

SKRÁTENÁ VERZIA ŠKÁLY REZILIENCIE – PSYCHOMETRICKÁ ANALÝZA PROSTREDNÍCTVOM IRT¹

MICHAL HAJDÚK², BARBORA MESÁROŠOVÁ², ANTON HERETIK²

Abstrakt: Škála reziliencie (Wagnild, Young, 1993) patrí v súčasnosti medzi najčastejšie používané metódy na skúmanie reziliencie ako osobnostnej črty. Výsledky výskumov poukazujú na veľmi dobré psychometrické charakteristiky škály. Cronbachovo α sa pohybuje od .72 do .94. Výsledky zahraničných aj lokálnych výskumov potvrdzujú dobrú konštruktívnu validitu škály. Pôvodná verzia obsahuje 25 položiek. Autori vytvorili aj skrátenú verziu, ktorá pozostáva zo 14 položiek. Cieľom výskumu bolo overiť psychometrické vlastnosti skrátenej verzie škály. Výskumný súbor tvorilo 609 uchádzačov o štúdium na Katedre psychológie na FiF UK, ktorí vyplnili kompletnú 25 položkovú verziu. Pri analýze sme porovnávali psychometrické ukazovatele klasickej teórie testov a taktiež sme pri analýzach aplikovali aj Teóriu odpovede na položku (IRT). Na odhad parametrov položiek sme použili Samejmovej model. Priemerná bodovo-biseriálna korelácia skrátenej verzie škály bola $r = .513$. Odhad úrovne reziliencie (prostredníctvom IRT) zo skrátenej verzie škály bol takmer identický ako pri použití 25 položkovej verzie ($R^2 = .932$). Obidve verzie škály merajú dostatočne presne na pomerne širokom intervale úrovne reziliencie a prinášajú najviac informácií o ľuďoch s podpriemernou až priemernou úrovňou reziliencie. Pôvodná 25 položková verzia meria trochu presnejšie, čo je spôsobené väčším počtom položiek. Z výsledkov vyplýva, že skrátená verzia Škály reziliencie je vhodný a dostatočne reliabilný nástroj, ktorý je časovo menej náročný ako pôvodná plná verzia Škály reziliencie.

Kľúčová slova: Reziliencie, psychometrická analýza, IRT

¹ Text vznikl na základě příspěvku prezentovaného na konferenci Psychologická diagnostika: Výzkum, prevence a poradenství, konané v Brně 19.–21. 6. 2013 (<http://psychodiagnostika.fss.muni.cz/2013>).

² Katedra psychológie, Filozofická fakulta UK, Gondova 2, 814 99 Bratislava

Úvod

Reziliencia

V súčasnosti sa skúmaniu reziliencie venuje veľké množstvo výskumov. Autori sa výrazne líšia v tom, ako na tento konštrukt nazerajú. Na jednej strane sú autori ako Wagnild a Young (1993), ktorí považujú rezilienciu za osobnostnú črtu. Markstrom et al. (2000) na druhej strane považujú rezilienciu za adaptívnu, voči stresu rezistentnú osobnostnú vlastnosť, ktorá dovoľuje jednotlivcovi prosperovať, napriek nepriaznivej životnej situácii. Zimmerman a Brennerová (2010) nazerajú na rezilienciu ako na dynamický proces, predstavujúci schopnosť udržať rovnováhu a adaptívne reagovať v stresujúcich podmienkach.

Ako vidíme, ani sami výskumníci nie sú jednotní v tom, ako nazerať a merať tento psychologický konštrukt. V súčasnosti existuje viacero dotazníkových metód merajúcich rezilienciu. V našom príspevku sa zameriame na skrátenú verziu Škály reziliencie (RS-14) (Wagnild, Young, 1993), ktorá meria rezilienciu ako osobnostnú črtu.

