 2005). Pre označenie položky ako odlišne fungujúcej je potrebné, aby boli splnené určité podmienky. Z uvedenej definície vyplýva prvá podmienka, ktorá sa týka miery meranej vlastnosti. Tá musí byť rovnaká, resp. približne rovnaká u porovnávaných jednotlivcov z dvoch skupín. Ak by táto podmienka nebola naplnená, nehovoríme o odlišnom fungovaní položky, ale o rozlišovacej schopnosti položky, t.j. schopnosti položky rozlíšiť rôznu mieru vlastnosti, čo je veľmi dôležitá a potrebná schopnosť položky. Taktiež prítomnosť alebo neprítomnosť odlišného fungovania položky nemožno zistiť jednoduchým porovnaním priemerného skóre alebo porovnaním celkového počtu kľúčových odpovedí na položku v jednej a druhej skupine. Rozdiely v priemernom skóre alebo počte kľúčových odpovedí môžu byť spôsobené reálnymi rozdielmi medzi skúmanými skupinami. Ďalšou dôležitou podmienkou je, že kritérium, podľa ktorého sme stanovili porovnávané skupiny, nesúvisí s tým, čo má daná položka merať (Clauser & Mazor, 2005; Osterlind & Everson, 2009). Fiktívnym príkladom môže byť testovanie extravenzie. Porovnávame fungovanie položky u mužov a žien. Pri položke, ktorá nevykazuje odlišné fungovanie sme zistili približne rovnakú pravdepodobnosť kľúčovej odpovede u mužov a žien pri rôznej miere extravenzie (avšak rovnakej miere u mužov a žien). Napríklad muži aj ženy s hrubým skóre 10 dosahujú približne 30% šancu kľúčovej odpovede. V prípade položky, ktorá funguje odlišne medzi týmito skupinami je pravdepodobnosť kľúčovej odpovede odlišná, napr.

30% u mužov, no 60% u žien napriek tomu, že porovnávame mužov a ženy s rovnakým hrubým skóre.

Pomenovanie kľúčová odpoveď možno používať pri výkonových testoch aj osobnostných dotazníkoch. Pri výkonových testoch je kľúčovou odpoveďou správna odpoveď, zatiaľ čo pri osobnostných dotazníkoch je to odpoveď, ktorá je bodovaná. V prípade položiek typu Likertovej škály je kľúčovou odpoveďou konkrétna hodnota škály. Zaužívaným spôsobom označovania skupín je referenčná a fokálna. Referenčná skupina je tá, u ktorej sa predpokladá zvýhodňovanie a fokálnou označujeme tú skupinu, u ktorej predpokladáme, že ju položka znevýhodňuje. Vo všeobecnosti ale nezáleží na tom, ktorú skupinu označíme ako referenčnú a fokálnu (Penfield & Camilli, 2006). Matematicky možno vyjadriť stav kedy položka nevykazuje odlišné fungovanie nasledujúcim vzorcom:

$$f(Y|\theta, G = R) = f(Y|\theta, G = F)$$

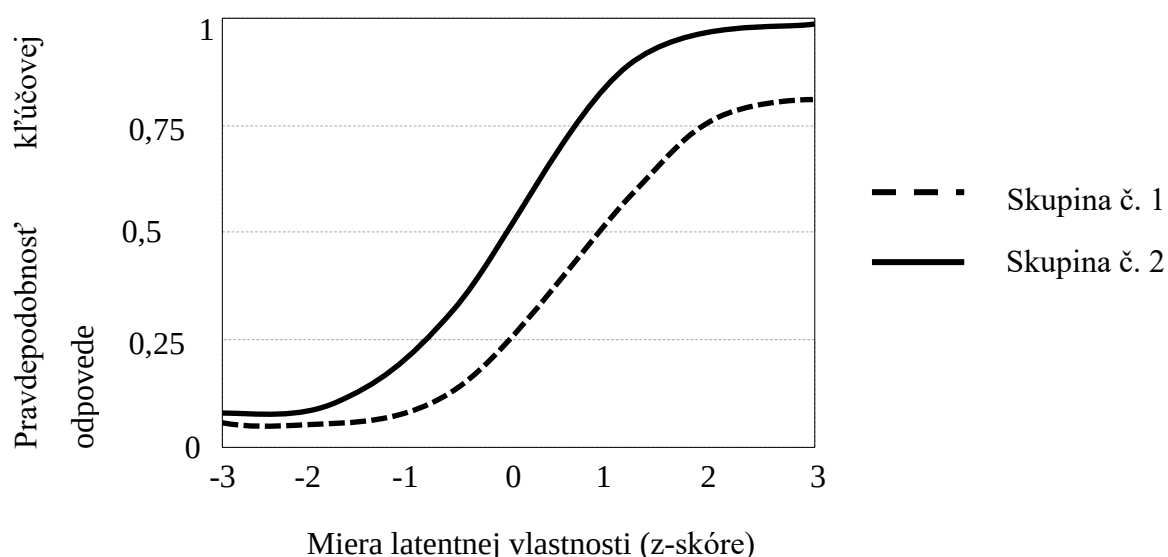
Tento vzorec vyjadruje situáciu, kedy je pravdepodobnosť kľúčovej odpovede (Y) pri určitej miere meranej (latentnej) vlastnosti (θ) rovnaká pre obe skupiny, referenčnú ($G = R$) aj fokálnu ($G = F$). Vyjadruje teda nulovú hypotézu, čiže neprítomnosť odlišného fungovania položky. Ak by tento vzorec neplatil, znamenalo by to, že jednotlivci z jednej skupiny (napr. fokálnej) majú nižšiu pravdepodobnosť správnej odpovede pri rovnakej miere meranej vlastnosti. V takomto prípade položka pravdepodobne zachytáva aj inú vlastnosť alebo schopnosť, v ktorej sa tieto skupiny líšia. Dôležité je upozorniť, že samotná prítomnosť odlišného fungovania položky automaticky neznamená, že je položka skresľujúca alebo nespravodlivá. Po identifikovaní odlišne fungujúcich položiek je potrebné analyzovať možný dôvod vzniku odlišného fungovania a posúdiť, či položka meria vlastnosti alebo schopnosti relevantné alebo irelevantné k pôvodnému zámeru (Holland & Thayer, 1988; Osterlind & Everson, 2009)

Takýmto spôsobom vysvetľuje vznik odlišného fungovania položiek aj multidimenzionálny model odlišného fungovania položky. Podľa tohto modelu vzniká odlišné fungovanie položky kvôli multidimenzionalite skúmaného konštruktu, ktorý bol pôvodne považovaný za jednodimenzionálny. V takomto prípade sú k primárnej dimenzii pričlenené aj ďalšie sekundárne dimenzie, ktoré môže položka zachytávať. Členovia dvoch skúmaných skupín sa v prípade odlišného fungovania položky nelíšia v miere primárnej dimenzie, to znamená, že majú (približne) rovnakú mieru meranej vlastnosti, no odlišujú sa v sekundárnej dimenzii, ktorá je potrebná pre riešenie položky (Roussos & Stout, 1996). Táto rozdielnosť v

miere sekundárnej dimenzie sa prejavuje rozdielnou pravdepodobnosťou kľúčovej odpovede medzi skupinami s rovnakou mierou primárnej dimenzie. Ako sme už uviedli, je ďalej potrebné posúdiť, či táto sekundárna dimenzia je alebo nie je v súlade so zámerom psychodiagnostického testu alebo dotazníku. V prípade, že je sekundárna dimenzia irelevantná k primárnej, môžeme hovoriť o skresľujúcej položke, čo môže mať dopad na spravodlivosť a validitu psychodiagnostickej metodiky.

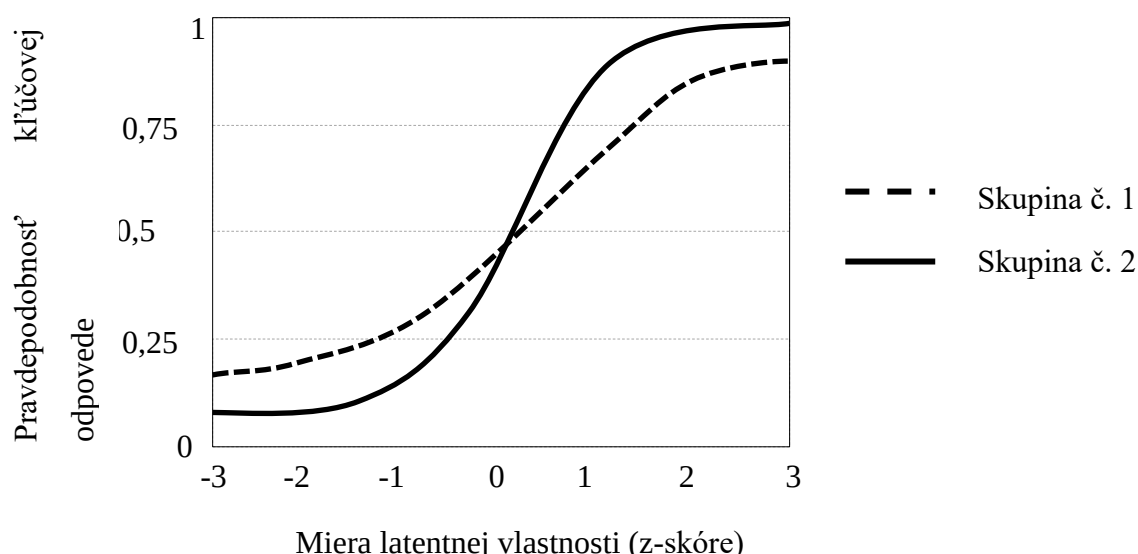
2.1. Uniformné a neuniformné DIF

Odlišné fungovanie položiek delíme na uniformné a neuniformné. Obrázok č. 2 znázorňuje príklad charakteristickej krivky položky s uniformným DIF. Vodorovná os znázorňuje mieru meranej vlastnosti, od nízkej (-3) cez strednú (0) až vysokú (3). Horizontálna os vyjadruje pravdepodobnosť kľúčovej odpovede na položku v rozmedzí od 0 po 1, resp. 0-100%. Plná a delená čiara znázorňuje charakteristickú krivku položky pre dve skupiny. V prípade uniformného DIF položka zvýhodňuje alebo znevýhodňuje jednu skupinu v rozsahu celej meranej (latentnej) vlastnosti. Obrázok 1 znázorňuje príklad položky, ktorá svojim odlišným fungovaním zvýhodňuje skupinu číslo 2. Tá má vyššiu pravdepodobnosť kľúčovej odpovede v rozsahu celej miery meranej (latentnej) vlastnosti. Napríklad pri priemernej miere vlastnosti (bod 0) je pravdepodobnosť kľúčovej odpovede pre skupinu č. 2 okolo 50%, zatiaľ čo pre skupinu č. 1 len okolo 25% (Zumbo, 1999; Osterlind, Everson, 2009).



Obrázok 1 Znázornenie uniformného odlišného fungovania položky

V prípade neuniformného DIF je taktiež pravdepodobnosť kľúčovej odpovede odlišná pri rovnakej miere meranej vlastnosti, no mení sa smer odlišného fungovania v závislosti od miery meranej vlastnosti. To znamená, že pri určitej miere vlastnosti položka zvyhodňuje jednu skupinu, no pri inej miere vlastnosti tú istú skupinu znevýhodňuje. Príklad v Obrázku 2 znázorňuje situáciu, kedy pri nízkej až priemernej miere vlastnosti položka zvyhodňuje skupinu č. 1, no pri priemernej až vysokej miere túto skupinu znevýhodňuje, resp. zvyhodňuje skupinu č. 2 (Zumbo, 1999; Osterlind, Everson, 2009).



Obrázok 2 Zobrazenie neuniformného odlišného fungovania položky

Zobrazené obrázky sú len príkladom, ako môžu vyzerat' charakteristické krivky položiek pri DIF. Pri reálnych dátach však môže nastať situácia, že pravdepodobnosť kľúčovej či správnej odpovede je bez štatisticky významného rozdielu pri niektorej úrovni meranej vlastnosti a DIF nastáva len v určitom pásme, t.j. krivky pre jednotlivé skupiny sa pri určitej úrovni meranej vlastnosti prekrývajú, pri inej sú od seba vzdialenejšie. Charakteristické krivky položiek sa môžu dokonca prekrížiť, no pokiaľ je rozdiel v pravdepodobnosti kľúčovej odpovede signifikantný len pre jednu zo skupín, stále ide o uniformné DIF.

2.2. Dôležitosť skúmania odlišného fungovania položiek

Skúmanie odlišného fungovania položiek sa začalo rozvíjať najmä v 60. a 70. rokoch 20. storočia v Spojených štátoch amerických v súvislosti s právnymi a spoločenskými zmenami týkajúcimi sa rasizmu a znevýhodňovania určitých skupín, avšak určité náznaky problematiky

odlišného fungovania položiek psychodiagnostických testov sa objavovali omnoho skôr. V spojitosti s testovaním inteligencie už Alfred Binet alebo Wiliam Stern upozorňovali na položky v intelligenčných testoch, ktorých správne vyriešenie je viazané na faktory, ktoré nemusia byť dôležité pre meraný konštrukt, ako napr. socio-ekonomická úroveň respondenta alebo domáca príprava (podľa Camilli, 2006). Podobné závery uvádzajú aj Eells et al. (in Camilli, 2006), ktorí skúmali viac ako 650 položiek z 8 intelligenčných testov. Tí v ich štúdiu publikovanej v roku 1951 neskúmali DIF, no zaoberali sa kultúrnym skreslením testových položiek. Identifikovali viacero položiek, pri ktorých je možnosť správnej odpovede ovplyvnená skôr špecifickým obsahom položiek ako rozumovou schopnosťou respondentov – ako príklad hovoria o položke, pre ktorej riešenie je potrebná určitá znalosť orchestrálnych hudobných nástrojov – deti, ktoré mali možnosť zúčastňovať sa takýchto hudobných koncertov majú vyššiu pravdepodobnosť správne vyriešiť položku ako tie, ktoré nikdy nevideli symfonický orchester (Eells et al. in Camilli, 2006).

Po spoločenských zmenách, ktoré prebiehali najmä v USA po roku 1960, týkajúcich sa rovnoprávnosti, a najmä uzákonením rovnakých podmienok pre všetkých obyvateľov bez ohľadu na pohlavie, rasu, vierovyznanie alebo pôvod v roku 1964, sa zvýšil záujem o boj proti diskriminácii. Snaha o spravodlivosť a zabránenie znevýhodňovaniu niektorých skupín sa postupne prejavila aj v psychodiagnostike. V roku 1999 boli tromi významnými organizáciami – Americkou asociáciou edukačného výskumu (American Educational Research Association, AERA), Americkou psychologickou asociáciou (American Psychological Association, APA) a Národným koncilom pre meranie v edukácii (National Council on Measurement in Education, NCME) – vytvorené štandardy pre pedagogické a psychologické testovanie, ktoré boli neskôr vydané aj u nás (AERA, APA, NCME, 2001). Štandardy obsahujú kapitolu, ktorá sa zaoberá spravodlivosťou pedagogického a psychologického testovania a taktiež predchádzaním skreslenia. Spravodlivý prístup v psychodiagnostike možno chápať tromi spôsobmi (AERA, APA, NCME, 2001):

- spravodlivý prístup ako neprítomnosť skreslenia – ide o skreslenie (angl. *bias*), ktoré vzniká v dôsledku nedostatkov psychodiagnostickej metódy alebo jej nesprávneho použitia. Dôsledkom takéhoto skreslenia v psychologickom testovaní je, že výsledné skóre nemá rovnaký význam pre členov rôznych skupín.
- spravodlivý prístup ako nestranné zaobchádzanie pri testovaní – v tomto význame chápeme spravodlivosť ako zabezpečenie primeraných podmienok psychodiagnostiky, rovnakú možnosť zoznámiť sa obsahom psychodiagnostickej metódy, rovnaký prístup

k dôležitým informáciám ohľadne metódy a pod. Z tohto dôvodu majú mnohé psychodiagnostické metódy stanovený štandardizovaný postup administrácie.

- spravodlivý prístup ako rovnosť výsledkov testovania – v tomto ponímaní je spravodlivosť chápaná tak, že odlišné skupiny testovaných osôb by mali dosahovať rovnaké výsledky psychodiagnostiky. Ako sme už v predchádzajúcej časti spomenuli, je dôležité, aby výsledok testovania nebol ovplyvnený príslušnosťou k určitej skupine, no rozdielne výsledky, napr. v podobe priemerného skóre, nie sú nutne dôkazom nespravodlivosti testovania, no môžu byť spôsobené reálnymi rozdielmi medzi týmito skupinami.

Potreba skúmania odlišného fungovania položiek sa odráža v prvom chápaní spravodlivého prístupu v psychodiagnostike. Ak položka vykazuje odlišné fungovanie, môže ísť o skreslenie, pretože pravdepodobnosť kľúčovej odpovede na ňu je vyššia pre jednu z testovaných skupín. Medzi najčastejšie sa vyskytujúce skresľujúce faktory patria rozličné socio-ekonomické podmienky, nerovnaká možnosť prístupu k informáciám, rozdielny spôsob výchovy, rozdielne záujmy a profesionálna orientácia medzi pohlaviami a iné (Einarsdóttir & Rounds, 2009; Sheppard et al., 2006). Vhodnosť a dôležitosť skúmania odlišného fungovania položiek je ukotvená aj v Štandardoch pre pedagogické a psychologické testovanie (AERA, APA, NCME, 2001, s. 87): „Pokiaľ spoľahlivé výskumné zistenia ukazujú, že v populácii testovaných osôb existuje v oblasti meranej testom rozlišovacie pôsobenie položiek [odlišné fungovanie položiek, pozn. autora] pre kategórie osôb vymedzené vekom, pohlavím, rasovou či etnickou alebo kultúrnou príslušnosťou, prítomnosťou postihnutia, alebo používaním jazykom, potom by autori testu mali ak je to možné vykonať príslušné výskumné štúdie. V nich by sa mali snažiť odhaliť a eliminovať tie aspekty usporiadania, obsahu a formátu testu, ktoré môžu skresľovať testové skóre u určitých osôb.“

Vážnosť dopadu odlišného fungovania položiek na spravodlivosť psychodiagnostickej metódy súvisí s počtom položiek, ktoré zvýhodňujú alebo znevýhodňujú určitú skupinu. Psychodiagnostická metodika môže obsahovať niekoľko položiek, ktoré zvýhodňujú jednu skupinu a zároveň niekoľko položiek, ktoré zvýhodňujú druhú skupinu. Z hľadiska celého testu alebo dotazníku sa odlišné fungovanie položiek navzájom ruší a metodika ako celok nezvýhodňuje ani jednu skupinu. Pri zvažovaní dôležitosti skúmania odlišného fungovania položiek je potrebné brať v úvahu aj kontext použitia psychodiagnostickej metodiky. Zaujímavý príklad uvádzajú Karami a Salmani Nodoushan (2011). Môže nastať situácia, že psychodiagnostický test obsahuje niekoľko položiek znevýhodňujúcich ženy. Z hľadiska celého testu je to zanedbateľný počet položiek, ktorý sa v celkovom skóre prejaví štatisticky

nevýznamným rozdielom napr. dvoch alebo troch bodov. Takýto test ale môže byť použitý v prijímacom konaní do zamestnania. Čo ak uchádzačka získa 59 bodov z potrebných 60 aby mohla postúpiť do ďalšieho kola? Z hľadiska kritéria neuspela, no je otázne, či by jeden potrebný bod nezískala, ak by test neobsahoval skresľujúce, odlišne fungujúce položky.

Ako sme v predchádzajúcom texte uviedli, s problematikou odlišného fungovania položiek je úzko spätá otázka validity psychodiagnostických testov a dotazníkov. Validita je spolu s reliabilitou a objektivitou dôležitou vlastnosťou psychodiagnostických testov a dotazníkov. Validita psychodiagnostických metodík má rôzne definície, no vo všeobecnosti vyjadruje schopnosť metodiky merať to, na čo bola skonštruovaná, resp. na čo je určená. Štandardy pre pedagogické a psychologické testovanie validitu definujú ako mieru, v ktorej empirické dôkazy a teórie podporujú interpretáciu testového skóre pri odporúčanom spôsobe použitia testu (AERA, APA, NCME, 2001). Na rozdiel od reliability, ktorá je pomerne ľahko vypočítateľná, je zisťovanie validity, ako predchádzajúca definícia napovedá, zložitejšie, založené na rôznych dôkazoch (Halama, 2011; John & Soto, 2007).

Prítomnosť DIF v psychodiagnostickej metodike môže poslúžiť ako indícia smerom k skúmaniu obsahu potencionálne problematickej položky. Ako sme pri definovaní odlišného fungovania položiek uviedli, DIF je spôsobené tým, že odpoveď na položku je ovplyvnená aj inými faktormi ako je zamýšľaná meraná vlastnosť a skúmané skupiny sa v tomto faktore odlišujú. Zdôrazňujeme, že prítomnosť DIF nemožno automaticky považovať za dôkaz problému s validitou metodiky – z hľadiska dimenzionality môžeme konštatovať, že položka nie je jednodimenzionálna, treba ale posúdiť, či to čo položka meria má alebo nemá súvis s tým, čo by mala merať. Potrebné je ďalej sa zaoberať obsahom položiek, identifikovaním možných príčin odlišného fungovania položiek a posúdením, či sú tieto príčiny ohrozujúce pre validitu metodiky alebo odlišné fungovanie položiek vzniká na základe reálnych rozdielov v čiastkových vlastnostiach, ktoré súvisia s meraným konštruktom a sú jeho dôležitou súčasťou. Limitáciou je, že položka môže merať niečo, na čo nie je určená, no pokiaľ sa skúmané skupiny neodlišujú v miere tejto vlastnosti, neprejaví sa to v podobe odlišného fungovania položky, no napriek limitáciám je skúmanie odlišného fungovania položky užitočným nástrojom pre skúmanie a zlepšovanie kvality, spravodlivosti a validity psychodiagnostických nástrojov (Walker, 2011).

Skúmať odlišné fungovanie položiek môže byť prínosné taktiež pri prekladaní a adaptácii zahraničných metód. Pri adaptácii metód sa vyskytujú viaceré problémy, položka nemusí mať rovnaký význam pre ľudí z odlišných krajín. Porovnaním fungovania položiek medzi respondentmi pôvodnej a preloženej metódy môžeme lepšie hodnotiť kvalitu prekladu a

vhodnosť používania položky (Ellis, 1989). Vďaka pokročilým štatistickým metódam, ktoré nevyžadujú stanovenie porovnávaných skupín možno skúmaním odlišného fungovania položiek rozšíriť poznanie ohľadom rozdielov medzi rôznymi skupinami respondentov. Nezostáva teda len pri zhodnotení spravodlivosti a validity metódy, no ďalej skúmame možné príčiny vzniku odlišného fungovania položky medzi identifikovanými skupinami (Zumbo, 2007).

2.3. Metódy skúmania odlišného fungovania položiek

V tejto časti uvádzame niektoré zo štatistických metód analýzy odlišného fungovania položiek. Momentálne sú dostupné rôzne metódy, ako aj ich modifikácie, ktoré sú založené na klasickej teórii testov ale i teórii odpovede na položku. Vzhľadom na množstvo metód, neposkytujeme vyčerpávajúci prehľad, ale skôr inšpiráciu pre čitateľov, ktorí majú o túto problematiku záujem.

Mantel-Haenszelova a Mantelova metóda

Túto neparametrickú metódu pre skúmanie odlišného fungovania položiek spopularizovali Holland a Thayer (1988). Analýza DIF funguje na základe porovnania pravdepodobnosti kľúčovej odpovede medzi dvomi skupinami v kontingenčnej tabuľke. Pre dodržanie podmienky skúmania DIF, že porovnávaní respondenti majú rovnakú alebo približne rovnakú mieru meranej vlastnosti, je potrebné súbor najskôr rozdeliť na viacero častí, zvyčajne 5 a viac, podľa úrovne diagnostikovanej vlastnosti, najčastejšie podľa hrubého skóre analyzovanej psychodiagnostickej metodiky. Všeobecný formát Mantel-Haenszelovej (M-H) kontingenčnej tabuľky pre dichotómne položky zobrazujeme nižšie (podľa Osterlind & Everson, 2009).

Všeobecný formát M-H kontingenčnej tabuľky

Položka i; Úroveň vlastnosti j			
Odpoveď	1 (kľúčová, správna)	0 (nesprávna)	Spolu
Referenčná skupina (R)	a_i	b_i	N_{Ri}
Cieľová skupina (F)	c_i	d_i	N_{Fi}
Spolu	N_1	N_0	N_T

Pri M-H procedúre testujeme nulovú hypotézu, ktorá hovorí, že pravdepodobnosť kľúčovej odpovede je rovnaká pre referenčnú a fokálnu skupinu. Pokiaľ sa táto

pravdepodobnosť štatisticky významne líši, položka vykazuje odlišné fungovanie pri danej úrovni meranej vlastnosti. Mantel-Haenszelova metóda je určená výhradne pre dichotómne položky, no upravená Mantelova metóda je použiteľná aj pre polytómne položky.

Metódy založené na teórii odpovede na položku

Ďalšou skupinou metód pre skúmanie DIF sú metódy založené na teórii odpovede na položku (angl. item response theory, ďalej ako IRT). Na rozdiel od klasickej teórie testov, pre ktorú nemá samotná položka väčší význam a je závislá od kontextu celku, IRT považuje samotnú položku za základ a test, resp. psychodiagnostickú metódu chápe ako súbor položiek. IRT pôvodne vznikla ako metóda pre analýzu dichotómnych položiek, no v dnešnej dobe disponuje metódami pre analýzu položiek s možnosťou výberu z viacerých možností alebo položiek typu Likertovej škály (Halama, 2011; Urbánek & Šimeček, 2001). Praktickým príkladom môže byť psychometrická analýza škál úzkosti na Slovensku od Urbánka a Sollára (2014) v ktorej sa skúmali odlišného fungovanie položiek prostredníctvom IRT.

Základom skúmania položiek v IRT je takzvaná charakteristická krivka položky (item characteristic curve – skratka ICC). Tá má tvar logistickej krivky, ktorú sme zobrazili v obrázku č. 1. Na základe tejto krivky môžeme zistiť (alebo predpovedať), akú má jednotlivец s určitou mierou vlastnosti, ktorá sa v IRT nazýva latentný rys, pravdepodobnosť odpovedať správne, resp. v smere meranej vlastnosti. Tieto krivky majú podobný tvar, no líšia sa od položky k položke. Poloha a sklon krivky závisí od modelu, ktorý použijeme. Pre dichotómne položky rozoznávame tri modely, ktoré sa líšia počtom parametrov. Základným modelom je jednoparametrový model v skratke 1PL (angl. one-parameter logistic model). Obsahuje len jeden parameter – obtiažnosť položky. Obtiažnosť položky vypovedá o tom, aká je miera latentného rysu pri 50% pravdepodobnosti správnej odpovede. 1PL model je pomerne jednoduchý a pre väčšinu psychodiagnostických metód je lepšie využiť dvojparametrový model 2PL. V ňom je okrem obtiažnosti položky aj parameter rozlišovacej účinnosti. Ten vyjadruje, ako dokáže položka rozlíšiť medzi rôznou mierou latentného rysu, inými slovami, ako sa zmena v úrovni latentného rysu prejaví na pravdepodobnosti kľúčovej odpovede. Na ICC sa tento parameter odrazí v jej sklone, resp. strmosti. Položky so strmejšou ICC majú vyššiu rozlišovaciu schopnosť. Posledným zo základných modelov je trojparametrový model, ktorý dopĺňa 2PL model o tretí parameter, nazývaný uhádnuteľnosť alebo pseudouhádnuteľnosť (angl. pseudo-guessing). Tento 3PL model je použiteľný pre položky psychodiagnostických testov, ktoré majú len jednu správnu odpoveď. Aj v prípade, že testovaná osoba nevie správnu odpoveď, má určitú pravdepodobnosť správnej odpovede. Tento parameter sa prejavuje vo výške, v ktorom ICC začína (Jelínek et al., 2011; Urbánek & Šimeček, 2001).

Analýza DIF je pomocou IRT možná dvomi spôsobmi, porovnaním parametrov položiek alebo porovnaním plochy oblasti medzi charakteristickými krivkami položiek (Jelínek et al., 2011). V prípade neprítomnosti DIF by mali byť parametre položky rovnaké pre referenčnú aj cieľovú skupinu. Neprítomnosť alebo prítomnosť DIF sa v tomto prípade overuje chí-kvadrát testom. Ak je preukázaný signifikantný rozdiel parametrov položky medzi dvomi skupinami, môžeme konštatovať prítomnosť DIF (Lord in Jelínek, Květon, Vobořil, 2011). Pri porovnávaní plochy oblasti medzi charakteristickými krivkami položky sa vychádza z predpokladu, že v prípade neprítomnosti DIF je táto plocha nulová, resp. čo najmenšia. Prítomnosť DIF je potvrdená na základe výslednej hodnoty plochy oblasti medzi krivkami, ktorá sa odvíja od zvolenej hladiny významnosti (Raju, 1988).

Metódy IRT pre skúmanie DIF polytómnych položiek vychádzajú z IRT modelov zameraných na polytómne položky, napr. model odstupňovaných odpovedí (Samejima, 1969), ktorý je rozšírením dvojparametrového modelu dichotómnych položiek. Skúmanie DIF polytómnych položiek tak prebieha porovnaním pravdepodobnosti určitej odpovede (bodu na škále) pri rovnakej miere latentnej vlastnosti medzi referenčnou a cieľovou skupinou. (Osterlind, Everson, 2009; Jelínek, Květon, Vobořil, 2011). Výhodou IRT pri analýze DIF je detailnejšie opísanie položky prostredníctvom jej parametrov v porovnaní s klasickou teóriou testov a tiež možnosť zrozumiteľného grafického znázornenia DIF. Možnou nevýhodou je vyššia zložitosť a náročnosť IRT metód.

Logistická regresia a ordinálna logistická regresia

Pomocou logistickej regresie, podobne ako pri Mantel-Haenszelovej procedúre, skúmame pravdepodobnosť kľúčovej odpovede. Neporovnávame ale pravdepodobnosť medzi skupinami, no skúmame, či je pravdepodobnosť kľúčovej odpovede predikovateľná príslušnosťou k referenčnej či fokálnej skupine (Swaminathan, Rogers, 1990, Zumbo, 1999). Logistická regresná analýza odlišného fungovania položky prebieha v troch krokoch, ktoré uvádzame na príklade dichotómnej položky. Keďže pravdepodobnosť kľúčovej odpovede na položku by mala byť závislá na miere meranej vlastnosti, prvým krokom je overenie základného modelu. Do logistickej regresie vložíme prvú podmienenú premennú vyjadrujúcu mieru meranej vlastnosti. Tento model možno zjednodušene vyjadriť nasledovne (Zumbo, 1999):

$$Y = b_0 + b_1 \text{ miera vlastnosti}$$

V druhom kroku do modelu pridáme premennú vyjadrujúcu príslušnosť k skupine. Takýmto spôsobom skúmame uniformné DIF.

$$Y = b_0 + b_1 \text{ miera vlastnosti} + b_2 \text{ skupina}$$

Napokon, v treťom kroku do modelu pridáme aj premennú, ktorá vyjadruje interakciu meranej vlastnosti a príslušnosti k skupine (HS*Skup). V tomto kroku overujeme uniformné a neuniformné DIF súčasne. Finálny model vyzerá nasledovne (Zumbo, 1999):

$$Y = b_0 + b_1 \text{ miera vlastnosti} + b_2 \text{ skupina} + b_3 \text{ miera vlastnosti} * \text{skupina}$$

Pre zistenie prítomnosti odlišného fungovania položky je ďalej potrebné porovnať výpovednosť modelu v troch uvedených krokoch. Zumbo (1999) navrhuje porovnávať prvý a tretí krok. Od hodnoty výpovednosti tretieho modelu odčítame hodnotu výpovednosti prvého modelu. Pri výslednej hodnote potom zisťujeme pravdepodobnosť rozdelenia chí-kvadrát pri dvoch stupňoch nezávislosti. Pokiaľ je chí-kvadrát test signifikantný, existuje podozrenie, že položka funguje odlišne, pretože výpovednosť modelu v ktorom je zahrnutá premenná vyjadrujúca príslušnosť k skupine a interakčná premenná, je signifikantne vyššia ako výpovednosť základného (prvého) modelu. Pokiaľ existuje podozrenie z DIF, je vhodné pokračovať ďalej a skúmať, či ide o uniformné alebo neuniformné DIF, a to podobným postupom, pri ktorom porovnáваме prvý a druhý model, následne druhý a tretí model pomocou chí-kvadrát testu s jedným stupňom voľnosti. Ak je tretí model signifikantne výpovednejší, ide pravdepodobne o neuniformné DIF, ak nie, ide o uniformné DIF. Pre zistenie veľkosti odlišného fungovania porovnáваме zmenu koeficientu logistickej regresie pseudo-R² (napr. Nagelkerkeho R²) medzi jednotlivými krokmi. Koeficienty logistickej regresie pseudo-R² však nemožno interpretovať ako koeficient determinácie pri lineárnej regresnej analýze. Z tohto dôvodu niektorí autori vytvorili normy pre interpretáciu (Jodoin & Gierl, 2001).

