

SOUBĚŽNÁ VALIDITA DOTAZNÍKU SPARO A MMPI-2

Hynek Cígler, Adéla Rudá

Abstrakt

V předcházející studii (Cígler & Rudá, 2021) se nám nepodařilo prokázat dostatečnou vnitřní konzistenci a faktorovou validitu dotazníku SPARO. Pro posouzení vhodnosti jeho využití v praxi jsou nicméně důležité ještě další psychometrické parametry. V tomto navazujícím výzkumu se proto pokusíme ověřit jeho souběžnou validitu oproti dotazníkové baterii MMPI-2. Její výhodou je mezinárodní popularita, díky níž se může opřít o řadu zahraničních studií. S využitím rozsáhlého vzorku policistů a uchazečů o práci u policie se nám nicméně nepodařilo souběžnou validitu uspokojivě demonstrovat. Nedostatečná je jak konvergentní, tak i divergentní validita hlavních komponent, obecných faktorů i skóre věrohodnosti (L) dotazníku SPARO vůči vybraným škálám MMPI-2, a to s výjimkou škály emocionální variabilnosti (EM). Vedlejším výstupem naší studie je ověření vnitřní konzistence a faktorové struktury jednotlivých škál dotazníku MMPI-2, protože jiné české empirické studie na toto téma neexistují. Musíme konstatovat, že jak reliabilita, tak i faktorová validita MMPI-2 rovněž není uspokojivá. Závěrem argumentujeme, že použití dotazníku SPARO v běžné praxi nelze v žádném případě doporučit. Opatrné použití dotazníku MMPI-2 je podle nás možné s ohledem na uspokojivé zahraniční výsledky, je však nezbytný další český výzkum, a to zejména po stránce prediktivní validity.

Klíčová slova: konvergentní validita; divergentní validita; SPARO; MMPI-2; Minnesota Multiphasic Personality Inventory; policejní sbor ČR; personální výběr

SIMULTANEOUS VALIDITY OF SPARO AND MMPI-2 QUESTIONNAIRES

Abstract

In the previous study (Cígler & Rudá, 2021), we were not able to confirm satisfactory internal consistency and factor validity of the SPARO questionnaire. However, further psychometric properties are also necessary to consider the questionnaire's validity and usefulness in practice. In this follow up study, we examined its criterion validity to MMPI-2, which is used globally and thus has broader empirical support. Despite that, we did not find any support for neither convergent nor divergent validity of the main components, general factor, and the credibility score (L) of SPARO questionnaire to selected MMPI-2 scales. The only exception was the emotional variability scale (EM). As a minor result, we also examined the internal consistency and factor validity of MMPI-2 scales as no other Czech study concerning the questionnaire exists. Neither reliability nor factor validity was satisfactory. Finally, we argue that using the SPARO questionnaire in regular practice cannot be recommended. Cautious use of the MMPI-2 is possible, regarding international validation studies, however, further Czech research is needed, especially in the predictive validity.

Keywords: *convergent validity; divergent validity; SPARO; MMPI-2; Minnesota Multiphasic Personality Inventory; Police of the Czech Republic; personal selection*

Došlo: 18. 7. 2020

Schváleno: 10. 11. 2020

Úvod

V našem předchozím článku (Cígler & Rudá, 2021) jsme popsali teoretická východiska a podobu dotazníku SPARO (Mikšík, 2004) a zejména jsme ověřili jeho vnitřní konzistenci a faktorovou validitu. Bohužel, naše výsledky praktické použití metody nepodpořily. Vnitřní konzistence všech skóru se pohybovala v rozmezí 0,32–0,87 s mediánem $Md = 0,71$, reliabilita zhruba poloviny škál nebyla uspokojivá (menší než 0,7). Mezi škály s neuspokojivou reliabilitou patřily dokonce hlavní komponenta Adjustační variabilnosti, AD ($r_{max} = 0,559$) a obecný faktor motorické hybnosti, MH ($r_{max} = 0,636$). Tento odhad je navíc korigovaný oproti případné vícedimenzionalitě škál; medián konvenčnějšího odhadu reliability prostřednictvím koeficientu alfa byl výrazně nižší, $Md = 0,61$. Většina škál totiž měla problematickou faktorovou strukturu, 38 % bylo jednoznačně vícedimenzionálních a korelace mezi položkami byly celkově velmi nízké.

Hlavní omezení konstruktové validity nicméně skýtaly silné vzájemné korelace čtyř hlavních komponent dotazníku, které by měly být podle teorie nezávislé, a rovněž nevyjasněná faktorová struktura dvou obecnějších faktorů. Zobecnitelnost těchto závěrů posiloval fakt, že námi pozorovaná korelační struktura škál dotazníku SPARO nebyla odlišná od originálních korelací reportovaných Mikšíkem (2001, 2004, 2009). Naše i originální vzorky respondentů tedy byly pravděpodobně vybrány ze stejné či alespoň podobné populace, a zmíněná omezení validity tak musela být přítomna již v originálních studiích. Ty totiž nereflektují, že příčinou nízkých pozorovaných korelací je nereliabilita škál, nikoli jejich nezávislost (Cígler & Rudá, 2021).

I přes uvedené nedostatky však stále dotazník SPARO může vykazovat dostatečnou souběžnou a prediktivní validitu. Cílem tohoto navazujícího výzkumu je proto ověření souběžné validity vybraných škál dotazníku SPARO vůči metodě MMPI-2. Vzhledem k nedostatku jiných českých empirických zdrojů ověříme rovněž i reliabilitu a faktorovou strukturu škál MMPI-2 podobně jako v případě dotazníku SPARO v předchozím článku (Cígler & Rudá, 2021). K 3. dubnu 2020 totiž jen sedm studií afilovaných českým výzkumníkům a indexovaných v databázích Web of Science nebo Scopus pracovalo s daty z MMPI (Farnikova et al., 2012; Hansmanová et al., 2007; Jensovsky et al., 2002; Kudela & Hansmanová, 2014; Líbalová et al., 2001; Obereignerů et al., 2010; Stancak et al., 1986), žádná z nich se však nezabývala jeho psychometrickými parametry. Termín „MMPI“ pak obsahovalo k 3. dubnu 2020 i 428 diplomových prací v databázi theses.cz; 26 závěrečných prací obsahovalo termíny MMPI i SPARO zároveň. I zde je však otázkou, kolik z nich se zabývá psychometrickými parametry.

Souběžná validita metody SPARO

Mikšík (2004) při popisu validity i reliability metody SPARO vychází ze starších zdrojů a analýz předchozích metod (IHAVEZ aj.). Tyto originální studie jsou nicméně nejenže nedohledatelné, ale navíc jsou zčásti založené i na předchozích verzích dotazníku s mírně odlišnými položkami, což omezuje jejich zobecnitelnost na aktuální verzi metody SPARO. Autor navíc necituje použité metody, což znemožňuje dohledání jejich konkrétních verzí, protože řada z těchto metod nebyla v době sběru dat standardizována v České republice. Je tedy pravděpodobné, že byly použity různé pracovní překlady a neoficiální varianty dotazníků s nejistými psychometrickými parametry. Zkratky jsou proto použity ve shodě s autorem bez další citace (Mikšík, 2004).

Mikšík (2004) reportuje prediktivní validitu některých typů osobnosti pro výkon detektivního povolání, dále korelace škál SPARO s inteligenčními testy IST, Raven a LOGUS a souběžnou

validitu s osobnostními dotazníky (16PF, Eysenckovy škály EOD B a DOPEN, FIRO-B a B-D-I), všechny jsou nicméně velmi nízké a působí nahodilým dojmem. Zajímavé jsou pouze korelace škál EM, UZ, EC a LS (negativní) s Eysenckovou škálou neuroticismu ($r > 0,65$) a korelace SE a NU s Eysenckovou škálou extravertze ($r < -0,40$). Ostatní korelace jen zřídka překračují hodnotu 0,4 (Mikšík, 2004). Řadu dalších studií bez jasného pojitka reportuje Mikšík v jiných publikacích (2007, 2009), nejsme však schopni z nich vyvodit žádné jasné závěry.

Tento pohled rozšiřují ojedinělé studie dalších autorů. Jurišová a Sarmány-Schuller (2013) uvádějí, že na vzorku 134 slovenských zdravotnických pracovníků klidný typ EM-RE-, tedy typ emocionálně stabilní a regulovaný, pozitivně souvisí se všemi maladaptivními způsoby rozhodování (hypervigilance, prokrastinace a přesun odpovědnosti), což je v rozporu s teorií dotazníku. U pozitivního způsobu rozhodování, tzv. vigilantního, se přitom neprokázaly souvislosti s komponentami a profily SPARO (Jurišová & Sarmány-Schuller, 2013).

Diplomová práce Prokopové (2011) obsahuje zajímavé souvislosti mezi úspěchem 173 uchazečů ve výběrovém řízení do Hasičského záchranného sboru a osobnostními typy, komponentami a škálami dotazníku SPARO. Naneštěstí šlo o poslední kolo výběrového řízení a uchazeči byli vybíráni právě s pomocí dotazníku SPARO; dotazníkové výsledky a výběrové kritérium jsou tedy shodné a reportované výsledky svědčí spíše o tom, že interpretace dotazníku u těchto konkrétních psychologů v záchranných sborech a jejich rozhodování o tom, kdo z uchazečů bude přijat, je v souladu s teoretickými východisky dotazníku. Studii tak nelze využít pro hodnocení validity metody.

Šucha a kol. (2017) srovnával řidiče s odebraným řidičským průkazem z důvodu řízení pod vlivem alkoholu s řidiči s bezproblémovou řidičskou praxí. Rozdíly mezi oběma skupinami na škálách dotazníku SPARO byly velmi slabé a spíše náhodné, což však platilo i o dotazníku NEO-PI-3.