Škála reziliencie

Škála reziliencie (RS-25) Wagnilda-Youngovej obsahuje 25 položiek, na ktoré respondent odpovedá na sedem stupňovej Likertovej škále. Rozsah škály je 25–175 bodov. Na základe skóre je možné respondentov rozdeliť na nízko, stredne a vysoko rezilientných (Wagnild, Young, 1993). Autori taktiež vypracovali skrátenú verziu škály, obsahujúcu 14 položiek. Položky majú formát 7 stupňovej Likertovej škály. Rozsah škály je od 14 do 98 bodov.

Teória odpovede na položku

Teória odpovede na položku (ďalej len IRT) predstavuje alternatívny prístup ku klasickej teórii testov. Predpoklady použitia IRT sa odlišujú od predpokladov klasickej teórie testov. Medzi predpoklady IRT patrí jednodimenzionalita testu, invariancia parametrov a lokálna nezávislosť. Detailnejšie sa IRT venujú na regionálnej úrovni napríklad Urbánek a Šimeček (2001) a Jelínek, Květon a Vobořil (2011). Psychometrickú analýzu Rosenbergovej škály sebaúcty prostredníctvom IRT na Slovensku publikoval Halama a Bieščad (2006). V príspevku priblížime len úplné základy IRT, ktoré považujeme za nevyhnutné pre porozumenie výsledkov a grafov.

Podstatou IRT je modelovanie vzťahu odpovede osoby na položku a meraného latentného konštraktu. IRT je možné aplikovať na výkonové testy, ako aj na dotazníkové metódy. V závislosti od použitého modelu sa odhaduje rôzny počet parametrov. Parameter a predstavuje rozlišovaciu schopnosť položky, ktorá predstavuje analógiu ku korelácii položky so zvyškom testu v klasickej teórii testov. Vyššie hodnoty rozlišovacej schopnosti znamenajú, že položka diskriminuje medzi osobami s vysokou a nízkou úrovňou meranej schopnosti. Parameter b predstavuje obťažnosť položky a c parameter (len pri výkonových testoch) predstavuje faktor hádania (uhádnuteľnosť položky – má to význam pri položkách s možnosťou voľby odpovede resp. výkonových

testoch). Spomínané parametre sú odhadované v procese kalibrácie položiek. Hodnoty parametrov majú vplyv na tvar *characteristickej krivky položky*. IRT predpokladá, že vzťah medzi úrovňou meranej schopnosti a pravdepodobnosťou kladnej odpovede na položku nie je lineárny, ale má tvar krivky (ogiva). Krivka sa zaznamenáva do grafu, kde na osi x je úroveň meranej schopnosti a na osi y pravdepodobnosť správnej odpovede. Z krivky je možné určiť pravdepodobnosť správnej odpovede pri danej úrovni meranej schopnosti (reziliencia, úzkostlivosť atď.). Úroveň schopnosti, pri ktorej je 50 % pravdepodobnosť voľby, predstavuje hodnotu parametra b . Pre každú osobu je v rámci analýzy odhadnutá aj úroveň jej schopnosti. Úroveň schopnosti a obťažnosti položiek je možné umiestniť na rovnakú škálu, čo umožňuje ich vzájomné porovnanie. Škála schopnosti má priemernú hodnotu $M=0$ a $SD=1$. Úroveň schopnosti sa označuje gréckym symbolom θ . Kladné hodnoty zodpovedajú vyšším úrovňam meranej schopnosti a vice versa (Hambleton, Swaminathan, Rogers, 1991).

IRT metódy na rozdiel od klasickej teórie testov nepredpokladajú, že presnosť merania je konštantná pre všetky úrovne meranej vlastnosti. Položky sa podľa IRT líšia v tom, o ktorých úrovniach meranej vlastnosti prinášajú najviac informácií. Niektoré položky merajú presnejšie vyššiu, resp. nižšiu úroveň meranej schopnosti. Pre každú položku je možné vypočítať množstvo informácií, ktoré prináša pre jednotlivé úrovne meranej vlastnosti. Analogicky vieme vypočítať informačnú funkciu aj pre celý test (súčet informačných funkcií všetkých položiek). Platí, že meranie je tým presnejšie, čím viac máme informácií o danej úrovni schopnosti (Embretson, Reise, 2000).