Skúmanie polytómnych položiek je možné vďaka ordinálnej logistickej regresii. Postup analýzy sa však z hľadiska postupu výrazne neodlišuje. Ako závislú premennú nastavíme analyzovanú položku a následne overujeme a porovnáваме 3 modely, podobne ako pri analýze dichotómnych položiek (Zumbo, 1999). Skúmať odlišné fungovanie položiek prostredníctvom logistickej regresie je možné v štandardných štatistických programoch (napr. SPSS). Metódu logistickej regresie možno vykonať aj pomocou niektorých balíkov v R (napr. difR, lordif). Niektorí autori (Crane et al., 2006) odporúčajú kombináciu logistickej regresie a teórie odpovede na položku, kedy ako premennú vyjadrujúcu meranú vlastnosť nepoužívame hrubé skóre ale odhad miery vlastnosti. Túto metódu využívame aj my pre skúmanie DIF vo výskumnej časti tejto monografie.

3. Invariantnosť merania

Invariantnosť merania, z anglického *measurement invariance*, je ďalším spôsobom posúdenia toho, či psychodiagnostický nástroj funguje rovnako, resp. podobne v rôznych skupinách prípadne meraniach. Tento prístup vieme využiť nielen pre posúdenie rovnosti merania medzi skupinami, ale aj medzi opakovanými meraniami, napríklad v prípade longitudinálneho výskumu alebo test-retest zmeny po intervencii. V posledných dvoch desaťročiach je tento prístup veľmi obľúbený v rámci psychometrických štúdií zameraných na validáciu nástrojov, no uplatnenie nachádza aj pri multikultúrnych výskumoch z rôznych oblastí psychológie (Putnick & Bornstein, 2016).

Na rozdiel od odlišného fungovania položiek, pri ktorom sa zaujímame o jednotlivé položky zvlášť, testovanie invariantnosti merania je zamerané na model s latentnými faktormi. Pre uskutočnenie analýz je najčastejšie využívaný prístup prostredníctvom konfirmačnej faktorovej analýzy (anglicky *confirmatory factor analysis*, skratka CFA), no existujú pokročilejšie analýzy využívajúce exploračnú faktorovú analýzu či kombináciu oboch v podobe exploračného štruktúrneho modelovania (Tóth-Király et al., 2017). V tejto monografii bližšie popíšeme len základný prístup založený na konfirmačnej faktorovej analýze. Pri konfirmačnej faktorovej analýze predpokladáme existenciu latentnej premennej, ktorú nemôžeme priamo merať, no táto latentná premenná sa manifestuje určitými spôsobmi, takzvanými manifestnými premennými, ktoré merať vieme. Základ analýzy vychádza z teoreticky stanoveného modelu. Ten sa pokúšame overiť prostredníctvom zistenia zhody s reálnymi dátami, ktoré sme vo výskume získali. Výsledkom analýzy sú koeficienty, ktoré nám indikujú ako dobre (alebo zle) sa model zhoduje s dátami. Vďaka týmto koeficientom vieme porovnávať rôzne modely medzi sebou (pozri tiež Sollár, 2014). Príkladom môže byť adaptácia a validácia slovenskej verzie Inventára Veľkej Päťky 2 (Halama et al., 2020), v ktorej sme vzájomne porovnávali viacero modelov s cieľom zistiť, ktorý má lepšiu zhodu s dátami. Tento princíp nachádza uplatnenie aj pri posudzovaní invariantnosti psychodiagnostickej metodiky.

Základným vyjadrením zhody modelu s dátami je chí-kvadrát test. Jeho nulovou hypotézou je, že model je v zhode s dátami. Pokiaľ je výsledok testu štatisticky významný, znamená to, že sa model štatisticky významne odlišuje od dát. Takýto prístup k posudzovaniu je však výrazne limitovaný tým, že pri väčšom počte respondentov je chí-kvadrát test veľmi citlivý aj na prakticky zanedbateľné rozdiely medzi modelom a dátami, čo vedie k častému odmietnutiu modelu. Z tohto dôvodu sú využívané takzvané alternatívne indexy zhody. Je ich viacero, no vo výskumoch sa najčastejšie uvádza komparatívny index zhody (*comparative fit*

index, CFI) a odmocnina priemerného štvorca chyby odhadu (*root mean square error of approximation*, RMSEA). Výhodou týchto indexov je, že nie sú závislé na veľkosti súboru, no môžu byť ovplyvnené komplexnosťou modelu (Putnick & Bornstein, 2016). Vo všeobecnosti sa model považuje za dobrý, ak je hodnota CFI väčšia ako 0,90 a hodnota RMSEA menšia ako 0,08. Hoci ide o indexy zhody, CFI vypovedá o tom, ako dobre model sedí na dáta (preto čím vyššia hodnota, tým lepšie), no RMSEA hovorí o tom, koľko chyby odhadu model obsahuje (preto čím nižšia hodnota, tým lepšie). Pri posudzovaní modelu môže nastať situácia, že na základe jedného indexu by sme model považovali za prijateľný, no na základe druhého nie. V takomto prípade je brať v úvahu najmä ten, ktorý je pre daný model výpovednejší (pozri Lai & Green, 2016).

3.1. Postup testovania invariantnosti modelu

Vandenberg a Lance (2000) vo svojej štúdii priniesli prehľad a syntézu literatúry ohľadom testovania invariantnosti merania. V rámci ich štúdie uvádzajú aj postup pre overenie invariantnosti. Ten sa skladá z troch, prípadne štyroch úrovní, ktoré sú využívané aj v dnešnej dobe (Putnick & Bornstein, 2016).

Na začiatok je potrebné vytvoriť latentný model, ktorý má akceptovateľnú zhodu s dátami. Po jeho vytvorení tento model postupne sprísňujeme a na základe zmeny posudzujeme dosiahnutie alebo nedosiahnutie určitej úrovne invariantnosti, konkrétne:

- 1) *Konfigurálna invariantnosť* – rovnaká forma: model testujeme samostatne pre skúmané skupiny. Pokiaľ model vykazuje dobrú zhodu s dátami, konštatujeme, že vzorec vzťahov medzi faktormi a položkami je v týchto skupinách rovnaký.
- 2) *Metrická invariantnosť* – rovnaké faktorové náboje: v tomto kroku nastavíme v skupinách rovnaké faktorové náboje (regresné koeficienty z latentného faktoru na položky). Pokiaľ sa zhoda modelu s dátami v tomto stupni nezhoršila, konštatujeme, že latentný faktor sýti položky rovnako v porovnávaných skupinách. Táto úroveň invariantnosti sa niekedy v literatúre označuje ako slabá (*weak*).
- 3) *Skalárna invariantnosť* – rovnaké úrovňové konštanty (angl. *intercept*): v tomto kroku sú priemery položiek nastavené ako rovnaké v porovnávaných skupinách, čo znamená, že pri rovnakej miere latentného faktora by mali mať respondenti z rôznych skupín rovnakú hodnotu priemeru položiek. Táto úroveň invariantnosti sa niekedy v literatúre označuje ako silná (*strong*).

- 4) *Invariantnosť reziduálov* – rovnaké chyby odhadu: v tomto kroku nastavíme, že variancia položky, ktorá nie je vysvetlená latentným faktorom (tzv. reziduál), je rovnaká v porovnávaných skupinách. Táto úroveň invariantnosti sa niekedy v literatúre označuje ako striktná (*strict*).

Posledná uvedená úroveň nie je nutná pri testovaní invariantnosti merania, najmä v prípade, ak je naším cieľom zabezpečenie spoľahlivého porovnávanie skupín v priemeroch latentných faktorov.

Posúdenie toho, či je dosiahnutá určitá úroveň invariantnosti, môže byť uskutočnené prostredníctvom porovnania dvoch modelov chí-kvadrát testom, keďže ide o vnorený resp. hierarchický model – stále ide o ten istý model, ktorý každou úrovňou viac špecifikujeme. Pokiaľ je výsledok chí-kvadrát testu štatisticky nevýznamný, konštatujeme, že zhoda modelu s dátami sa výrazne nezmenila a daná úroveň invariantnosti modelu bola dosiahnutá. Avšak, tak ako v prípade posúdenia zhody celkového modelu, je aj v tomto prípade problémom prílišná citlivosť chí-kvadrát testu. Z tohto dôvodu sa možno zamerať skôr na zmenu v alternatívnych indexoch zhody, napr. CFI a RMSEA. Stále však nie je jasné, aké zhoršenie zhody modelu s dátami je tolerovateľné. Chen (2007) uvádza ako kritérium maximálne zníženie CFI a 0,01 spolu s nárastom RMSEA o 0,015 a SRMR o 0,03 v prípade metrickej invariantnosti a 0,015 v prípade skalárnej invariantnosti a invariantnosti reziduálov. Naproti tomu, Meade et al. (2008) na základe analýz simulovaných dát uvádzajú omnoho striktnejšie kritéria (pokles CFI maximálne o 0,002) a odporúčajú využitie aj necentrálneho indexu zhody NCI. Toto kritérium je však veľmi prísne a v praxi sa najčastejšie používa uvedené kritérium zmeny CFI a RMSEA približne o 0,01 (Putnick & Bornstein, 2016).

Čo v prípade, ak sa zhoda modelu výraznejšie zhorší? Môžeme konštatovať, že sme nedosiahli danú úroveň invariantnosti modelu a analýzu ukončiť pri poslednej dosiahnutej. Druhou možnosťou je skúmať, ktoré parametre modelu spôsobujú najvýraznejšie zhoršenie zhody modelu s dátami a tieto parametre uvoľniť pre porovnávané skupiny. Ak napríklad identifikujeme, že sa faktorový náboj jednej položky výrazne líši medzi niektorými skupinami, nastavíme model metrickej invariantnosti tak, že tento náboj nebude rovnaký medzi skupinami. Pokiaľ takto upravený model nebude výrazne horší od modelu na predchádzajúcej úrovni invariantnosti, môžeme skonštatovať čiastočnú invariantnosť. Práve takéto situácie sú zaujímavé pre zhodnotenie obsahu psychodiagnostickej metodiky – podobne ako pri skúmaní odlišného fungovania položiek sa v tomto momentne môžeme zamyslieť nad tým, prečo to medzi skupinami nefunguje rovnako. Skúmanie čiastočnej invariantnosti je z nášho pohľadu vhodnejšie ako ukončenie analýzy pri výraznejšom zhoršení zhody modelu. Výhodou je, že po

dosiahnutí čiastočnej úrovne invariantnosti, môžeme pokračovať na ďalšiu úroveň so zvyšnými položkami a diskutovať možné príčiny nerovnosti parametrov. Slabou stránkou takéhoto prístupu je určenie počtu parametrov, ktoré môžeme uvoľniť tak, aby sa stále dalo hovoriť o čiastočnej invariantnosti modelu. Toto rozhodnutie zostáva na zvážení výskumníkov. V základe by mali byť parametre väčšiny položiek sýtených jedným faktorom nastavené ako rovnaké v porovnávaných skupinách, aby sme mohli hovoriť o čiastočnej invariantnosti modelu (Vandenberg & Lance, 2000). Dôležité je podotknúť, že toto pravidlo by malo platiť pre každý jeden latentný faktor, ak je ich v modeli viac.

4. Päťfaktorový model osobnosti

Päťfaktorové osobnostné inventáre patria k najčastejšie využívaným v rámci psychodiagnostiky osobnosti, a to najmä vo výskume. V psychológii osobnosti majú pomerne dlhú a zaužívanú tradíciu. Vznik päťfaktorového modelu nezačal snahou o vytvorenie niečoho nového, šlo skôr o snahu zjednotiť existujúce metodiky do zmysluplného a ľahšie komunikovateľného celku, ktorý môže nájsť uplatnenie vo výskumoch a praxi. Psychológovia využívali mnoho metodík pre skúmanie osobnosti, problém bol však v tom, že mnohé metódy s rovnakým názvom skúmali iné oblasti, a podobne naopak, mnohé odlišné metódy boli určené na skúmanie rovnakej oblasti (John et al., 2008).

Za základ päťfaktorového modelu možno považovať lexikálnu hypotézu, ktorá hovorí o tom, že všetky dôležité osobnostné charakteristiky sú prirodzene používané v jazyku. Analýze týchto termínov, resp. adjektív popisujúcich osobnosť človeka, sa venovali napríklad Allport a Odbert (1936), ktorí identifikovali takmer 18000 termínov používaných v anglickom jazyku, ktoré potom kategorizovali. Vytvorili štyri hlavné kategórie adjektív, ktoré popisujú (1) osobnostné črty, (2) dočasné stavy a nálady, (3) spoločenské hodnotenie a poslednou kategóriou (4) – zmiešanou, rozmanitou – boli tie adjektíva, ktoré nie je možné jednoznačne zaradiť do predchádzajúcich kategórií, no týkali sa napríklad fyzických charakteristík, ktoré sú spojené s psychologickými črtami a pod. Takéto množstvo osobnostných charakteristík, či deskriptívnych adjektív neskôr zredukoval Cattell do podoby dnes známeho dotazníka 16PF – 16-faktorového osobnostného dotazníka (Cattell, Eber, Tatsouka in John, Newman, Soto, 2008). Neskôr boli tieto faktory zredukované na menší počet, čím vznikli modely podobné tomu, ktorý je používaný dnes. Dôležitým medzníkom je NEO-PI (angl. NEO Personality Inventory), teda NEO osobnostný inventár, ktorý vznikol po zjednodušení Cattellovho 16PF dotazníka. NEO-PI zachytával 3 faktory, z ktorých je odvodený názov, neuroticizmus, extraverziu, otvorenosť voči skúsenosti, ku ktorým neskôr pribudli faktory prítvornosť a svedomitosť (Costa, McCrae in John, Newman, Soto, 2008). V dnešnej dobe používané dotazníky piatich osobnostných faktorov sú napríklad NEO-FFI (NEO Five Factor Inventory od Costu a McCraeho), ktorý na Slovensku preložili a adaptovali (Ruisel a Halama, 2007) alebo NEO-PI-R (revidovaný NEO osobnostný inventár) používaný v Českej republike od Hřebíčkovéj (2004). V zahraničí je momentálne najnovšia druhá verzia Inventára Veľkej Päťky (angl. Big Five Inventory 2 – BFI-2; Soto & John, 2017), ktorý je adaptovaný aj na Slovensku (Halama et al., 2020).

Napriek rôznemu počtu a obsahu položiek v rôznych verziách týchto dotazníkov, môžeme priblížiť 5 základných faktorov, ktoré tieto dotazníky merajú: neuroticizmus,

extraverziu, otvorenosť voči skúsenosti, prívetivosť a svedomitosť. Neuroticizmus zachytáva tendenciu prežívať negatívne alebo neprijemné emócie, napríklad úzkosť, depresiu alebo hnev. Vyššie skóre v tomto faktore vyjadruje vyššiu častosť prežívania týchto emócií a taktiež emocionálnu labilitu. Naopak jednotlivci, ktorí dosahujú nízke skóre sú emocionálne stabilní, majú pokojnú osobnosť, aj keď môžu byť považovaní za nezaujatých a pod. Extraverzia hovorí o energickosti a spoločenskosti osobnosti. Jednotlivci s nízkym skóre extravenzie sú rezervovaní a vyhľadávajú skôr samotu a pokoj. Vyššia miera extravenzie je spojená s túžbou po spoločnosti a pozornosti okolia. Faktor otvorenosti voči skúsenosti zachytáva tendenciu k vyhľadávaniu určitej nezávislosti osobnosti, vyhľadávaniu a oceňovaniu umenia či rôznych netradičných zážitkov. Jednotlivci s vyššou otvorenosťou voči skúsenosti majú tendenciu k nezávislosti a kreativitě, no na druhú stranu môžu mať nízku toleranciu voči bežným „denným“ povinnostiam a rutine. Prívetivosť zjednodušene vyjadruje priateľskosť osobnosti. Jednotlivci s vyššou mierou prívetivosti sú priateľskí, nápomocní, no môžu byť považovaní za submisívnych. Nízka miera prívetivosti je spojená skôr so súperením či vyhľadávaním interpersonálnych konfliktov. Posledným faktorom je svedomitosť, ktorá zachytáva tendenciu osobnosti k efektívnosti a precíznosti alebo naopak voľnejšiemu prístupu, ktorý možno nazvať určitou bezstarostnosťou. Vyššia svedomitosť je spojená so sebadisciplínou, plánovaním a dodržiavaním plánov, no príliš vysoká miera môže viesť k nízkej flexibilitě. Jednotlivci s nízkou mierou svedomitosti môžu byť spontánni, no negatívnou stránkou je ich nižšia spoľahlivosť (Ruisel, Halama, 2007).

V nasledujúcej kapitole približujeme Inventár Veľkej Päťky 2. V ďalších dvoch častiach uvádzame výsledky výskumov, ktoré sa vzťahom piatich faktorov osobnosti a pohlavia a veku, no upriamujeme pozornosť na výskumy, ktoré skúmali odlišné fungovanie položiek a invariantnosť merania päťfaktorových osobnostných inventárov a dotazníkov.

4.1. Inventár Veľkej Päťky 2

Jednou z najnovších metodík pre skúmanie piatich veľkých faktorov osobnosti je Inventár Veľkej Päťky 2 (Big Five Inventory 2, BFI-2; Soto & John, 2017), v slovenskej verzii od Halamu et al. (2020). Inventár pozostáva zo 60 položiek, ktoré sú výrokmami o sebe. Respondenti vyjadrujú mieru svojho súhlasu alebo nesúhlasu s týmito tvrdeniami prostredníctvom 5-bodovej škály Likertovho typu od „*Veľmi nesúhlasím*“, cez „*Mierne nesúhlasím*“, „*Neutrálne, nemám názor*“, „*Mierne súhlasím*“ až po „*Veľmi súhlasím*“. Inventár má spoločný začiatok „*Som niekto, kto...*“, ktorý pokračuje jednotlivými položkami. Silnou

stránkou tohto inventára je hierarchická štruktúra. Inventár dokáže merať päť veľkých faktorov osobnosti, ktoré sú zložené z troch podfaktorov, nazvaných ako facety. Extraverzia je zložená zo sociability (napr. „*je spoločenský, rád trávi čas s inými ľuďmi*“), asertivity (napr. „*má asertívnu osobnosť, vyjadruje svoje názory*“) a energickej úrovne (napr. „*je málokedy pre niečo zapálený*“). Prívetivosť zahŕňa facety súcitu (napr. „*je súcitný, má dobré srdce*“), zdvorilosti (napr. „*je úctivý, správa sa k iným úctivo*“) a dôvery (napr. „*má v povahe odpúšťať*“). Svedomitosť zahŕňa organizovanosť (napr. „*má sklon byť chaotický*“), produktivitu (napr. „*má sklon byť lenivý*“) a zodpovednosť (napr. „*je spoľahlivý a dôsledný*“). Negatívna emocionalita (známa tiež ako neurotizmus) sa skladá z úzkosti (napr. „*je uvoľnený, dobre zvláda stres*“), depresie (napr. „*ostáva optimistický napriek prekážkam*“) a emocionálnej nestálosti (napr. „*je náladový, máva výkyvy nálady*“). Otvorenosť pozostáva z intelektuálnej zvedavosti (napr. „*je zvedavý na mnoho rozličných vecí*“), estetickej senzitivity (napr. „*je očarený umením, hudbou alebo literatúrou*“) a kreatívnej predstavivosti (napr. „*je vynaliezavý, hľadá šikovné spôsoby ako robiť veci*“). Užívateľ inventára takto môže osobnosť skúmať globálnejšie prostredníctvom domén alebo detailnejšie prostredníctvom facetov. Druhou silnou stránkou je vyrovnaný počet položiek v jednotlivých doménach a facetoch ako aj to, že presne polovica položiek je formulovaná v smere meranej črty a polovica v opačnom (reverzná položka). Takýmto spôsobom je redukovaný bezobsahovo súhlasný a nesúhlasný štýl odpovedania.

4.2. Päť veľkých faktorov osobnosti v súvislosti s pohlavím – odlišné fungovanie položiek a invariantnosť merania

Rozdiely medzi mužmi a ženami sú častým záujmom výskumov z rôznych oblastí psychológie. Tento trend je jasne viditeľný aj pri množstve výskumov, ktoré skúmali rozdiely medzi pohlaviami v piatich veľkých faktoroch osobnosti. Schmitt et al. (2008) skúmali rozdiely medzi pohlaviami v 55 kultúrach z celého sveta, celkovo na vyše ako 17000 respondentoch. Ich zistenia poukazujú na pomerne všeobecný trend. Ženy dosahujú vyššiu mieru extravenzie, prívetivosti, svedomitosti a negatívnej emocionality (neurotizmus). Ruisel a Halama (2007) na slovenskej populácii zistili mierne vyššie hodnoty u žien pri všetkých faktoroch okrem extravenzie.

V posledných rokoch sme sa o rozdiely medzi mužmi zaujímali aj v rámci validizácie slovenskej verzie Inventára Veľkej Päťky 2 (Soto & John, 2017). Výsledky poukázali na štatisticky významne vyššiu mieru všetkých piatich faktorov u žien. V extravenzii sme zistili malý rozdiel. Z hľadiska subškál ženy dosahovali vyššiu mieru sociability a energickosti, no

nelíšili sa v miere asertivity. Rozdiely v prívetivosti boli na úrovni malého až stredne veľkého. Najväčší rozdiel bol v miere súcitu a zdvorilosti, rozdiel v miere dôvery bol na úrovni veľmi malého. Rozdiel vo svedomitosti bol na úrovni malého. Najväčší, stredne veľký rozdiel bol zodpovednosti, rozdiel v produktivite bol o niečo menší a v organizácii len veľmi malý. Rozdiel v miere negatívnej emocionality bol veľmi malý, v prípade depresie a emocionálnej rozladenosti dokonca štatisticky nevýznamný. Stredne veľký rozdiel bol v prípade úzkosti. Rovnako, stredne veľký rozdiel bol zistený v miere otvorenosti, pričom na úrovni subškál sa štatisticky významný rozdiel zistil len pri estetickú senzitivite, ktorý bol na úrovni stredne veľkého.

Naproti tomu v ďalšom výskume zameranom na validáciu krátkych verzií Inventára Veľkej Päťky 2 (Kohút et al., 2020a) sme u žien zistili štatisticky významne vyššie hodnoty len v prípade prívetivosti a negatívnej emocionality. Rozdiely v oboch faktoroch boli len malé. Na úrovni subškál bol štatisticky významný rozdiel zistený pri 11 z 15 subškálach. Veľmi malý rozdiel v smere vyšších hodnôt u mužov bol zistený v asertivite, intelektuálnej zvedavosti a kreatívnej predstavivosti. Na úrovni malých rozdielov boli ženy viac súcitné, zdvorilé, dôverujúce a zodpovedné, a dosahovali vyššiu mieru úzkosti, depresie, emocionálnej nestálosti a estetickú senzitivity.

Podobné výsledky uvádzajú aj Hřebíčková et al. (2020), ktorí v rámci adaptácie a validácie českej verzie tohto inventára informujú o rozdieloch medzi mužmi a ženami. Výskumný súbor tvorili mladí ľudia, konkrétne študenti stredných a vysokých škôl s vekovým rozpätím 15 až 26 rokov. Zistili štatisticky významne vyššiu mieru prívetivosti a negatívnej emocionality, avšak rozdiel v negatívnej emocionalite bol stredne veľký. Pri ostatných faktoroch štatisticky významné rozdiely nezistili. Nad rámec subškál prívetivosti a negatívnej emocionality, uvádzajú štatisticky významne vyššiu asertivitu a intelektuálnu zvedavosť u mužov, a zodpovednosť a estetickú senzitivitu u žien. Z hľadiska veľkosti sú tieto rozdiely veľmi malé až malé.

Originálni autori Inventára Veľkej Päťky 2, Soto a John (2020) zistili veľmi malé rozdiely v extravertzii a svedomitosti, malé rozdiely v prívetivosti a negatívnej emocionalite. Otvorenosť sa štatisticky významne nelíšila. Na úrovni subškál bola situácia rôznorodejšia, no stále na úrovni veľmi malých až malých rozdielov. Muži a ženy sa štatisticky významne nelíšili v asertivite, depresii a kreatívnej predstavivosti. Zaujímavým zistením je, že hoci sa ženy a muži nelíšili v miere otvorenosti, na úrovni subškál zistili mierne vyššiu intelektuálnu zvedavosť u mužov a mierne vyššiu estetickú zvedavosť u žien. Uvedené zistenia naznačujú pomerne stabilný trend v medzipohlavných rozdieloch v prívetivosti a negatívnej emocionalite,

ale variabilitu vzhľadom na rozdiely v ostatných faktoroch. Zjavná je tiež potreba diferencovať jednotlivé subškály, keďže povaha rozdielov môže byť diametrálne odlišná.

Vzhľadom na relatívnu novosť Inventára Veľkej Päťky 2 a najmä jeho adaptácie v iných krajinách, sme zatiaľ nenašli výskumy, ktoré by sa zaoberali odlišným fungovaním položiek tohto inventára, no už sú dostupné zistenia dvoch štúdií, ktoré sa zaoberajú jeho invariantnosťou merania medzi mužmi a ženami. Shchebetenko et al. (2020) sa v rámci validizácie ruskej verzie BFI-2 zamerali aj na invariantnosť tohto inventára medzi mužmi a ženami. Pre overenie faktorovej štruktúry využili štatistickú metódu exploračného štruktúrneho modelovania (Asparouhov & Muthén, 2009). Ich model bol zložený z piatich korelovaných latentných faktorov. Každý faktor primárne sýtil príslušných 12 položiek, no bolo umožnené, aby sýtil aj položky z iných faktorov. Pri položkách, ktoré sú koncepčne v rámci jednej subškály, boli povolené korelácie reziduálov. Navyše, pre identifikovanie a začlenenie takzvaného bezobsahového súhlasného odpovedového štýlu (angl. *acquiescence*, napr. Havan et al., 2022) pridali aj druhý faktor, ktorý nutne pozitívne sýtil všetky položky a nekoreloval so žiadnym iným faktorom. Takýto model je teoreticky podložený a v rámci ich analýz vykazoval najlepšiu zhodu s dátami. Následne, postupným špecifikovaním modelu dospeli k záveru, že ruská verzia BFI-2 je invariantná až na striktnnej úrovni, to jest rovnosti reziduálov položiek. Na základe tohto zistenia konštatujú, že rozdiely medzi mužmi a ženami vo faktoroch reflektujú reálne rozdiely v osobnostných faktoroch a nie sú artefaktom merania. Toto zistenie sa potvrdilo aj pri druhých dátach v ich výskume. Autori uvádzajú, že ženy majú výrazne vyššiu negatívnu emocionalitu (najmä úzkosť), stredne vyššiu prívetivosť, mierne nižšiu extraverziu (spôsobené nižšou asertivitou), nelíšili sa vo svedomitosti a rozdiel v otvorenosti bol len úrovni subškál, pričom mali vyššiu estetickú senzitivitu, no nižšiu kreatívnu predstavivosť a intelektuálnu zvedavosť (Shchebetenko et al., 2020).

Gallardo-Pujol et al., (v tlači) vo validačnej štúdii španielskej verzie BFI-2 zistili, že ženy dosahujú štatisticky vyššiu úroveň extravenzie (sociability, energickosti, nie však asertivity), prívetivosti a jej subškál, svedomitosti (produktivity, zodpovednosti, nie však organizovanosti), negatívnej emotionality (najmä úzkosti, emocionálnej nestability, nie však depresie). Tieto rozdiely sú na úrovni veľmi malých až malých. Muži a ženy sa nelíšili v otvorenosti, no ženy dosahovali mierne vyššie hodnoty estetickej senzitivity a muži kreatívnej predstavivosti. Autori sa v rámci štúdie zaujímali o invariantnosť merania, pričom využívajú alternatívny model konfirmačnej faktorovej analýzy. Pre každý faktor BFI-2 vytvorili samostatný model troch korelovaných subškál, ktoré sýtili príslušné položky. Rovnako ako v predchádzajúcom prípade, tento model obsahoval aj faktor bezobsahového súhlasného

odpovedového štýlu. Z výsledkov vyplynulo, že všetkých päť faktorov má invariantnú štruktúru a faktorové náboje, no len svedomitosť, negatívna emocionalita a otvorenosť dosiahli skalárnu invariantnosť. Striktná úroveň invariantnosti nebola autormi skúmaná. Autori konštatujú, že výsledky porovnávania hodnôt extravenzie a prívetivosti u mužov a žien môžu byť skreslené. Hoci je to vzhľadom na povahu štúdie pochopiteľné, je škoda, že autori nepokračovali k stanoveniu čiastočnej skalárnej invariantnosti, vďaka čomu by mohli byť identifikované položky, a teda aj obsahy, ktoré najviac prispievajú k nerovnosti merania.

Na Slovensku sme skúmali odlišné fungovanie položiek päťfaktorových inventárov v našej dizertačnej práci, avšak vo vedeckom časopise boli doposiaľ publikované len výsledky analýzy NEO päťfaktorového osobnostného inventára NEO-FFI (Kohút & Halama, 2019). V tomto výskume sme identifikovali viacero položiek, ktoré fungovali odlišne, avšak len časť z nich bola identifikovaná v zhode dvoch metód analýzy DIF a celkovo 6 (10 %), ktoré boli identifikované všetkými tromi použitými metódami. Dve z týchto položiek sú zamerané na meranie neurotizmu a ich obsah sa týkal pociťovania strachu a ťažkostí so zvládaním strachu. Pri oboch položkách dosahovali ženy vyššiu hodnotu pri rovnakej miere črty v porovnaní s mužmi. Ďalšie dve sú z faktoru otvorenosti a zisťujú záujem o cudzokrajné jedlá a otvorenosť voči abstraktným myšlienkam. Pri týchto položkách boli zvýhodnení muži, to jest, pri rovnakej miere otvorenosti bola priemerná odpoveď na tieto položky vyššia u mužov ako u žien. Posledné dve položky merali svedomitosť. Obsah prvej sa týkal dosahovania stanovených cieľov a táto položka zvýhodňovala mužov. Druhá bola zameraná na mieru, v akej sa možno na osobu spoľahnúť a zvýhodňovala ženy. Možnou príčinou vzniknutého odlišného fungovania niektorých identifikovaných položiek je široký záber faktorov NEO-FFI. Tento inventár vznikol výberom položiek z rôznych subškál revidovaného NEO osobnostného inventára, v ktorých sa môžu muži a ženy líšiť, no tento rozdiel nie je výrazný pri celom faktore. Efekt však môže zohrávať aj reálny rozdiel v určitom konkrétnom správaní, či spoločenská žiadúcnosť určitého správania mužov a žien, ktoré položka zachytáva (Kohút & Halama, 2019).

Gogniat et al. (2022) skúmali medzipohlavné odlišné fungovanie položiek v NEO-FFI u ľudí v neskorej dospelosti. Pre analýzu využili ordinálnu logistickú regresiu s využitím IRT a Monte Carlo simuláciami. Identifikovali viacero položiek, ktoré vykazovali štatisticky významné DIF, avšak len 2 z nich malo mieru efektu výrazne vyššiu ako pri simulovaných súboroch bez DIF. Autori na základe svojich zistení konštatujú, že NEO-FFI je invariantné u starších mužov a žien. Limitáciou je však pomerne nízky počet respondentov, ako aj to, že nie každý respondent odpovedal na všetky položky, kvôli čomu boli niektoré položky vyradené z analýzy. Tento fakt výrazne limituje zovšeobecnenie zistení.

Smith (2002) vo svojom článku predstavuje dôležitosť skúmanie odlišného fungovania položiek, ktorú ilustruje na doméne otvorenosť v prvej verzii Inventára Veľkej Päťky. Podľa stanovených kritérií identifikovala 1 položku, pri ktorej odpovedali muži v smere vyšších hodnôt. Táto položka sa týka rutínnej práce a je reverzná, to znamená, že muži menej preferujú takúto prácu. Ženy odpovedali súhlasnejšie pri 3 položkách, týkajúcich sa hrania s myšlienkami, hudby, umenia a literatúry a estetických zážitkov. Autorka identifikované odlišné fungovanie položiek interpretuje v zmysle prirodzených rozdielov, pričom poukazuje na dôležitosť skúmania odlišného fungovania položiek najmä v spojitosti s rozšírením našich znalostí ohľadom meraného konštruktu.