K dispozici je rovněž i srovnání dotazníku Washingtonského psychosociálního dotazníku pro záchvatová onemocnění (WPSI) s předchozí verzí SPARO, dotazníkem VAROS (Preiss & Haas, 1997). Bohužel vzorek byl velmi malý ($n = 39$) a statisticky významné a teoreticky podložené, avšak slabé korelace byly pozorovány pouze mezi škálou emocionální variabilnosti (EM) a schopností se přizpůsobit a celkovou psychosociální výkonností. Dále byla pozorována korelace škály extremity (EX) s většinou obsahových škál WPSI, což však není ve shodě s teorií.

Celkově lze shrnout, že souběžná validita dotazníku SPARO byla podpořena pouze v případě emocionální variabilnosti (EM), a i to jen částečně.

Minnesota Multiphasic Personality Inventory

Vývoj dotazníku MMPI (Minnesota Multiphasic Personality Inventory) začal na konci 30. let 20. století v reakci na kritiku teoretického přístupu ke konstrukci metod, a tedy i jejich nízké prediktivní validitě během první světové války (Nichols, 2011).

Položky MMPI byly vybírány z počáteční banky více než 1000 položek podle schopnosti diskriminovat mezi vybranými kontrastními skupinami; první verze metody s 550 položkami (Hathaway & McKinley, 1942) obsahovala osm klinických škál (Hs, D, Hy, Pd, Mf, Pa, Pt, Sc) (Hathaway, McKinley, 1942). Postupně byly přidávány další klinické škály a MMPI se brzy rozšířilo nejen v klinické, ale rovněž i organizační nebo forenzní psychologii. Spolu s rozšířením metody přicházela také její kritika a v 70. letech vznikla potřeba revize metody, především nových norem a

přeformulování některých položek. Reakcí byla druhá revize MMPI-2 s 567 položkami (z toho 107 nově přidaných a 90 odstraněných), s novými normami a škálami, později byly přidány i tzv. restrukturalizované škály (A. Tellegen et al., 2003) a dnes je MMPI-2 celosvětově rozšířenou metodou adaptovanou do více než 20 jazyků. Autorem české adaptace je Karel Netík (Butcher et al., 2002), druhé české vydání obsahuje rovněž i restrukturalizované škály (Butcher et al., 2013).

Metoda nabízí 10 standardních klinických škál (viz Přílohu 1, tab. 1.1). K přesnější interpretaci některých z nich (D, Hy, Pd, Pa, Sc, Ma) lze využít Harrisovy-Lingoesovy subškály a Si škály. Wiener-Harmonovy škály rozlišují u pěti klinických škál (D, Hy, Pd, Pa, Ma) tzv. skrytou a zjevnou subškálu, které zachycují skryté a zjevné identifikátory psychopatologie. Na základě jejich posouzení můžeme usuzovat na tendenci jevit se lépe („fake-good“) nebo hůře („fake-bad“). Ke klinickým škálám (kromě Mf a Si) bylo dále vytvořeno devět restrukturovaných klinických škál (tzv. RC škály) odvozených na základě faktorových a obsahových analýz, které mohou sloužit jako dílčí nebo ověřovací škály standardních klinických škál; jsou totiž zacílené na primární faktor, tzv. „jádrovou komponentu“, příslušné klinické škály (Butcher et al., 2013, p. 46). Pouze restrukturovaná škála RCd – demoralizace, která se vztahuje k celkovému pesimismu a pocitu nepohody, nemá protějšek ve standardních škálách (Friedman et al., 2015; Greene, 2000).

MMPI-2 obsahuje rovněž validizační škály: škála „nemohu říci“ („?“) udává počet vynechaných odpovědí, škála L ukazuje na míru podávání společensky žádoucích odpovědí, škála F identifikuje snahu záměrně zkreslit výsledek testu a tendenci odpovídat neobvykle, konečně pak K-škála ukazuje na defenzivitu respondenta. Ze vzájemné konstelace těchto tří škál usuzujeme na nadhodnocení či podhodnocení psychopatologie v protokolu. Dalšími novými validizačními škálami jsou například VRIN detekující inkonzistenci ve smyslu náhodného odpovídání a TRIN, která rozpoznává tendenci odpovídat převážně souhlasně či nesouhlasně. (Friedman et al., 2015; Greene, 2000)

Obsahové položky jsou tvořeny 15 škálami (např. ANX – úzkost, ANG – zloba, LSE – nízká sebeúcta) a jejich komponentami (Friedman et al., 2015; Greene, 2000). U položek sytících tyto škály je přitom zřejmé, na co se ptají, a výsledky tedy poskytují informace o tom, co respondent chce sdělit. Škály PSY-5 se zaměřují spíše na sledování osobnostních rysů než symptomatiky. Ukazují jak na chronicitu problému, tak na osobnostní rysy, které mohou indikovat osobnostní psychopatologii. MMPI-2 obsahuje i seznamy tzv. kritických položek, kterým se však v naší studii nevěnujeme. Celkem MMPI-2 poskytuje přes 130 různých skóre vyjádřených v tzv. uniformních T-skórech (Auke Tellegen & Ben-Porath, 1992) za účelem korekce zešíkmení a percentilové srovnatelnosti napříč nimi.

Psychometrické parametry MMPI-II

Dle manuálu k české verzi metody obsahuje český standardizační vzorek z roku 1999 až 2002 soubor 325 žen a 325 mužů vybraných srovnatelně s původním americkým vzorkem (Butcher et al., 2013). Je však nutné zmínit, že původní verze MMPI byla před rokem 2000 v českém klinickém prostředí nejpoužívanějším diagnostickým nástrojem pro posouzení psychopatologie (Svoboda et al., 2004), ačkoli byla dostupná pouze v různých experimentálních verzích a nebyla oficiálně vydána ani standardizována (Butcher et al., 2002).

Přestože se český manuál validitě příliš nevěnuje, existuje rozsáhlá databáze validizačních výzkumů pro různé účely, a to i v policejním kontextu (Sellbom et al., 2007); jejich shrnutí však přesahuje možnosti tohoto článku a zájemce proto odkazujeme na internetové stránky původního

distributora¹. Jak jsme již zmínili, české validizační studie chybějí; to je ostatně i hlavní výtkou recenze české adaptace (Čurdová & Pourová, 2019), která MMPI-2 hodnotí jako vhodné pro opatrné použití zkušenými uživateli.

Výjimkou jsou diplomové práce Hajného (2018), týkající se invariance měření napříč různými skupinami respondentů, a Bednárové (2016), která podpořila souběžnou validitu validizačních škál L, K a S vůči dotazníku Balanced Inventory of Desirable Responding (BIDR), nikoli však vůči dotazníku přehánění (OCQ-25). Protože jsme diplomové práce nevyhledávali cíleně, je možné, že podobných absolventských studií je dostupné větší množství.

Český manuál metody (Butcher et al., 2013) uvádí vnitřní konzistenci škál na americkém vzorku, která se pohybuje v rozmezí 0,34–0,87 u klinických škál, 0,68–86 u obsahových škál a 0,24–0,90 u doplňkových škál. Test-retest reliabilita nabývá hodnot 0,54–0,93 pro klinické a 0,77–0,91 pro obsahové škály, u doplňkových škál jsou koeficienty obdobné. Reliabilita restrukturalizovaných škál je obdobná nebo vyšší. Většina skóre tedy má podle těchto údajů dobrou nebo uspokojivou reliabilitu.

Reliabilita české verze v manuálu (Butcher et al., 2013) není uvedena. Její odhad nicméně uvádí diplomová práce Hajného (2018) na dvou vzorcích policistů a na vzorku z klinické praxe. Napříč těmito vzorky se vnitřní konzistence pohybovala v rozmezí 0,84–0,97 (validizační škály L, F a K), 0,80–0,98 (klinické škály) a 0,90–0,98 (RC škály). Hajný (2018) dále reportuje výsledky konfirmačních analýz založených na teorii odpovědi na položku (IRT). Zatímco standardní klinické škály se nezdají být uspokojivě jednodimenzionální, validizační škály (kromě K-škály) a RC škály jsou nejenže jednodimenzionální, ale zpravidla vykazují alespoň metrickou (a zpravidla skalární) invarianci měření napříč policejním a klinickým vzorkem.

Výzkumné cíle

Cílem naší studie je ověřit souběžnou validitu metod SPARO a MMPI-II v nezávislé studii na českém vzorku. Využili jsme toho, že se dotazník SPARO některým administrátorům z řad policejních psychologů nejevil dostatečně validním, a to zejména z důvodu nízké odolnosti vůči sociální desirabilitě. Vznikla proto potřeba jej doplnit nebo úplně nahradit jinou metodou, která by byla jednak méně náchylná ke stylizaci, jednak lépe pokrývala kritéria výběru pro vstup do policie. Zvolena byla právě metoda MMPI-2, souběžně administrovaná s dotazníkem SPARO. Tím vznikl vzorek, který používáme v této studii.

Protože u použité české verze MMPI-2 jsou nedostatečně prozkoumané psychometrické parametry, našim vedlejším cílem je za tímto účelem poprvé ověřit vnitřní konzistenci a faktorovou strukturu MMPI-2 stejně jako v případě metody SPARO v naší předcházející studii (Cígler & Rudá, 2021). I zde využijeme data z reálné diagnostiky v kontextu personálního výběru a ověřování profesní způsobilosti u členů policejního sboru České republiky, půjde však o jiný vzorek. Naše výzkumné otázky a hypotézy lze rozdělit do dvou oblastí. První ověřuje psychometrické parametry MMPI-2 (reliabilitu a vnitřní strukturu škál) jako „etalonu“ pro navazující ověření souběžné validity, na které se zaměřuje druhá oblast.