Prostredníctvom IRT je možné analyzovať nielen výkonové testy, ale aj dotazníkové metódy. Pri analýze dotazníkov sa používa najmä Samejmovej model, pričom sa pre každú položku odhaduje parameter a a niekoľko b parametrov (prahov), ktoré sú analógiou k obťažnosti pri dichotómnych položkách. Ak analyzujeme polytómne položky (škály), tak počet b parametrov položiek predstavuje počet stupňov Likertovej škály – 1. Napríklad v prípade 4 stupňovej škály sa odhadujú 3 prahové koeficienty. Hodnota koeficientu vyjadruje miesto na škále meranej vlastnosti, pri ktorej si respondent vyberie s 50 % pravdepodobnosťou kategóriu väčšiu ako prah. V prípade spomínanej škály ide o výber medzi 1 vs. 2, 3, 4. Zvyšné prahové koeficienty sú odhadnuté analogicky (1, 2 vs. 3, 4; 1, 2, 3 vs. 4).

Cieľom výskumu bolo overiť, prostredníctvom použitia IRT a klasickej teórie testov, psychometrické charakteristiky RS-14.

Parciálnymi cieľmi sú:

1. Preskúmať faktorovú štruktúru RS-14.
2. Preskúmať psychometrické vlastnosti RS-14.
3. Prostredníctvom IRT metód preskúmať vlastnosti RS-14 a zhodnotiť možnosti jej využitia v porovnaní s pôvodnou škálou.

Metódy

Priebeh výskumu

V júni 2012 sa 609 uchádzačov o štúdium psychológie na Katedre psychológie, FiF UK v Bratislave podrobilo sérii testov (všeobecný spoločný test pre celú FiF UK, špeciálne testy pre uchádzačov o štúdium psychológie) zároveň uchádzači vyplnili aj 25 položkovú verziu Škály reziliencie, ktorá bola doplnená aj o 10 položiek zo Social Desirability Scale.

Analýza dát

Položkovú analýzu, faktorovú a Paralelnú analýzu sme robili v programe SPSS. Kalibráciu položiek a odhad schopnosti prostredníctvom IRT sme realizovali v programe X Calibre 4.1.8.

Výskumný súbor

Výskumný súbor tvorilo 609 uchádzačov o štúdium psychológie. RS-25 vyplnilo 142 mužov a 467 žien. Presný vek respondentov sa nám z administratívnych dôvodov nepodarilo získať.

Výsledky

Deskriptívna štatistika a reliabilita

Priemerné skóre v RS-14 bolo $M = 77.81$ a $SD = 9.63$. Cronbachovo α dosiahlo hodnotu $\alpha = .840$, čo svedčí o veľmi dobrej vnútornej konzistencii škály. Všetky položky mali dostatočne vysokú koreláciu so zvyškom testu. Priemerná hodnota korelácie so zvyškom testu bola $r = .513$. Priemerné hodnoty a point-biserial položiek sú uvedené v tabuľke 1.

Tabuľka 1. Odhad parametrov položiek

Položka	M	p-bis	a	b1	b2	b3	b4	b5	b6	x²	p
2	5.874	.557	1.21	-4.73	-3.84	-3.42	-2.31	-1.07	0.90	40.50	0.4926
6	6.146	.507	1.13	-4.41	-4.16	-3.59	-2.58	-1.30	-0.04	49.44	0.1718
7	4.626	.432	0.91	-5.79	-4.22	-2.33	-0.29	1.62	3.69	70.27	0.0030
8	5.533	.534	1.19	-4.96	-4.03	-2.72	-1.39	-0.39	1.17	46.90	0.2434
9	5.344	.570	1.24	-4.11	-3.57	-2.68	-1.52	-0.01	1.74	32.79	0.8160
10	5.231	.629	1.40	-4.29	-2.85	-1.79	-0.97	0.07	1.18	50.08	0.1563
13	5.425	.423	0.91	-5.22	-4.32	-2.92	-1.48	-0.25	1.30	44.86	0.3132
14	5.541	.526	1.15	-4.91	-4.08	-2.72	-1.45	-0.32	1.05	34.53	0.7522
15	5.564	.506	1.14	-5.18	-4.47	-3.39	-1.68	-0.32	1.26	56.81	0.0512
16	6.110	.406	0.86	-5.85	-5.48	-4.34	-3.04	-1.65	0.15	37.11	0.6443
17	5.159	.589	1.32	-3.72	-3.04	-2.18	-0.79	0.15	1.39	47.17	0.2349
18	5.804	.418	0.93	-5.21	-4.61	-4.00	-2.48	-0.86	0.87	61.95	0.0189
21	6.394	.468	1.24	-4.33	-4.07	-3.54	-2.41	-1.60	-0.82	42.58	0.4030
23	5.589	.611	1.39	-4.28	-4.10	-3.04	-1.80	-0.32	1.35	40.50	0.4928