Táto monografia rozširuje doterajšie zistenia ohľadom rovnosti merania piatich faktorov osobnosti medzi mužmi a ženami prostredníctvom skúmania odlišného fungovania položiek a invariantnosti merania na dvoch retestových súboroch zo všeobecnej dospeléj populácie, ako aj na súbore mladých dospelých. Vďaka použitiu viacerých súborov je umožnená komparácia a sumarizácia výsledkov v snahe identifikovať položky, ktoré stabilne vykazujú odlišné fungovanie medzi mužmi a ženami. Päť faktorov osobnosti bolo vo všetkých súboroch meraných prostredníctvom Inventára Veľkej Päťky 2.

4.3. Päť veľkých faktorov osobnosti v súvislosti s vekom – odlišné fungovanie položiek a invariantnosť merania

Podobne ako pri pohlaví, aj súvislosť piatich veľkých faktorov osobnosti s vekom je častou skúmanou problematikou. Zaujímavý výskum uskutočnili napríklad Soto et al. (2011), ktorí skúmali rozdiely v osobnostných faktoroch u ľudí vo veku od 10 do 65 rokov. Ich výskum síce nie je longitudinálny ale prierezový, no na veľmi veľkom súbore vyše jeden a štvrt' milióna respondentov. Pre meranie osobnosti použili prvú verziu Inventára Veľkej Päťky. Z hľadiska svedomitosti zistili najnižšiu úroveň v období adolescencie, ktorá sa následne so zvyšujúcim vekom zvyšovala v období rannej dospelosti a tento trend pokračoval aj do neskorej dospelosti aj keď len v miernejšom tempe. Podobný trend zistili aj pri prívetivosti, ten však nebol tak výrazný. Z hľadiska neurotizmu sa vývinový trend pomerne líšil u mužov a žien. U mužov mal miernu tendenciu k poklesu s narastajúcim vekom, avšak rozdiel medzi rannou a neskorou dospelosťou bol pomerne malý. U žien bola situácia výrazne odlišná – ženy dosahovali všeobecne vyššiu mieru neurotizmu a to najmä v období adolescencie, mladej a strednej dospelosti. Vo vyššom veku sa ich miera neurotizmu približovala k hodnotám mužov, čo sa prejavilo na trende k poklesu tejto črty s pribúdajúcim vekom. Extraverzia vykazovala pomerne

stabilný trend vynímajúc prechod z detského veku do adolescencie, kde nastal pomerne rázny pokles. Otvorenosť prechodom z neskorého detského veku do adolescencie klesá, avšak následne narastá v období vynárajúcej sa (mladej) dospelosti a tento trend v miernejšej podobe pokračuje aj v období strednej a neskoršej dospelosti.

Zistenie zo Slovenska prinášajú napríklad Ruisel a Halama (2007), ktorí zistili štatisticky významné vzťahy medzi vekom a jednotlivými faktormi. Neurotizmus má veľmi slabý trend vekom klesať, podobne tak aj otvorenosť či extravergia, pri ktorej je však tento trend o niečo výraznejší. Pri prívetivosti autori uvádzajú veľmi slabý pozitívny vzťah s vekom, ktorý však zistili len u mužov. Svedomitosť má tendenciu mierne narastať s vekom a to najmä u žien. Halama et al. (2020) zistili veľmi slabý pozitívny vzťah medzi vekom a extraverziou, pričom je tento vzťah podobný aj pri subškálach sociability a asertivity avšak prakticky absentuje pri miere energickosti. Prívetivosť s vekom súvisí mierne pozitívne a tento vzťah je konzistentný aj na úrovni subškál. Podobná situácia je tiež v prípade svedomitosti. Negatívna emocionalita má slabý trend vekom klesať, najvýraznejšie (avšak stále len veľmi slabo) v prípade emocionálnej nestability. Otvorenosť vekom mierne narastá, no tento trend je na úrovni subškál prítomný len v prípade estetickej senzitivity, pri kreativite či intelektuálnej zvedavosti je zanedbateľný. Podobné zistenia prináša aj výskum Kohúta et al. (2020a), no zistené trendy sú vo všeobecnosti ešte slabšie prípadne neprítomné.

Podobne, ako pri porovnávaní mužov a žien, je aj pri skúmaní súvislosti medzi piatimi faktormi osobnosti a vekom potrebné brať v úvahu rovnosť merania. Otázkou, či sú zistené rozdiely medzi mladšími a staršími dospelými len artefaktom merania alebo skutočným trendom sa zaoberali Nye et al. (2016). Vo svojej štúdii pristupujú k ekvivalencii merania prostredníctvom analýzy založenej na teórii odpovede na položku a konfirmačnej faktorovej analýzy. Analýzou dospeli k niekoľkým položkám, ktoré preukázali nerovnosť merania medzi mladšími a staršími ľuďmi. Odlišné fungovanie týchto položiek je však len slabé a dopad odlišne fungujúcich položiek na zistenú povahu rozdielov medzi mladšími a staršími je zanedbateľný. Na základe týchto zistení autori konštatujú, že zistená vyššia svedomitosť a nižší neurotizmus u starších dospelých je spôsobený reálnymi rozdielmi a nie artefaktom merania v podobe odlišného fungovania položiek medzi skupinami.

Cieľom tejto monografie je prispieť k poznaniu v tejto oblasti prostredníctvom skúmania odlišného fungovania položiek slovenskej verzie Inventára Veľkej Päťky 2 medzi mladými a staršími dospelými.

5. Výskumné súbory a analýza dát

V tomto výskume pracujeme s viacerými výskumnými súbormi. Pre skúmanie test-retest stability výsledkov analýzy odlišného fungovania položiek a invariantnosti merania používame dva longitudinálne súbory.

Prvý longitudinálny súbor pochádza zo štúdie Kohúta et al. (2020b). Prvý zber bol uskutočnený v novembri 2018 pre účely štúdie Kohút et al. (2020a) a po pol roku boli respondenti opätovne prizvaní k účasti vo výskume. Tento longitudinálny súbor pozostáva zo 414 respondentov, ktorí boli v čase prvého zberu vo veku od 18 do 75 rokov ($M = 46,23$; $SD = 14,36$). Zastúpenie mužov ($N = 239$; 57,7%) bolo mierne vyššie ako žien ($N = 175$; 42,3%), no je pomerne vyrovnané. Priemerný vek mužov bol 45,03 roku ($SD = 13,87$) a žien bol 47,86 ($SD = 14,89$).

Druhý longitudinálny súbor vychádza z toho istého prvého zberu ako vyššie menovaný, no druhý zber bol uskutočnený samostatne po necelých dvoch rokoch v septembri 2020. Súbor pozostáva zo 443 respondentov, pri prvom zbere vo veku od 18 po 76 rokov ($M = 44,88$; $SD = 14,95$). Zastúpenie mužov ($N = 213$; 48,1%) a žien ($N = 230$; 51,9%) je vyrovnané. Priemerný vek mužov bol 44,89 roku ($SD = 13,98$) a žien bol 44,86 ($SD = 15,82$).

Pre skúmanie odlišného fungovania položiek a invariantnosti merania u mužov a žien v období vynárajúcej sa dospelosti používame súbor pozostávajúci z 600 respondentov vo veku medzi 18 až 29 rokmi ($M = 25,20$; $SD = 2,96$). Zastúpenie mužov ($N = 300$) a žien ($N = 300$) je vyrovnané. Priemerný vek mužov bol 25,22 roku ($SD = 3,15$) a žien bol 25,18 ($SD = 2,76$). Tento súbor mladých dospelých sme následne spojili so súborom z výskumu Kohúta et al. (2020a), z ktorého sme najskôr odstránili respondentov vo veku do 30 rokov, tak aby nebol možný duplicitný záznam, to znamená, že respondenti sú unikátni. Výsledný súbor tvorilo 1260 respondentov (51% mužov), 600 vo veku vynárajúcej sa dospelosti a 660 starších dospelých.

Všetky spomenuté zbery prebiehali prostredníctvom online dotazníka. Respondenti boli k výskumu prizvaní výskumnými agentúrami a boli za svoju účasť odmenení. Všetci respondenti boli informovaní o účeloch výskumu, súhlasili s účasťou ako aj využitím získaných dát pre ďalšie výskumné účely. Pri každom zbere boli v snahe zaručiť vyššiu kvalitu dát použité kontrolné položky na pozornosť. Pokiaľ respondent nesprávne zodpovedal na kontrolnú položku, bola jeho účasť vo výskume ukončená a jeho odpovede neboli použité.

Pre analýzu dát sme využili program Rstudio (RStudio Team, 2020). V ďalších dvoch častiach približujeme analýzy použité v rámci nášho výskumu.

5.1. Analýza odlišného fungovania položiek BFI-2

Odlišné fungovanie položiek sme analyzovali prostredníctvom R balíka *lordif* (Choi et al., 2011). Tento balík pre analýzu odlišného fungovania položiek využíva princíp ordinálnej logistickej regresie. Ako sme už v kapitole 1.3. uviedli, pri ordinálnej logistickej regresii predikujeme odpoveď na položku prostredníctvom miery vlastnosti (napr. sumárne skóre), premennej vyjadrujúcej príslušnosť k skupine (napr. pohlavie – muž/žena) a interakciou týchto dvoch premenných. Každá premenná je pridaná v samostatnom modeli, čo umožňuje porovnať, ktorý model vysvetľuje najviac variácie a tak identifikovať, či ide o uniformné alebo neuniformné odlišné fungovanie položiek. Výhodou balíka *lordif* je, že namiesto použitia hrubého skóre využíva IRT odhad miery vlastnosti. Ten je navyše doplnený o iterácie, vďaka čomu je tento odhad miery prepočítavaný a očisťovaný od položiek, ktoré sú v niektorom z opakovaní prepočtu označené ako odlišne fungujúce. Pokiaľ sú výsledky v dvoch za sebou idúcich iteráciách rovnaké, to jest, rovnaké položky sú identifikované ako vykazujúce odlišné fungovanie, analýza je ukončená. Slabou stránkou akýchkoľvek testov založených na testovaní štatistickej významnosti nulovej hypotézy je, že v prípade veľkého výskumného súboru môžeme získať štatisticky významný výsledok aj pri prakticky zanedbateľnom jave. Rovnako tak aj v prípade testovania odlišného fungovania položiek môžeme zistiť štatisticky významný výsledok napriek tomu, že odlišné fungovanie konkrétnej položky je prakticky zanedbateľné a nemá dopad na celkové skóre, či v praktickej rovine, dopad na validitu psychodiagnostickej metodiky alebo spravodlivosť testovania. Zhodnotenie závažnosti odlišného fungovania položky je pomerne komplikované z dôvodu rôznorodých indikátorov, ktoré nemožno jednoducho intuitívne interpretovať ako napríklad Cohenovo *d*. Balík *lordif* však ponúka dva spôsoby, ako možno posúdiť závažnosť odlišného fungovania položky.

Prvým z nich je grafické znázornenie charakteristickej krivky položky samostatne pre porovnávané skupiny. Takýmto spôsobom možno posúdiť, ako výrazne sa skupiny líšia pri odpovedi na položku pri rovnakej miere vlastnosti. Porovnať možno aj charakteristickú krivku testu, vďaka čomu možno posúdiť, ako sa líši zistené skóre medzi skupinami pri rovnakej odhadovanej miere vlastnosti. Hoci nezískame konkrétne hodnoty, vizuálne zhodnotenie je ľahko pochopiteľné a napomáha zhodnoteniu závažnosti odlišného fungovania.

Druhý spôsob je založený na porovnaní výsledkov analýzy reálnych dát s výsledkami analýzy simulovaných dát. Balík *lordif* umožňuje Monte-Carlo simuláciu dát, ktoré sú pomerovo podobné našim reálnym dátam avšak neobsahujú odlišne fungujúce položky. Automaticky je simulované ľubovoľne veľké množstvo takýchto súborov (napr. 1000) a pri

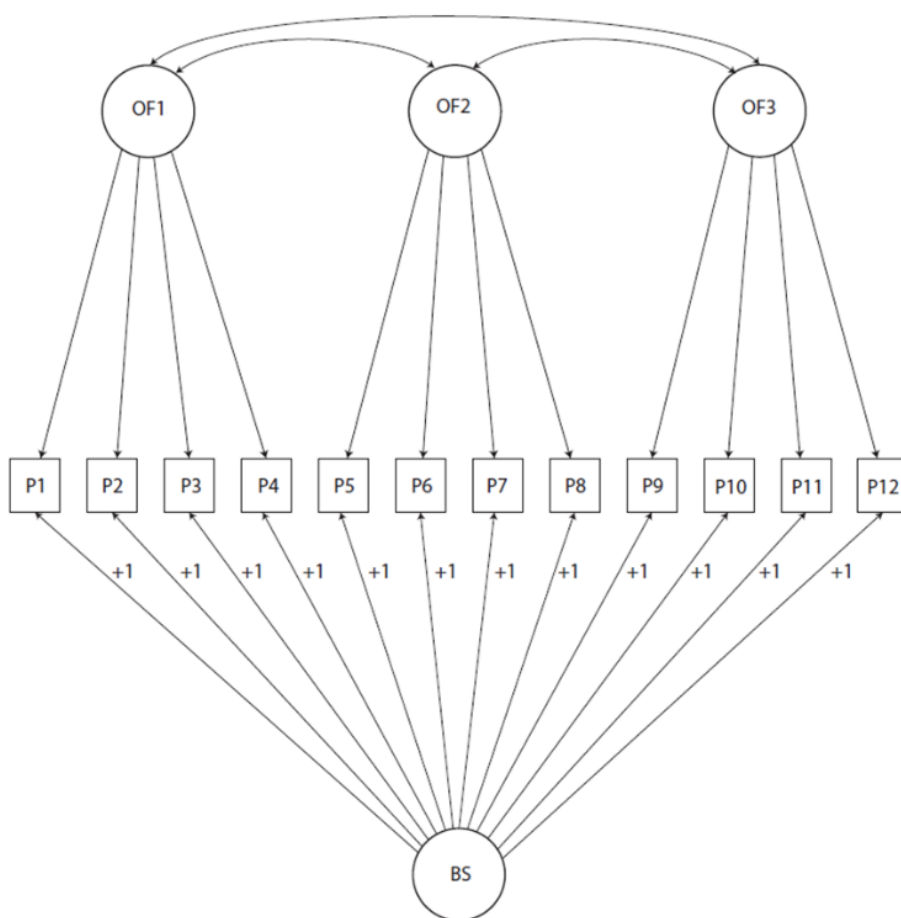
každom súbore je uskutočnená analýza odlišného fungovania položiek. Výsledky týchto analýz – štatistická významnosť alebo pseudo- R^2 koeficienty, nám ponúkajú informáciu o pomeroch v dátach neobsahujúcich odlišne fungujúce položky. Takýmto spôsobom získame distribúciu výsledkov analýz. Vďaka tomu si môžeme zvoliť kritickú hodnotu, podľa ktorej hodnotíme významnosť odlišného fungovania položky v našom súbore. Pri štandardnom prístupe by sme si ako kritérium pre identifikáciu odlišného fungovania mohli zvoliť konvenčne zaužívané štatistickú významnosť pri hladine alfa je rovné 0,05. Ak by sme pri niektorej položke zistili nižšiu hodnotu signifikancie (napr. 0,02), prišli by sme k záveru, že položka vykazuje štatisticky významné odlišné fungovanie. Pri simulovaných dátach si môžeme zvoliť kritérium, napríklad 1% najextrémnejších hodnôt a zistiť kritickú hodnotu. Môžeme napríklad zistiť, že 99. percentil štatistickej významnosti simulovaných analýz je 0,01 – túto hodnotu môžeme použiť pri posudzovaní odlišného fungovania v našich dátach a tak položka, ktorú by sme pri konvenčnej hladine alfa označili ako odlišne fungujúcu, bude pri tomto kritériu štatisticky nevýznamná. Pre posúdenie však nemusíme využiť len štatistickú významnosť ale rovnakým princípom aj pseudo- R^2 koeficienty.

V našom výskume sme pri analýze odlišného fungovania položiek využili nominálnu hodnotu $\alpha = 0,01$. Pri položkách Identifikované položky sme následne porovnali voči hodnote 99. percentilu McFadden pseudo- R^2 zo simulovaných dát pri využití 1000 replikácií. Odlišné fungovanie položiek sme posudzovali z užšieho hľadiska faciet, to jest, do analýz vstupovali len štyri položky. Každá hodnota Likertovej škály položky musela byť zastúpená v dátach aspoň päťkrát (aspoň päť respondentov muselo odpovedať danou hodnotou), inak bola zlúčená s ďalšou hodnotou.

5.2. Analýza invariantnosti merania BFI-2

Invariantnosť merania sme skúmali prostredníctvom konfirmačnej faktorovej analýzy v R balíku *lavaan* (Rosseel, 2012). Invariantnosť merania sme skúmali v osobnostných doménach s použitím modelu troch korelovaných latentných faktorov (vyjadrujúcich facety danej domény) a faktora, ktorý vyjadruje bezobsahový súhlas (*acquiescence*). Tento faktor nekoreluje s ostatnými a všetky položky sýti nutne pozitívne. Tento model je odporúčaný pre faktorovú analýzu BFI-2 (Halama et al., 2020; Soto & John, 2017). Príklad modelu znázorňujeme v Obrázku 3. Pre výpočet sme použili robustný odhad maximálnej vierohodnosti (angl. *robust maximum likelihood - MLR*).

V rámci analýz sme sa zaujímali o konfiguračnú, metrickú a skalárnu invariantnosť. Ak model dosahoval aspoň prijateľné hodnoty alternatívnych indexov zhody modelu s dátami ($CFI > 0,80$ a $RMSEA < 0,10$), prijali sme konfiguračnú úroveň invariantnosti. Následne sme nastavili rovnosť faktorových nábojov v oboch skupinách. Ďalej sme pokračovali nastavením rovnosti úrovňových konštánt (*intercept*). Pokiaľ bol medzi jednotlivými modelmi pokles CFI maximálne o 0,10 a zvýšenie RMSEA maximálne o 0,015, považovali sme úroveň invariantnosti za potvrdenú. Pri posúdení sme štatistickú významnosť chí-kvadrát testu nebrali v úvahu. Pokiaľ sa určitá úroveň invariantnosti nepotvrdila, pokračovali sme ďalej identifikáciou problematickej položky tým, že sme postupne uvoľňovali jednotlivé parametre a prepočítavali upravený model. Zistili sme, ako uvoľnenie jednotlivých položiek zmení hodnotu CFI a RMSEA, a testovali sme čiastočnú invariantnosť tým, že sme uvoľnili najviac problematickú položku. Pokiaľ takto upravený model vykazoval dostatočnú hodnotu CFI a RMSEA pre prijatie invariantnosti, analýzu sme ukončili, ak nie, pokračovali sme ďalej uvoľnením ďalšej položky.



Obrázok 3 Latentný model použitý pri analýze invariantnosti merania (OF – osobnostná faceta, BS – bezobsahový súhlas)

6. Invariantnosť merania a odlišné fungovanie položiek BFI-2 medzi mužmi a ženami

V nasledujúcich častiach uvádzame výsledky analýz zameraných na skupinu mužov a žien. Vzhľadom na použitie viacerých výskumných súborov sú výsledky rozdelené do samostatných častí na základe domén BFI-2.

6.1. Extraverzia

Výsledky konfirmačnej faktorovej analýzy ukázali na všeobecne dobrú zhodu modelu s dátami. Vo všetkých súboroch bola dosiahnutá konfiguračná a metrická úroveň invariantnosti. Extraverzia, resp. jej tri facety sa výrazne nelíšia vo faktorovej štruktúre alebo faktorových nábojoch u mužov a žien. Podobne je tomu aj v prípade úrovňových konštánt s výnimkou druhého zberu prvého retestového súboru. V tomto prípade sa pri zafixovaní úrovňových konštánt u mužov a žien zhoršila zhoda modelu s dátami, výraznejšie najmä CFI. Ďalším skúmaním sme zistili, že najproblematickejšou položkou je položka číslo 56 z facety Energická úroveň: „*Som niekto, kto prejavuje veľa nadšenia.*“ Pri tejto položke majú ženy vyššiu hodnotu úrovňovej konštanty (3,57) v porovnaní s mužmi (3,28), teda napriek rovnakej miere energickosti odpovedajú súhlasnejšie. Uvoľnenie rovnosti úrovňovej konštanty pri tejto položke však nezlepšilo hodnotu CFI na uspokojivú hodnotu, preto sme pokračovali uvoľnením položky číslo 21 z facety Asertivita: „*Som niekto, kto je dominantný, spáva sa ako vodca.*“ Pri tejto položke je hodnota úrovňovej konštanty vyššia u mužov (2,86) v porovnaní so ženami (2,50), teda muži s týmto vyjadrením súhlasia viac ako ženy pri rovnakej miere asertivity. Uvoľnením tejto položky sme dosiahli zníženie CFI oproti metrickej úrovni len o 0,006 a teda sme potvrdili parciálnu skalárnu invariantnosť. Hodnoty CFI a RMSEA a ich zmenu pri rôznych úrovniach invariantnosti uvádzame v Tabuľke 1.

Tabuľka 1 Výsledky konfirmačnej faktorovej analýzy pre doménu extravertia

Úroveň invariantnosti	90% KI					
	CFI	RMSEA	Spodný	Horný	Δ CFI	Δ RMSEA
Retestový súbor 1 - prvý zber						
Konfiguračná	0,933	0,055	0,037	0,072		
Metrická	0,931	0,053	0,036	0,070	0,001	0,002
Skalárna	0,925	0,054	0,037	0,069	0,007	-0,001
Retestový súbor 1 - druhý zber						
Konfiguračná	0,952	0,044	0,022	0,062		
Metrická	0,953	0,042	0,019	0,060	-0,001	0,002

Tabuľka 1 Výsledky konfirmačnej faktorovej analýzy pre doménu extravérzia

Úroveň invariantnosti	CFI	RMSEA	90% KI		Δ CFI	Δ RMSEA
			Spodný	Horný		
Skalárna	0,927	0,050	0,033	0,066	0,027	-0,009
Parciálna skalárna	0,947	0,043	0,022	0,060	0,006	-0,001
Retestový súbor 2 - prvý zber						
Konfiguračná	0,955	0,045	0,025	0,062		
Metrická	0,955	0,044	0,024	0,060	0,001	0,002
Skalárna	0,950	0,044	0,026	0,060	0,005	-0,001
Retestový súbor 2 - druhý zber						
Konfiguračná	0,930	0,060	0,045	0,076		
Metrická	0,927	0,059	0,044	0,074	0,003	0,001
Skalárna	0,929	0,056	0,041	0,071	-0,002	0,003
Súbor mladých dospelých						
Konfiguračná	0,933	0,065	0,053	0,077		
Metrická	0,932	0,063	0,051	0,074	0,001	0,002
Skalárna	0,923	0,065	0,053	0,076	0,009	-0,002

Poznámka, KI - konfidenčný interval. Zmena CFI a RMSEA je počítaná odčítaním od nižšej úrovne invariantnosti. Hodnoty CFI a RMSEA indikujúce výrazné zhoršenie zhody modelu s dátami sú zvýraznené hrubým,

Výsledky analýzy odlišného fungovania položiek priniesli nesúrodé zistenia, podobne ako výsledky konfirmačnej faktorovej analýzy. Výsledky z prvého retestového súboru uvádzame v Tabuľke 2. Identifikovali sme jednu položku zo Sociability a jednu položku z Energetickej úrovne, ktoré vykazovali odlišné fungovanie. Obe položky vykazovali DIF pri hladine významnosti menšej ako 0,05 avšak len v čase 2 pri nami zvolenej úrovni 0,01. Hodnoty zmeny pseudo- R^2 boli pri oboch položkách v čase 2 vyššie ako pri simulovaných dátach bez odlišného fungovania, čo naznačuje prítomnosť významného DIF nad rámec posúdenia štatistickej významnosti, no na druhú stranu, iba položka 56 vykazuje zmenu vyššiu ako 0,02, ktorú možno považovať za určitú hranicu medzi praktickou zanedbateľnosťou a miernou významnosťou (Cohen in Choi et al., 2011).

Tabuľka 2 Výsledky analýzy DIF vo facetách extravérzie v prvom retestovom súbore

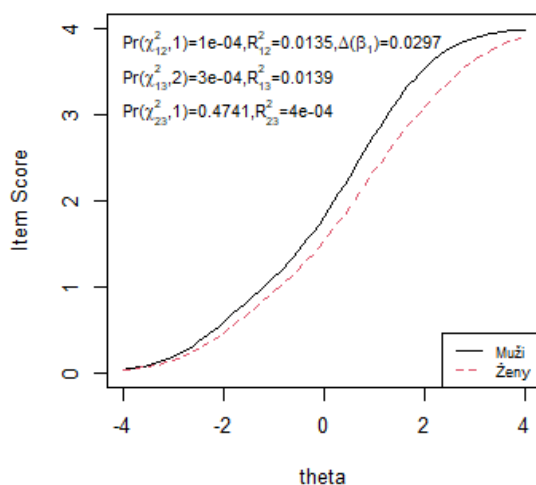
Položka	Retest súbor 1 zber 1			Retest súbor 1 zber 2		
	Δ McFadden pseudo- R^2			Δ McFadden pseudo- R^2		
	U	NU	C	U	NU	C
Sociabilita						
1 je spoločenský, rád trávi čas s inými ľuďmi,	0,000	0,000	0,000	0,001	0,004	0,004
16R má sklon byť tichý,	0,001	0,000	0,001	0,001	0,001	0,002
31R je niekedy hanblivý, uzavretý,	0,004	0,001	0,005	0,014	0,000	0,014
46 je zhovorčivý,	0,001	0,002	0,003	0,001	0,001	0,002
Asertivita						
6 má asertívnu osobnosť, vyjadruje svoje názory,	0,001	0,002	0,004	0,003	0,002	0,005

Tabuľka 2 Výsledky analýzy DIF vo facetách extravenzie v prvom retestovom súbore

Položka	Retest súbor 1 zber 1			Retest súbor 1 zber 2		
	Δ McFadden	pseudo- R^2		Δ McFadden	pseudo- R^2	
	U	NU	C	U	NU	C
21 je dominantný, spáva sa ako vodca,	0,002	0,000	0,002	0,002	0,003	0,005
36R pokladá za ťažké ovplyvňovať ľudí,	0,002	0,002	0,004	0,000	0,000	0,000
51R preferuje, aby iní robili rozhodnutia,	0,000	0,001	0,001	0,001	0,000	0,001
Energická úroveň						
11R je málokedy pre niečo zapálený,	0,000	0,000	0,000	0,002	0,000	0,002
26R je menej aktívny než ostatní ľudia,	0,000	0,003	0,003	0,001	0,000	0,001
41 je plný energie,	0,002	0,000	0,002	0,000	0,004	0,004
56 prejavuje veľa nadšenia,	0,003	0,004	0,007	0,028	0,011	0,039

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

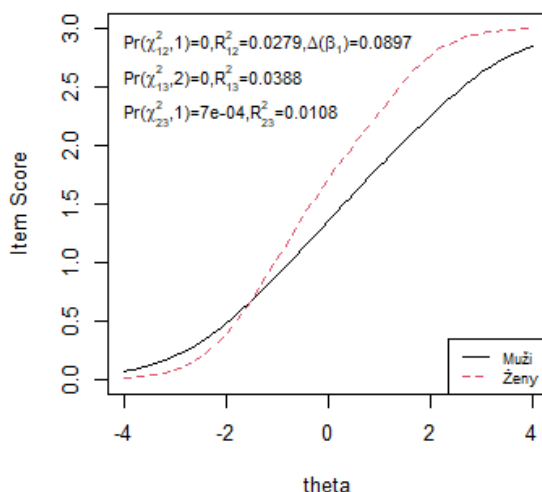
Pri položke 31 „*Som niekto, kto je niekedy hanblivý, uzavretý.*“ dosahujú muži vyššie hodnoty pri rovnakej miere sociability v porovnaní so ženami. Ide však o reverzne skórovanú položku, takže v tomto znení, *ženy vyjadrujú vyšší súhlas s touto položkou*, teda s tým, že sú niekedy hanblivé a uzavreté. Obrázok 3 znázorňuje charakteristickú krivku položky 31 samostatne pre mužov a ženy (položka je reverzná). Rozdiel je pomerne malý, badateľný najmä pri priemernej až vyššej miere sociability.



Obrázok 3 Charakteristické krivky položky 31 (retest súbor 1, zber 2)

Pri položke 56 „*Som niekto, kto prejavuje veľa nadšenia.*“ sme podobne ako pri CFA identifikovali uniformné odlišné fungovanie položiek v smere vyššieho súhlasu u mužov. V prípade tejto položky je však štatisticky významné aj neuniformné odlišné fungovanie, ktoré je však pomerne zanedbateľné. Charakteristické krivky položky znázorňujeme v Obrázku 4.

Ako je možné vidieť, pri veľmi nízkej miere energetickej úrovne, muži vyjadrujú mierne vyšší súhlas s položkou, avšak už pri podpriemernej úrovni sa smer otočil a vyšší súhlas vyjadrujú ženy. Odpovede boli pri tejto položke zlúčené do 4 kategórií.



Obrázok 4 Charakteristické krivky položky 56 (retest súbor 1, zber 2)

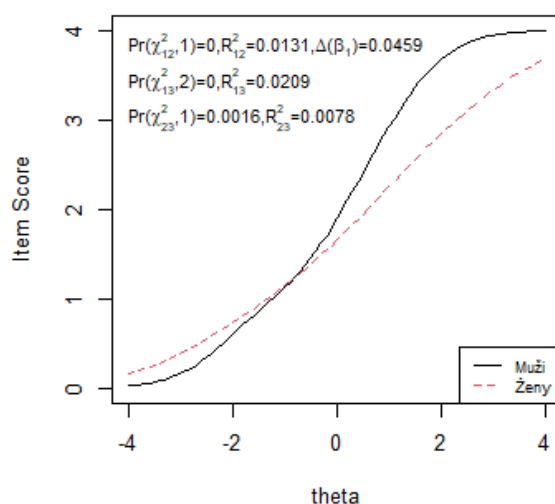
Výsledky z druhého retestového súboru poukázali na štatisticky významné uniformné aj neuniformné odlišné fungovanie položky 31, ktoré je však prítomné len v čase 1. Výsledky uvádzame v Tabuľke 3.

Tabuľka 3 Výsledky analýzy DIF vo facetách extravenzie v druhom retestovom súbore

Položka	Retest súbor 2 zber 1			Retest súbor 2 zber 2		
	Δ McFadden	pseudo- R^2		Δ McFadden	pseudo- R^2	
	U	NU	C	U	NU	C
Sociabilita						
1 je spoločenský, rád trávi čas s inými ľuďmi,	0,001	0,000	0,001	0,000	0,003	0,003
16R má sklon byť tichý,	0,001	0,000	0,001	0,000	0,000	0,000
31R je niekedy hanblivý, uzavretý,	0,013	0,008	0,021	0,000	0,000	0,000
46 je zhovorčivý,	0,000	0,001	0,002	0,000	0,002	0,002
Asertivita						
6 má asertívnu osobnosť, vyjadruje svoje názory,	0,003	0,000	0,003	0,000	0,000	0,000
21 je dominantný, spáva sa ako vodca,	0,000	0,001	0,002	0,000	0,000	0,000
36R pokladá za ťažké ovplyvňovať ľudí,	0,001	0,001	0,002	0,001	0,001	0,002
51R preferuje, aby iní robili rozhodnutia,	0,000	0,000	0,000	0,001	0,002	0,003
Energická úroveň						
11R je málokedy pre niečo zapálený,	0,002	0,000	0,002	0,002	0,002	0,003
26R je menej aktívny než ostatní ľudia,	0,001	0,004	0,005	0,000	0,000	0,001
41 je plný energie,	0,000	0,000	0,001	0,000	0,001	0,001
56 prejavuje veľa nadšenia,	0,001	0,002	0,002	0,002	0,000	0,002

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

Charakteristické krivky tejto položky z retestového súboru 2 čas 1 zobrazuje Obrázok 5. V tomto prípade je štatisticky významné aj neuniformné odlišné fungovanie. Pri veľmi nízkej sociabilite ženy dosahujú mierne vyššie hodnoty (vyjadrujú nižší súhlas) v porovnaní s mužmi no zhruba v oblasti priemernej sociability sa situácia mení a ženy dosahujú nižšie hodnoty (vyjadrujú vyšší súhlas) v porovnaní s mužmi.



Obrázok 5 Charakteristické krivky položky 31 (retest súbor 2, zber 1)

Výsledky analýzy v súbore mladých dospelých (Tabuľka 4) poukázali na uniformné odlišné fungovanie už identifikovaných položiek 31 a 56, ako aj položky 41 „*Som niekto, kto, je plný energie.*“ ktorá v retestových súboroch nevykazovala štatisticky významné odlišné fungovanie. Muži pri tejto položke vyjadrujú vyšší súhlas, avšak len pri podpriemernej energetickej úrovni. V oblasti priemeru a vyšších hodnôt sa odlišné fungovanie stráca. Charakteristické krivky identifikovaných položiek v súbore mladých dospelých zobrazujeme v Obrázkoch 6 až 8. Vo všetkých troch prípadoch, no najmä v prípade položiek 41 a 56 je dôležité zdôrazniť, že hodnota McFaddenovho pseudo- R^2 je veľmi nízka, na úrovni prakticky zanedbateľného odlišného fungovania.