¹ <http://mmpi.umn.edu/>

Psychometrické parametry MMPI-2 jako kritéria pro validizaci SPARO

Výzkumná otázka 1. *Je dotazník MMPI-2 dostatečně reliabilním nástrojem pro individuální diagnostiku?*

Kritéria hodnocení reliability jsou shodná s předchozí studií (Cígler & Rudá, 2021).

Výzkumná otázka 2. *Jsou jednotlivé škály MMPI-2 jednodimenzionální?*

Tato otázka se týká primárně restrukturalizovaných klinických škál (RC), které byly s ohledem (nejen) na jednodimenzionální strukturu konstruovány (Cheng, 2006; Hoelzle & Meyer, 2008; A. Tellegen et al., 2003). Ověříme ji však u všech škál.

H1: Jednotlivé škály MMPI-2 jsou jednodimenzionální.

Souběžná validita SPARO a MMPI-II

Výzkumná otázka 3. *Odpovídají vzájemné vztahy škál SPARO a MMPI-II vztahům očekávaným na základě teorie?*

Jednou z kritik důrazu na vnitřní konzistenci je fakt, že její odhad je vždy jen spodní hranicí skutečné vnitřní konzistence. Zejména v případě multidimenzionálních, mnohofasetových konstruktů je potenciálně možné, aby i škála s nízkým odhadem vnitřní konzistence měřila latentní rys přesně, a díky tomu měla i vysokou prediktivní či souběžnou validitu. V takovém případě by však neměl platit teoretický vztah reliability a validity, popisovaný např. „koeficientem proti oslabení“ (Spearman, 1904); tedy s klesající reliabilitou by neměla klesat i validita. Proto za účelem ověření možnosti věcné interpretace škál nezávisle na odhadované vnitřní konzistenci stanovujeme dvě hypotézy:

H7: Vnitřní konzistence škál dotazníku SPARO pozitivně souvisí s jejich souběžnou validitou vůči škálám dotazníku MMPI-2.²

H8: Vnitřní konzistence škál dotazníku MMPI-2 pozitivně souvisí s jejich souběžnou validitou vůči škálám dotazníku SPARO.

Kromě toho jsme na základě obsahového popisu jednotlivých škál (Butcher et al., 2002; Friedman et al., 2015; Mikšík, 2004) stanovili dílčí hypotézy o vztahu validizační škály, základních komponent a obecnějších faktorů metody SPARO s klinickými a obsahovými škálami dotazníku MMPI-2. Tyto hypotézy vznikly vzájemným obsahovým porovnáním obsahu škál a praktickou úvahou.

H9: **L skór věrohodnosti podávaných výpovědí** dotazníku SPARO pozitivně koreluje se standardními validizačními škálami MMPI-2:

H9a: se škálou L, tedy zkreslení odpovědí ve snaze projevit společensky žádoucí kvality

H9b: se škálou K, určené k detekování defenzivního postoje.

² Číslování hypotéz je společné s předcházejícím článkem (Cígler & Rudá, 2021), který obsahuje hypotézy H2–H6. Je tomu z důvodu přehlednosti společného analytického skriptu ve společném on-line supplementu.

- H10: **Kognitivní variabilnost (KO)**, která se vyznačuje tíhnutím ke změně a vyhledáváním dynamických a intenzivních vnějších podnětů:
H10a: pozitivně koreluje se škálou **Hypomanie (Ma)**, která souvisí s impulzivitou a hyperaktivitou.
H10b_i–xv: nebude korelovat s jednotlivými **obsahovými škálami MMPI-II**.
- H11: **Emocionální variabilnost (EM)**, jejíž vysoké hodnoty značí emocionální vzrušivost, tenzi i euforii:
H11a: pozitivně koreluje se standardní klinickou škálou **Psychastenie (Pt)**, která ukazuje na úzkosti, nerozhodnost, sebepochybnosti a problémy s koncentrací.
H11b: pozitivně koreluje s obsahovou škálou **Úzkost (ANX)**, kde respondenti sami reportují úzkosti a nervozitu.
- H12: **Regulační variabilnost (RE)**, ukazující na nízké sebeovládání a nepředvídaní důsledků svého chování:
H12a: pozitivně souvisí s klinickou škálou **Hypomanie (Ma)**, která detekuje osoby hyperaktivní, impulzivní a s obtížemi s oddálením uspokojení.
H12b: pozitivně souvisí s obsahovou škálou **Zloba (ANG)**, která značí potíže se ztrátou sebekontroly ve vzteku, netrpělivost.
- H13: **Adjustační variabilnost (AD)**, která při vysokých hodnotách ukazuje na schopnost se přizpůsobit, zatímco při nízkých na tendenci přizpůsobovat si:
H13a: pozitivně souvisí s **Hysterií (Hy)**, která identifikuje osoby, které se samy považují za dobře přizpůsobené, v zátěži však selhávají a bývají popisovány jako egocentričtí a manipulativní.
H13b: negativně souvisí se škálou **Psychopatie (Pd)**, která detekuje osoby nepředvídatelné, sociálně nekonformní a konfliktní osoby, které okolí vidí jako společenské a sociálně obratné.
H13c: pozitivně koreluje s obsahovou škálou **Cynismus (CYN)**, sledující oblast vnějších agresivních tendencí a ukazující na hostilní nastavení vůči druhým lidem.
- H14: **Obecná hladina psychické vzrušivosti (PV)**, která značí vysokou situační vzrušivost a tíhnutí k dynamickým interakcím:
H14a: pozitivně (a spíše slabě) souvisí se škálou **Psychastenie (Pt)**, která poukazuje na vyšší emocionalitu
H14b: pozitivně (a spíše slabě) souvisí se škálou **Hypomanie (Ma)**, která značí tendenci k vyhledávání dynamických podnětů.
H14c_i–xv: nesouvisí s jednotlivými **obsahovými škálami MMPI-2**.
- H15: **Motorická hybnost (MH)**, značící vyhledávání změn při nižší sebe-regulaci a emocionální a adjustační rigiditě:
H15a: pozitivně souvisí se škálou **Hypomanie (Ma)**, která detekuje osoby s tendencí k hyperaktivitě.
H15b_i–xv: nesouvisí s jednotlivými **obsahovými škálami MMPI-2**.

Metoda

Výzkumný vzorek a procedura

Pro účely této studie byla využita data získaná v průběhu reálné testové situace, poskytnutá Policejním prezidiem České republiky. Kontext sběru dat je popsán v předchozí studii (Cígler & Rudá, 2021).

Vzorek tvoří $N_{MMPI} = 834$ osob testovaných v letech 2013–2014 ve výběrové situaci, z toho 613 mužů (74 %) a 221 žen (26 %). Průměrný věk byl 24,3 roku ($SD = 5,4$). 97 % vzorku tvořili uchazeči (civilisté) o služební poměr v policii, 1,0 % příslušníci jiných bezpečnostních sborů přestupující k policii a 2 % příslušníci policie testovaní v rámci dalšího interního výběru. Všechny tyto osoby byly testované za účelem zjišťování osobnostní způsobilosti, 35 % s výsledkem způsobilý k výkonu služby, 65 % s výsledkem nezpůsobilý. Data obsahují souhrnné skóry škál testu SPARO a MMPI-2 a dále pak položky MMPI-2 pro každého respondenta.

Všechna data byla získána během supervidované elektronické administrace na počítači, případně ve formě tužka–papír počítačem vyhodnocené. Použitými metodami jsou SPARO (Mikšík, 2004) a MMPI-2 ve druhém vydání s restrukturalizovanými škálami (Butcher et al., 2013). Podle distributorů obou metod jsou papírové a počítačové formy dotazníků ekvivalentní. Rozdíly mezi muži a ženami jsou zanedbány (vyjma škály Mf) vzhledem k tomu, že struktura obou metod by měla být pro obě pohlaví velmi podobná (Butcher et al., 2013; Mikšík, 1992, 2004, 2009). Navíc i námi pozorované rozdíly byly spíše zanedbatelné; 95 % rozdílů mezi muži a ženami (Kendallové tau) na všech škálách MMPI-2 a SPARO bylo v rozmezí $\pm 0,2$. Větší efekty než $\pm 0,3$ pak byly v MMPI-2 dle očekávání pozorovány jen u škál maskulinita–feminita (Mf, $\tau = 0,58$), maskulinní (GM, $\tau = -0,37$) a feminní role (GF, $\tau = 0,56$).

Škála maskulinita–feminita (Mf) dotazníku MMPI-II byla problematická i z toho důvodu, že některé položky jsou skórovány rozdílně (pozitivně vs. negativně) pro muže a ženy. U této škály proto provádíme analýzu separátně pro obě pohlaví a anotujeme je jako MfM (pro muže), resp. MfF (pro ženy).

Statistická analýza dat

Veškeré analýzy byly provedeny v prostředí R ver. 3.6.2 (R Core Team, 2020) s využitím několika knihoven (Jorgensen et al., 2020; Lakens, 2017; Mangiafico, 2020; Phillips, 2017; Revelle, 2019; Rosseel, 2012; Signorell & et al., 2020; Torchiano, 2019; Wickham & Bryan, 2019). Pro všechny analýzy byla vzhledem k velikosti vzorků a počtu testů zvolena přísnější hladina spolehlivosti $\alpha = 0,01$. Analytický skript, doplňkové soubory a další materiály jsou veřejně dostupné na <https://doi.org/10.17605/OSF.IO/CNU58>. Nemáme svolení sdílet zdrojová data, proto je na uvedeném odkazu k dispozici pouze jejich struktura a deskriptivní statistiky včetně korelačních matic (Pearsonových i tetrachorických).