Analýza prostredníctvom IRT

V rámci analýz sme použili Samajimovej model. Aby sme mohli porovnať skrátenú verziu škály (porovnanie odhadu reziliencie), najskôr sme urobili odhad parametrov položiek a úrovne schopnosti pre RS-25 a osobitne aj pre RS-14. Na overenie vhodnosti použitého modelu sme použili chí-kvadrát test. Použitý model celkovo vystihuje dáta ($\chi^2(574)=655.489$; $p=.01$; $-2LL=21424$). Napriek signifikantným výsledkom chí-kvadrát testu je pomer χ^2/df prijateľný, čo svedčí o vhodnosti použitia nami zvoleného modelu. Pri analýze RS-25 sme identifikovali niekoľko položiek, ktoré mali mimoriadne nízku rozlišovaciu schopnosť. Žiadna z týchto položiek ale nepatrí zároveň do skrátenej verzie škály. Odhad úrovne schopností respondentov sme realizovali prostredníctvom metódy maximálnej vierohodnosti.

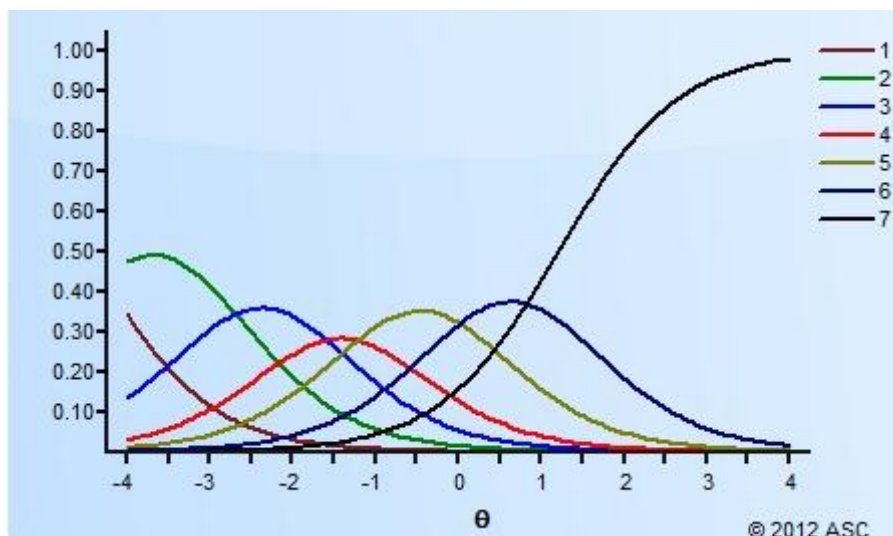
Pred použitím IRT na analýzu testu, resp. dotazníka, je nevyhnutné overiť predpoklad, že nástroj meria len „jeden konštrukt“. Jednodimenzionalitu sme overovali prostredníctvom viacerých štatistických metód. Najskôr sme použili Analýzu hlavných komponentov so šikmou rotáciou (Direct Oblimin). Prvý faktor vysvetľoval 29.42 % variancie dát. Dvojfaktorové riešenie vysvetľuje 39.07 %. Ako druhú metódu na overenie jednodimenzionality sme použili Paralelnú analýzu (O'Connor, 2000). Metóda spočíva v porovnávaní nameraných vlastných hodnôt faktorov/komponentov s hodnotami, ktoré sú odvodené z väčšieho počtu dátových súborov (N = 1000)