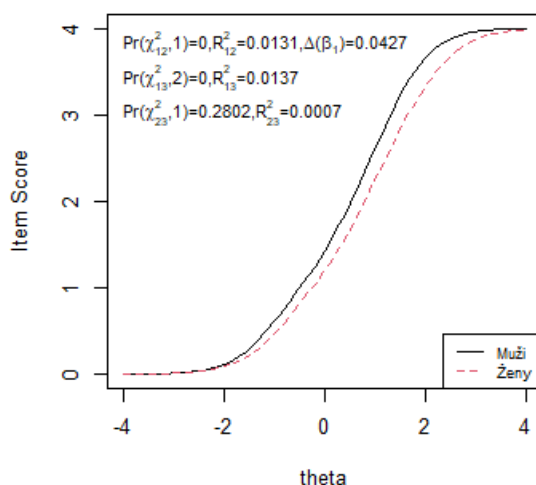
Tabuľka 4 Výsledky analýzy DIF vo facetách extravenzie v súbore mladých dospelých

Položka	Δ McFadden pseudo- R^2		
	U	NU	C
Sociabilita			
1 je spoločenský, rád trávi zber s inými ľuďmi,	0,000	0,001	0,001
16R má sklon byť tichý,	0,000	0,000	0,000
31R je niekedy hanblivý, uzavretý,	0,013	0,001	0,014
46 je zhovorčivý,	0,000	0,001	0,001
Asertivita			

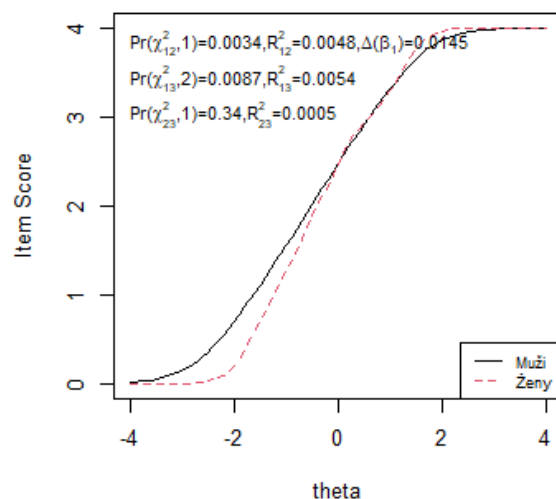
Tabuľka 4 Výsledky analýzy DIF vo facetách extravenzie v súbore mladých dospelých

Položka	Δ McFadden pseudo- R^2		
	U	NU	C
6 má asertívnu osobnosť, vyjadruje svoje názory,	0,004	0,000	0,004
21 je dominantný, spáva sa ako vodca,	0,000	0,000	0,000
36R pokladá za ťažké ovplyvňovať ľudí,	0,001	0,000	0,001
51R preferuje, aby iní robili rozhodnutia,	0,000	0,002	0,002
Energická úroveň			
11R je málokedy pre niečo zapálený,	0,001	0,001	0,002
26R je menej aktívny než ostatní ľudia,	0,001	0,004	0,004
41 je plný energie,	0,005	0,001	0,005
56 prejavuje veľa nadšenia,	0,006	0,000	0,006

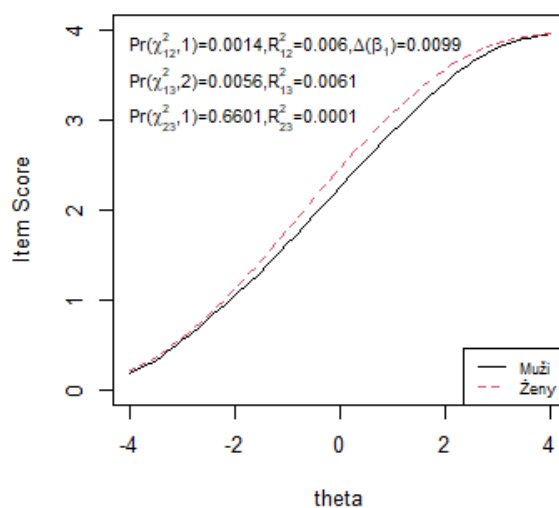
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.



Obrázok 6 Charakteristické krivky položky 31 (súbor mladých dospelých)



Obrázok 7 Charakteristické krivky položky 41 (súbor mladých dospelých)



Obrázok 8 Charakteristické krivky položky 56 (súbor mladých dospelých)

Z hľadiska výsledkov získaných na všeobecnej dospeléj populácii a populácii mladých dospelých nedokážeme jednoznačne potvrdiť rovnosť merania faciet extravenzie BFI-2, no na základe nekonzistencie výsledkov sa domnievame, že prípadná nerovnosť je len minimálna a nemá prakticky významný efekt na výsledky porovnávania mužov a žien v miere sociability, asertivity a energetickej úrovne, a teda extravenzie ako takej.

6.2. Prívetivosť

Podobne ako v prípade extravenzie sa aj pri prívetivosti podarilo prostredníctvom CFA dosiahnuť konfiguračnú aj metrickú úroveň invariantnosti merania vo všetkých súboroch. Skalárna úroveň bola dosiahnutá vo všetkých súboroch s výnimkou prvého retestového súboru v čase 2. Zmena CFI pri fixovaní úrovňových konštánt položiek u mužov a žien spôsobilo výraznejšiu CFI. Ďalšou analýzou sme zistili, že uvoľnenie položky 37: „*Som niekto, kto je niekedy hrubý k iným.*“ spôsobí zmenu CFI na prijateľnú úroveň, čím sme dosiahli parciálnu skalárnu invariantnosť. Táto položka je určená na meranie zdvorilosti a je reverzne formulovaná. V prekódovanej podobe majú muži nižšiu hodnotu úrovňovej konštanty v porovnaní so ženami (3,56 / 4,02), čo znamená, že vyjadrujú vyšší súhlas touto položkou. Výsledky konfirmačnej faktorovej analýzy uvádzame v Tabuľke 5.

Tabuľka 5 Výsledky konfirmačnej faktorovej analýzy pre doménu prívetivosť

Úroveň invariantnosti	CFI	RMSEA	90% KI		Δ CFI	Δ RMSEA
			Spodný	Horný		
Retestový súbor 1 - prvý zber						
Konfiguračná	0,902	0,067	0,052	0,081		
Metrická	0,901	0,064	0,050	0,078	0,001	0,003
Skalárna	0,895	0,064	0,050	0,077	0,007	0,000
Retestový súbor 1 - druhý zber						
Konfiguračná	0,881	0,075	0,062	0,089		
Metrická	0,885	0,071	0,057	0,084	-0,005	0,005
Skalárna	0,861	0,075	0,062	0,088	0,024	-0,004
Parciálna skalárna	0,878	0,071	0,058	0,084	0,007	0,000
Retestový súbor 2 - prvý zber						
Konfiguračná	0,909	0,065	0,050	0,080		
Metrická	0,904	0,064	0,050	0,078	0,004	0,001
Skalárna	0,899	0,064	0,050	0,077	0,005	0,000
Retestový súbor 2 - druhý zber						
Konfiguračná	0,965	0,043	0,022	0,061		
Metrická	0,969	0,039	0,016	0,057	-0,004	0,004
Skalárna	0,963	0,041	0,020	0,058	0,005	-0,002
Súbor mladých dospelých						
Konfiguračná	0,897	0,069	0,057	0,081		
Metrická	0,897	0,066	0,055	0,078	0,000	0,003
Skalárna	0,898	0,063	0,052	0,075	-0,001	0,003

Poznámka. KI - konfidenčný interval. Zmena CFI a RMSEA je počítaná odčítaním od nižšej úrovne invariantnosti. Hodnoty CFI a RMSEA indikujúce zhoršenie zhody modelu s dátami sú zvýraznené hrubým.

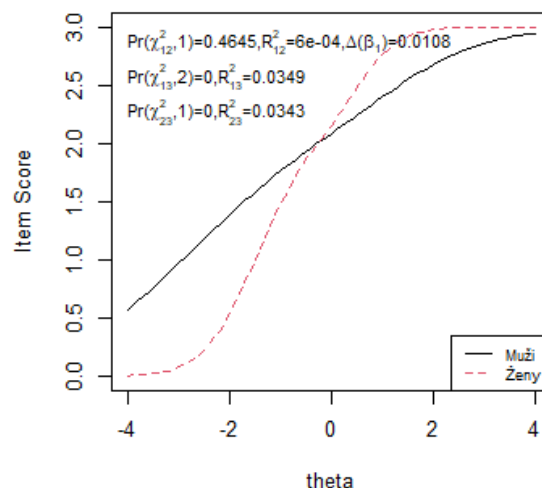
Analýza odlišného fungovania položiek v prvom zbere prvého retestového súboru neidentifikovala žiadne položky, no situácia sa líši v druhom zbere, v ktorom 2 položky vykazujú štatisticky významné uniformné, 1 položka neuniformné a 1 položka zmiešané odlišné fungovanie. Výsledky uvádzame v Tabuľke 6.

Tabuľka 6 Výsledky analýzy DIF vo facetách prítvornosti v prvom retestovom súbore

Položka	Retest súbor 1 zber 1			Retest súbor 1 zber 2		
	Δ McFadden pseudo- R^2			Δ McFadden pseudo- R^2		
	U	NU	C	U	NU	C
Súcit						
2 je súcitný, má dobré srdce.	0,005	0,002	0,007	0,008	0,001	0,009
17R cíti málo pochopenia a súcitu k iným.	0,000	0,002	0,003	0,000	0,000	0,000
32 je nápomocný a nezištný k iným.	0,005	0,001	0,005	0,001	0,034	0,035
47R môže byť chladný a ľahostajný.	0,000	0,000	0,000	0,000	0,001	0,001
Zdvorilosť						
7 je úctivý, správa sa k iným úctivo.	0,001	0,000	0,001	0,022	0,006	0,028
22R vyvoláva spory s inými.	0,000	0,001	0,001	0,001	0,000	0,001
37R je niekedy hrubý k iným.	0,003	0,000	0,003	0,026	0,006	0,033
52 je slušný, zdvorilý voči iným.	0,003	0,000	0,003	0,002	0,000	0,003
Dôvera						
12R má sklon hľadať chyby v ostatných.	0,002	0,001	0,003	0,002	0,000	0,002
27 má v povahe odpúšťať.	0,002	0,000	0,002	0,004	0,002	0,005
42R je podozrievavý voči úmyslom iných.	0,000	0,000	0,000	0,010	0,000	0,011
57 si myslí o ľuďoch to najlepšie.	0,000	0,001	0,001	0,000	0,000	0,000

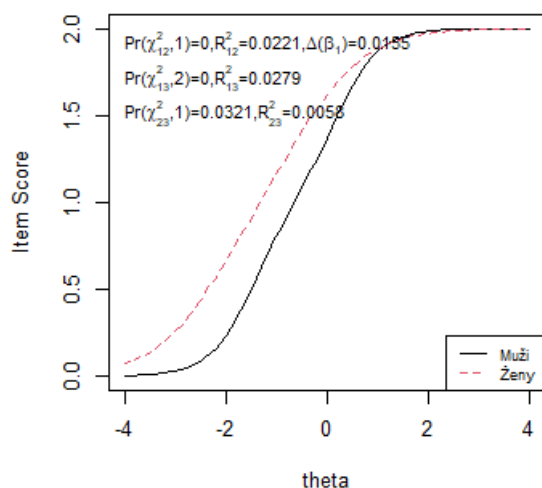
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

Neuniformné odlišné fungovanie sme zistili pri položke 32 „*Som niekto, kto je nápomocný a nezištný k iným.*“ Pri nízkej miere súcitu muži odpovedajú viac súhlasne v porovnaní so ženami, a naopak od priemernej miery vyššie odpovedajú súhlasnejšie ženy. Charakteristické krivky položky sú zobrazené v Obrázku 9. Odpovede boli pri tejto položke zlúčené do 4 kategórií.



Obrázok 9 Charakteristické krivky položky 32 (retest súbor 1, zber 2)

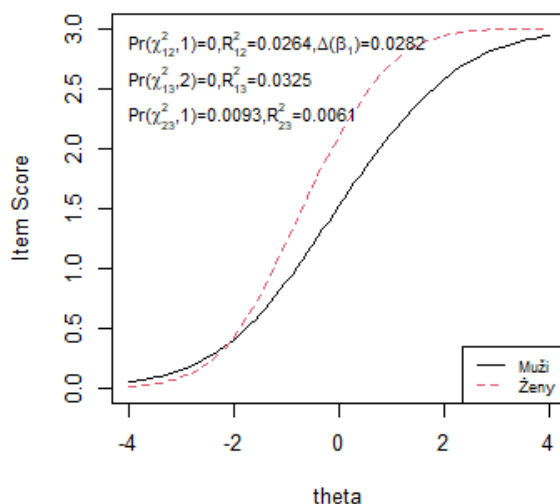
Položka 7 „*Som niekto, kto je úctivý, správa sa k iným úctivo.*“ vykazuje odlišné fungovanie v smere vyššieho súhlasu u žien pri rovnakej miere zdvorilosti. Tento rozdiel je prítomný najmä pri nízkej miere zdvorilosti. Charakteristické krivky položky sú zobrazené v Obrázku 10. Odpovede boli pri tejto položke zlúčené do 3 kategórií.



Obrázok 10 Charakteristické krivky položky 7 (retest súbor 1, zber 2)

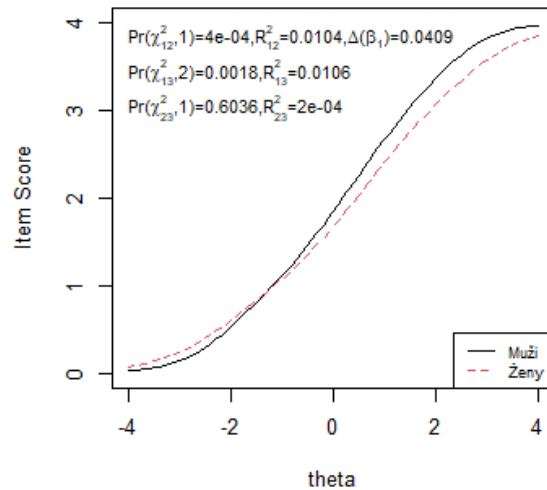
Druhou identifikovanou položkou z tejto facety je reverzne formulovaná položka 37 „*Som niekto, kto je niekedy hrubý k iným.*“ Táto položka bola identifikovaná aj prostredníctvom CFA. Analýza prostredníctvom ordinálnej regresnej analýzy však poukázala aj na štatisticky významné neuniformné odlišné fungovanie položky, ktoré je však prakticky zanedbateľné. Z Obrázku 11 je zrejmé, že pri veľmi nízkej úrovni zdvorilosti, muži dosahujú mierne vyššie

hodnoty (odpovedajú menej súhlasne), no tento rozdiel je minimálny a prevažná väčšina odlišného fungovania je v smere nižšieho súhlasu u žien pri rovnakej miere zdvorilosti. Odpovede boli pri tejto položke zlúčené do 4 kategórií.



Obrázok 11 Charakteristické krivky položky 37 (retest súbor 1, zber 2)

Poslednou identifikovanou položkou z prvého retestového súboru je reverzne formulovaná položka 42 „*Som niekto, kto je podozrievavý voči úmyslom iných.*“ merajúca dôveru. Pri tejto položke dosahujú muži vyššie hodnoty pri rovnakej miere dôvery, to znamená, že ženy viac súhlasia, to jest, vyjadrujú vyššiu podozrievavosť. Grafické znázornenie charakteristických kriviek poukazuje aj na neuniformné odlišné fungovanie, to však nevyšlo štatisticky významné. Celkovo je odlišné fungovanie tejto položky zanedbateľné. Charakteristické krivky zobrazujeme v Obrázku 12.



Obrázok 12 Charakteristické krivky položky 42 (retest súbor 1, zber 2)

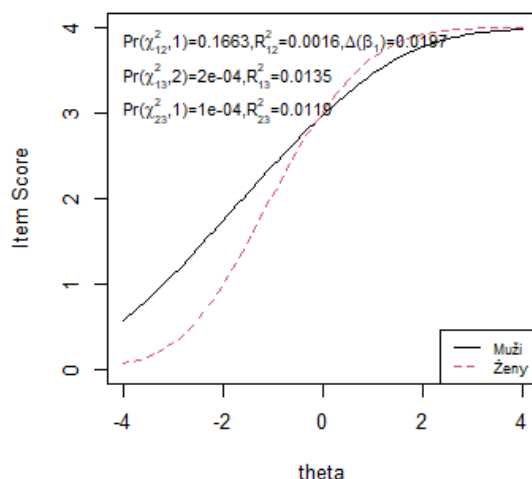
Výsledky z druhého retestového súboru uvádzame v Tabuľke 7. Výsledky z prvého zberu nepotvrdili odlišné fungovanie žiadnej položky z prvého retestového súboru, no boli identifikované iné položky, ktoré sa však nepotvrdili v druhom zbere. Položka 32, ktorú sme identifikovali v prvom retestovom súbore (zber 2), bola identifikovaná aj pri druhom zbere druhého retestového súboru, no miera odlišného fungovania je výrazne nižšia, na úrovni prakticky zanedbateľného.

Tabuľka 7 Výsledky analýzy DIF vo facetách prívetivosti v druhom retestovom súbore

Položka	Retest súbor 2 zber 1			Retest súbor 2 zber 2		
	Δ McFadden pseudo- R^2			Δ McFadden pseudo- R^2		
	U	NU	C	U	NU	C
Súcit						
2 je súcitný, má dobré srdce.	0,005	0,004	0,009	0,000	0,002	0,002
17R cíti málo pochopenia a súcitu k iným.	0,002	0,012	0,014	0,001	0,002	0,003
32 je nápomocný a nezištný k iným.	0,001	0,000	0,001	0,010	0,002	0,012
47R môže byť chladný a ľahostajný.	0,000	0,000	0,000	0,000	0,000	0,001
Zdvorilosť						
7 je úctivý, správa sa k iným úctivo.	0,001	0,001	0,002	0,001	0,000	0,001
22R vyvoláva spory s inými.	0,000	0,001	0,001	0,001	0,000	0,002
37R je niekedy hrubý k iným.	0,002	0,001	0,003	0,001	0,002	0,003
52 je slušný, zdvorilý voči iným.	0,018	0,001	0,019	0,000	0,001	0,001
Dôvera						
12R má sklon hľadať chyby v ostatných.	0,002	0,001	0,002	0,000	0,000	0,000
27 má v povahe odpúšťať.	0,010	0,001	0,011	0,003	0,000	0,003
42R je podozrievavý voči úmyslom iných.	0,000	0,001	0,002	0,004	0,000	0,004
57 si myslí o ľuďoch to najlepšie.	0,002	0,000	0,002	0,000	0,000	0,001

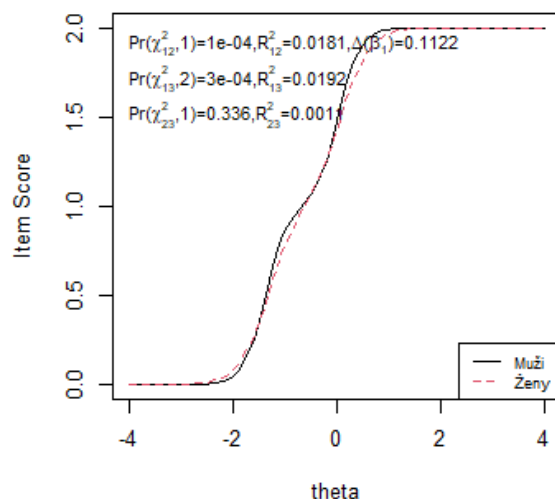
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF významná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

Reverzne formulovaná položka 17 z facety súcitu „*Som niekto, kto cíti málo pochopenia a súcitu k iným.*“ vykazuje neuniformné odlišné fungovanie. Pri nízkej miere súcitu dosahujú ženy nižšie skóre položky, teda vyjadrujú vyšší súhlas. V oblasti priemernej a nadpriemernej úrovne súcitu sa situácia mení a ženy dosahujú vyššie skóre, čo znamená, že pri rovnakej miere súcitu majú muži vyššiu tendenciu súhlasiť. Charakteristické krivky tejto položky zobrazujeme v Obrázku 13.



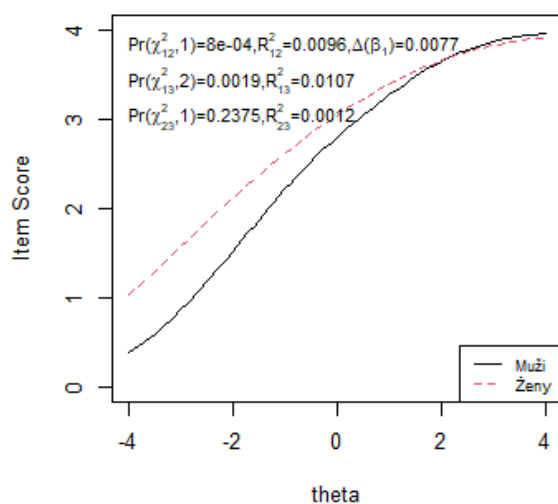
Obrázok 13 Charakteristické krivky položky 17 (retest súbor 2, zber 1)

Položka 52 z facety zdvorilosť „*Som niekto, kto je slušný, zdvorilý voči iným.*“ vykazuje uniformné odlišné fungovanie avšak ide s najväčšou pravdepodobnosťou o štatistický artefakt. V tejto položke boli odpovede zlúčené do 3 kategórií, keďže nebol dostatočný počet odpovedí *Veľmi nesúhlasím* a *Mierne nesúhlasím*, boli tieto kategórie zlúčené spolu s neutrálnou kategóriou. Charakteristické krivky tejto položky (Obrázok 14) sa pomerne prelínajú, preto usudzujeme, že táto položka nevykazuje odlišné fungovanie.



Obrázok 14 Charakteristické krivky položky 52 (retest súbor 2, čas 1)

Posledná identifikovaná položka 27 „*Som niekto, kto má v povahe odpúšťať.*“ Vykazuje mierne, prakticky zanedbateľné uniformné odlišné fungovanie. Pri tejto položke majú ženy vyššiu tendenciu súhlasiť s jej obsahom v porovnaní s mužmi s rovnakou mierou dôvery. Rozdiel je prítomný najmä pri veľmi nízkej a nízkej dôvere, pri priemernej a najmä vyššej dôvere sa rozdiel stráca. Charakteristické krivky tejto položky zobrazujeme v Obrázku 15.



Obrázok 15 Charakteristické krivky položky 27 (retest súbor 2, zber 1)

Oproti retestovým súborom sme v súbore mladých dospelých neidentifikovali ani jednu položku vykazujúcu štatisticky významné odlišné fungovanie. Výsledky uvádzame v Tabuľke 8.

Tabuľka 8 Výsledky analýzy DIF vo facetách prítvornosti v súbore mladých dospelých

Položka	Δ McFadden pseudo- R^2		
	U	NU	C
Súcit			
2 je súcitný, má dobré srdce.	0,002	0,000	0,002
17R cíti málo pochopenia a súcitu k iným.	0,001	0,000	0,001
32 je nápomocný a nezištný k iným.	0,003	0,000	0,003
47R môže byť chladný a ľahostajný.	0,001	0,000	0,001
Zdvorilosť			
7 je úctivý, správa sa k iným úctivo.	0,002	0,001	0,002
22R vyvoláva spory s inými.	0,000	0,000	0,000
37R je niekedy hrubý k iným.	0,000	0,002	0,003
52 je slušný, zdvorilý voči iným.	0,000	0,000	0,001
Dôvera			
12R má sklon hľadať chyby v ostatných.	0,000	0,000	0,001
27 má v povahe odpúšťať.	0,001	0,001	0,001
42R je podozrievavý voči úmyslom iných.	0,000	0,001	0,001
57 si myslí o ľuďoch to najlepšie.	0,000	0,000	0,000

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF významná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

6.3. Svedomitosť

Výsledky invariantnosti merania v doméne svedomitosti uvádzame v Tabuľke 9. V prípade svedomitosti sme v každom súbore potvrdili úplnú konfiguračnú a metrickú invariantnosť, no ani v jednom retestovom súbore nebola dosiahnutá úplná skalárna invariantnosť – tá bola dosiahnutá len v súbore mladých dospelých. Ďalšou analýzou sme identifikovali rovnakú problematickú položku, a to reverznú položku 3 z facety organizovanosti „*Som niekto, kto má sklon byť chaotický.*“ Pri tejto položke dosahujú vyššie hodnoty muži, to znamená, že ženy majú vyššiu tendenciu súhlasiť s jej obsahom, napriek rovnakej miere organizovanosti v porovnaní s mužmi. Úrovňové konštanty u mužov sa pohybovali medzi 3,59 až 3,79, zatiaľ čo u žien medzi 3,13 až 3,35. Aj keď uvoľnenie tejto položky bolo indikované ako najprínosnejšie, musíme upozorniť, že pri druhom zbere prvého retestového súboru bola s prakticky rovnakým prínosom identifikovaná aj neskôr spomenutá položka 48. Po uvoľnení položky 3 sa zmena CFI znížila na prijateľnú úroveň pri prvom retestovom súbore a pri prvom zbere retestového súboru, no v druhom zbere táto zmena nebola dostatočná. Ďalšou analýzou sme identifikovali reverzne formulovanú položku 48 „*Som niekto, kto zanecháva neporiadok, neupratuje.*“ Pri tejto položke dosahovali vyššie hodnoty úrovňovej konštanty ženy (4,17) v porovnaní s mužmi (3,78), to znamená, že muži v základe vyjadrovali väčší súhlas s touto

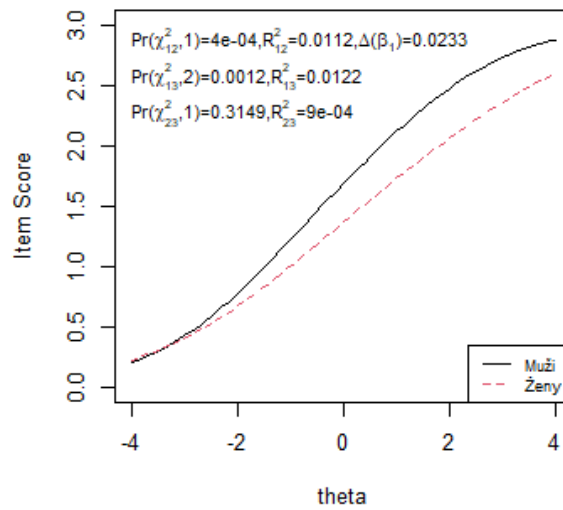
položkou. Po uvoľnení tejto položky sa zhoda modelu s dátami vrátila na hodnotu metrickej úrovne, a teda bola dosiahnutá parciálna skalárna invariantnosť.

Tabuľka 9 Výsledky konfirmačnej faktorovej analýzy pre doménu svedomitost'

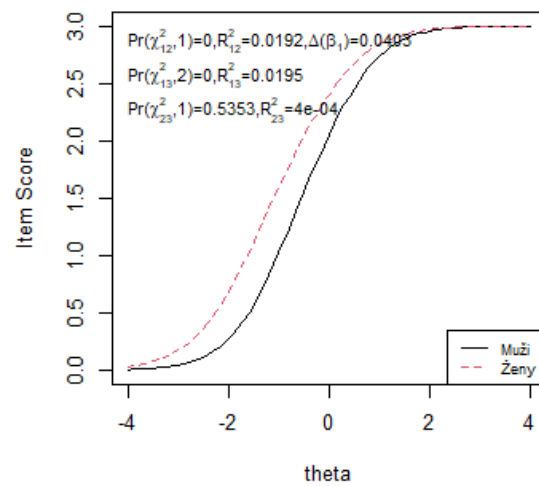
Úroveň invariantnosti	CFI	RMSEA	90% KI		Δ CFI	Δ RMSEA
			Spodný	Horný		
Retestový súbor 1 - prvý zber						
Konfiguračná	0,951	0,060	0,044	0,074		
Metrická	0,953	0,056	0,040	0,070	-0,002	0,004
Skalárna	0,934	0,064	0,050	0,077	0,019	-0,008
Parciálna skalárna	0,948	0,057	0,042	0,071	0,005	-0,001
Retestový súbor 1 - druhý zber						
Konfiguračná	0,942	0,066	0,051	0,082		
Metrická	0,943	0,063	0,048	0,078	-0,001	0,003
Skalárna	0,932	0,066	0,052	0,080	0,010	-0,003
Parciálna skalárna	0,938	0,064	0,049	0,078	0,005	-0,001
Retestový súbor 2 - prvý zber						
Konfiguračná	0,953	0,056	0,040	0,071		
Metrická	0,952	0,054	0,038	0,069	0,000	0,002
Skalárna	0,934	0,061	0,047	0,075	0,018	-0,007
Parciálna skalárna	0,951	0,053	0,037	0,067	0,001	0,001
Retestový súbor 2 - druhý zber						
Konfiguračná	0,965	0,050	0,033	0,066		
Metrická	0,963	0,050	0,033	0,065	0,003	0,000
Skalárna	0,931	0,065	0,052	0,079	0,031	-0,015
Parciálna skalárna	0,963	0,048	0,032	0,063	0,000	0,002
Súbor mladých dospelých						
Konfiguračná	0,966	0,047	0,032	0,062		
Metrická	0,968	0,044	0,028	0,058	-0,002	0,004
Skalárna	0,962	0,047	0,032	0,060	0,007	-0,003

Poznámka. KI - konfidenčný interval. Zmena CFI a RMSEA je počítaná odčítaním od nižšej úrovne invariantnosti. Hodnoty CFI a RMSEA indikujúce zhoršenie zhody modelu s dátami sú zvýraznené hrubým.

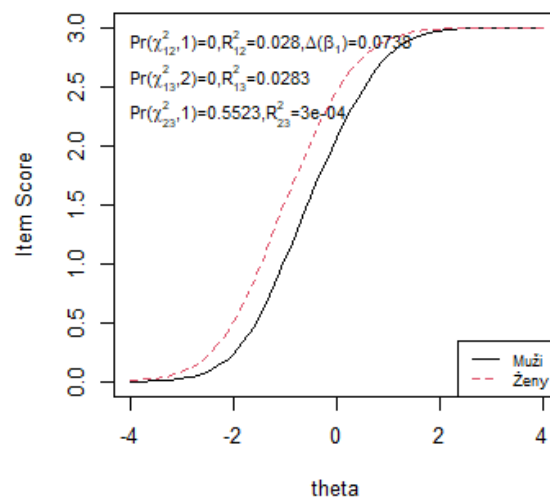
Prostredníctvom ordinálnej logistickej regresie sme v prvom zbere prvého retestového súboru potvrdili uniformné odlišné fungovanie položky 3, avšak silnejšie odlišné fungovanie vykazovala položka 48. V druhom zbere tohto retestového súboru položka 3 nevyšla ako štatisticky významne odlišne fungujúca dokonca ani pri hladine významnosti menšej ako 0,05, no položka 48 bola opätovne potvrdená. Charakteristické krivky týchto položiek prezentujeme v Obrázkoch 16 až 18. Odpovede boli pri týchto položkách zlúčené do 4 kategórií.



Obrázok 16 Charakteristické krivky položky 3 (retest súbor 1, zber 1)

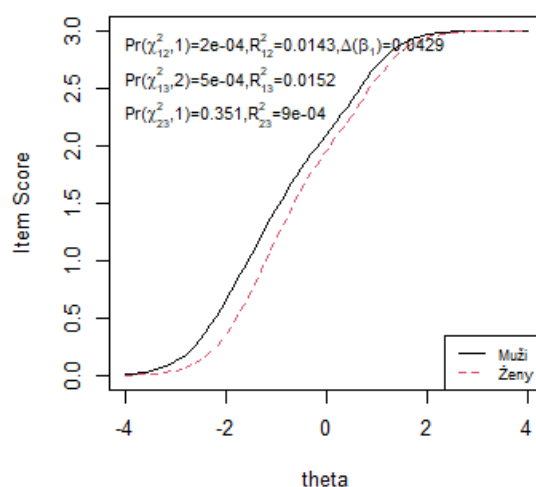


Obrázok 17 Charakteristické krivky položky 48 (retest súbor 1, zber 1)



Obrázok 18 Charakteristické krivky položky 48 (retest súbor 1, zber 2)

Okrem tejto položky bola v tomto zbere identifikovaná aj položka 38 „*Som niekto, kto je efektívny, veci dokončí.*“ z facety produktivita. Táto položka vykazuje veľmi malé uniformné odlišné fungovanie v smere vyšších hodnôt u mužov. Tento trend je pomerne konzistentný od podpriemernej až po nadpriemernú úroveň produktivity. Charakteristické krivky tejto položky zobrazujeme v Obrázku 19. Odpovede boli pri tejto položke zlúčené do 4 kategórií. Výsledky analýzy uvádzame v Tabuľke 10.