Před zahájením analýz jsme zkontrolovali, zda námi použité zařazení položek do škál je správné. Narazili jsme na dílčí nesrovnalosti, které by však neměly ovlivnit výsledky analýz. Podrobnosti jsou uvedeny v příloze 2.

Reliabilita, faktorová validita. Odhad reliability a faktorové struktury dotazníku MMPI-2 je shodný s postupem, který jsme zvolili v případě dotazníku SPARO, a je proto podrobně popsán v předcházející studii (Cígler & Rudá, 2021). V případě počtu vynechaných položek (škála Nemohu

řící, „?“) však nebylo smysluplné odhadovat faktorovou strukturu a při odhadu reliability jsme museli zvolit specifický postup, protože značnou část položek nevynechal žádný z respondentů a běžné postupy tak selhaly. Z toho důvodu jsme provedli 10 000 náhodných rozdělení položek na dvě stejně velké poloviny a pro každou spočítali odhad split-half reliability s pomocí Guttmanovy λ_4 . Průměr z těchto odhadů jsme použili namísto koeficientu α (který lze chápat jako průměr všech možných split-half reliabilit), nejnižší pak jako Revellovo β a nejvyšší jako λ_4 a r_{max} (koeficient r_{glb} byl vynechán). Pro popis a interpretaci koeficientů viz (Cígler & Rudá, 2021), případně další literaturu (Benton, 2015; Guttman, 1945; Revelle & Zinbarg, 2009).

Vztah reliability a validity (H7, H8). Pro každou škálu dotazníku SPARO (resp. MMPI-2) sestavíme regresní model, kde závislou proměnnou bude daná škála a prediktorem všechny škály MMPI-2 (resp. SPARO). Souvislost neadjustovaných vysvětlených rozptylů z těchto modelů s vnitřní konzistencí příslušných škál (r_{max}) následně ověříme pomocí Spearmanova koeficientu pořadové korelace.

Dílčí hypotézy (H9–H15). V případě hypotéz, kde předpokládáme existenci vztahu dvou škál, použijeme korelační koeficient a oboustranný test hypotéz s hladinou spolehlivosti $\alpha = 0,01$. V případě hypotéz, kde naopak předpokládáme neexistenci vztahu dvou škál (např. H6b nebo H10b), použijeme dva paralelní postupy.

Prvním z nich je test ekvivalence TOST (Rose et al., 2018) odhadnutý s pomocí funkce `dataTOSTr` z balíčku `TOSTER` (Lakens, 2017). Tento postup se zaměřuje vždy na dvě škály a ověřuje, zda je jejich vzájemná korelace slabší než zvolená kritická hodnota. S pomocí power-analýzy poskytované funkcí `powerTOSTr` ze stejného balíčku, zvolené hladiny signifikance $\alpha = 0,01$, síle testu $1-\beta = 0,95$ a počtu respondentů $N_{MMPI} = 834$, jsme odhadli kritické hodnoty jako $\pm 0,148$; po zaokrouhlení využíváme hranice $\pm 0,150$. Jinými slovy, hypotézu budeme považovat za podpořenou, pokud pozorované korelace budou signifikantně ($\alpha = 0,01$) ležet v rozmezí $\pm 0,15$. Tento test opakujeme pro každou kriteriální škálu separátně.

Druhý zvolený postup je „globální“ a ověřuje, zda zvolený skór nesouvisí se všemi skóry (typicky obsahovými škálami) dotazníku MMPI-2. Za tímto účelem využijeme strukturní model, v němž budou všechny kriteriální škály dotazníku MMPI-2 vzájemně volně korelovat, avšak jejich souvislosti s testovanou škálou dotazníku SPARO budou zafixována na nulu. Způsob interpretace a odhadu modelu pak bude založen na běžných indexech dobré shody; pro podrobný popis viz šestou hypotézu (H6) v předcházejícím článku (Cígler & Rudá, 2021).

Výsledky

Deskriptivní statistiky

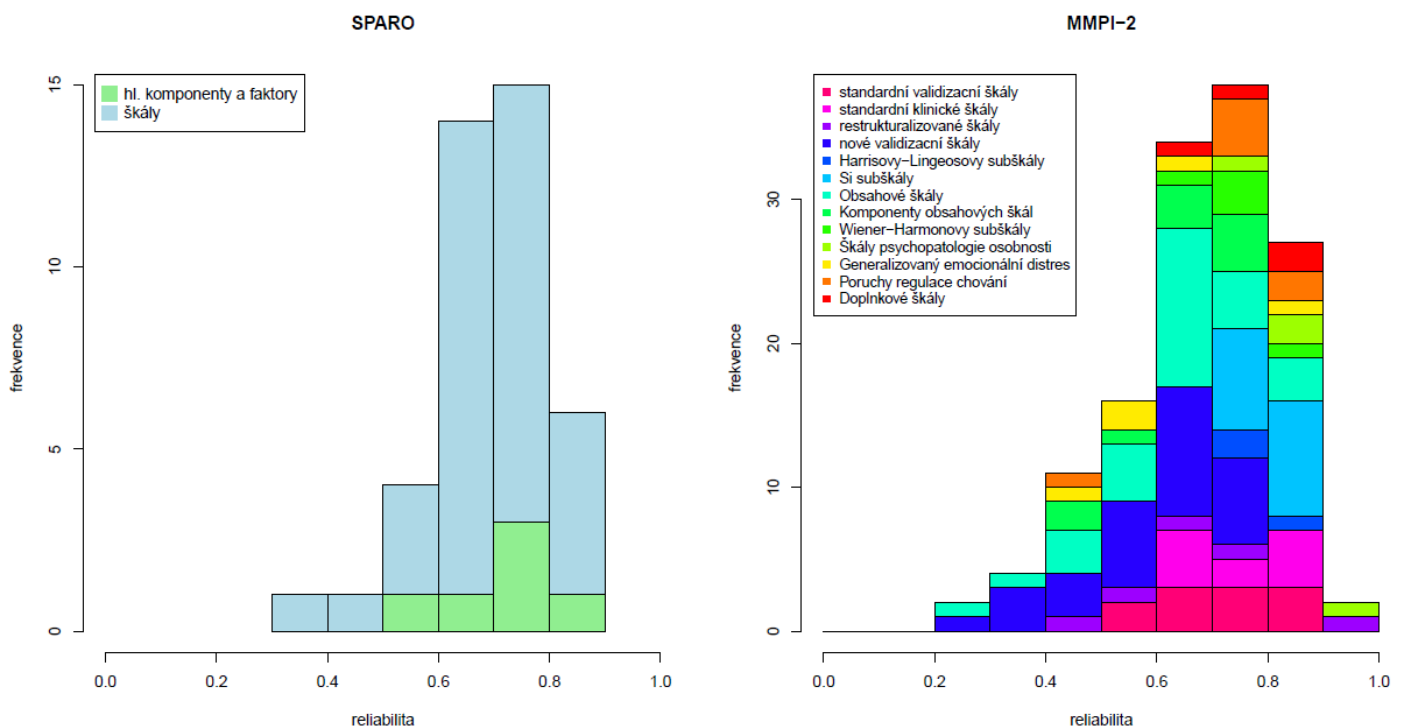
Deskriptivní statistiky a matice Pearsonových i tetrachorických korelací položek jsou vzhledem ke svému rozsahu uvedeny v on-line supplementu. Položky dotazníku MMPI-2 byly spíše obtížnější, průměr popularit položek (podílu souhlasných odpovědí) byl 0,34 ($SD = 0,32$, $Md = 0,22$). Na položky 489 (problém s alkoholem a drogami) a 506 (úvahy o sebevraždě) dokonce nikdo neodpověděl souhlasně – zřejmě z důvodu sociální žádoucnosti – a byly proto vyřazeny z datové matice při odhadu reliabilit.

Celkově chybělo 0,19 % dat MMPI-2, počet se však lišil napříč respondenty. Průměrný počet vynechaných položek jedním respondentem byl 1,1 položky (0,2 %; $SD = 2,5$). Jen 5 % respondentů

mělo více než 6 (1,1 %) vynechaných položek, maximum bylo 18 (3,2 %). Rovněž napříč položkami se lišilo množství vynechaných dat; nejčastěji byla vynechána položka 524 týkající se sebevraždy ($n = 33$), hned první položka týkající se technických časopisů ($n = 24$), položka 345 ($n = 22$), 43 ($n = 20$) či 560 ($n = 18$). Hrubé skóry MMPI-2 byly po vzoru originálního postupu spočteny jako součet položek bez ohledu na chybějící data. Vzhledem k povaze testování nechyběla žádná data dotazníku SPARO.

Extremita položek činila problém při odhadu tetrachorické korelační matice. Velmi mnoho bivariačních frekvenčních tabulek obsahovalo nuly (žádný z respondentů nevyužil některou z možných kombinací odpovědí na dvě položky), což může zásadním způsobem zkreslit odhad tetrachorických korelací. Deskriptivní statistiky všech položek MMPI-2, společně s jejich Pearsonovou a tetrachorickou korelační maticí, jsou součástí on-line supplementu. Obrázek 1 předkládá rozložení popularit položek.

Deskriptivní statistiky hlavních škál MMPI-2 jsou uvedeny níže (společně s vnitřní konzistencí) v Tabulka 1. Deskriptivní statistiky všech škál a jejich histogramy jsou k dispozici v on-line supplementu.