s rovnakým počtom respondentov a premenných, ktoré ale obsahujú náhodné čísla. Faktor alebo komponent zostáva v analýze, ak je jeho vlastná hodnota vyššia ako 95. percentil distribúcie vlastnej hodnoty odvodennej z náhodných dát. Paralelná analýza sa používa pri rozhodovaní o tom, koľko faktorov zachovať. Výsledky taktiež potvrdili, že prvý faktor je dominantný.

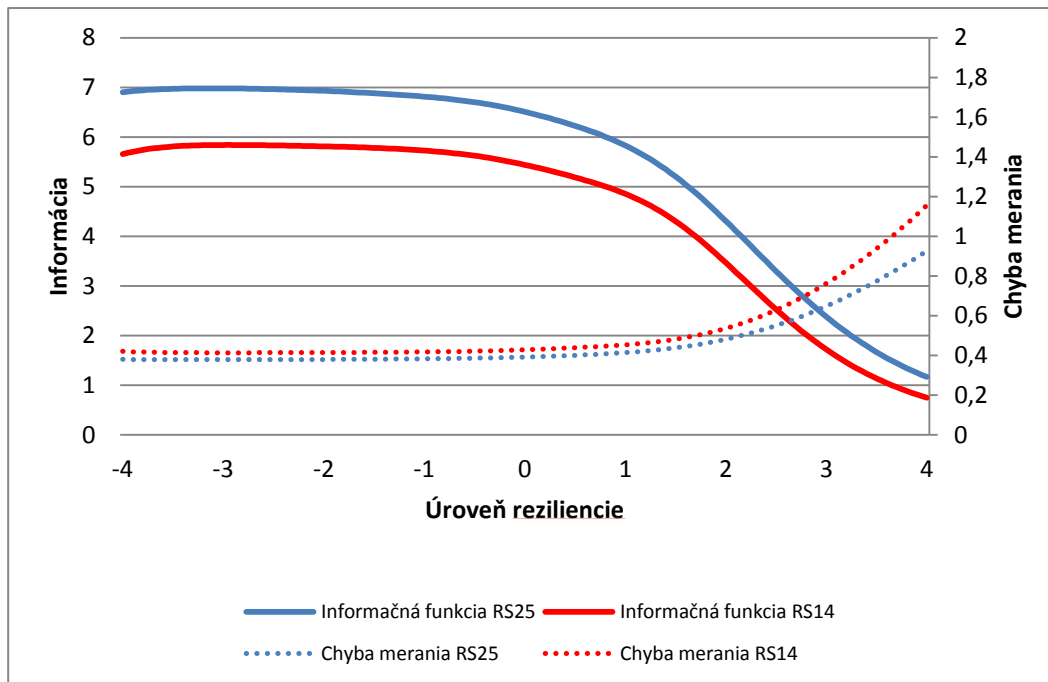
Parametre položiek

Najvyššiu diskriminačnú schopnosť (a parameter) mala položka č. 10 (Som rozhodný/(-á) (odhodlaný/(-á).) (pozri obrázok 1). Najmenej dokázala rozlišovať medzi vysoko a nízko rezilientnými uchádzačmi položka č. 16 (Dokážem zvyčajne nájsť niečo, na čom sa dá zasmiať.). V tabuľke 1 vidíme, že väčšina odhadnutých prahových koeficientov má záporné hodnoty. Nakoľko má škála meraného rysu fixovanú priemernú hodnotu na 0, tieto hodnoty poukazujú, že položky v škále sú pomerne „jednoduché“. Jednoduchosť v tomto prípade znamená, že aj pomerne slabo rezilientní uchádzači si vyberali vysoké hodnoty na Likertovej škále. V tabuľke 1 uvádzame priemerné hodnoty pre jednotlivé položky.