Obrázok 19 Charakteristické krivky položky 38 (retest súbor 1, zber 2)

Tabuľka 10 Výsledky analýzy DIF vo facetách svedomitosti v prvom retestovom súbore

Položka	Retest súbor 1 zber 1			Retest súbor 1 zber 2		
	Δ McFadden pseudo-R ²			Δ McFadden pseudo-R ²		
	U	NU	C	U	NU	C
Organizovanosť						
3R má sklon byť chaotický.	0,011	0,001	0,012	0,003	0,000	0,003
18 je systematický, rád udržiava veci v poriadku.	0,001	0,000	0,001	0,001	0,000	0,001
33 udržiava svoje veci usporiadané a zorganizované.	0,001	0,000	0,001	0,001	0,001	0,002
48R zanecháva neporiadok, neupratuje.	0,019	0,000	0,020	0,028	0,000	0,028
Produktivita						
8R má sklon byť lenivý.	0,002	0,003	0,005	0,001	0,000	0,001
23R má problém začať nejakú úlohu.	0,001	0,001	0,001	0,000	0,000	0,000
38 je efektívny, veci dokončí.	0,000	0,000	0,000	0,014	0,001	0,015
53 je vytrvalý, pracuje pokiaľ nie je úloha dokončená.	0,000	0,000	0,000	0,000	0,003	0,003
Zodpovednosť						
13 je spoľahlivý a dôsledný.	0,000	0,000	0,000	0,003	0,002	0,005
28R môže byť trochu nedbanlivý.	0,000	0,001	0,001	0,000	0,001	0,001
43 je dôveryhodný, dá sa na neho vždy spoľahnúť.	0,002	0,002	0,004	0,000	0,000	0,000
58R sa niekedy správa nezodpovedne.	0,001	0,000	0,002	0,005	0,003	0,007

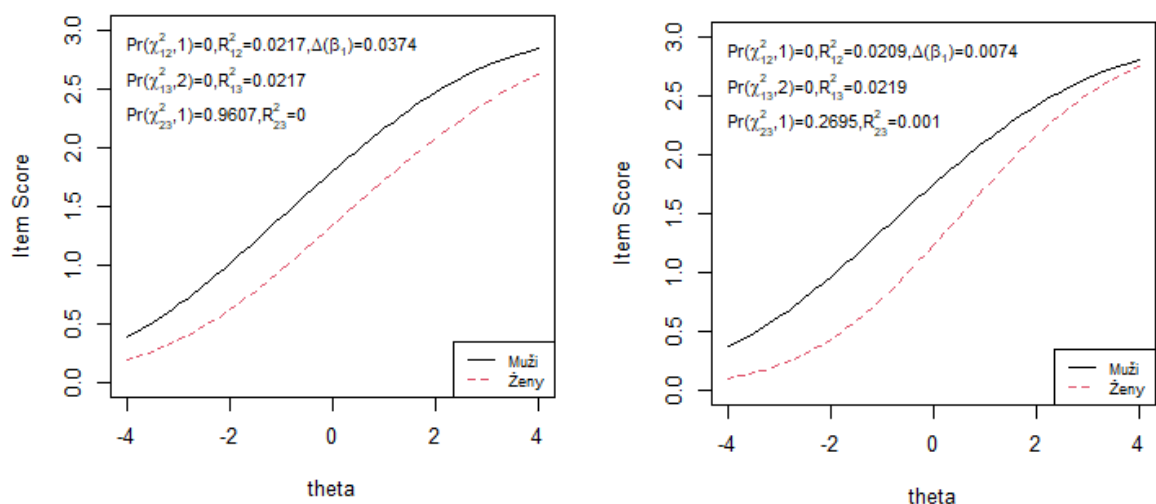
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

V druhom retestovom súbore sa potvrdilo odlišné fungovanie položiek 3 a 48 v oboch zberoch, aj keď je výrazná variabilita výsledkov – položka 48 v prvom zbere vyšla prakticky zanedbateľne odlišne fungujúca, zatiaľ čo v druhom zbere bola hodnota pseudo- R^2 výrazne vyššia. Predtým identifikovaná položka 38 sa nepotvrdila ako odlišne fungujúca. Výsledky z druhého retestového súboru uvádzame v Tabuľke 11. Charakteristické krivky položiek 3 a 48 zobrazujem v Obrázku 20 a 21.

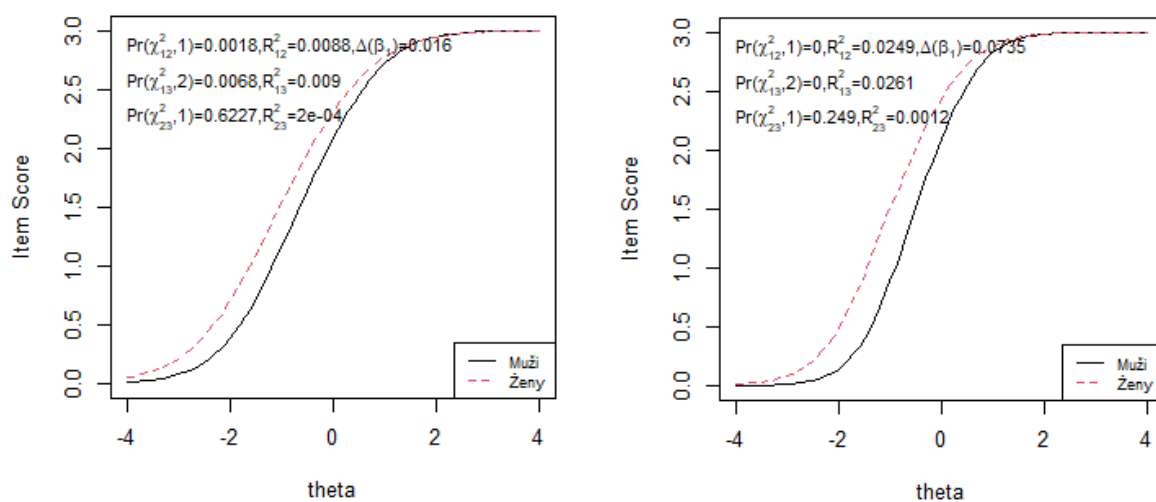
Tabuľka 11 Výsledky analýzy DIF vo facetách svedomitosti v druhom retestovom súbore

Položka	Retest súbor 2 zber 1			Retest súbor 2 zber 2		
	Δ McFadden pseudo- R^2			Δ McFadden pseudo- R^2		
	U	NU	C	U	NU	C
Organizovanosť						
3R má sklon byť chaotický.	0,022	0,000	0,022	0,021	0,001	0,022
18 je systematický, rád udržiava veci v poriadku.	0,000	0,000	0,000	0,001	0,002	0,003
33 udržiava svoje veci usporiadané a zorganizované.	0,000	0,000	0,001	0,002	0,001	0,003
48R zanecháva neporiadok, neupratuje.	0,009	0,000	0,009	0,025	0,001	0,026
Produktivita						
8R má sklon byť lenivý.	0,000	0,002	0,002	0,000	0,000	0,000
23R má problém začať nejakú úlohu.	0,000	0,000	0,000	0,000	0,001	0,002
38 je efektívny, veci dokončí.	0,001	0,001	0,001	0,001	0,000	0,001
53 je vytrvalý, pracuje pokiaľ nie je úloha dokončená.	0,000	0,000	0,001	0,000	0,000	0,000
Zodpovednosť						
13 je spoľahlivý a dôsledný.	0,000	0,000	0,000	0,000	0,001	0,001
28R môže byť trochu nedbanlivý.	0,000	0,000	0,000	0,001	0,002	0,002
43 je dôveryhodný, dá sa na neho vždy spoľahnúť.	0,000	0,002	0,002	0,000	0,000	0,000
58R sa niekedy správa nezodpovedne.	0,000	<i>0,004</i>	0,004	0,000	0,000	0,000

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.



Obrázok 20 Charakteristické krivky položky 3 (retest súbor 2, zber 1 a 2)



Obrázok 21 Charakteristické krivky položky 48 (retest súbor 2, zber 1 a 2)

Výsledky zo súboru mladých dospelých uvádzame v Tabuľke 12. Nepotvrdili sme odlišné fungovanie položiek zistených v retestových súboroch, no identifikovali sme 2 položky vykazujúce prevažne neuniformné odlišné fungovanie.

Tabuľka 12 Výsledky analýzy DIF vo facetách svedomitosti v súbore mladých dospelých

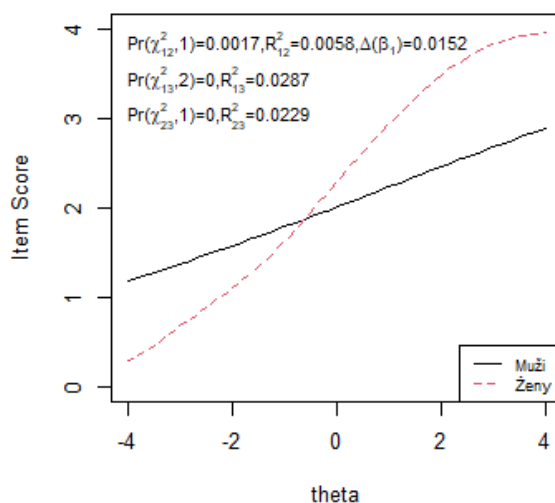
Položka	Δ McFadden pseudo- R^2		
	U	NU	C
Organizovanosť			
3R má sklon byť chaotický.	0,003	0,000	0,003
18 je systematický, rád udržiava veci v poriadku.	0,001	0,000	0,002
33 udržiava svoje veci usporiadané a zorganizované.	0,000	0,000	0,000
48R zanecháva neporiadok, neupratuje.	0,003	0,000	0,003
Produktivita			

Tabuľka 12 Výsledky analýzy DIF vo facetách svedomitosti v súbore mladých dospelých

Položka	Δ McFadden pseudo- R^2		
	U	NU	C
8R má sklon byť lenivý.	0,001	0,001	0,002
23R má problém začať nejakú úlohu.	0,000	0,000	0,000
38 je efektívny, veci dokončí.	0,000	0,000	0,000
53 je vytrvalý, pracuje pokiaľ nie je úloha dokončená.	0,001	0,001	0,001
Zodpovednosť			
13 je spoľahlivý a dôsledný.	0,001	0,001	0,002
28R môže byť trochu nedbanlivý.	0,006	0,023	0,029
43 je dôveryhodný, dá sa na neho vždy spoľahnúť.	0,002	0,000	0,003
58R sa niekedy správa nezodpovedne.	0,002	0,015	0,017

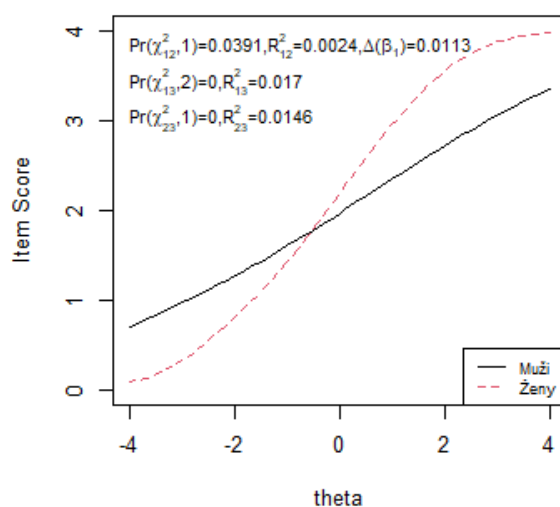
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

Prvou z nich je reverzná položka 28 „*Som niekto, kto môže byť trochu nedbanlivý.*“ Pri tejto položke sme zistili štatisticky významné uniformné aj neuniformné odlišné fungovanie. Charakteristickú krivku tejto položky zobrazuje Obrázok 22. Podľa tvaru krivky u mužov je zrejmé, že položka pomerne nedostatočne rozlišuje medzi nízkou a vysokou mierou zodpovednosti. Pri nízkej miere dosahujú vyššie hodnoty muži, to znamená, že s touto položkou súhlasia menej ako ženy a situácia sa mení približne v oblasti priemernej zodpovednosti.



Obrázok 22 Charakteristické krivky položky 28 (súbor mladých dospelých)

Podobná situácia je aj pri druhej reverznej položke z facety Zodpovednosť. Položka 58 „*Som niekto, kto sa niekedy správa nezodpovedne*“ má veľmi podobný tvar charakteristickej krivky u mužov a rovnaký charakter odlišného fungovania (Obrázok 23).



Obrázok 23 Charakteristické krivky položky 58 (súbor mladých dospelých)

6.4. Negatívna emocionalita

V rámci domény negatívnej emocionality sme pri všetkých dátových súboroch dosiahli metrickú úroveň invariantnosti, no výsledky konfirmačnej faktorovej analýzy ohľadom skalárnej invariantnosti sa líšia. V prvom retestovom súbore sa pri nastavení rovnakých úrovňových konštánt zhoršila zhoda modelu s dátami, čo sa prejavilo výrazným znížením CFI. V prvom zbere sme identifikovali dve položky, ktorých uvoľnenie nám umožnilo konštatovať parciálnej skalárnej invariantnosti. Prvá z nich je položka 34 „*Som niekto, kto sa veľa znepokojuje.*“ z facety Úzkosť. Pri tejto položke sme zistili vyššiu hodnotu úrovňovej konštanty u žien (3,31) ako u mužov (2,98). Druhá identifikovaná položka je reverzná položka 4 je z rovnakej facety: „*Som niekto, kto je uvoľnený, dobre zvláda stres.*“ Pri tejto položke majú muži vyššiu hodnotu úrovňovej konštanty (2,97), čo znamená, že ženy (2,75) majú vyššiu tendenciu na túto položku odpovedať súhlasne. V druhom zbere sa však tieto zistenia nepotvrdili a ako problematickú sme identifikovali reverzne formulovanú položku 9 „*Som niekto, kto ostáva optimistický napriek prekážkam.*“ z facety Depresia. Vyššiu hodnotu úrovňovej konštanty sme zistili u mužov (2,77), čo znamená, že ženy (2,38) na túto položku odpovedajú vyšším súhlasom pri rovnakej miere depresie. Druhý retestový súbor nepotvrdil v prvom súbore identifikované položky. V prvom zbere sme identifikovali ako problematickú reverznú položku 29 z facety Emocionálna nestálosť: „*Som niekto, kto je emocionálne stabilný, neznepokojí sa tak ľahko.*“ Pri tejto položke odpovedali súhlasnejšie muži, keďže majú nižšiu hodnotu úrovňovej konštanty

v porovnaní so ženami (2,73 / 3,04). V dátach z druhého zberu sa nám podarilo potvrdiť úplnú skalárnu invariantnosť. Výsledky uvádzame v Tabuľke 13.

Tabuľka 13 Výsledky konfirmačnej faktorovej analýzy pre doménu negatívna emocionalita

Úroveň invariantnosti	90% KI					
	CFI	RMSEA	Spodný	Horný	Δ CFI	Δ RMSEA
Retestový súbor 1 - prvý zber						
Konfiguračná	0,938	0,070	0,055	0,085		
Metrická	0,940	0,066	0,052	0,080	-0,002	0,004
Skalárna	0,919	0,074	0,061	0,087	0,021	-0,008
Parciálna skalárna	0,931	0,069	0,055	0,083	0,009	-0,003
Retestový súbor 1 - druhý zber						
Konfiguračná	0,970	0,051	0,033	0,068		
Metrická	0,973	0,047	0,028	0,063	-0,003	0,004
Skalárna	0,956	0,058	0,042	0,072	0,017	-0,011
Parciálna skalárna	0,964	0,053	0,036	0,068	0,009	-0,006
Retestový súbor 2 - prvý zber						
Konfiguračná	0,958	0,055	0,038	0,070		
Metrická	0,961	0,051	0,034	0,066	-0,003	0,004
Skalárna	0,946	0,057	0,042	0,071	0,015	-0,007
Parciálna skalárna	0,953	0,054	0,038	0,068	0,008	-0,003
Retestový súbor 2 - druhý zber						
Konfiguračná	0,990	0,029	0,000	0,050		
Metrická	0,984	0,035	0,001	0,054	0,006	-0,006
Skalárna	0,981	0,036	0,011	0,054	0,002	-0,001
Súbor mladých dospelých						
Konfiguračná	0,966	0,051	0,037	0,065		
Metrická	0,965	0,050	0,036	0,063	0,001	0,001
Skalárna	0,957	0,054	0,041	0,066	0,008	-0,004

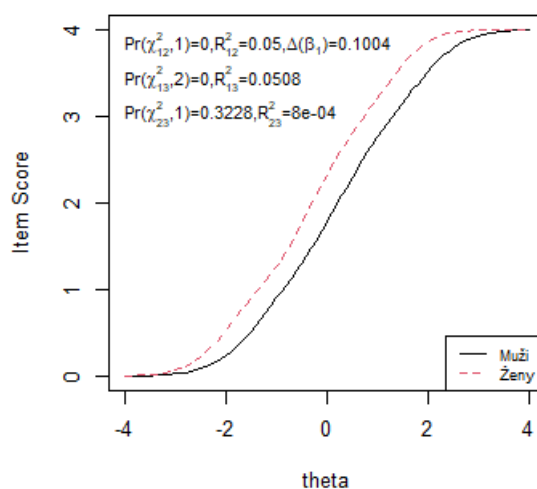
Poznámka. KI - konfidenčný interval. Zmena CFI a RMSEA je počítaná odčítaním od nižšej úrovne invariantnosti. Hodnoty CFI a RMSEA indikujúce zhoršenie zhody modelu s dátami sú zvýraznené hrubým.

Výsledky analýzy odlišného fungovania položiek sú rovnako nekonzistentné. V prvom zbere prvého retestového súboru sme identifikovali položky 34 a 49 ako uniformne odlišne fungujúce a položku 29 ako neuniformne odlišne fungujúcu. Z hľadiska pseudo- R^2 položka 34 vykazuje najsilnejšie odlišné fungovanie z pomedzi všetkých položiek identifikovaných v našom výskume. Napriek tomu sa odlišné fungovanie tejto položky nepotvrdilo v druhom zbere tohto retest súboru, čo platí aj pre položky 29 a 49. Výsledky uvádzame v Tabuľke 14. Charakteristické krivky položiek zobrazujeme v Obrázkoch 24 až 26.

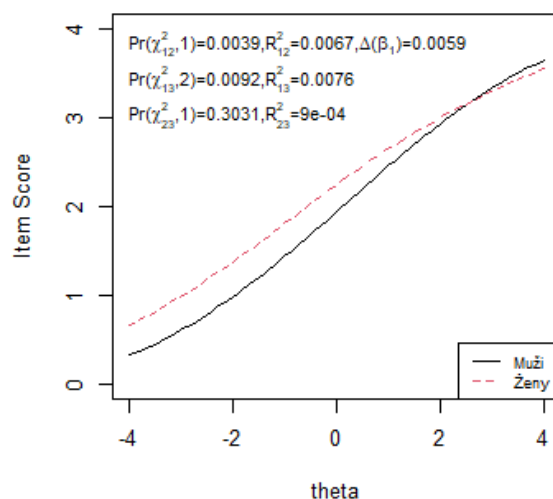
Tabuľka 14 Výsledky analýzy DIF vo facetách negatívnej emocionality v prvom retestovom súbore

	Retest súbor 1 zber 1			Retest súbor 1 zber 2		
	Δ McFadden pseudo- R^2			Δ McFadden pseudo- R^2		
Položka	U	NU	C	U	NU	C
Úzkosť						
4R je uvoľnený, dobre zvláda stres.	0,000	0,000	0,000	0,003	0,001	0,004
19 môže bývať napätý.	0,000	0,001	0,001	0,010	0,001	0,011
34 sa veľa znepokojuje.	0,050	0,001	0,051	0,001	0,000	0,001
49R zriedkakedy cíti úzkosť alebo obavy.	0,007	0,001	0,008	<i>0,004</i>	0,000	0,004
Depresia						
9R ostáva optimistický napriek prekážkam.	0,001	0,000	0,001	<i>0,005</i>	0,001	<i>0,006</i>
24R sa cíti sebaistý, je spokojný so sebou.	0,001	0,000	0,001	0,002	0,000	0,003
39 sa často cíti smutný.	0,002	0,000	0,002	0,002	0,002	0,004
54 má sklon cítiť sa sklúčený, skleslý.	0,001	0,000	0,001	0,002	0,000	0,002
Emocionálna nestálosť						
14 je náladový, máva výkyvy nálady.	0,003	0,001	0,003	0,000	0,000	0,000
29R je emocionálne stabilný, neznepokojí sa tak ľahko.	0,016	0,023	0,040	0,002	0,000	0,002
44R má svoje emócie pod kontrolou.	<i>0,005</i>	0,003	<i>0,008</i>	0,001	0,000	0,001
59 je temperamentný, ľahko sa rozčúli.	0,000	0,000	0,000	0,000	0,000	0,000

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

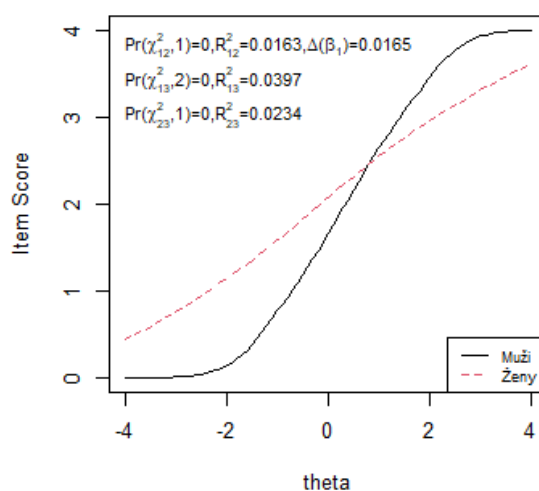


Obrázok 24 Charakteristické krivky položky 34 (retest súbor 1, zber 1)



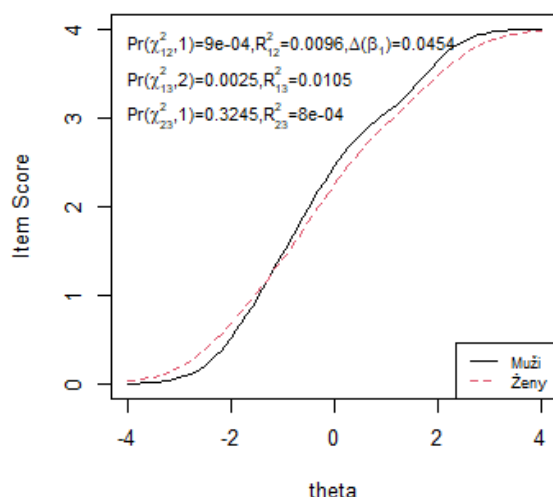
Obrázok 25 Charakteristické krivky položky 49 (retest súbor 1, zber 1)

Položka 29 vykazuje neuniformné odlišné fungovanie, pričom pri nízkej miere emocionálnej nestálosti ženy menej súhlasia s touto položkou, no pri vyššej miere sa smer otáča, a v porovnaní s mužmi vyjadrujú vyšší súhlas.



Obrázok 26 Charakteristické krivky položky 29 (retest súbor 1, zber 1)

V druhom zbere bola identifikovaná položka 19 z facety Úzkosť „*Som niekto, kto môže bývať napätý.*“, pričom muži majú mierne vyššiu tendenciu súhlasiť. Odlišné fungovanie tejto položky je však zanedbateľné, čo vidno aj na charakteristických krivkách, ktoré sú pomerne blízko pri sebe (Obrázok 27).



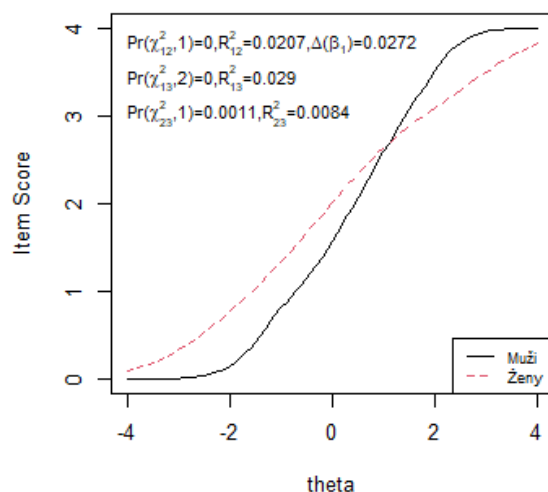
Obrázok 27 Charakteristické krivky položky 19 (retest súbor 1, zber 2)

Výsledky z prvého zberu druhého retestového súboru potvrdili odlišné fungovanie položky 29 (Obrázok 28), ktoré je v tomto prípade prevažne uniformné, no táto položka nevykazovala významné odlišné fungovanie v druhom zbere. V druhom zbere bola identifikovaná položka 59. Výsledky uvádzame v Tabuľke 15.

Tabuľka 15 Výsledky analýzy DIF vo facetách negatívnej emocionality v druhom retestovom súbore

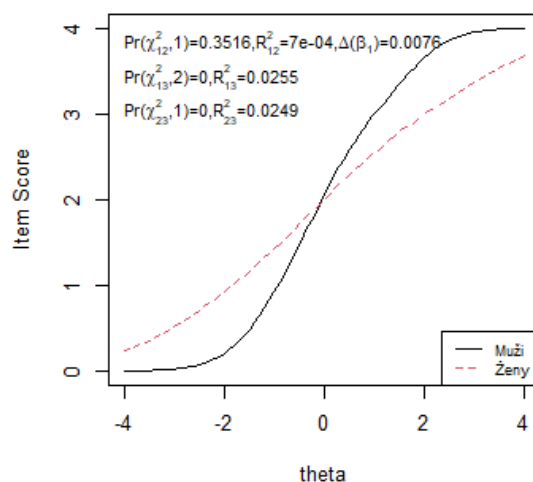
Položka	Retest súbor 2 zber 1			Retest súbor 2 zber 2		
	Δ McFadden pseudo- R^2			Δ McFadden pseudo- R^2		
	U	NU	C	U	NU	C
Úzkosť						
4R je uvoľnený, dobre zvláda stres.	0,003	0,000	0,003	0,000	0,001	0,001
19 môže bývať napätý.	0,003	0,000	0,003	0,000	0,000	0,000
34 sa veľa znepokojuje.	0,002	0,001	0,003	0,000	0,001	0,001
49R zriedkakedy cíti úzkosť alebo obavy.	0,003	0,003	0,006	0,004	0,001	0,004
Depresia						
9R ostáva optimistický napriek prekážkam.	0,001	0,000	0,001	0,000	0,002	0,002
24R sa cíti sebaistý, je spokojný so sebou.	0,000	0,000	0,000	0,000	0,000	0,001
39 sa často cíti smutný.	0,000	0,000	0,001	0,000	0,000	0,000
54 má sklon cítiť sa sklúčený, skleslý.	0,000	0,001	0,001	0,000	0,001	0,001
Emocionálna nestálosť						
14 je náladový, máva výkyvy nálady.	0,000	0,001	0,001	0,001	0,003	0,004
29R je emocionálne stabilný, neznepokojí sa tak ľahko.	0,021	0,008	0,029	0,005	0,000	0,005
44R má svoje emócie pod kontrolou.	0,000	0,001	0,001	0,001	0,000	0,001
59 je temperamentný, ľahko sa rozčúli.	0,000	0,000	0,000	0,001	0,025	0,026

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.



Obrázok 28 Charakteristické krivky položky 29 (retest súbor 2, zber 1)

V Obrázku 29 zobrazujeme charakteristické krivky Položky 59 „*Som niekto, kto je temperamentný, ľahko sa rozčúli.*“ merajúcej emocionálnu nestálosť. Táto položka doposiaľ nebola identifikovaná v iných súboroch a vykazuje len neuniformné odlišné fungovanie. Pri nízkej miere emocionálnej nestability ženy vyjadrujú vyšší súhlas s touto položkou v porovnaní s mužmi, no od priemernej úrovne tejto facety je situácia opačná.



Obrázok 29 Charakteristické krivky položky 59 (retest súbor 2, zber 2)

U mladých dospelých sme v doméne Negatívna emocionalita identifikovali až tri položky z facety Úzkosť. Dve z nich (4 a 49) vykazujú uniformné odlišné fungovanie, zatiaľ čo tretia položka 19 neuniformné. Okrem toho bola ako štatisticky významne, uniformne odlišne fungujúca aj reverzná položka 44. Vo všetkých prípadoch je odlišné fungovanie

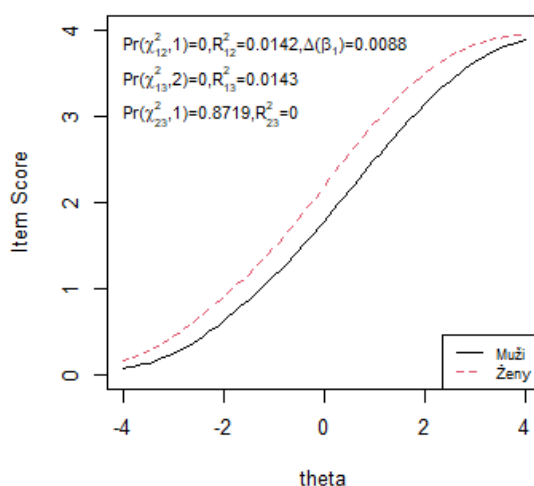
z hľadiska miery efektu na úrovni prakticky zanedbateľného, najmä však pri posledných dvoch menovaných položkách. Výsledky analýzy uvádzame v Tabuľke 16.

Tabuľka 16 Výsledky analýzy DIF vo facetách negatívnej emocionality v súbore mladých dospelých

Položka	Δ McFadden pseudo- R^2		
	U	NU	C
Úzkosť			
4R je uvoľnený, dobre zvláda stres.	0,014	0,000	0,014
19 môže bývať napätý.	0,001	0,005	0,005
34 sa veľa znepokojuje.	0,001	0,001	0,002
49R zriedkakedy cíti úzkosť alebo obavy.	0,008	0,003	0,011
Depresia			
9R ostáva optimistický napriek prekážkam.	0,002	0,001	0,003
24R sa cíti sebaistý, je spokojný so sebou.	0,000	0,000	0,000
39 sa často cíti smutný.	0,000	0,000	0,000
54 má sklon cítiť sa sklúčený, skleslý.	0,001	0,000	0,001
Emocionálna nestálosť			
14 je náladový, máva výkyvy nálady.	0,002	0,000	0,002
29R je emocionálne stabilný, neznepokojí sa tak ľahko.	0,000	0,000	0,000
44R má svoje emócie pod kontrolou.	0,006	0,000	0,006
59 je temperamentný, ľahko sa rozčúli.	0,000	0,001	0,001

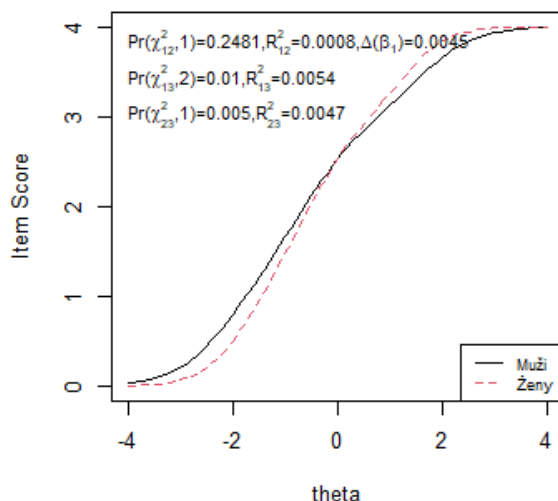
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

Charakteristické krivky reverzne hodnotenej položky 4 „Som niekto, kto je uvoľnený, dobre zvláda stres.“ zobrazujeme v Obrázku 30. V rozmedzí celej miery úzkosti dosahujú ženy vyššie hodnoty položky, čo znamená, že napriek rovnakej miere úzkosti, muži majú vyššiu tendenciu súhlasiť s obsahom tejto položky.



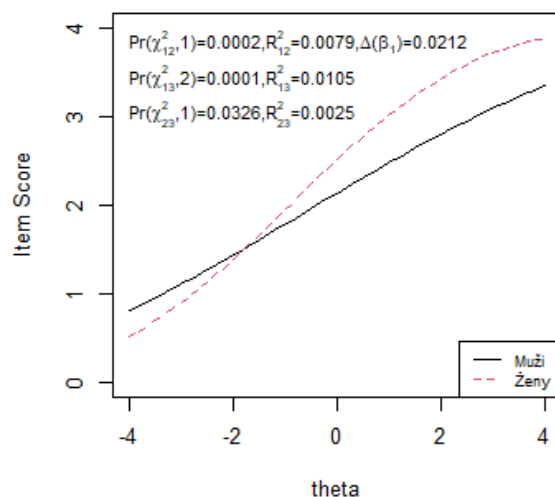
Obrázok 30 Charakteristické krivky položky 4 (súbor mladých dospelých)

Položka 19 bola identifikovaná aj v druhom zbere prvého retestového súboru, kde však vykazovala uniformné odlišné fungovanie. V prípade súboru mladých dospelých vykazuje veľmi slabé, prakticky zanedbateľné neuniformné odlišné fungovanie. Pri nízkej miere úzkosti muži udávajú mierne vyššiu mieru napätosti a situácia je opačná pri vyššej miere úzkosti. Charakteristické krivky zobrazujeme v Obrázku 31.