Obrázek 1:

Popularita položek a rozložení reliability jednotlivých škál dotazníku MMPI-2

Psychometrické parametry MMPI-2

Vnitřní konzistence

Reliability nejčastěji používaných škál dotazníku MMPI-2 jsou uvedeny v Tabulka 1 společně s deskriptivami celkových skóre (tabulka všech škál je uvedena v on-line supplementu). Celkem 67 (50 %) všech škál mělo neakceptovatelnou reliabilitu, 38 (28 %) akceptovatelnou, 27 (20 %) dobrou a jen 2 škály (1 %) výbornou. Dohromady pět z deseti hlavních klinických škál mělo neakceptovatelnou reliabilitu, 3 akceptovatelnou a jen 2 dobrou. Pro rozložení reliabilit viz Obrázek 1. Pokud vezmeme v úvahu jen nejčastěji používané skóre, tedy standardní a restrukturalizované klinické škály, standardní a nové validizační škály a obsahové škály, pak 13 (29 %) z nich mělo neakceptovatelnou úroveň reliability, 14 (31 %) akceptovatelnou, 17 (38 %) dobrou a jediná (2 %) výbornou vnitřní konzistenci.

Vnitřní konzistence všech skóre MMPI-2 se pohybovala v rozmezí 0,27–0,90 s mediánem 0,69. Reliabilita podle typu škál byla následující: standardní klinické škály $Md = 0,72$ (rozmezí hodnot 0,59–0,87); standardní validizační škály $Md = 0,80$ (0,68–0,89); restrukturalizované klinické škály $Md = 0,70$ ($Md = 0,65$ –0,86); nové validizační škály $Md = 0,63$ (0,43–0,90); obsahové škály $Md = 0,80$ (0,72–0,90); komponenty obsahových škál $Md = 0,64$ (0,27–0,87); doplňkové škály $Md = 0,72$ (0,44–0,90); generalizovaný emocionální distres $Md = 0,86$ (0,72–0,90); Si subškály $Md = 0,79$ (0,72–0,81); škály psychopatologie osobnosti $Md = 0,72$ (0,68–0,88); Harrisovy–Lingeosovy subškály $Md = 0,63$ (0,29–0,75); poruchy regulace chování $Md = 0,57$ (0,43–0,89); Wienerovy–Harmonovy subškály $Md = 0,68$ (0,45–0,78).

Rozdíl mezi vnitřní konzistencí všech skóre dotazníku SPARO (Cígler & Rudá, 2021) a MMPI-2 (tato studie) ověřený Mann-Whitneyho testem nebyl statisticky významný, $p(175) = 0,745$, se zanedbatelnou velikostí účinku, $r = 0,025$. Celkově se nelišilo ani rozložení reliabilit napříč oběma testy, Kolmogorův–Smirnovův test $D = 0,111$, $p = 0,782$.

Tabulka 1: Deskriptivy a reliability dotazníku MMPI-2 (standardní a restrukturalizované klinické škály, standardní a nové validizační škály a obsahové škály)

	M	SD	Md	min	max	Sk	α	λ_4	r_{glb}	β	nf	r_{max}	hodnocení r_{max}
<i>Standardní validizační škály</i>													
?	1,04	2,50	0	0	18	3,56	0,833	0,891		0,720		0,891	dobré
L	6,00	2,76	6	0	14	0,21	0,704	0,765	0,788	0,515	8	0,788	akceptovatelné
F	3,44	2,30	3	0	19	1,40	0,486	0,597	0,684	0,290	27	0,684	neakceptovatelné
K	19,28	4,22	20	6	29	-0,37	0,746	0,810	0,815	0,599	13	0,815	dobré
<i>Nové validizační škály</i>													
Fb	1,12	1,39	1	0	11	2,08	0,519	0,632	0,718	0,331	17	0,718	akceptovatelné
Fp	1,54	1,32	1	0	6	0,85	0,279	0,459	0,523	-0,125	13	0,523	neakceptovatelné
VRIN	4,38	2,35	4	0	13	0,70	0,406	0,530	0,627	0,259	31	0,627	neakceptovatelné
TRIN	8,19	1,27	8	4	14	0,38	0,161	0,326	0,426	-0,123	11	0,426	neakceptovatelné
S	33,23	7,64	33	9	49	-0,19	0,866	0,901	0,903	0,764	20	0,903	výborné

	M	SD	Md	min	max	Sk	α	λ_4	r_{glb}	β	nf	r_{max}	hodnocení r_{max}
<i>Standardní klinické škály</i>													
Hs	3,11	2,48	3	0	14	1,19	0,598	0,680	0,722	0,447	15	0,722	akceptovatelné
D	16,74	3,41	17	7	32	0,44	0,397	0,561	0,640	-0,443	25	0,640	neakceptovatelné
Hy	19,59	3,65	20	8	32	0,01	0,452	0,598	0,616	-0,222	26	0,616	neakceptovatelné
Pd	15,37	2,98	15	3	28	0,36	0,251	0,452	0,598	-0,469	22	0,598	neakceptovatelné
Mf (m)	22,81	3,74	22	11	37	0,28	0,602	0,707	0,745	0,333	24	0,745	akceptovatelné
(ž)	31,70	3,81	32	24	42	0,15	0,588	0,749	0,763	0,276	24	0,763	akceptovatelné
Pa	8,25	2,44	8	2	16	0,21	0,234	0,470	0,593	-0,796	18	0,593	neakceptovatelné
Pt	5,95	4,52	5	0	28	1,55	0,806	0,854	0,874	0,721	21	0,874	dobré
Sc	6,82	4,65	6	0	30	1,44	0,766	0,823	0,859	0,621	34	0,859	dobré
Ma	15,54	3,54	15	6	31	0,44	0,444	0,587	0,637	0,001	20	0,637	neakceptovatelné
Si	22,36	6,42	22	6	48	0,86	0,742	0,813	0,848	0,443	27	0,848	dobré
<i>Restrukturalizované validizační škály</i>													
FBS	9,41	3,04	9	1	21	0,15	0,539	0,712	0,679	-0,258	17	0,712	akceptovatelné
RCd	1,83	2,31	1	0	19	2,81	0,739	0,809	0,849	0,617	11	0,849	dobré
RC1	2,36	2,02	2	0	12	1,23	0,536	0,637	0,651	0,355	12	0,651	neakceptovatelné
RC2	4,40	2,20	4	0	12	0,36	0,540	0,633	0,662	0,341	8	0,662	neakceptovatelné
RC3	6,98	3,48	7	0	15	-0,19	0,811	0,855	0,865	0,696	7	0,865	dobré
RC4	1,57	1,61	1	0	10	1,31	0,526	0,648	0,690	0,279	10	0,690	neakceptovatelné
RC6	1,09	1,35	1	0	8	1,71	0,525	0,620	0,677	0,275	8	0,677	neakceptovatelné
RC7	3,02	3,12	2	0	18	1,44	0,785	0,846	0,853	0,651	11	0,853	dobré
RC8	1,97	1,89	2	0	11	1,35	0,595	0,679	0,707	0,408	9	0,707	akceptovatelné
RC9	8,52	4,33	8	0	25	0,59	0,763	0,835	0,841	0,664	12	0,841	dobré
<i>Obsahové škály</i>													
ANX	3,12	2,89	2	0	18	1,31	0,739	0,821	0,828	0,512	11	0,828	dobré
FRS	2,88	2,54	2	0	14	1,21	0,693	0,778	0,791	0,439	10	0,791	akceptovatelné
OBS	2,76	2,63	2	0	14	1,19	0,740	0,792	0,814	0,648	8	0,814	dobré
DEP	3,11	2,70	3	0	21	1,79	0,684	0,764	0,811	0,514	15	0,811	dobré
HEA	3,99	2,79	3	0	17	1,07	0,622	0,712	0,734	0,456	16	0,734	akceptovatelné
BIZ	2,02	1,95	2	0	10	1,22	0,604	0,712	0,722	0,344	12	0,722	akceptovatelné
ANG	3,63	2,53	3	0	15	0,89	0,705	0,773	0,795	0,580	8	0,795	akceptovatelné
CYN	10,80	5,26	11	0	23	-0,11	0,858	0,893	0,892	0,749	9	0,893	dobré
ASP	6,44	3,34	6	0	17	0,17	0,742	0,824	0,818	0,436	10	0,824	dobré
TPA	5,89	2,81	6	0	18	0,56	0,672	0,754	0,739	0,529	7	0,754	akceptovatelné
LSE	2,93	2,59	2	0	17	1,50	0,700	0,773	0,780	0,489	11	0,780	akceptovatelné
SOD	5,70	3,35	5	0	20	1,19	0,740	0,814	0,851	0,528	11	0,851	dobré
FAM	2,20	2,49	1	0	17	1,70	0,727	0,804	0,789	0,564	11	0,804	dobré
WRK	4,13	3,27	3	0	21	1,65	0,750	0,807	0,831	0,608	14	0,831	dobré
TRT	3,63	2,89	3	0	18	1,12	0,696	0,781	0,788	0,547	11	0,788	akceptovatelné