Obrázok 1. Charakteristická krivka položka č. 10



Na obrázku 2 vidíme informačné funkcie pre RS-25 a RS-14. Prerušovanou čiarou je v obrázku zaznačená chyba merania. Ako sme uviedli vyššie, pri IRT sa vypočíta chyba merania pre každú úroveň meranej schopnosti. Na rozdiel od klasickej teórie testov, kde je predpoklad, že chyba merania je konštantná. Z uvedeného obrázku vyplýva, že obidva testy merajú najpresnejšie respondentov s priemernou, resp. podpriemernou úrovňou reziliencie. Pri tejto úrovni reziliencie je aj najnižšia chyba merania. Približne od úrovne $\theta = 1$ (1 SD nad priemerom) začína stúpať nepresnosť merania. Odhadnutá úroveň z RS-14 je viac menej identická s odhadom z RS-25. Koeficient determinácie dosiahol hodnotu $R^2 = .932$.

Obrázok 2. Informačná funkcia testu a chyba merania**Diskusia**

Z výsledkov je zrejmé, že skrátená verzia Škály reziliencie má veľmi dobré psychometrické vlastnosti. Hodnota Cronbachovho α je takmer identická ako pri pôvodnej verzii (.838 resp. .840 pre RS-25). Priemerná hodnota korelácie položiek so zvyškom testu je pri skrátenej verzii dokonca ešte vyššia ako pri celej škále. Odhad schopnosti prostredníctvom IRT pri použití skrátenej verzie je takmer identický ($R^2=.932$) s odhadom reziliencie pri použití kompletnej škály. Chyba merania (podľa IRT) je prirodzene nižšia pri celej škále. RS-14 meria ale dostatočne presne na pomerne širokom intervale úrovni reziliencie.

Výhodou použitia skrátenej verzie je hlavne skrátenie času administrácie. Najmä pri výskumoch, kde sa používa celá batéria dotazníkov je mimoriadne dôležité, aby bol respondent dostatočne motivovaný a ochotný pozorne vyplniť všetky dotazníky. Práve použitie overených skrátených verzií škál môže byť veľmi užitočné pri takomto type výskumov.

Prezentované výsledky sú len parciálnym výstupom väčšieho výskumného projektu. V príspevku sme sa zamerali najmä na štruktúru škály, jej reliabilitu a z časti na položkovú analýzu. Do budúcnosti by bolo ešte potrebné overiť aj validitu skrátenej verzie. Či už konkurenčnú validitu, súbežne administrovať Conon-Davidson Resilience Scale (Conon, Davidson, 2003) alebo konštruktovú validitu. Z predchádzajúcich výskumov je zrejmé, že vyššia miera reziliencie by mala súvisieť s napríklad vyšším dispozičným optimizmom, vyššou sebaúctou a pod. (Tusaie et al., 2007).

Z analýzy charakteristických kriviek jednotlivých položiek sme usúdili, že pri niektorých položkách boli niektoré stupne Likertovej škály málo, resp. takmer vôbec zastúpené. Dochádzalo k tomu, že s narastajúcou úrovňou schopnosti nestúpala dostatočne pravdepodobnosť, že si respondenti vyberú na škále vyššiu hodnotu. Otázkou zostáva, či by nebolo vhodné do budúcnosti overiť, ako by sa správala škála, ak by sme použili nie 7 ale napríklad 5 stupňovú škálu.

Dve položky vykazujú rozdielne fungovanie ako predpokladá model (misfit). Zistené výsledky si vysvetľujeme tým, že pravdepodobne pri týchto položkách si niektoré stupne Likertovej škály (hodnoty 1–3) vyberalo pomerne málo respondentov, čo mohlo spôsobiť spomínané výsledky. Pre presnejší odhad parametrov by bolo potrebné ešte rozšíriť súbor.