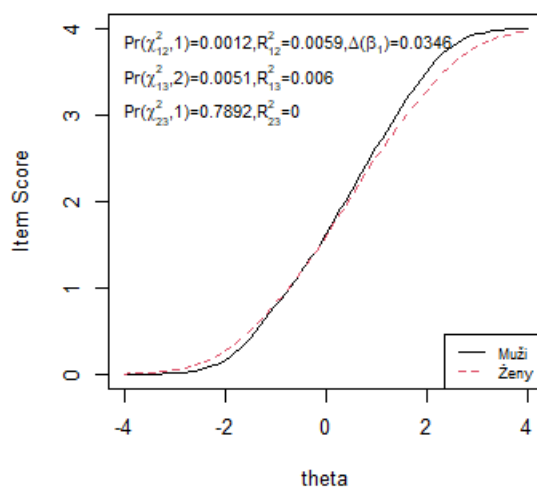
Obrázok 31 Charakteristické krivky položky 19 (súbor mladých dospelých)

Rovnako, položka 49 bola identifikovaná aj v prvom zbere prvého retestového súboru. Podobne ako v retestovom súbore, aj u mladých dospelých dosahujú ženy vyššie hodnoty, čo znamená, že muži majú tendenciu viac súhlasiť. Na rozdiel od retestového súboru je však tento trend prítomný pri vyššej miere úzkosti. Odlišné fungovanie však môže byť spôsobené aj tým, že položka celkovo nefunguje ideálne u mužov. Charakteristické krivky uvádzame v Obrázku 32.



Obrázok 31 Charakteristické krivky položky 49 (súbor mladých dospelých)

Reverzná položka 44 „*Som niekto, kto má svoje emócie pod kontrolou.*“ nebola v retestových súboroch identifikovaná ako štatisticky významne odlišne fungujúca pri hladine $p < 0,01$, no v prvom zbere prvého retestového súboru bola signifikancia na úrovni menšej ako 0,05. V súbore mladých dospelých je hodnota významnosti nižšia, napriek tomu, že praktická veľkosť DIF je mierne nižšia. Tento rozdiel je najpravdepodobnejšie spôsobený vyšším počtom respondentov. Zanedbateľnosť odlišného fungovania je zrejmá aj na vzdialenosti charakteristických kriviek položky, ktoré uvádzame v Obrázku 32.



Obrázok 32 Charakteristické krivky položky 44 (súbor mladých dospelých)

6.5. Otvorenosť

Výsledky analýzy invariantnosti merania vo facetách otvorenosti uvádzame v Tabuľke 17. Zistenia sú v tomto prípade konzistentné. V každom zo zberov oboch retestových súborov bola dosiahnutá konfiguračná a metrická úroveň invariantnosti, no v žiadnom z nich nebola dosiahnutá úplná skalárna invariantnosť. V súbore mladých dospelých sa však podarilo dosiahnuť aj túto úroveň. Ďalšou analýzou sme v retestových súboroch identifikovali tú istú problematickú položku, a to reverznú položku 50 z facety Estetická senzitivita: „*Som niekto, kto považuje poéziu a divadlo za nudné.*“ Úrovňové konštanty sú nižšie u mužov (3,26; 3,31; 3,27; 3,21) ako u žien (3,75; 3,74; 3,78; 3,66), čo znamená, že muži pri tejto položke vyjadrujú vyšší súhlas v porovnaní so ženami. Pre zaujímavosť sme aj v súbore mladých dospelých analyzovali, aký prínos k zhode modelu s dátami by malo uvoľnenie jednotlivých položiek a rovnako ako v retestových súboroch, by bolo najprínosnejšie uvoľnenie položky 50.

Tabuľka 17 Výsledky konfirmačnej faktorovej analýzy pre doménu otvorenosť

Úroveň invariantnosti	90% KI					
	CFI	RMSEA	Spodný	Horný	Δ CFI	Δ RMSEA
Retestový súbor 1 - prvý zber						
Konfiguračná	0,919	0,068	0,052	0,083		
Metrická	0,921	0,064	0,049	0,079	-0,002	0,004
Skalárna	0,903	0,069	0,054	0,083	0,018	-0,004
Parciálna skalárna	0,917	0,064	0,049	0,078	0,004	0,000
Retestový súbor 1 - druhý zber						
Konfiguračná	0,945	0,061	0,044	0,077		
Metrická	0,939	0,061	0,045	0,076	0,006	0,000
Skalárna	0,922	0,066	0,052	0,081	0,017	-0,005
Parciálna skalárna	0,933	0,062	0,047	0,077	0,006	-0,001
Retestový súbor 2 - prvý zber						
Konfiguračná	0,959	0,047	0,027	0,064		
Metrická	0,953	0,048	0,030	0,065	0,006	-0,001
Skalárna	0,939	0,053	0,037	0,068	0,014	-0,005
Parciálna skalárna	0,956	0,045	0,027	0,061	-0,003	0,003
Retestový súbor 2 - druhý zber						
Konfiguračná	0,968	0,044	0,022	0,062		
Metrická	0,965	0,044	0,023	0,061	0,003	0,000
Skalárna	0,954	0,049	0,032	0,065	0,012	-0,005
Parciálna skalárna	0,966	0,042	0,022	0,059	-0,001	0,002
Súbor mladých dospelých						
Konfiguračná	0,914	0,066	0,054	0,079		
Metrická	0,917	0,062	0,050	0,074	-0,003	0,004
Skalárna	0,908	0,064	0,052	0,075	0,009	-0,001

Tabuľka 17 Výsledky konfirmačnej faktorovej analýzy pre doménu otvorenosť

90% KI						
Úroveň invariantnosti	CFI	RMSEA	Spodný	Horný	Δ CFI	Δ RMSEA
<i>Poznámka.</i> KI - konfidenčný interval. Zmena CFI a RMSEA je počítaná odčítaním od nižšej úrovne invariantnosti. Hodnoty CFI a RMSEA indikujúce zhoršenie zhody modelu s dátami sú zvýraznené hrubým.						

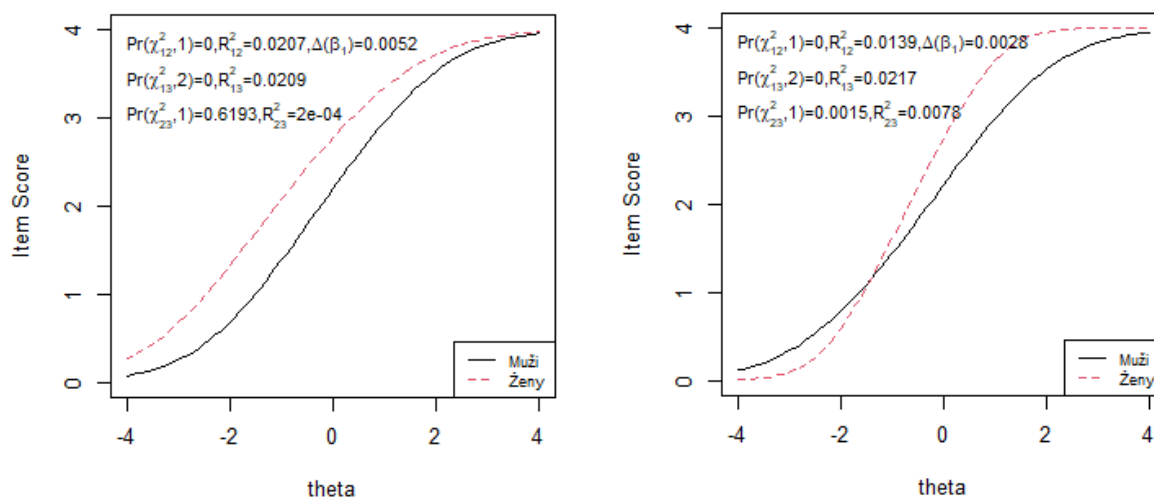
Výsledky analýzy odlišného fungovania položiek sú v tomto prípade pomerne konzistentné v zmysle identifikovania položky 50, no boli identifikované aj dve iné položky, ktoré však nie sú konzistentné medzi zbermi ani retestovými súbormi. Výsledky z prvého retestového súboru uvádzame v Tabuľke 18.

Tabuľka 18 Výsledky analýzy DIF vo facetách otvorenosti v prvom retestovom súbore

Položka	Retest súbor 1 zber 1			Retest súbor 1 zber 2		
	Δ McFadden pseudo-R ²			Δ McFadden pseudo-R ²		
	U	NU	C	U	NU	C
Estetická senzitivita						
5R má málo umeleckých záujmov.	0,000	0,002	0,003	0,002	0,000	0,002
20 je očarený umením, hudbou alebo literatúrou.	0,001	0,001	0,002	0,000	0,003	0,004
35 si cení umenie a krásu.	0,001	0,001	0,003	0,000	0,004	0,004
50R považuje poéziu a divadlo za nudné.	0,021	0,000	0,021	0,014	0,008	0,022
Intelektuálna zvedavosť						
10 je zvedavý na mnoho rozličných vecí.	0,000	0,000	0,001	0,003	0,000	0,003
25R sa vyhýba intelektuálnym, filozofickým diskusiám.	0,002	0,000	0,002	0,015	0,000	0,015
40 rozmýšľa komplexne, do hĺbky.	0,002	0,000	0,002	0,000	0,003	0,004
55R má malý záujem o abstraktné myšlienky.	0,001	0,000	0,001	0,000	0,001	0,001
Kreatívna predstavivosť						
15 je vynaliezavý, hľadá šikovné spôsoby ako robiť veci.	0,001	0,000	0,001	0,000	0,000	0,000
30R je málo tvorivý.	0,001	0,003	0,004	0,000	0,000	0,000
45R má problém predstavovať si veci.	0,000	0,001	0,001	0,000	0,001	0,001
60 je originálny, prichádza s novými nápadmi.	0,000	0,005	0,005	0,001	0,002	0,002

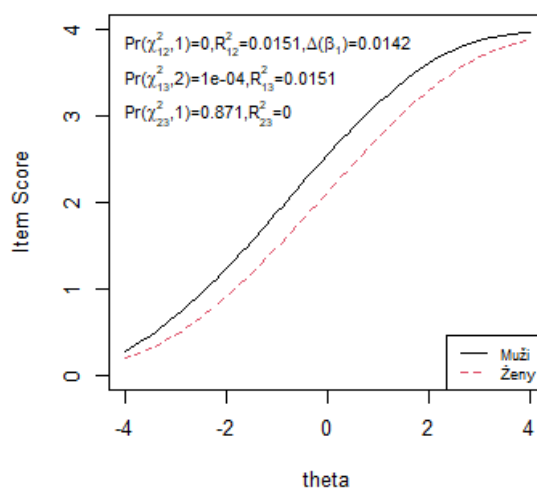
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo-R² pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

V Obrázku 33 znázorňujeme charakteristické krivky položky 50 z prvého a druhého zberu prvého retestového súboru. V prípade prvého zberu ide o jednoznačne uniformné odlišné fungovanie, jednak na základe výsledkov uvedených v Tabuľke 14 ako aj vizuálnou inšpekciou charakteristických kriviek. V druhom zbere výsledky naznačujú aj prítomnosť neuniformného odlišného fungovania, ktoré je však štatisticky nevýznamné pri hladine alfa menšej ako 0,01. Zaujímavé je, že v prípade prvého zberu bolo odlišné fungovanie najväčšie pri nižšej miere estetickej senzitivity, v druhom zbere je trend v tejto oblasti opačný a odlišné fungovanie je prítomné pri priemernej až vyššej estetickej senzitive.



Obrázok 33 Charakteristické krivky položky 50 (retest súbor 1, zber 1 a 2)

Druhou identifikovanou v prvom retestovom súbore bola reverzná položka 25 z facety Intelektuálna zvedavosť: „*Som niekto, kto sa vyhýba intelektuálnym, filozofickým diskusiam.*“ Pri tejto položke dosahujú muži vyššie hodnoty, teda ženy s touto položkou viac súhlasia pri rovnakej miere intelektuálnej zvedavosti. Odlišné fungovanie je podľa charakteristických kriviek (Obrázok 34) prítomné pri celom rozsahu tejto črty.



Obrázok 34 Charakteristické krivky položky 25 (retest súbor 1, zber 2)

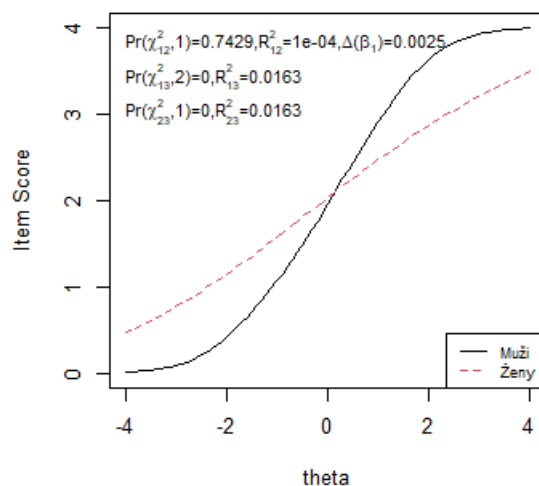
V Tabuľke 19 uvádzame výsledky z druhého retestového súboru. Opätovne bola identifikovaná položka 50 v oboch zberoch, pričom pribudla aj ďalšia položka z facety Estetická senzitivita, ktorá však bola identifikovaná len v prvom zbere.

Tabuľka 19 Výsledky analýzy DIF vo facetách otvorenosti v druhom retestovom súbore

Položka	Retest súbor 2 zber 1			Retest súbor 2 zber 2		
	Δ McFadden	pseudo- R^2		Δ McFadden	pseudo- R^2	
	U	NU	C	U	NU	C
Estetická senzitivita						
5R má málo umeleckých záujmov.	0,000	0,016	0,016	0,001	0,003	0,004
20 je očarený umením, hudbou alebo literatúrou.	0,000	0,002	0,002	0,001	0,003	0,004
35 si cení umenie a krásu.	0,001	0,000	0,001	0,000	0,000	0,000
50R považuje poéziu a divadlo za nudné.	0,018	0,001	0,019	0,012	0,000	0,012
Intelektuálna zvedavosť						
10 je zvedavý na mnoho rozličných vecí.	0,000	0,000	0,000	0,000	0,000	0,000
25R sa vyhýba intelektuálnym, filozofickým diskusiám.	0,001	0,000	0,001	0,001	0,000	0,001
40 rozmýšľa komplexne, do hĺbky.	0,001	0,001	0,002	0,000	0,002	0,002
55R má malý záujem o abstraktné myšlienky.	0,000	0,001	0,001	0,002	0,000	0,002
Kreatívna predstavivosť						
15 je vynaliezavý, hľadá šikovné spôsoby ako robiť veci.	0,000	0,002	0,002	0,000	0,000	0,000
30R je málo tvorivý.	0,000	0,001	0,001	0,000	0,001	0,001
45R má problém predstavovať si veci.	0,001	0,000	0,001	0,000	0,001	0,001
60 je originálny, prichádza s novými nápadmi.	0,000	0,003	0,003	0,000	0,001	0,001

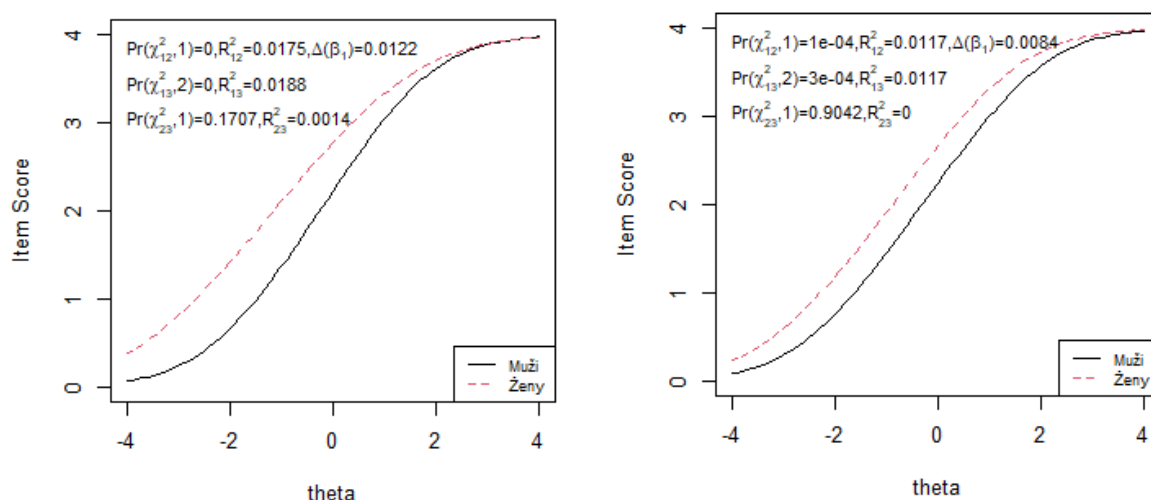
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

Reverzná položka 5 „Som niekto, kto má málo umeleckých záujmov.“ vykazuje štatisticky významné neuniformné odlišné fungovanie. Charakteristické krivky zobrazujeme v Obrázku 35. Pri nízkej miere estetickej senzitivity dosahujú ženy vyššie hodnoty, čo znamená, že muži majú tendenciu viac súhlasiť s obsahom položky. Trend sa však pri priemernej úrovni otáča a ženy dosahujú nižšie hodnoty, to znamená, že s obsahom položky viac súhlasia v porovnaní s mužmi.



Obrázok 35 Charakteristické krivky položky 5 (retest súbor 2, zber 1)

Charakteristické krivky už identifikovanej položky 50 zobrazujeme v Obrázku 36. Na základe vizuálneho posúdenia je možné vidieť veľmi podobný trend v oboch zberoch, aj keď v druhom zbere je tento trend o niečo slabší (aj na základe pseudo- R^2). Trend je podobný tiež tomu v prvom zbere prvého retestového súboru, preto sa domnievame, že v druhom zbere prvého retestového súboru ide skôr o chybu odhadu modelu.



Obrázok 36 Charakteristické krivky položky 50 (retest súbor 2, zber 1 a 2)

Výsledky analýzy odlišného fungovania položiek faciet Otvorenosti u mladých dospelých uvádzame v Tabuľke 20. Jedinou identifikovanou položkou je reverzná položka 50. Tento výsledok je konzistentný z výsledkami analýz na retestových súboroch. Charakter odlišného fungovania tejto položky je veľmi podobný tomu v prvom zbere prvého a druhého retestového súboru, to jest, najväčší rozdiel medzi hodnotami položky je najmä v oblasti podpriemernej estetickej senzitivity. Charakteristické krivky zobrazujeme v Obrázku 37.

Tabuľka 20 Výsledky analýzy DIF vo facetách otvorenosti v súbore mladých dospelých

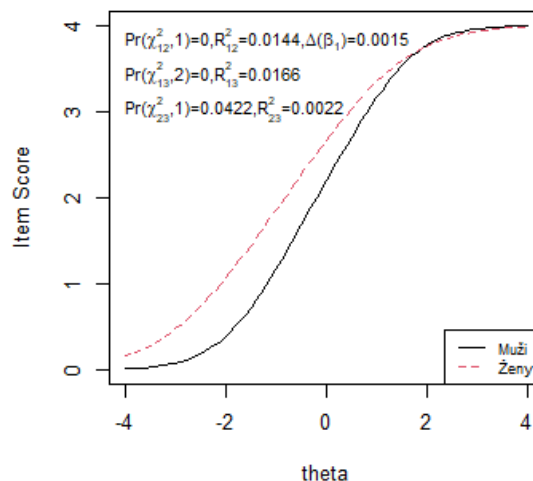
Položka	Δ McFadden pseudo- R^2		
	U	NU	C
Estetická senzitivita			
5R má málo umeleckých záujmov.	0,000	0,000	0,000
20 je očarený umením, hudbou alebo literatúrou.	0,002	0,001	0,003
35 si cení umenie a krásu.	0,002	0,001	0,003
50R považuje poéziu a divadlo za nudné.	0,014	0,002	0,017
Intelektuálna zvedavosť			
10 je zvedavý na mnoho rozličných vecí.	0,000	0,000	0,000

25R	sa vyhýba intelektuálnym, filozofickým diskusiám.	0,001	0,000	0,001
40	rozmyšľa komplexne, do hĺbky.	0,000	0,000	0,000
55R	má malý záujem o abstraktné myšlienky.	0,000	0,000	0,000

Kreatívna predstavivosť

15	je vynaliezavý, hľadá šikovné spôsoby ako robiť veci.	0,001	0,000	0,001
30R	je málo tvorivý.	0,000	0,001	0,001
45R	má problém predstavovať si veci.	0,002	0,000	0,002
60	je originálny, prichádza s novými nápadmi.	0,000	0,001	0,001

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.



Obrázok 37 Charakteristické krivky položky 50 (súbor mladých dospelých)

7. Invariantnosť merania a odlišné fungovanie položiek BFI-2 podľa veku

V tejto časti uvádzame výsledky analýz zameraných na invariantnosť merania a odlišné fungovanie položiek Inventára Veľkej Päťky 2 u dospelých vo veku vynárajúcej sa dospelosti (do 30 rokov) a staršími (30 rokov a viac).

V Tabuľke 21 uvádzame výsledky konfirmačnej faktorovej analýzy všetkých domén BFI-2. Vo všetkých doménach sa potvrdila konfiguračná a metrická úroveň invariantnosti, no len v doméne Negatívna emocionalita a Otvorenosť aj skalárna úroveň. V doméne Extraverzia je pokles CFI po nastavení rovnakých úrovňových konštánt hraničný, no napriek tomu sme sa rozhodli preskúmať, ktorá položka sa najviac podieľa na zhoršení zhody modelu s dátami. Identifikovali sme položku 16 z facety Sociabilita: „*Som niekto, kto má sklon byť tichý.*“ Táto položka je opačne formulovaná. V už prekódovanej podobe majú pri tejto položke respondenti vo veku vynárajúcej sa dospelosti vyššiu hodnotu úrovňovej konštanty (3,16) v porovnaní s ostatnými (2,86), čo naznačuje ich tendenciu menej súhlasiť s obsahom. Po uvoľnení tejto položky sa zhoda modelu s dátami zlepšila a bola dosiahnutá parciálna skalárna úroveň invariantnosti.

V doméne Prívetivosť sme identifikovali problematickú položku 37 „*Som niekto, kto je niekedy hrubý k iným.*“ Táto reverzná položka je z facety Zdvorilosť. Zistili sme, že v prekódovanej podobe je vyššia hodnota úrovňovej konštanty u starších dospelých (3,75) v porovnaní s vynárajúcimi sa dospelými (3,40), čo znamená, že vynárajúci sa dospelí majú vyššiu tendenciu súhlasiť s tým, že bývajú hrubý k iným. Po uvoľnení tejto položky bola dosiahnutá parciálna skalárna úroveň invariantnosti.

Na záver, v doméne Svedomitosť bola identifikovaná reverzná položka 28 „*Som niekto, kto môže byť trochu nedbanlivý.*“ z facety Zodpovednosť. Pri tejto položke majú vynárajúci sa dospelí vyššiu hodnotu úrovňovej konštanty (3,44) v porovnaní so staršími (3,14), čo znamená, že starší vyjadrujú vyššiu mieru nedbanlivosti. Po uvoľnení tejto položky sa zlepšila zhoda modelu s dátami a konštatujeme čiastočnú skalárnu invariantnosť.

Tabuľka 21 Výsledky konfirmačnej faktorovej analýzy na základe vekových skupín

Úroveň invariantnosti	CFI	RMSEA	90% KI		Δ CFI	Δ RMSEA
			Spodný	Horný		
Extraverzia						
Konfiguračná	0,942	0,053	0,045	0,062		
Metrická	0,942	0,051	0,043	0,059	0,000	0,002
Skalárna	0,933	0,053	0,046	0,061	0,010	-0,002
Parciálna skalárna	0,940	0,051	0,043	0,058	0,002	0,000
Prívetivosť						

Tabuľka 21 Výsledky konfirmačnej faktorovej analýzy na základe vekových skupín

Úroveň invariantnosti	CFI	RMSEA	90% KI		Δ CFI	Δ RMSEA
			Spodný	Horný		
Konfiguračná	0,919	0,061	0,053	0,069		
Metrická	0,916	0,059	0,052	0,067	0,003	0,002
Skalárna	0,904	0,061	0,054	0,069	0,012	-0,002
Parciálna skalárna	0,914	0,058	0,051	0,066	0,002	0,001
Svedomitosť						
Konfiguračná	0,969	0,045	0,036	0,053		
Metrická	0,971	0,042	0,033	0,050	-0,002	0,003
Skalárna	0,959	0,048	0,040	0,055	0,012	-0,006
Parciálna skalárna	0,966	0,044	0,036	0,052	0,005	-0,002
Negatívna emocionalita						
Konfiguračná	0,958	0,055	0,047	0,063		
Metrická	0,954	0,055	0,047	0,063	0,004	0,000
Skalárna	0,950	0,055	0,048	0,063	0,004	0,000
Otvorenosť						
Konfiguračná	0,940	0,055	0,047	0,064		
Metrická	0,940	0,053	0,045	0,061	-0,001	0,003
Skalárna	0,932	0,054	0,047	0,062	0,008	-0,001

Poznámka. KI - konfidenčný interval. Zmena CFI a RMSEA je počítaná odčítaním od nižšej úrovne invariantnosti. Hodnoty CFI a RMSEA indikujúce zhoršenie zhody modelu s dátami sú zvýraznené hrubým.

7.1. Extraverzia

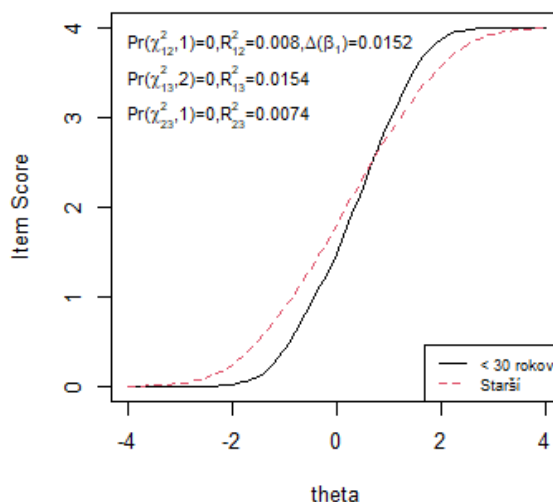
Ďalej uvádzame výsledky odlišného fungovania položiek jednotlivých faciet. Výsledky z domény Extraverzia uvádzame v Tabuľke 22. Vo facete Sociabilita sme celkovo identifikovali až 3 položky vykazujúce odlišné fungovanie, z toho 1 čisto uniformné a dve položky zmiešané. Vo facete Asertivita nastala kuriózna situácia, kedy boli identifikované všetky 4 položky ako odlišne fungujúce pri $p < 0,01$ aj $p < 0,001$. V rámci prvej iterácie bola označené položky 6 a 51. Následne, po prečistení miery vlastnosti od odlišne fungujúcich položiek boli označené všetky štyri položky. Ich odlišné fungovanie je však veľmi slabé, prakticky zanedbateľné. V tabuľke uvádzame výsledky, ktoré sme získali bez ďalšieho postupu očisťovania miery vlastnosti. Ako možno vidieť, najvyššia hodnota zmeny pseudo- R^2 je len 0,004, čo je hodnota veľmi vzdialená hodnote 0,02, ktorú môžeme považovať za hranicu slabého odlišného fungovania. V poslednej facete bola identifikovaná jedna položka, ktorá vykazuje uniformné odlišné fungovanie.

Tabuľka 22 Výsledky analýzy DIF podľa veku vo facetách extravenzie

Položka	Δ McFadden pseudo- R^2		
	U	NU	C
Sociabilita			
1 je spoločenský, rád trávi zber s inými ľuďmi.	0,001	0,001	0,002
16R má sklon byť tichý.	0,008	0,007	0,015
31R je niekedy hanblivý, uzavretý.	0,027	0,009	0,035
46 je zhovorčivý.	0,014	0,001	0,015
Asertivita			
6 má asertívnu osobnosť, vyjadruje svoje názory.	0,001	0,003	0,004
21 je dominantný, spáva sa ako vodca.	0,001	0,000	0,002
36R pokladá za ťažké ovplyvňovať ľudí.	0,001	0,001	0,002
51R preferuje, aby iní robili rozhodnutia.	0,004	0,000	0,004
Energická úroveň			
11R je málokedy pre niečo zapálený.	0,000	0,000	0,000
26R je menej aktívny než ostatní ľudia.	0,007	0,000	0,007
41 je plný energie.	0,000	0,000	0,000
56 prejavuje veľa nadšenia.	0,000	0,001	0,001

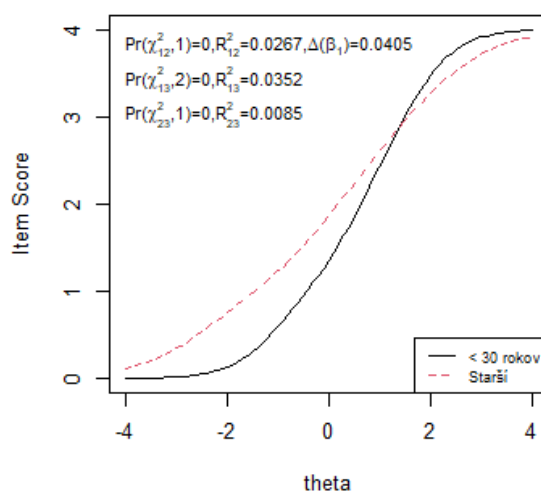
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF významná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

Obrázok 38 znázorňuje charakteristické krivky reverznej položky 16. Táto položka bola identifikovaná ako problematická aj pri analýze invariantnosti merania prostredníctvom konfirmačnej faktorovej analýzy. Zaujímavé je, že v tomto prípade vykazuje zmiešané odlišné fungovanie, resp. ako významné sa ukázali oba typy odlišného fungovania, avšak oba typy sú veľmi slabé. Pri nízkej miere sociability dosahujú starší respondenti vyššie hodnoty položky, tento smer sa však mení pri vyššej miere sociability.



Obrázok 38 Charakteristické krivky položky 16

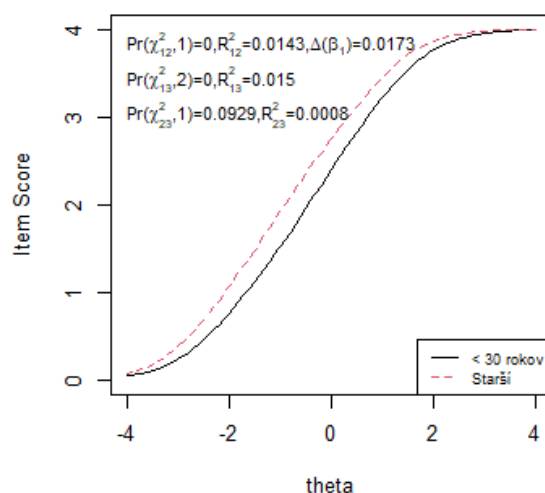
Druhá položka z facety Sociabilita, ktorá vykazovalo najväčšie odlišné fungovanie je reverzná položka 31 „*Som niekto, kto je niekedy hanblivý, uzavretý.*“. Táto položka vykazuje prevažne uniformné no prítomné je aj štatisticky významné neuniformné odlišné fungovanie, ktoré je však výrazne slabšie. Pri tejto položke majú vyššie hodnoty starší respondenti, čo znamená, že vyjadrujú nižší súhlas v porovnaní s vynárajúcimi sa dospelými. Tento trend je prítomný v oblasti nízkej a priemernej sociability, pri vyššej miere je trend opačný. Charakteristické krivky položky zobrazuje Obrázok 39.



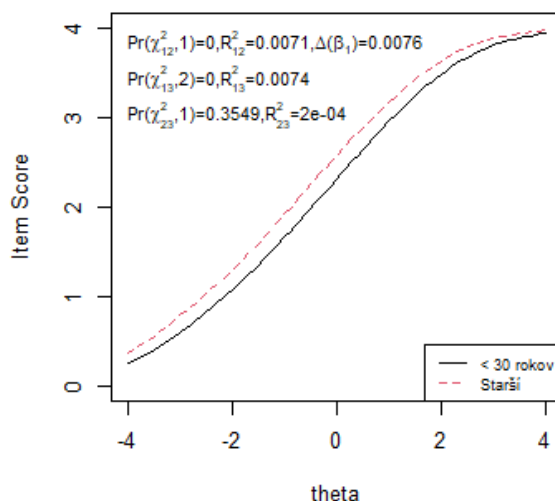
Obrázok 39 Charakteristické krivky položky 31

Posledná identifikovaná položka z tejto facety je 46 „*Som niekto, kto je zhovorčivý.*“. Charakteristické krivky uvádzame v Obrázku 40. Starší ľudia majú vyššiu tendenciu k súhlasu, najmä v rozsahu podpriemernej až nadpriemernej sociability.