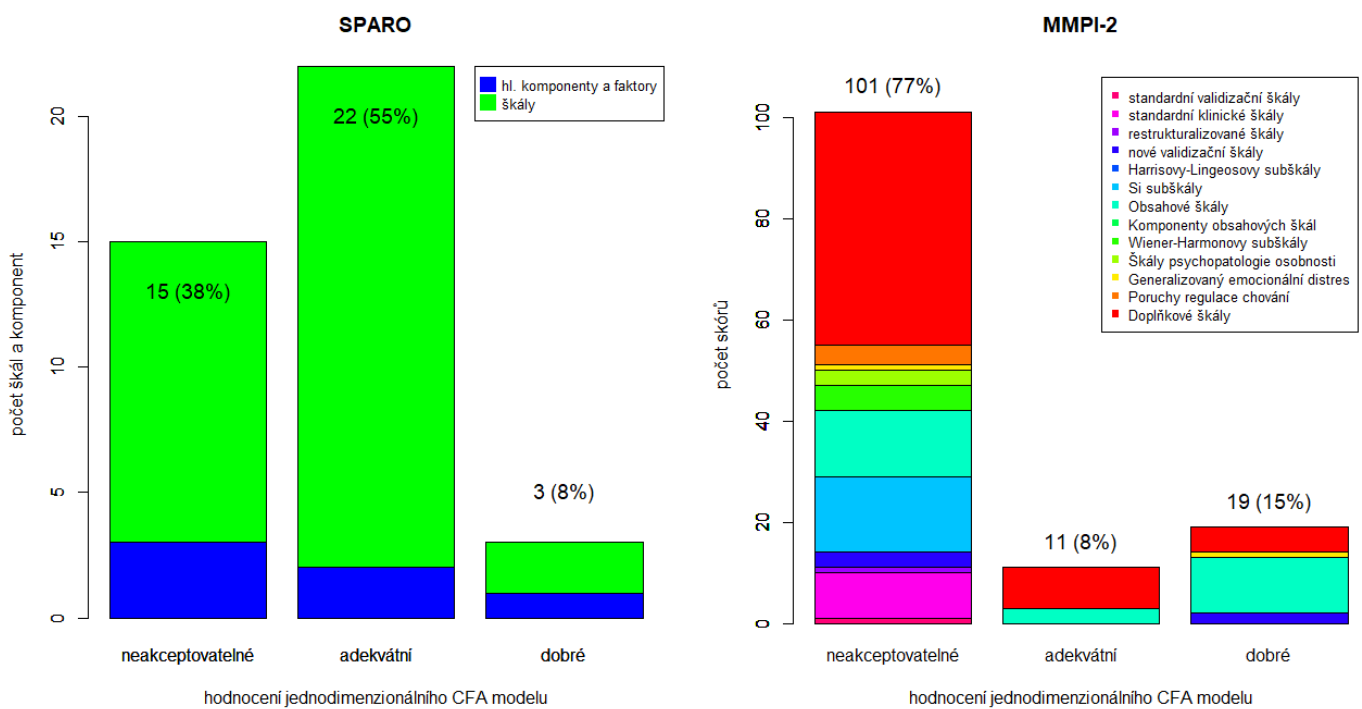
$N = 834$. *Poznámka:* M – průměr; SD – směrodatná odchylka; Md – medián; min, max – minimální a maximální pozorovaný skóre; Sk – zešíkmení škály; α – Cronbachova alfa; λ_4 – maximální split-half reliabilita; r_{glb} – koeficient GLB; β – nejnižší split-half reliabilita; nf – počet faktorů využitých pro odhad koeficientů GLB a β ; r_{max} – vyšší z koeficientů λ_4 a r_{glb} . Deskriptivy a reliability všech škál dotazníku MMPI-2 jsou uvedeny v on-line supplementu.

H1: Jednodimenzionalita škál MMPI-2

Během odhadu konfirmačních modelů jsme pozorovali záporný reziduální rozptyl jedné z položek u škál PA, SC2, FRS1 a HEA1. Vždy jsme jej omezili na hodnotu větší než 0,01 a model odhadli znovu; vliv na shodu s daty byl nicméně minimální.

Obrázek 2 předkládá hodnocení shody jednodimenzionálních modelů s daty pro dotazník SPARO (Cígler & Rudá, 2021) a MMPI-2, v Tabulka 2 pak jsou výsledky pro klinické, validizační a RC škály. Podrobné výsledky pro všechny škály jsou součástí on-line supplementu.

U většiny škál MMPI-2 (101; 77 %) jednodimenzionální model popsal data neuspokojivě. Pouze v 11 případech (8 %) byl model adekvátní a v 19 (15 %) dobrý. Je patrné, že převážně šlo o obsahové a doplňkové škály. Kromě RCd nebylo možné žádnou z restrukturalizovaných klinických škál, které by na základě teoretických východisek měly být jednodimenzionální, hodnotit jako jednofaktorovou (u škály cynismu, RC3, byla shoda s daty hraniční).



Obrázek 2:
Jednodimenzionalita škál obou dotazníků

Tabulka 2:

Ukazatele shody jednodimenzionálních CFA modelů s daty u základních validizačních, klinických a RC škál

	χ^2	df	p	TLI	RMSEA	[95% CI]	SRMR	RMSEA _{null}	hodnocení
<i>Validizační škály</i>									
L	151,4	90	0	0,963	0,029	[0,02; 0,036]	0,077	0,148	adekvátní
F	2435,2	1710	0	0,717	0,023	[0,02; 0,025]	0,166	0,042	neakceptovatelné
K	858,8	405	0	0,869	0,037	[0,033; 0,04]	0,091	0,101	neakceptovatelné
<i>Klinické škály</i>									
HS	699,2	464	0	0,665	0,025	[0,021; 0,028]	0,128	0,043	neakceptovatelné
D	2282,1	1539	0	0,683	0,024	[0,022; 0,026]	0,125	0,043	neakceptovatelné
HY	2340,6	1710	0	0,782	0,021	[0,019; 0,023]	0,123	0,045	neakceptovatelné
PD	1880,6	1175	0	0,704	0,027	[0,025; 0,029]	0,117	0,049	neakceptovatelné
MfM	2671,5	1484	0	0,416	0,036	[0,034; 0,038]	0,131	0,047	neakceptovatelné
MfF	1778,8	1484	0	0,363	0,031	[0,025; 0,036]	0,168	0,039	neakceptovatelné
PA	3255,6	740	0	0,332	0,064	[0,062; 0,066]	0,278	0,078	neakceptovatelné
PT	1405,3	1080	0	0,879	0,019	[0,016; 0,022]	0,129	0,055	neakceptovatelné
SC	4170,2	2925	0	0,696	0,023	[0,021; 0,024]	0,169	0,041	neakceptovatelné
MA	1442,4	989	0	0,817	0,023	[0,021; 0,026]	0,097	0,055	neakceptovatelné
SI	5249,1	2277	0	0,562	0,04	[0,038; 0,041]	0,126	0,060	neakceptovatelné
<i>Restrukturalizované klinické škály</i>									
RCd	302,3	252	0,016	0,961	0,015	[0,007; 0,022]	0,107	0,078	dobré
RC1	511,0	324	0	0,560	0,026	[0,022; 0,031]	0,137	0,040	neakceptovatelné
RC2	233,0	119	0	0,779	0,034	[0,027; 0,04]	0,109	0,072	neakceptovatelné
RC3	297,7	90	0	0,952	0,053	[0,046; 0,059]	0,080	0,240	neakceptovatelné
RC4	312,9	189	0	0,908	0,028	[0,022; 0,033]	0,142	0,092	neakceptovatelné
RC6	316,0	119	0	0,779	0,045	[0,039; 0,051]	0,154	0,095	neakceptovatelné
RC7	411,2	252	0	0,932	0,028	[0,023; 0,032]	0,116	0,106	neakceptovatelné
RC8	307,7	135	0	0,779	0,039	[0,033; 0,045]	0,124	0,083	neakceptovatelné
RC9	1093,3	350	0	0,782	0,05	[0,047; 0,054]	0,099	0,108	neakceptovatelné

Pozn.: U škály maskulinismu-feminismu jsou výsledky zvlášť pro muže (MfM) a ženy (MfF).

CI – interval spolehlivosti. *RMSEA_{null}* – *RMSEA* základního modelu (s předpokladem nekorelovaných položek). Pokud je menší než 0,158, je index *TLI* neinformativní (Kenny et al., 2014).

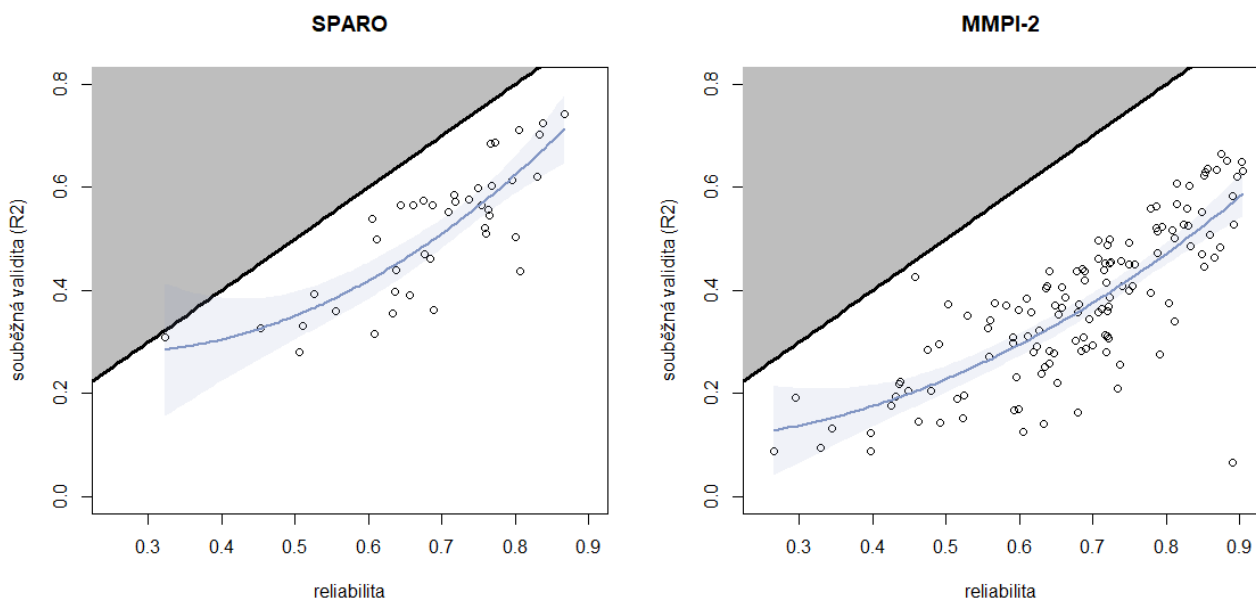
Souběžná validita

Pearsonovy a Spearmanovy korelační matice všech škál obou dotazníků jsou uvedeny v online supplementu.