Odhad parametrov položiek, úrovne reziliencie (prostredníctvom IRT) a taktiež hrubé skóre, mohli byť ovplyvnené/skreslené aj situáciou, počas ktorej uchádzači vyplňali dotazník. Pravdepodobne dochádzalo k miernemu nadhodnocovaniu vlastnej rezilientnosti. V prospech tohto názoru svedčia aj získané skóre v škále sociálnej žiadúcnosti, ktorým sa ale v tejto štúdiu nevenujeme.

V nasledujúcich výskumoch je potrebné overiť psychometrické vlastnosti škály aj na iných vzorkách ako uchádzači o štúdium, resp. v menej kompetitívnej situácii, ako je účasť na prijímacom pohovore na vysokú školu.

Výskum vznikol vďaka podpore grantového projektu VEGA č. 1/0392/11. Národnému ústavu certifikovaných meraní vzdelávania ďakujeme za možnosť využiť softvér X Calibre 4.1.8.

Literatúra

- Connor, K.M., & Davidson, J.R.T., (2003). Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and anxiety*, 18:76–82.
- Embretson, S.E., & Reise, S.P. (2000). *Item Response Theory for Psychologists*. Mahwah-London: Lawrence Erlbaum Associates.
- Halama, P., & Bieščad, M. (2006). Psychometrická analýza Rosenbergovej škály sebahodnotenia s použitím metód klasickej teórie testov (CTT) a teórie odpovede na položku (IRT). *Československá psychologie*, 50, 6, 588-603.
- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: SAGE.
- Jelínek, M., Květoň, P., & Vobořil, D. (2011). Testování v psychologii – Teorie odpovědi na položku a počítačové adaptivní testování. Praha : Portál.

M. Hajdúk, B. Mesárošová, & A. Heretik:

Skrátená verzia škály reziliencie – psychometrická analýza prostredníctvom IRT

- Markstrom, C.A. et al. (2000). Resiliency, social support and coping in rural low-income Appalachian adolescents from two racial groups. *Journal of Adolescence*, 23, 693 - 703.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, 32, 396-402.
- Tusaie, K., Puskar, K., & Sereika, S.M. (2007). A Predictive and Moderating Model of Psychosocial Resilience in Adolescents. *Journal of Nursing Scholarship, First Quarter*, 54 - 60.
- Urbánek, T. & Šimeček, M. (2001). Teorie odpovědi na položku. *Československá psychologie*, 45, 428-441.
- Wagnild, G., & Young, H. (1993). Development and psychometric evaluation of the resilience scale. *Journal of Nursing Measurement*, 1(2), 165-178.
- Zimmerman, M. A., Brenner, A.B. (2010). Resilience in adolescence: overcoming neighborhood disadvantage. In J.W. Reich, A.J. Zautra, J.S. Hall, (Eds.), *Handbook of adult resilience*. The Guilford Press, New York. 283-308.

Hajdúk, M., Mesárošová, B., & Heretik, A. (2015): Shortened version of Resilience scale – psychometric analysis using IRT

Resilience Scale (Wagnild, Young, 1993) is currently one of the most frequently used questionnaire for research about resilience as a personality traits. The research findings have showed very good psychometric characteristics of the scale. Cronbach's α ranging from .72 to .94. The results of foreign and local studies confirmed good construct validity. The original version contains 25 items. The authors have developed a shortened version consisting only of 14 items. The aim of the research was to investigate the psychometric properties of shortened version of the scale. Sample consisted of 609 applicants for study at the Department of Psychology at the Faculty of Arts UK, who filled the entire 25 item version. In the analysis, we compared the psychometric properties of Classical Test Theory and Item Response Theory (IRT). We used Samejima's model. The average point-biserial correlation of shortened version scale was $r = .513$. Estimated level of resilience (through IRT) with a shortened version of the scale was almost identical to estimate from 25 item version ($R^2 = .932$). Both scales measure with sufficient accuracy on relatively wide range level of resilience. Both scales bring the most information about people with below-average to average level of resilience. The original 25 item version measured more accurately, due to a larger number of items. Shortened version of Resilience Scale is truly suitable and sufficiently reliable tool to measure resilience, which is less time consuming than the original full version scale.

Keywords: Resilience, Psychometric analysis, IRT