Vo facete Energická úroveň sme identifikovali uniformne odlišne fungujúcu položku 26 „*Som niekto, kto je menej aktívny než ostatní ľudia.*“, ktorá je reverzne formulovaná. Vyššie hodnoty položky dosahujú starší dospelí, čo znamená, že vynárajúci sa dospelí majú vyššiu tendenciu súhlasiť s obsahom tejto položky. Stále však platí, že tento rozdiel je veľmi malý, čo je vidieť aj na vzdialenosti charakteristických kriviek položky (Obrázok 41).



Obrázok 40 Charakteristické krivky položky 46



Obrázok 41 Charakteristické krivky položky 26

7.2. Prívetivosť

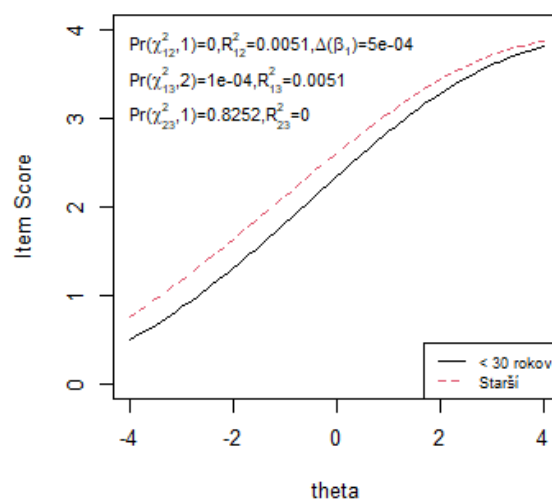
Výsledky analýzy z domény Prívetivosť uvádzame v Tabuľke 23. Celkovo sme identifikovali len dve položky, jednu z facety Súcit a jednu z facety Zdvorilosť. Reverzná položka 47 „*Som niekto, kto môže byť chladný a ľahostajný.*“ vykazuje veľmi slabé uniformné odlišné fungovanie v smere vyšších hodnôt (nižšieho súhlasu) u starších dospelých (Obrázok 42). Položka 37 bola identifikovaná aj prostredníctvom konfirmačnej faktorovej analýzy. Z hľadiska odlišného fungovania vykazuje zmiešané, no prevažne uniformné odlišné fungovanie. Pri veľmi nízkej miere zdvorilosti majú vyššiu tendenciu k súhlasu starší dospelí, no tento trend je zanedbateľný, pričom od priemernej zdvorilosti majú starší dospelí vyššie

hodnoty položky, čo znamená, že s obsahom menej súhlasia. Charakteristické krivky tejto položky znázorňujeme v Obrázku 43.

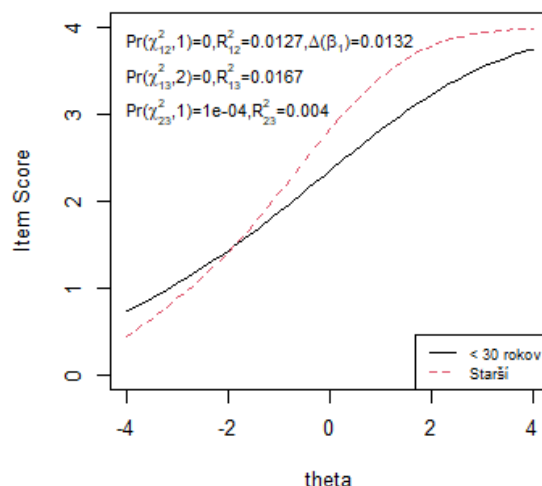
Tabuľka 23 Výsledky analýzy DIF podľa veku vo facetách prívetivosti

Položka	Δ McFadden pseudo- R^2		
	U	NU	C
Súcit			
2 je súcitný, má dobré srdce.	0,001	0,002	0,002
17R cíti málo pochopenia a súcitu k iným.	0,001	0,000	0,001
32 je nápomocný a nezištný k iným.	0,000	0,000	0,000
47R môže byť chladný a ľahostajný.	0,005	0,000	0,005
Zdvorilosť			
7 je úctivý, správa sa k iným úctivo.	0,000	0,000	0,000
22R vyvoláva spory s inými.	0,000	0,000	0,000
37R je niekedy hrubý k iným.	0,013	0,004	0,017
52 je slušný, zdvorilý voči iným.	0,000	0,000	0,000
Dôvera			
12R má sklon hľadať chyby v ostatných.	0,000	0,000	0,000
27 má v povahe odpúšťať.	0,000	0,000	0,000
42R je podozrievavý voči úmyslom iných.	0,001	0,000	0,001
57 si myslí o ľuďoch to najlepšie.	0,001	0,000	0,001

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF významná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.



Obrázok 42 Charakteristické krivky položky 47



Obrázok 43 Charakteristické krivky položky 37

7.3. Svedomitosť

Vo facetách domény Svedomitosť sme identifikovali vyššie množstvo odlišne fungujúcich položiek, všetky z nich však vykazujú len prakticky zanedbateľnú úroveň. Výsledky uvádzame v Tabuľke 24.

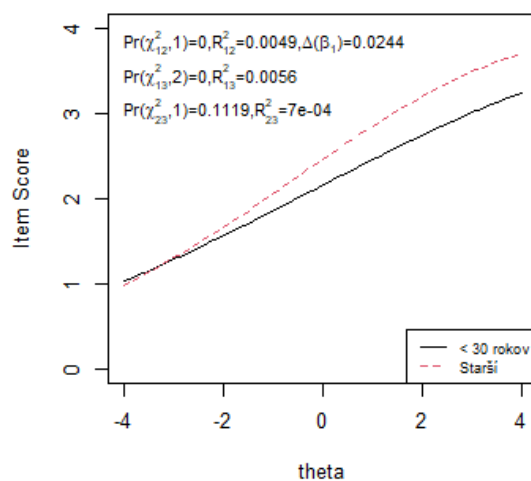
Tabuľka 24 Výsledky analýzy DIF podľa veku vo facetách svedomitosti

Položka	Δ McFadden pseudo-R ²		
	U	NU	C
Organizovanosť			
3R má sklon byť chaotický.	0,005	0,001	0,006
18 je systematický, rád udržiava veci v poriadku.	0,000	0,000	0,000
33 udržiava svoje veci usporiadané a zorganizované.	0,000	0,000	0,000
48R zanecháva neporiadok, neupratuje.	0,001	0,000	0,001
Produktivita			
8R má sklon byť lenivý.	0,007	0,000	0,007
23R má problém začať nejakú úlohu.	0,006	0,000	0,006
38 je efektívny, veci dokončí.	0,005	0,000	0,005
53 je vytrvalý, pracuje pokiaľ nie je úloha dokončená.	0,002	0,000	0,002
Zodpovednosť			
13 je spoľahlivý a dôsledný.	0,000	0,000	0,000
28R môže byť trochu nedbanlivý.	0,000	0,001	0,001
43 je dôveryhodný, dá sa na neho vždy spoľahnúť.	0,001	0,000	0,001
58R sa niekedy správa nezodpovedne.	0,004	0,000	0,004

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo-R² pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

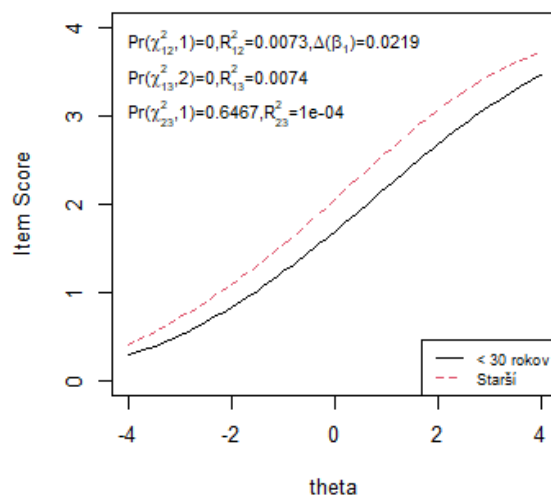
Vo facete Organizovanosť je to reverzná položka 3 „Som niekto, kto má sklon byť chaotický.“ Pri tejto položke dosahujú vyššie hodnoty starší dospelí, čo znamená, že majú nižšiu

tendenciu s obsahom súhlasiť v porovnaní s vynárajúcimi sa dospelými s rovnakou mierou organizovanosti. Charakteristické krivky zobrazujeme v Obrázku 44.

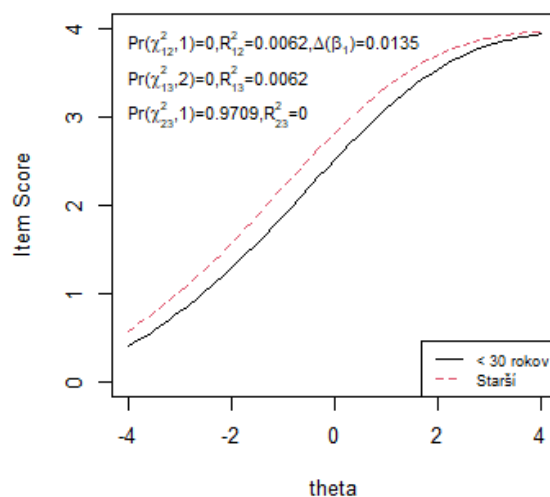


Obrázok 44 Charakteristické krivky položky 3

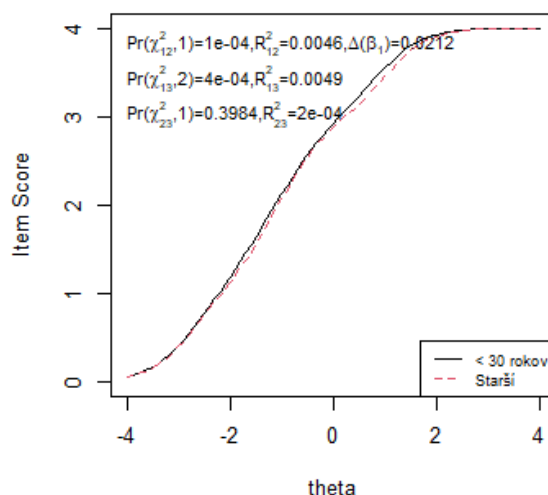
Vo facete Produktivita boli identifikované až tri položky, konkrétne položka 8 „*Som niekto, kto má sklon byť lenivý.*“, položka 23 „*Som niekto, kto má problém začať nejakú úlohu.*“ a položka 38 „*Som niekto, kto je efektívny, veci dokončí.*“. Pri prvých dvoch zmienených položkách dosahujú vyššie hodnoty starší dospelí, čo znamená, že s obsahom menej súhlasia. Odlišné fungovanie tretej položky je z hľadiska vizuálneho posúdenia charakteristických kriviek neprítomné, teda ide najskôr o problém výpočtu, práve z dôvodu očisťovania miery vlastnosti – odlišné fungovanie tejto položky bolo posúdené len na základe hodnôt zvyšnej položky 53. Charakteristické krivky položiek zobrazujeme v Obrázkoch 45 až 47.



Obrázok 45 Charakteristické krivky položky 8

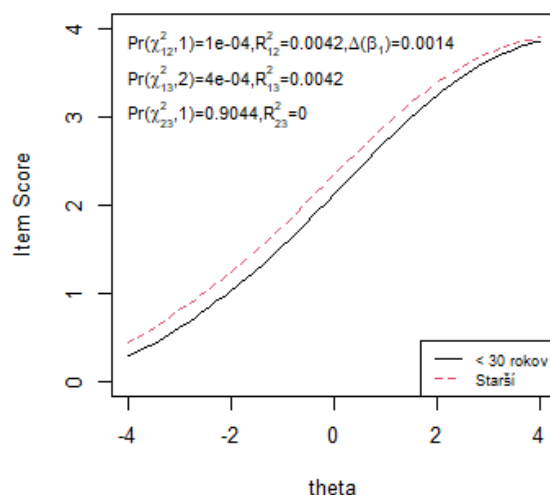


Obrázok 46 Charakteristické krivky položky 23



Obrázok 47 Charakteristické krivky položky 38

Vo facete Zodpovednosť sme identifikovali reverznú položku 58 „*Som niekto, kto sa niekedy správa nezodpovedne.*“ Zaujímavé je, že v tomto súbore doposiaľ identifikované položky v rámci konfirmačnej faktorovej boli identifikované aj prostredníctvom logistickej regresie, no v tomto prípade to tak nie je. Aj v tomto prípade však dosahujú vyššiu hodnotu položky starší dospelí, čo znamená, že menej súhlasia s vlastnou nezodpovednosťou v porovnaní s vynárajúcimi sa dospelými s rovnakou mierou zodpovednosti. Avšak, tento rozdiel vo fungovaní položky je prakticky zanedbateľný, čo je možné vidieť aj na vzdialenosti charakteristických kriviek položky (Obrázok 48).



Obrázok 48 Charakteristické krivky položky 58

7.4. Negatívna emocionalita

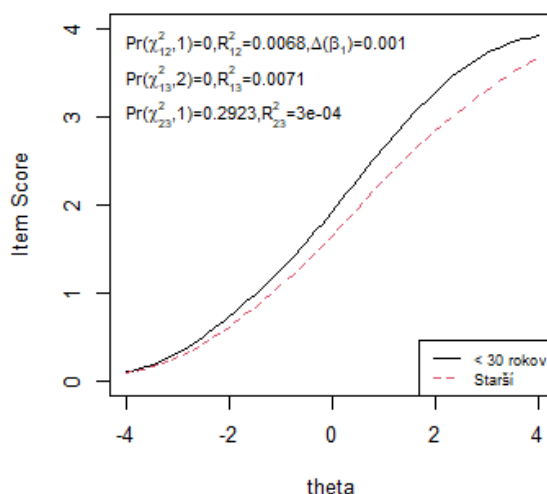
Výsledky z domény Negatívna emocionalita poukazujú na dve uniformne odlišne fungujúce položky vo facete Úzkosť a dve neuniformne fungujúce položky vo facete Emocionálna nestálosť. Odlišné fungovanie všetkých položiek je však prakticky nevýznamné. Výsledky uvádzame v Tabuľke 25. Pri týchto výsledkoch nie je zhoda s výsledkom konfirmačnej faktorovej analýzy, ktorá nepreukázala žiadnu problematickú položku.

Tabuľka 25 Výsledky analýzy DIF podľa veku vo facetách negatívnej emocionality

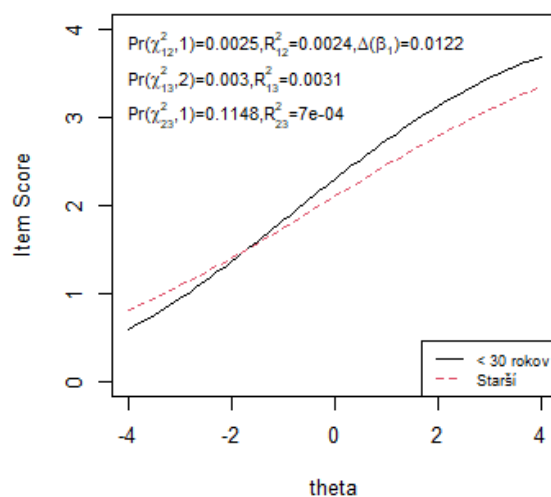
Položka	Δ McFadden pseudo- R^2		
	U	NU	C
Úzkosť			
4R je uvoľnený, dobre zvláda stres.	0,007	0,000	0,007
19 môže bývať napätý.	0,000	0,001	0,002
34 sa veľa znepokojuje.	0,000	0,000	0,000
49R zriedkakedy cíti úzkosť alebo obavy.	0,002	0,001	0,003
Depresia			
9R ostáva optimistický napriek prekážkam.	0,000	0,002	0,002
24R sa cíti sebaistý, je spokojný so sebou.	0,000	0,001	0,001
39 sa často cíti smutný.	0,000	0,000	0,000
54 má sklon cítiť sa sklúčený, skleslý.	0,000	0,000	0,000
Emocionálna nestálosť			
14 je náladový, máva výkyvy nálady.	0,001	0,000	0,001
29R je emocionálne stabilný, neznepokojí sa tak ľahko.	0,000	0,006	0,006
44R má svoje emócie pod kontrolou.	0,000	0,002	0,002
59 je temperamentný, ľahko sa rozčúli.	0,002	0,005	0,007

Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF signifikantná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

Obe položky merajúce úzkosť sú reverzné. Pri položke 4 „*Som niekto, kto je uvoľnený, dobre zvláda stres.*“ dosahujú vynárajúci sa dospelí vyššie hodnoty, čo znamená, že starší dospelí s obsahom tejto položky viac súhlasia. Rovnako je tomu aj v prípade položky 49 „*Som niekto, kto zriedkakedy cíti úzkosť alebo obavy.*“ Táto položka na základe zmeny pseudo- R^2 vykazuje skoro nulové odlišné fungovanie, avšak je signifikantné a väčšie ako hodnota 99. percentilu zo simulovaných dát (0,002). Zaujímavé je tiež, že charakteristické krivky položky poukazujú na vyššiu mieru odlišného fungovania, dokonca s prítomným neuniformným fungovaním, čo naznačuje skôr problém pri výpočte a odhade miery vlastnosti, ako odlišné fungovanie položky. Charakteristické krivky týchto položiek znázorňujeme v Obrázkoch 49 a 50.

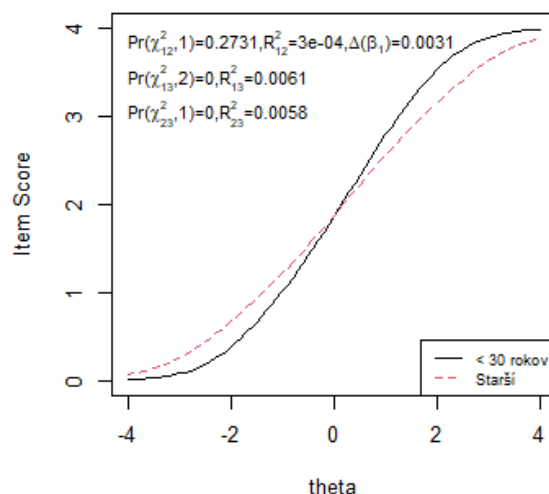


Obrázok 49 Charakteristické krivky položky 4

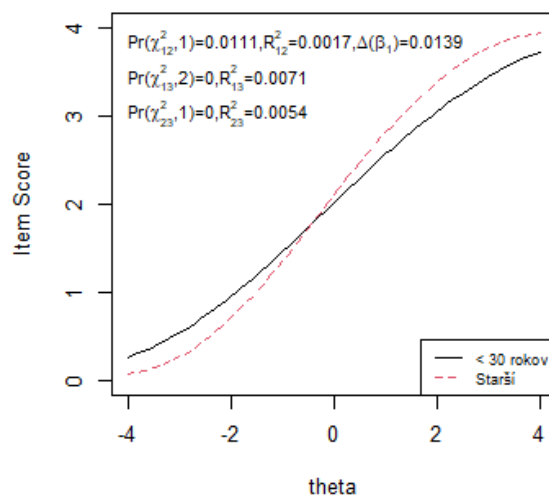


Obrázok 50 Charakteristické krivky položky 49

Štatisticky významné neuniformné odlišné fungovanie vykazovali položky 29 „*Som niekto, kto je emocionálne stabilný, neznepokojí sa tak ľahko.*“ a 59 „*Som niekto, kto je temperamentný, ľahko sa rozčúli.*“ Pri oboch položkách vyjadrovali starší dospelí väčší súhlas s obsahom pri podpriemernej úrovni emocionálnej nestálosti, no menší súhlas pri nadpriemernej úrovni. Charakteristické krivky zobrazujeme v Obrázkoch 51 a 52.



Obrázok 51 Charakteristické krivky položky 29



Obrázok 52 Charakteristické krivky položky 59

7.5. Otvorenosť

Výsledky z faciet domény Otvorenosť poukázali na celkovo 4 odlišne fungujúce položky, 1 z facety Estetická senzitivita, 2 z facety Intelektuálna zvedavosť a 1 z facety

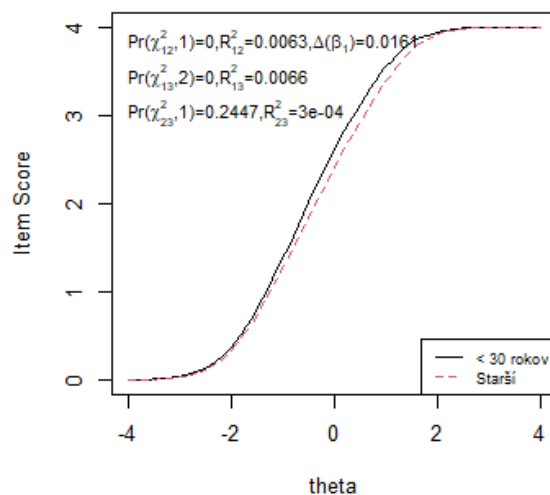
Kreatívna predstavivosť. Opäť, všetky položky vykazujú síce štatisticky významné, no prakticky zanedbateľné odlišné fungovanie a rovnako ako v prípade Negatívnej emocionality sa teda nezhodujú s výsledkom konfirmačnej faktorovej analýzy. Výsledky uvádzame v Tabuľke 26.

Tabuľka 26 Výsledky analýzy DIF podľa veku vo facetách otvorenosti

Položka	Δ McFadden pseudo- R^2		
	U	NU	C
Estetická senzitivita			
5R má málo umeleckých záujmov.	0,000	0,001	0,001
20 je očarený umením, hudbou alebo literatúrou.	0,006	0,000	0,007
35 si cení umenie a krásu.	0,000	0,000	0,000
50R považuje poéziu a divadlo za nudné.	0,000	0,001	0,001
Intelektuálna zvedavosť			
10 je zvedavý na mnoho rozličných vecí.	0,003	0,003	0,005
25R sa vyhýba intelektuálnym, filozofickým diskusiam.	0,000	0,007	0,007
40 rozmýšľa komplexne, do hĺbky.	0,001	0,002	0,003
55R má malý záujem o abstraktné myšlienky.	0,000	0,001	0,002
Kreatívna predstavivosť			
15 je vynaliezavý, hľadá šikovné spôsoby ako robiť veci.	0,000	0,000	0,000
30R je málo tvorivý.	0,005	0,000	0,005
45R má problém predstavovať si veci.	0,000	0,001	0,001
60 je originálny, prichádza s novými nápadmi.	0,000	0,001	0,002

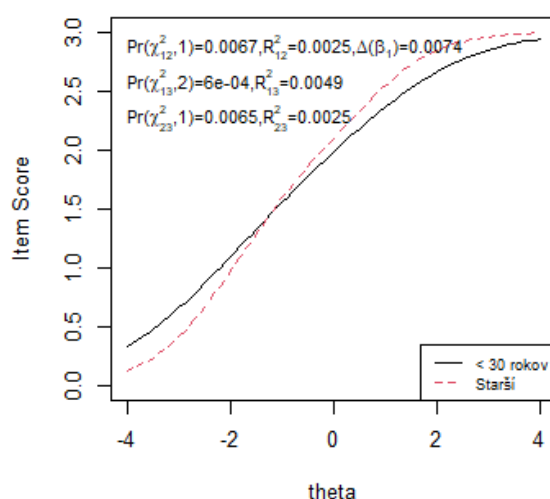
Poznámka. Hodnoty predstavujú zmenu McFaddenovho pseudo- R^2 pri danom type odlišného fungovania položky. U - uniformné DIF, NU - neuniformné DIF, C - celkové DIF. Miera DIF významná pri $p < 0,05$ je zvýraznená kurzívou a $p < 0,01$ je zvýraznená hrubým. R - položka je reverzne bodovaná.

Charakteristické krivky položky 20 „Som niekto, kto je očarený umením, hudbou alebo literatúrou.“ zobrazujeme v Obrázku 53. Ako je možno vidieť, vzdialenosť medzi krivkami je úplne minimálna, pri nízkej miere črty sa dokonca plne prekrývajú.

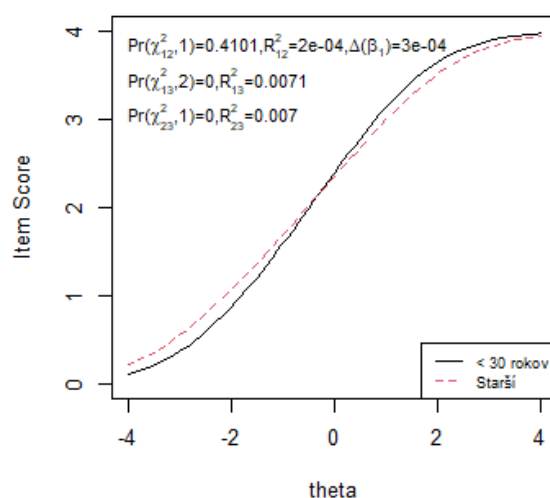


Obrázok 53 Charakteristické krivky položky 20

V Obrázkoch 54 a 55 prezentujeme charakteristické krivky položiek 10 a 25. Položka 25 „*Som niekto, kto je zvedavý na mnoho rozličných vecí.*“ vykazuje odlišné fungovanie v smere vyšších hodnôt u vynárajúcich sa dospelých pri nízkej miere intelektuálnej zvedavosti, no naopak v prípade priemernej a nadpriemernej miere tejto črty. Pri reverznej položke 25 „*Som niekto, kto sa vyhýba intelektuálnym, filozofickým diskusiám.*“ sú charakteristické krivky veľmi blízko pri sebe, čo rovnako ako číselný výsledok naznačuje zanedbateľnú mieru odlišného fungovania. Z hľadiska trendu, pri nízkej intelektuálnej zvedavosti s obsahom viac súhlasia vynárajúci sa dospelí, zatiaľ čo pri nadpriemernej úrovni vyjadrujú vyšší súhlas starší dospelí.

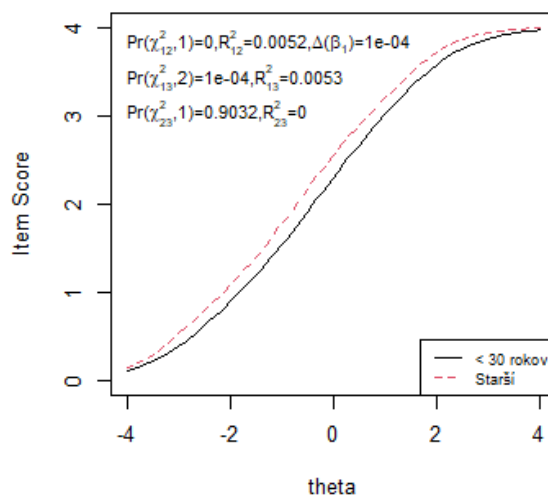


Obrázok 54 Charakteristické krivky položky 10



Obrázok 55 Charakteristické krivky položky 25

Poslednou identifikovanou položkou je reverzná položka 30 „*Som niekto, kto je málo tvorivý.*“ Tak ako v ostatných prípadoch, odlišné fungovanie je len prakticky zanedbateľnej veľkosti. V celej miere kreatívnej predstavivosti dosahujú starší dospelí vyššie hodnoty položky, čo znamená, že vynárajúci sa dospelí majú vyššiu tendenciu súhlasiť s obsahom. Charakteristické krivky tejto položky znázorňuje Obrázok 56.



Obrázok 56 Charakteristické krivky položky 30

8. Zhrnutie a diskusia výsledkov

Cieľom tejto výskumnej monografie, ktorá nadväzuje na náš predchádzajúci výskum (Kohút & Halama, 2017; Kohút & Halama, 2019; Kohút, 2018), bolo zodpovedať jednoduchú otázku, či meriame to isté, u mužov a žien, u vynárajúcich sa dospelých a starších respondentov. „Tým istým“ sme v tomto prípade mali na mysli osobnostné črty veľkej päťky, ktoré sme merali prostredníctvom Inventára Veľkej Päťky 2. Hoci je otázka jednoduchá, odpovedať je pomerne zložité.

8.1. Problémy pri skúmaní rovnosti merania

Existuje množstvo rôznych prístupov ako posúdiť, či daný nástroj meria „to isté“ v rôznych skupinách. V tejto monografii sme využili prístup založený na konfirmačnej faktorovej analýze, ktorý sa zaužívané nazýva invariantnosť merania a prístup založený na ordinálnej logistickej regresii s využitím teórie odpovede na položku, ktorý sa špecificky zameriava na odlišné fungovanie položiek. V našom predchádzajúcom výskume (Kohút, 2018) zameranom len na odlišné fungovanie položiek medzi mužmi a ženami sme použili aj iné metódy analýzy odlišného fungovania položiek. Celkovo sme v tomto výskume identifikovali pomerne veľké množstvo odlišne fungujúcich položiek (27-52 %) v štyroch analyzovaných osobnostných dotazníkoch. Toto číslo môže navodzovať dojem, že tieto osobnostné dotazníky nefungujú správne, a teda *nemeriame „to isté.“* Komplikovanosť tejto problematiky spočíva v tom, akú metódu analýzy zvolíme. Rôzne metódy majú svoje silné ale aj slabé stránky, ktoré ovplyvňujú ich presnosť a spoľahlivosť pri reálnych (a nedokonalých) dátach. Okrem toho, akú metódu analýzy zvolíme, záleží aj na kritériu – teda nejakej hodnote výsledku, kedy usúdime, že dochádza k narušeniu rovnosti merania. Hoci sú pri niektorých metódach zaužívané normy, ich prísnosť je rôznorodá. Príkladom môže byť posúdenie invariantnosti merania – môžeme sa zamerať na štatistické porovnanie modelov, prostredníctvom chí-kvadrát testu, ktorý je však veľmi striktný a môže dochádzať k chybe I. typu, teda zisteniu invariantnosti, ktorá je prakticky zanedbateľná, teda nemá dopad na praktické využitie výsledkov merania. Možným východiskom je posúdenie zmeny alternatívnych indikátorov zhody modelu s dátami – tento prístup sme využili aj v našom výskume. Aj v tomto prípade je však otázne, aké kritéria nastaviť a prebiehajúca diskusia naznačuje, že odpoveď nie je jednoznačná (Putnick & Bornstein, 2016). V predchádzajúcom výskume (Kohút, 2018) sme zvolili dostupné kritéria pre označenie položky ako odlišne fungujúcej, no pre limitáciu chyby I. typu sme sa zaoberali najmä položkami, ktoré boli označené aspoň dvomi z troch použitých metód skúmania ako prakticky

nie zanedbateľné. Pri takomto prístupe sa počet položiek vykazujúcich odlišné fungovanie výrazne znížil, napríklad z 31 na 9 v NEO-FFI, z 12 na 3 v BFI alebo z 25 na 9 v BFI-2. Tento jav poukazuje na nezhodu metód, kritérií hodnotenia závažnosti DIF a problematickosť vyvodenia záverov na základe použitia jednej metódy analýzy.

8.2. Konzistentnosť identifikovaných položiek

V tejto monografii sme sa na problematiku rovnosti merania u mužov a žien pozreli nie tak z hľadiska konzistencie rôznych metód analýzy ale najmä z hľadiska konzistencie identifikovaných položiek v čase a rôznych súboroch. Pokiaľ je rovnosť merania o psychodiagnostickej metodike, očakávali by sme, že budú identifikované rovnaké problematické položky. Pre získanie ešte robustnejších výsledkov sme použili dva retestové súbory – teda súbory tých istých respondentov, ktoré boli merané v rôznom čase. Celkovo sme identifikovali 26 (43,3 %) položiek, ktoré boli problematické z hľadiska invariantnosti merania alebo vykazovali akékoľvek odlišné fungovanie. Dvadsaťštyri položiek vykazovalo odlišné fungovanie položiek. Desať položiek bolo identifikovaných prostredníctvom konfirmačnej faktorovej analýzy, 8 z nich bolo identifikovaných aj prostredníctvom logistickej regresie. Väčšina z nich však nebola konzistentná jednak medzi metódami analýzy, ani medzi súbormi a dokonca ani medzi zbermi u tých istých respondentov. Jedným z dôvodov môže byť použitie kritéria štatistickej významnosti. Použitá metóda analýzy odlišného fungovania položiek môže byť pri tomto kritériu veľmi citlivá aj na drobné rozdiely medzi porovnávanými skupinami. Tento trend je pomerne viditeľný pri našich výsledkoch analýzy odlišného fungovania položiek medzi mladšími a staršími dospelými. Tento súbor bol väčší ako súbory, na ktorých sme skúmali rovnosť merania u mužov a žien, čo sa odrazilo na tom, že štatisticky významne odlišne fungujúce položky boli identifikované aj pri veľmi nízkej, prakticky úplne zanedbateľnej miere odlišného fungovania. Pre zúženie okruhu položiek sa môžeme zamerať len na tie, ktoré aspoň v jednom zo súborov vykazovali celkové odlišné fungovania na úrovni aspoň slabo prakticky významného. Pri kritériu zmeny McFaddenovho pseudo- R^2 o 0,02 sa počet položiek identifikovaných prostredníctvom logistickej regresie znížil z 24 na 12. Tento trend je podobný tomu z predchádzajúceho výskumu (Kohút, 2018). Avšak, aj napriek použitiu prísnejšieho kritéria, je väčšina položiek nekonzistentná medzi súbormi.