H7, H8: Souvislost reliability a validity

Souvislost validity a vnitřní konzistence byla velmi silná. Reliabilita jednotlivých škál dotazníku SPARO souvisela s jejich souběžnou validitou (rozptylem vysvětleným všemi skóry MMPI-2), Spearmanovo $\rho_{H7} = 0,783$, $p < 0,001$, a stejně tak reliabilita škál MMPI-2, $\rho_{H8} = 0,779$, $p < 0,001$.

Vzájemný vztah reliability a souběžné validity pro obě metody je zobrazen na obrázku 3, a to včetně maximální teoretické validity. Outlierem v případě metody MMPI-2 s vysokou reliabilitou, avšak nízkou validitou je škála „nemohu říci“ (?), tedy počet vynechaných položek.



Obrázek 3:

Vztah reliability a souběžné validity. Oblast vlevo nahoře vymezuje teoreticky možné hodnoty souběžné validity

H9: věrohodnost podávaných výpovědí (L)

H9a: Věrohodnost podávaných výpovědí (L) v rozporu s hypotézou nekoreluje s L škálou MMPI-2, $r(832) = -0,030$, $p = 0,391$.

H9b: Věrohodnost podávaných výpovědí (L) škála v rozporu s hypotézou nekoreluje s K škálou, $r(832) = 0,053$, $p = 0,128$.

Hypotézy H9 nebyly podpořeny.

H10: Kognitivní variabilnost (KO)

H10a. Hlavní komponenta KO dotazníku SPARO (kognitivní variabilnost) v souladu s hypotézou koreluje se škálou MMPI Hypomanie (MA), $r(832) = 0,380$, $p < 0,001$, přestože korelace je relativně slabá.

H10b. Jak je patrné z Tabulka 3 (a v grafické podobě v online supplementu), kognitivní variabilnost v rozporu s hypotézou statisticky významně a pozitivně korelovala s většinou obsahových škál MMPI-2 ($p < 0,01$). Výjimkou byly škály zájem o vlastní zdraví (HEA; $r = -0,135$), nízká sebeúcta (LSE; $r = -0,010$), poruchy pracovního výkonu (WRK; $r = 0,050$) a negativní postoje k terapii (TRT; $r = 0,062$), se kterými kognitivní variabilnost korelovala signifikantně slaběji než 0,15 ($p < 0,01$). Všechny obsahové škály predikovaly KO dohromady středně silně, $adjR^2 = 0,169$ a úsekový model s regresními koeficienty nastavenými na nulu nepopsal data uspokojivě, $\chi^2(15) = 174,8$, $p < 0,001$, $TLI = 0,831$, $RMSEA = 0,111$ s $90\%CI = [0,097, 0,126]$, $SRMR = 0,050$.

Hypotéza nebyla spolehlivě podpořena, korelace Kognitivní variabilnosti s obsahovými škálami byla věcně zanedbatelná pouze ve 4 případech (27 %), zatímco v 11 případech (73 %) se statisticky významně lišila od 0.

Hypotézy H10 celkově nebyly podpořeny.

Tabulka 3:

Souvislost Kognitivní variabilnosti (KO) s obsahovými škálami MMPI-2 (H10b)

	NHST			TOST	
	r	[99% CI]	p	p.lower	p.upper
ANX	0,143	[0,055; 0,23]	0,000	0,000	0,424
FRS	-0,135	[-0,221; -0,046]	0,000	0,326	0,000
OBS	0,146	[0,058; 0,232]	0,000	0,000	0,456
DEP	0,108	[0,019; 0,195]	0,002	0,000	0,108
HEA***	0,001	[-0,088; 0,091]	0,968	0,000	0,000
BIZ	0,267	[0,182; 0,348]	0,000	0,000	1,000
ANG	0,178	[0,09; 0,263]	0,000	0,000	0,797
CYN	0,187	[0,1; 0,272]	0,000	0,000	0,865
ASP	0,137	[0,048; 0,223]	0,000	0,000	0,349
TPA	0,253	[0,167; 0,334]	0,000	0,000	0,999
LSE***	-0,010	[-0,1; 0,079]	0,762	0,000	0,000
SOD	-0,133	[-0,219; -0,044]	0,000	0,307	0,000
FAM	0,143	[0,054; 0,229]	0,000	0,000	0,416
WRK**	0,050	[-0,04; 0,138]	0,153	0,000	0,002
TRT**	0,062	[-0,027; 0,15]	0,073	0,000	0,005

Pozn.: NHST – test nulové hypotézy (korelace je odlišná od 0); TOST – výsledky TOST testu; r – Pearsonův korelační koeficient; CI – 99% interval spolehlivosti; p – p-hodnota pro korelační koeficient; p.lower, p.upper – signifikance spodního a horního okraje intervalu spolehlivosti. Pokud jsou obě p-hodnoty menší než $p < 0,02$, pak korelace na hladině spolehlivosti $\alpha = 0,01$ neleží mimo zvolený interval $\pm 0,15$.

** TOST signifikantní na $\alpha = 0,01$; *** TOST signifikantní na $\alpha = 0,001$. Pro vizuální podobu viz on-line supplement.

H11: Emocionální variabilnost (EM)

H11a: Emocionální variabilnost (EM) v souladu s hypotézou silně a pozitivně koreluje se škálou Psychastenie (Pt), $r(832) = 0,727, p < 0,001$.

H11b: Emocionální variabilnost (EM) v souladu s hypotézou silně a pozitivně koreluje se škálou Úzkost (ANX), $r(832) = 0,633, p < 0,001$.

Hypotézy H11 byly podpořeny, a to s vysokou velikostí efektu.

H12: Regulační variabilnost (RE)

H12a: Regulační variabilnost (RE) v souladu s hypotézou spíše slabě, ale pozitivně koreluje se škálou Hypomanie (Ma), $r(832) = 0,359, p < 0,001$.

H12b: Regulační variabilnost (RE) v souladu s hypotézou středně silně a pozitivně koreluje se škálou Zloba (ANG), $r(832) = 0,448, p < 0,001$.

Hypotézy H12 byly podpořeny, velikosti efektů jsou však spíše slabé.

H13: Adjustační variabilnost (AD)

H13a: Adjustační variabilnost (AD) v rozporu s hypotézou koreluje se škálou Hysterie (Hy) negativně, $r(832) = -0,292, p < 0,001$.

H13b: Adjustační variabilnost (AD) v rozporu s hypotézou nekoreluje se škálou Psychopatie (Pd), $r(832) = -0,059, p = 0,086$.

H13c: Adjustační variabilnost (AD) v souladu s hypotézou koreluje se škálou Cynismu (CYN), avšak jen slabě, $r(832) = 0,318, p < 0,001$.

Hypotézy H13 z větší části nebyly podpořeny, část výsledků je v přímém rozporu s hypotézami.

H14: Obecná hladina psychické vzrušivosti (PV)

H14a: Škála Obecné hladiny psychické vzrušivosti (PV) v souladu s hypotézou pozitivně a středně silně koreluje se škálou Psychastenie (Pt), $r(832) = 0,477, p < 0,001$.

H14b: Škála Obecné hladiny psychické vzrušivosti (PV) v souladu s hypotézou pozitivně a středně silně koreluje se škálou Hypomanie (Ma), $r(832) = 0,428, p < 0,001$.

H14c: Jak je patrné z Tabulka 4 (a v grafické podobě v online-supplementu), v rozporu s hypotézou Obecná hladina psychické vzrušivosti (PV) pozitivně a zpravidla středně silně korelovala s většinou obsahových škál; jedinou výjimkou byla Sociální nepohoda (SOD), kde korelace statisticky významně neležela vně intervalu $\pm 0,15$. Všechny obsahové škály dohromady predikovaly skór faktoru PV středně silně, $adjR^2 = 0,408$, a úsekový model s regresními koeficienty nastavenými na nulu popsal data velmi špatně, $\chi^2(15) = 401,3, p < 0,001, TLI = 0,541, RMSEA = 0,187$ s $90\%CI = [0,171, 0,203], SRMR = 0,127$.

Hypotézy H14 byly podpořeny jen částečně.