8.3. Meriame to isté vzhľadom na pohlavie?

Zo všetkých identifikovaných položiek vykazujú 3 pomerne dobrú konzistenciu. Položka 3 z facety Organizovanosť (doména Svedomitosť) „*Som niekto, kto má sklon byť chaotický.*“ bola identifikovaná ako problematická prostredníctvom konfirmačnej faktorovej analýzy v oboch zberoch oboch retestových súborov. Uniformné odlišné fungovanie vykazovala v oboch zberoch druhého retestového súboru a prvom zbere prvého súboru. V druhom zbere prvého retestového súboru ani v súbore vynárajúcich sa dospelých identifikovaná nebola. Z rovnakej facety je aj druhá pomerne konzistentná položka 48 „*Som niekto, kto zanecháva neporiadok, neupratuje.*“ Táto položka bola identifikovaná prostredníctvom konfirmačnej faktorovej analýzy síce len v druhom zbere prvého retest súboru, no vykazovala prakticky nie zanedbateľné uniformné odlišné fungovanie v oboch zberoch oboch retestových súborov. Keďže sú tieto položky z jednej facety, sú si obsahovo blízke, no napriek tomu vykazujú opačný smer odlišného fungovania. Ženy viac vyjadrujú súhlas s vlastnou chaotickosťou a muži viac súhlasia s tým, že zanechávajú neporiadok a neupratujú. Hoci sú si položky blízke, prvá z nich sa zameriava na chaotickosť inak povedané neusporiadanosť, ktorú možno chápať ako opak systematickosti. Táto položka bola identifikovaná aj v našej dizertačnej práci, v rámci ktorej sme skúmali aj pochopenie položky respondentmi. Pochopenie tejto položky bolo v zmysle neporiadku, nedostatku systematickosti a organizovanosti, či už vo vlastných veciach, myšlienkach či správaní. Výsledky naznačili aj vnímaný pozitívny význam chaotickosti, v zmysle spontánnosti jednotlivca. Rozdiely v pochopení položky však neboli zistené (Kohút, 2018). Možným vysvetlením odlišného fungovania môžu byť zistenia ohľadom vyššej tendencie k systemizácii u mužov (Baron-Cohen et al., 2003) alebo zistenia ohľadom nižšej racionality žien z oblasti behaviorálnej ekonómie (García-Gallego et al., 2012; Kuhn & Holling, 2009). Odlišné fungovanie druhej položky môže byť spôsobené rozdielom v zapájaní sa mužov a žien do domácich prác, a to ako kvantitatívny¹ tak aj kvalitatívny. Hoci sa podľa niektorých výskumov rozdiel medzi rodmi v tejto oblasti znižuje, muži žijúci s partnerkou sa častejšie v domácnostiach zameriavajú napr. na údržbu a opravu spotrebičov a fyzicky náročnejšie činnosti spojené s potrebou väčšej fyzickej sily, zatiaľ čo iné rutinné domáce práce sú skôr doménou žien, čo je spojené aj s kratším časom, ktorý muži týmto činnostiam venujú (napr. Bianchi et al., 2012; Treas & Tai, 2016). Táto položka bola rovnako identifikovaná aj v predchádzajúcom výskume (Kohút, 2018). Zaujímavé je, že tieto

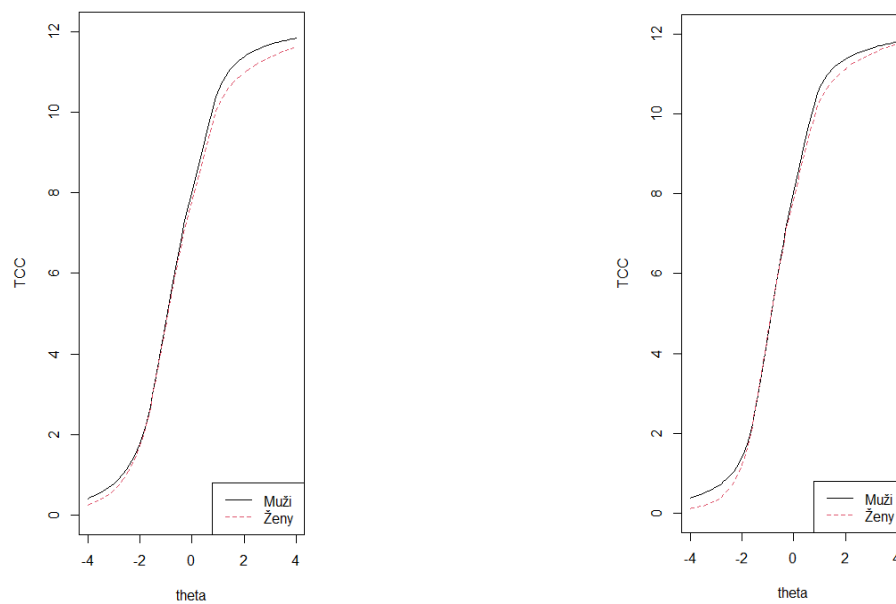
¹ Na základe výsledkov prieskumu Európskeho inštitútu pre rodovú rovnosť sa muži výrazne menej zapájajú do starostlivosti o domácnosť a prípravy pokrmov:
https://eige.europa.eu/gender-statistics/dgs/indicator/ta_timeuse_hswk__eqls_famcareactiv__cook

položky neboli identifikované v súbore mladých dospelých, čo môže viesť k diskusii o určitej spoločenskej žiadúcnosti odpovedí u starších mužov a žien či generačnému posunu.

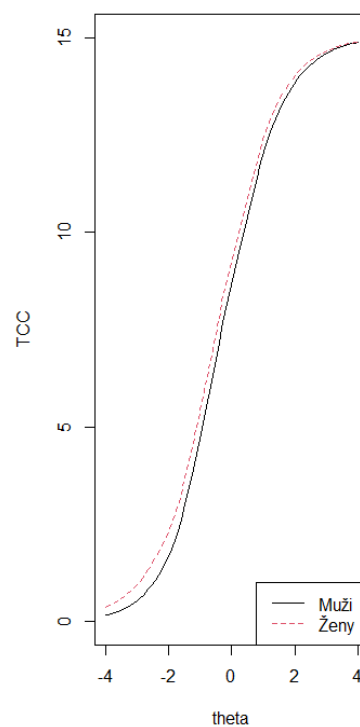
Posledná položka, ktorá vykazovala najvyššiu stabilitu odlišného fungovania, je položka 50 z facety Estetická senzitivita domény Otvorenosť: „*Som niekto, kto považuje divadlo a poéziu za nudné.*“ Táto položka bola identifikovaná ako problematická prostredníctvom konfirmačnej faktorovej analýzy v oboch zberoch oboch retestových súboroch. Prakticky nie zanedbateľné odlišné fungovanie bolo však identifikované len v prvom retestovom súbore, no bola identifikovaná ako štatisticky významná pri zmene McFaddenovho pseudo-R² väčšej ako 0,01 aj v druhom retestovom súbore a súbore vynárajúcich sa dospelých. Pri rovnakej miere estetickej senzitivity majú muži vyšší sklon s touto položkou súhlasiť v porovnaní so ženami. Možným problémom tejto položky je jej špecifickosť, keďže ide o jedinú položku z tejto facety, ktorá menuje špecifické formy umenia. Je možné, že muži vyhľadávajú iné formy umenia pre saturovanie ich potreby po estetickom zážitku. Táto položka bola rovnako identifikovaná aj v predchádzajúcom výskume. V rámci kvalitatívneho skúmania sa nezistil výrazný rozdiel v pochopení alebo procese odpovedania na položku, no respondenti vnímali problematickosť odpovede vzhľadom na rozdielne menované formy umenia, ktoré sú spojené v jednej položke (Kohút, 2018).

Meriame teda to isté? Hoci bolo identifikovaných mnoho problematických položiek, prikláňame sa k odpovedi áno – výsledky z rôznych súborov sa líšia a konzistencia identifikovaných problematických položiek je veľmi nízka, najmä s prihliadnutím na podobnosť retestových súborov. Okrem toho, je potrebné brať v úvahu aj dopad odlišného fungovania položiek na testové skóre, ktoré je vo väčšine prípadov zanedbateľné (podobne ako odlišné fungovanie položiek). V Obrázku 57 zobrazujeme charakteristické krivky facety Organizovanosť. Naľavo je zobrazená krivka z prvého zberu druhého retestového súboru, v ktorom bola identifikovaná len položka 3 ako vykazujúca prakticky nie zanedbateľné DIF. Napravo je zobrazená krivka z druhého zberu druhého retestového súboru. V tomto súbore vykazovali obe položky odlišné fungovanie na úrovni prakticky nie zanedbateľného. Napriek rozdielom v pomeroch odlišného fungovania položiek, sú tieto krivky veľmi podobné, no čo je dôležitejšie, rozdiely v testovom skóre pri rovnakej miere organizovanosti sú prevažne identické. V tomto prípade sa opačný trend odlišného fungovania položiek vzájomne ruší. Obrázok 58 zobrazuje charakteristickú krivku Estetickej senzitivity z prvého zberu prvého retestového súboru. V tomto prípade faceta neobsahuje iné odlišne fungujúce položky, a uniformné odlišné fungovanie položky 50 je najsilnejšie zo všetkých súborov. V tomto

prípade je možné vidieť, že sa krivky u mužov a žien neprekrývajú, ich vzájomná vzdialenosť je však minimálna.



Obrázok 57 Príklad charakteristických kriviek facety Organizovanosť



Obrázok 58 Príklad charakteristickej krivky facety Estetická senzitivita

8.4. Meriame to isté vzhľadom na vek?

Celkovo sme pri skúmaní rovnosti merania medzi vynárajúcimi sa a staršími dospelými identifikovali 24 (40 %) problematických položiek. Prostredníctvom konfirmačnej faktorovej analýzy to boli 3 položky, z toho 2 boli identifikované aj prostredníctvom ordinálnej logistickej regresie. Vysoký počet identifikovaných odlišne fungujúcich položiek súvisí s použitím kritéria štatistickej významnosti – kvôli vysokému počtu respondentov bola analýza príliš citlivá aj na drobné rozdiely, čo sa prejavilo tým, že všetky položky s výnimkou jednej preukazujú prakticky zanedbateľnú úroveň odlišného fungovania. Spomenutou výnimkou je položka 31 z facety Sociabilita domény Extraverzia, ktorá hovorí o vlastnej hanblivosti a uzavretosti. Napriek rovnakej miere sociability mali vynárajúci sa dospelí vyššiu tendenciu súhlasiť s touto položkou. Možným vysvetlením sú vývinové zmeny v hanblivosti počas dospelosti. Longitudinálna štúdia od Dennissen et al., (2007) poukázala na pokles hanblivosti od detského veku k vynárajúcej sa dospelosti. Tento trend má súvis s celkovým vyzrievaním osobnosti ale aj osvojovaním si nových sociálnych rolí, napríklad zamestnaním sa. Zistenia prierezovej štúdie na súbore vyše 500 tisíc britských respondentov od van Zalk et al., (2017) ukázali, že tento trend pokračuje ďalej do neskorších fáz dospelosti, a teda priemerná hanblivosť mierne klesá s pribúdajúcim vekom. Odlišné fungovanie položky 31 je najmä uniformné a najväčší rozdiel je pri podpriemernej úrovni sociability. Predpokladáme, že v období vynárajúcej sa dospelosti je nízka sociabilita ľudí viac spojená s ich hanblivosťou, zatiaľ čo vo vyššom veku možno nezaujmom o stretávanie sa s druhými. Napriek tomu, že táto položka vykazovala z hľadiska pseudo- R^2 najvyššiu mieru odlišného fungovania, nebola identifikovaná prostredníctvom konfirmačnej faktorovej analýzy.

Prostredníctvom konfirmačnej faktorovej analýzy sme vo facete Sociabilita identifikovali inú položku vykazujúcu uniformné odlišné fungovanie, a to položku 16, ktorá hovorí o sklone k tichosti. S touto položkou vyjadrovali väčší súhlas starší dospelí, čo naznačuje, že napriek rovnakej miere sociability majú mladí ľudia tendenciu byť menej tichí, čo môže súvisieť aj s odlišným fungovaním ďalšej položky z tejto facety, ktorá hovorí o zhovorčivosti. Z ďalších domén sme prostredníctvom konfirmačnej faktorovej analýzy identifikovali položku 37 z facety Zdvorilosť (Prívetivosť), ktorá hovorí o hrubosti k iným. S touto položkou vyjadrujú vyšší súhlas vynárajúci sa dospelí, čo môže súvisieť so zvýšenou agresivitou voči druhým v období životných tranzícií, akou je aj prechod do dospelosti (Liu et al., 2013). Posledná identifikovaná položka z facety Zodpovednosť domény Svedomitosť

hovorí o miernej nedbanlivosti osoby. Táto položka však nebola identifikovaná prostredníctvom ordinálnej lineárnej regresie.

Odpovedať na položenú otázku ohľadom rovnosti merania medzi vynárajúcimi sa dospelými a staršími respondentmi nie je jednoduché a z povahy tohto výskumu je kriticky povedané, nemožné. Hlavnou príčinou nášho záveru nie je tak počet identifikovaných položiek, no najmä poučenie zo skúmania rovnosti merania medzi mužmi a ženami. Vzhľadom na zistenú nekonzistentnosť výsledkov pri viacerých súboroch, je náš výskum rovnosti medzi vekovými skupinami len pilotnou sondou do tejto problematiky. Z tohto dôvodu hodnotíme, že sú potrebné viaceré výskumy, ktoré by skúmali rovnosť merania v súvislosti s vekom jednak použitím viacerých metód analýzy ale aj viacerých, ideálne retestových súborov, vďaka čomu by bolo možné získať robustnejšie závery. Domnievame sa, že sa s vysokou pravdepodobnosťou vyskytuje veľmi malé množstvo položiek v Inventári Veľkej Päťky 2, ktoré v nie zanedbateľnej miere fungujú odlišne v súvislosti s vekom, no rovnako sa domnievame, že dopad týchto položiek na rovnosť merania a validitu inventára je len malý. V tejto fáze poznania však ide len o našu domnienku, vytvorenú na základe výsledkov našich výskumov prezentovaných v tejto monografii, našej dizertačnej práci (Kohút, 2018) a skúseností z publikovaných výskumných štúdií (Kohút & Halama, 2017; Kohút & Halama, 2019). Vzhľadom na validitu metodiky je dôležité tiež si uvedomiť, že nie každá položka vykazujúca odlišné fungovanie musí *nutne* ohrozovať validitu psychodiagnostickej metodiky. Odlišné fungovanie položiek sa dá odôvodniť aj viacdimezionalitou skúmaného konštrukt, t.j. viacerými faktormi, ktoré spolu úzko súvisia, no zároveň je možné ich samostatne identifikovať, a porovnávané skupiny sa v týchto nuansách konštruktu reálne líšia. Odstránenie takýchto položiek môže znamenať zužovanie meraného konštrukt a vytváranie *umelej* rovnosti medzi porovnávanými skupinami.

8.5. Limitácie výskumu a odporúčania pre ďalšie výskumy a prax

Možnou limitáciou nášho výskumu je použitie metódy analýzy odlišného fungovania, ktorá pre výpočet používa model vychádzajúci z teórie odpovede na položku. Tento odhad vlastnosti môže byť skreslený kvôli nízkemu počtu položiek vo facetách. Podľa štúdie od Dai et al., (2021) je pre optimálne fungovanie tohto modelu potrebných 5 položiek, zatiaľ čo v facety Inventára Veľkej Päťky 2 obsahujú len 4 položky.

Našimi odporúčaniami pre ďalšie výskumy v oblasti rovnosti merania je najmä použitie viacerých metód analýzy a použitie retestových súborov pre limitáciu náhodne

identifikovaných položiek. Kľúčová je tiež voľba kritéria pre uznanie položky ako problematickej, hoci si uvedomujeme, že je najskôr potrebný pokrok v oblasti štatistiky a psychometrie, ktorý by záujemcom o túto problematiku poskytol jasnejšie kritéria, ktoré je vhodné zvoliť.

Za potrebné považujeme uviesť aj všeobecnejšie odporúčania pre psychologické výskumy a prax. Na základe výsledkov odporúčame používateľom psychodiagnostických metodík – či už vo výskume alebo psychologickej praxi, vyššiu obozretnosť pri používaní psychodiagnostických metodík. Používatelia by mali byť detailne oboznámení s obsahom použitých metodík – nielen *na akú* oblasť je metodika zameraná, ale aj *ako* je táto oblasť operacionalizovaná v podobe jednotlivých položiek. Obozretnosť je dôležitá aj pri interpretácii testových skóre, pričom je vhodné zohľadniť konkrétne obsahy položiek, ako aj pomer rôznych tém v položkách, ktoré metodika obsahuje. Na strane autorov nových psychodiagnostických metodík vnímame za potrebné osvojenie si konceptu odlišného fungovania položiek v procese prípravy položiek a pri overovaní psychometrických vlastností týchto metodík. Pri príprave položiek je dôležité samotné uvedomenie si existencie odlišného fungovania položiek, ktoré môže dopomôcť pri rozhodovaní o obsahoch, formuláciách, ale aj zastúpení jednotlivých položiek.

Záver

Na základe prezentovaných výsledkov hodnotíme, že Inventár Veľkej Päťky 2 je z hľadiska rovnosti merania mužov a žien vhodný pre použitie, keďže nevykazuje výraznú mieru odlišného fungovania medzi týmito skupinami vo všeobecnej populácii alebo populácii vynárajúcich sa dospelých. Rovnaký záver môžeme vyvodiť aj v prípade rovnosti merania medzi vynárajúcimi sa dospelými a staršími dospelými, aj keď vzhľadom na použitie len jedného súboru je tento záver značne limitovaný.

Použitá literatura

- AERA, A. A. (2001). *Standardy pro pedagogické a psychologické testování*. Testcentrum.
- Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, 47(1), i–171. <https://doi.org/10.1037/h0093360>
- Asparouhov, T., & Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 397–438. <https://doi.org/10.1080/10705510903008204>
- Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1430), 361–374. <https://doi.org/10.1098/rstb.2002.1206>
- Berthelsen, H., Hakanen, J. J., Kristensen, T. S., Lönnblad, A., & Westerlund, H. (2016). A Qualitative Study on the Content Validity of the Social Capital Scales in the Copenhagen Psychosocial Questionnaire (COPSOQ II).
- Bianchi, S. M., Sayer, L. C., Milkie, M. A., & Robinson, J. P. (2012). Housework: Who Did, Does or Will Do It, and How Much Does It Matter? *Social Forces*, 91(1), 55–63. <https://doi.org/10.1093/sf/sos120>
- Bumberry, W., Oliver, J. M., & McClure, J. N. (1978). Validation of the Beck Depression Inventory in a university population using psychiatric estimate as the criterion. *Journal of Consulting and Clinical Psychology*, 46(1), 150–155. <https://doi.org/10.1037/0022-006X.46.1.150>
- Camilli, G. (2006). Test fairness. In *Educational measurement* (pp. 221–256).
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). **lordif**: An R Package for Detecting Differential Item Functioning Using Iterative Hybrid Ordinal Logistic Regression/Item Response Theory and Monte Carlo Simulations. *Journal of Statistical Software*, 39(8). <https://doi.org/10.18637/jss.v039.i08>
- Clauser, B. E., & Mazor, K. M. (2005). Using Statistical Procedures to Identify Differentially Functioning Test Items. *Educational Measurement: Issues and Practice*, 17(1), 31–44. <https://doi.org/10.1111/j.1745-3992.1998.tb00619.x>

- Crane, P. K., Gibbons, L. E., Jolley, L., & van Belle, G. (2006). Differential Item Functioning Analysis With Ordinal Logistic Regression Techniques. *Medical Care*, 44(Suppl 3), S115–S123. <https://doi.org/10.1097/01.mlr.0000245183.28384.ed>
- Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 43(3), 245–265. <https://doi.org/10.1348/0144665031752934>
- Dai, S., Vo, T. T., Kehinde, O. J., He, H., Xue, Y., Demir, C., & Wang, X. (2021). Performance of Polytomous IRT Models With Rating Scale Data: An Investigation Over Sample Size, Instrument Length, and Missing Data. *Frontiers in Education*, 6. <https://doi.org/10.3389/educ.2021.721963>
- Dennissen, J. J. A., Asendorpf, J. B., & van Aken, M. A. G. (2007). Childhood Personality Predicts Long-Term Trajectories of Shyness and Aggressiveness in the Context of Demographic Transitions in Emerging Adulthood. *Journal of Personality*, 76(1), 67–100. <https://doi.org/10.1111/j.1467-6494.2007.00480.x>
- Downing, S. M. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education*, 37(9), 830–837. <https://doi.org/10.1046/j.1365-2923.2003.01594.x>
- Einarsdóttir, S., & Rounds, J. (2009). Gender bias and construct validity in vocational interest measurement: Differential item functioning in the Strong Interest Inventory. *Journal of Vocational Behavior*, 74(3), 295–307. <https://doi.org/10.1016/j.jvb.2009.01.003>
- Ellis, B. B. (1989). Differential item functioning: Implications for test translations. *Journal of Applied Psychology*, 74(6), 912–921. <https://doi.org/10.1037/0021-9010.74.6.912>
- Gallardo-Pujol, D., Rouco, V., Cortijos-Bernabeu, A., Ocejja, L., Soto, C., & John, O. (2021). Factor structure, gender invariance, measurement properties and short forms of the Spanish adaptation of the Big Five Inventory-2 (BFI-2). PsyArXiv. <https://doi.org/10.31234/osf.io/nxr4q>
- García-Gallego, A., Georgantzís, N., & Jaramillo-Gutiérrez, A. (2012). Gender differences in ultimatum games: Despite rather than due to risk attitudes. *Journal of Economic Behavior & Organization*, 83(1), 42–49. <https://doi.org/10.1016/j.jebo.2011.06.012>
- Gogniat, M. A., Rodriguez, V. J., Granros, M., Jean, K. R., Robinson, T. L., & Miller, L. S. (2022). Differential Item Functioning: An Examination of the NEO-FFI by Sex in Older Adults. *SAGE Open*, 12(1), 215824402210866. <https://doi.org/10.1177/21582440221086607>

- Halama, P. (2011). *Princípy psychologické diagnostiky* (2. doplnené vydanie). Filozofická fakulta Trnavskej univerzity v Trnave.
- Halama, P., Kohút, M., Soto, C. J., & John, O. P. (2020). Slovak Adaptation of the Big Five Inventory (BFI-2): Psychometric Properties and Initial Validation. *Studia Psychologica*, 62(1). <https://doi.org/10.31577/sp.2020.01.792>
- Havan, P., Kohút, M., & Halama, P. (2022). Domain and method-specificity of acquiescence estimated in different measures. *Personality and Individual Differences*, 187, 111437. <https://doi.org/10.1016/j.paid.2021.111437>
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In *Test validity*. (pp. 129–145). Lawrence Erlbaum Associates, Inc. <https://doi.org/10.1037/14047-004>
- Hřebíčková, M. (2004). *Pětifaktorový model v psychologii osobnosti. Přístupy, diagnostika, uplatnění*. Grada.
- Hrebickova, M., Jelinek, M., Kveton, P., Benkovic, A., Botek, M., Sudzina, F., Soto, C. J., & John, O. P. (2020). Big Five Inventory 2 (BFI-2): Hierarchical model. *CESKOSLOVENSKA PSYCHOLOGIE*, 64(4), 437–460.
- Jelínek, M., Květon, P., & Vobořil, D. (2011). *Testování v psychologii. Teorie odpovědi na položku a počítačové adaptivní testování*. Grada Publishing.
- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type I Error and Power Rates Using an Effect Size Measure With the Logistic Regression Procedure for DIF Detection. *Applied Measurement in Education*, 14(4), 329–349. https://doi.org/10.1207/S15324818AME1404_2
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In *Handbook of personality: Theory and research*, 3rd ed. (pp. 114–158). The Guilford Press.
- John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In *Handbook of research methods in personality psychology*. (pp. 461–494). The Guilford Press.
- Karami, H., & Salmani Nodoushan, M. A. (2011). Differential Item Functioning (DIF): Current problems and future directions. *International Journal of Language Studies*, 5, 133–142.
- Kohút, M. (2018). *Odlišné fungovanie položiek u mužov a žien v Eysenckovom osobnostnom dotazníku a vybraných päťfaktorových inventároch*. Filozofická fakulta, Trnavská univerzita v Trnave.

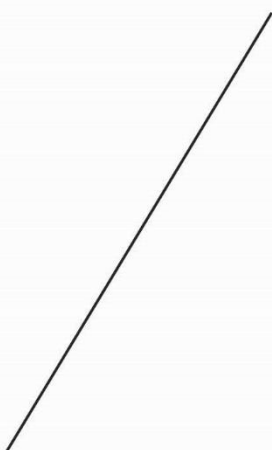
- Kohut, M., & Halama, P. (2017). Gender uniform differential item functioning (DIF) in Eysenck Personality Questionnaire: comparison of three analysis methods. *CESKOSLOVENSKA PSYCHOLOGIE*, 61(6), 559–574.
- Kohút, M., & Halama, P. (2019). GENDER DIFFERENTIAL ITEM FUNCTIONING IN SLOVAK VERSION OF NEO FIVE FACTOR INVENTORY. *Československá Psychologie*, 63(1), 71–83.
- Kohut, M., Halama, P., Soto, C. J., & John, O. P. (2020). PSYCHOMETRIC PROPERTIES OF SLOVAK SHORT AND EXTRA-SHORT FORMS OF BIG FIVE INVENTORY-2 (BFI-2). *CESKOSLOVENSKA PSYCHOLOGIE*, 64(5), 550–563.
- Kuhn, J.-T., & Holling, H. (2009). Gender, reasoning ability, and scholastic achievement: A multilevel mediation analysis. *Learning and Individual Differences*, 19(2), 229–233. <https://doi.org/10.1016/j.lindif.2008.11.007>
- Lai, K., & Green, S. B. (2016). The Problem with Having Two Watches: Assessment of Fit When RMSEA and CFI Disagree. *Multivariate Behavioral Research*, 51(2–3), 220–239. <https://doi.org/10.1080/00273171.2015.1134306>
- LIU, J., LEWIS, G., & EVANS, L. (2013). Understanding aggressive behaviour across the lifespan. *Journal of Psychiatric and Mental Health Nursing*, 20(2), 156–168. <https://doi.org/10.1111/j.1365-2850.2012.01902.x>
- Lund, R., Nielsen, L. S., Henriksen, P. W., Schmidt, L., Avlund, K., & Christensen, U. (2014). Content validity and reliability of the Copenhagen social relations questionnaire. *Journal of Aging and Health*, 26(1), 128–150. <https://doi.org/10.1177/0898264313510033>
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, 93(3), 568–592. <https://doi.org/10.1037/0021-9010.93.3.568>
- Miller, J. D., McCain, J., Lynam, D. R., Few, L. R., Gentile, B., MacKillop, J., & Campbell, W. K. (2014). A comparison of the criterion validity of popular measures of narcissism and narcissistic personality disorder via the use of expert ratings. *Psychological Assessment*, 26(3), 958–969. <https://doi.org/10.1037/a0036613>
- Nye, C. D., Allemand, M., Gosling, S. D., Potter, J., & Roberts, B. W. (2016). Personality Trait Differences Between Young and Middle-Aged Adults: Measurement Artifacts or Actual Trends? *Journal of Personality*, 84(4), 473–492. <https://doi.org/10.1111/jopy.12173>
- Osterlind, S., & Everson, H. (2009). *Differential Item Functioning*. SAGE Publications, Inc. <https://doi.org/10.4135/9781412993913>

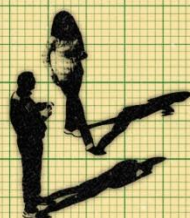
- Penfield, R. D., & Camilli, G. (2006). 5 *Differential Item Functioning and Item Bias* (pp. 125–167). [https://doi.org/10.1016/S0169-7161\(06\)26005-X](https://doi.org/10.1016/S0169-7161(06)26005-X)
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. In *Developmental Review* (Vol. 41, pp. 71–90). Mosby Inc. <https://doi.org/10.1016/j.dr.2016.06.004>
- Raju, N. S. (1988). The area between two item characteristic curves. *Psychometrika*, 53(4), 495–502. <https://doi.org/10.1007/BF02294403>
- Říčan, P. (1977). *Úvod do psychometrie*. Psychodiagnostické a didaktické testy.
- Rosseel, Y. (2012). **lavaan**: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2). <https://doi.org/10.18637/jss.v048.i02>
- Roussos, L., & Stout, W. (1996). A Multidimensionality-Based DIF Analysis Paradigm. *Applied Psychological Measurement*, 20(4), 355–371. <https://doi.org/10.1177/014662169602000404>
- RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. <http://www.rstudio.com/>
- Ruisel, I., & Halama, P. (2007). *NEO päťfaktorový osobnostný inventár (podľa NEO Five Factor Inventory P. T. Costu a R. R. McCraeho)*. Testcentrum-Hogrefe.
- Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, 94(1), 168–182. <https://doi.org/10.1037/0022-3514.94.1.168>
- Shchebetenko, S., Kalugin, A. Y., Mishkevich, A. M., Soto, C. J., & John, O. P. (2020). Measurement Invariance and Sex and Age Differences of the Big Five Inventory–2: Evidence From the Russian Version. *Assessment*, 27(3), 472–486. <https://doi.org/10.1177/1073191119860901>
- Sheppard, R., Han, K., Colarelli, S. M., Dai, G., & King, D. W. (2006). Differential Item Functioning by Sex and Race in the Hogan Personality Inventory. *Assessment*, 13(4), 442–453. <https://doi.org/10.1177/1073191106289031>
- Smith, L. L. (2002). On the Usefulness of Item Bias Analysis to Personality Psychology. *Personality and Social Psychology Bulletin*, 28(6), 754–763. <https://doi.org/10.1177/0146167202289005>
- Sollár, T. (2014). *Empirická verifikácia štrukturálnych modelov psychologických konštruktov*. S.A.P. Slovak Academic Press, s.r.o.

- 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(1), 117–143. <https://doi.org/10.1037/pspp0000096>
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age Differences in Personality Traits From 10 to 65: Big Five Domains and Facets in a Large Cross-Sectional Sample. *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY*, 100(2), 330–348. <https://doi.org/10.1037/a0021717>
- St-Onge, C., Young, M., Eva, K. W., & Hodges, B. (2017). Validity: one word with a plurality of meanings. *Advances in Health Sciences Education*, 22(4), 853–867. <https://doi.org/10.1007/s10459-016-9716-3>
- Tóth-Király, I., Bőthe, B., Rigó, A., & Orosz, G. (2017). An Illustration of the Exploratory Structural Equation Modeling (ESEM) Framework on the Passion Scale. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01968>
- Treas, J., & Tai, T. (2016). Gender Inequality in Housework Across 20 European Nations: Lessons from Gender Stratification Theories. *Sex Roles*, 74(11–12), 495–511. <https://doi.org/10.1007/s11199-015-0575-9>
- Urbánek, T., & Šimeček, M. (2001). Teorie odpovědí na položku. *Československá Psychologie*, 45(5), 428–440.
- Urbánek, T., & Sollár, T. (2014). *Psychometrické analýzy škál na meranie úzkosti*. Univerzita Konštantína Filozofa v Nitre.
- van Zalk, N., Lamb, M. E., & Jason Rentfrow, P. (2017). Does Shyness Vary According to Attained Social Roles? Trends Across Age Groups in a Large British Sample. *Journal of Personality*, 85(6), 830–840. <https://doi.org/10.1111/jopy.12291>
- Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, 3(1), 4–70. <https://doi.org/10.1177/109442810031002>
- Walker, C. M. (2011). What's the DIF? Why Differential Item Functioning Analyses Are an Important Part of Instrument Development and Validation. *Journal of Psychoeducational Assessment*, 29(4), 364–376. <https://doi.org/10.1177/0734282911406666>
- Willis, G. (2005). *Cognitive Interviewing*. SAGE Publications, Inc. <https://doi.org/10.4135/9781412983655>

- Zumbo, B. D. (1999). *A 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.*
- Zumbo, B. D. (2007). Three Generations of DIF Analyses: Considering Where It Has Been, Where It Is Now, and Where It Is Going. *Language Assessment Quarterly*, 4(2), 223–233.
<https://doi.org/10.1080/15434300701375832>

© Michal Kohút 2022
Filozofická fakulta TU v Trnave





© Michal Kohút 2022
Filozofická fakulta TU v Trnave
ISBN 978-80-568-0507-7