Tabulka 4:
Souvislost Obecné hladiny psychické vzrušivosti (PV) s obsahovými škálami MMPI-2 (H14c)

	NHST			TOST	
	r	[99% CI]	p	p.lower	p.upper
ANX	0,436	[0,361; 0,506]	0,000	0,000	1,000
FRS	0,131	[0,042; 0,218]	0,000	0,000	0,289
OBS	0,478	[0,406; 0,544]	0,000	0,000	1,000
DEP	0,356	[0,276; 0,431]	0,000	0,000	1,000
HEA	0,222	[0,135; 0,305]	0,000	0,000	0,984
BIZ	0,428	[0,352; 0,498]	0,000	0,000	1,000
ANG	0,495	[0,425; 0,560]	0,000	0,000	1,000
CYN	0,371	[0,291; 0,445]	0,000	0,000	1,000
ASP	0,504	[0,434; 0,567]	0,000	0,000	1,000
TPA	0,518	[0,450; 0,580]	0,000	0,000	1,000
LSE	0,264	[0,179; 0,345]	0,000	0,000	1,000
SOD***	-0,017	[-0,106; 0,072]	0,617	0,000	0,000
FAM	0,356	[0,276; 0,432]	0,000	0,000	1,000
WRK	0,388	[0,309; 0,461]	0,000	0,000	1,000
TRT	0,366	[0,286; 0,441]	0,000	0,000	1,000

Pozn.: NHST – test nulové hypotézy (korelace je odlišná od 0); TOST – výsledky TOST testu; r – Pearsonův korelační koeficient; CI – 99% interval spolehlivosti; p – p-hodnota pro korelační koeficient; p.lower, p.upper – signifikance spodního a horního okraje intervalu spolehlivosti. Pokud jsou obě p-hodnoty menší než $p < 0,02$, pak korelace na hladině spolehlivosti $\alpha = 0,01$ neleží mimo zvolený interval $\pm 0,15$.

** TOST signifikantní na $\alpha = 0,01$; *** TOST signifikantní na $\alpha = 0,001$. Pro vizuální podobu viz on-line supplement.

H15: Motorická hybnost (MH)

H15a: Škála Motorické hybnosti (MH) v souladu s hypotézou pozitivně, avšak jen slabě koreluje se škálou Hypomanie (Ma), $r(832) = 0,148$, $p < 0,001$.

H15b: Škála Motorické hybnosti (MH) v rozporu s hypotézou ve většině případů (10; 67 %) negativně a statisticky významně souvisela s obsahovými škálami MMPI-2, ačkoli velikosti efektů byly celkově nízké. Výjimkou byly souvislosti se škálou Bizarní psychické aktivity (BIZ), Zloba (ANG), Antisociální chování (ASP), Chování typu A (TPA) a Problémy v rodině (FAM), se kterými MH nekorelovala silněji než $\pm 0,15$.

Podrobné výsledky jsou v tabulce 5, nebo v grafické podobě v on-line supplementu.

Všechny obsahové škály dohromady predikovaly skór faktoru PV středně silně, $adjR^2 = 0,183$, a úsekový model s regresními koeficienty nastavenými na nulu nepopsal data uspokojivě, $\chi^2(15) = 190,2$, $p < 0,001$, $TLI = 0,815$, $RMSEA = 0,116$ s $90\%CI = [0,102, 0,131]$, $SRMR = 0,052$.

Hypotézy H15 byly podpořeny částečně.

Tabulka 5: Souvislost Motorické hybnosti (MH) s obsahovými škálami MMPI-2 (H15b)

	NHST			TOST	
	r	[99% CI]	p	p.lower	p.upper
ANX	-0,121	[-0,208; -0,032]	0,000	0,199	0,000
FRS	-0,289	[-0,369; -0,206]	0,000	1,000	0,000
OBS	-0,171	[-0,256; -0,083]	0,000	0,731	0,000
DEP	-0,163	[-0,249; -0,075]	0,000	0,655	0,000
HEA	-0,090	[-0,178; -0,001]	0,009	0,040	0,000
BIZ***	-0,023	[-0,112; 0,067]	0,512	0,000	0,000
ANG***	-0,001	[-0,09; 0,088]	0,974	0,000	0,000
CYN	-0,102	[-0,19; -0,013]	0,003	0,080	0,000
ASP**	-0,063	[-0,152; 0,026]	0,068	0,006	0,000
TPA***	-0,010	[-0,099; 0,079]	0,773	0,000	0,000
LSE	-0,235	[-0,318; -0,149]	0,000	0,995	0,000
SOD	-0,256	[-0,337; -0,171]	0,000	0,999	0,000
FAM***	-0,018	[-0,107; 0,072]	0,612	0,000	0,000
WRK	-0,188	[-0,272; -0,1]	0,000	0,867	0,000
TRT	-0,187	[-0,271; -0,099]	0,000	0,862	0,000

Pozn.: NHST – test nulové hypotézy (korelace je odlišná od 0); TOST – výsledky TOST testu; r – Pearsonův korelační koeficient; CI – 99% interval spolehlivosti; p – p-hodnota pro korelační koeficient; p.lower, p.upper – signifikance spodního a horního okraje intervalu spolehlivosti. Pokud jsou obě p-hodnoty menší než $p < 0,02$, pak korelace na hladině spolehlivosti $\alpha = 0,01$ neleží mimo zvolený interval $\pm 0,15$.

** TOST signifikantní na $\alpha = 0,01$; *** TOST signifikantní na $\alpha = 0,001$.

Diskuze

Tato studie rozšiřuje informace o reliabilitě a faktorové validitě metody SPARO (Mikšík, 2004) z bezprostředně předcházejícího výzkumu (Cígler & Rudá, 2021) a nově přináší poznatky o souběžné validitě dotazníku. Obě studie jsou podle nám dostupných informací prvním nezávislým pokusem validizovat dotazník SPARO; kromě toho jde o první tuzemskou studii (vyjma několika diplomových prací), která se zaměřila na psychometrické parametry české adaptace inventáře MMPI-2 (Butcher et al., 2002, 2013). Výsledky nejsou vstřícné ani vůči metodě SPARO, ani vůči MMPI-2.

Psychometrické parametry MMPI-2

Reliabilitu dotazníku MMPI-2 můžeme v českém prostředí srovnávat pouze s diplomovou prací Hajného (2018). Ať už vezmeme celý jeho výzkumný soubor, nebo jen podsoubor policistů, naše vlastní odhady jsou výrazně nižší. Jsou nicméně srovnatelné s parametry dotazníku v amerických vzorcích tak, jak jsou reportovány v českém manuálu (Butcher et al., 2013); výjimkou jsou standardní klinické škály, kde jsme pozorovali vnitřní konzistenci poněkud vyšší. Je však nutné vzít v úvahu, že náš odhad s pomocí koeficientů r_{glb} a λ_4 adekvátně pracuje s vícedimenzionální strukturou škál a porušením předpokladu tau-ekvivalence, která je právě v případě MMPI-2 výrazná. Pokud tedy

vezmeme v úvahu odhad vnitřní konzistence prostřednictvím koeficientu Cronbachova alfa (který byl zřejmě použit v americkém vzorku), pozorujeme naopak hodnoty nižší a lze tedy očekávat, že česká verze dotazníku MMPI-2 je o něco méně reliabilnější než americký originál. Příčinou mohou být dopady českého překladu na fungování položek, případně pak omezení variability ve specifickém policejním vzorku při srovnání s heterogennějšími americkými vzorky. Obdobný rozdíl se však nemusel projevit u dotazníku SPARO (Cígler & Rudá, 2021), protože Mikšík (1992, 2001, 2004, 2009) stejně jako my nepoužíval reprezentativní výběry z populace, ale dostupné vzorky v kontextu pracovní-psychologického vyšetření, které jsou srovnatelné s naším vlastním vzorkem.

Možná příčina rozdílu mezi našimi výsledky a výsledky Hajného (2018) na obdobných policejních vzorcích může být způsobena rovněž postupem odhadu. Hajný pro odhad jím reportované McDonalduvy celkové omegy (ω_t) a Cronbachovy alfy (α) použil matici tetrachorických korelací (M. Hajný, osobní komunikace, 11. květen 2018). Zdá se však, že zároveň nebyla aplikována příslušná korekce (Green & Yang, 2009), nutná pro interpretaci koeficientu tradičním způsobem. Jeho výsledky tak zřejmě odpovídají spíše tzv. ordinálnímu koeficientu alfa, resp. omega (Gadermann et al., 2012; Zumbo et al., 2007), a nelze je tedy interpretovat tradičním způsobem jako reliabilitu součtu položek (Chalmers, 2018).

Reliabilita je přitom nezbytnou, avšak nikoli postačující podmínkou validity (Lord & Novick, 1968). Všechny odhady vnitřní konzistence jsou vždy založeny na sdíleném rozptylu mezi položkami dané škály a veškeré běžné odhady reliability (Cronbachovo alfa i námi použité vhodnější koeficienty λ_4 a r_{GLB}) jsou tedy pouze spodní hranicí reliability, resp. nejnižší možnou reliabilitou (Revelle & Zinbarg, 2009; Sijtsma, 2009). Je tomu tak proto, že unikátní, s ostatními položkami nesdílený rozptyl, nemusí být výhradně chybou měření, ale může obsahovat i nějaký systematický zdroj variability, který může validitu zajišťovat. To může (společně s porušeným předpokladem tau-ekvivalence koeficientu alfa, který odhad vnitřní konzistence dále podhodnocuje) být např. příčinou toho, že test–retest reliabilita škál MMPI-2 je zpravidla vyšší než jejich vnitřní konzistence (Wise et al., 2010), ačkoli vzhledem k očekávané fluktuaci rysů v čase bychom očekávali opak.

Z našich výsledků je nicméně patrný velmi silný vztah mezi vnitřní konzistencí škál dotazníku SPARO a jejich souběžnou validitou se škálami MMPI-2 (a naopak). Jinými slovy: čím vyšší reliabilita, tím silněji škála souvisí s nezávislými škálami druhého dotazníku. Domníváme se, že tento jev lze do jisté míry zobecnit i na tendenci škál souviset s jinými než dotazníkovými kritérii. Zdá se nepravděpodobné, aby škály s nízkou reliabilitou mohly predikovat libovolně zvolené kritérium (pracovní úspěch, schopnost vykonávat zaměstnání aj.), když nedokáží predikovat ani jiné dotazníkové skóry. To je nutné mít na zřeteli při interpretaci neuspokojivě nízké reliability příslušných škál.

Jedinou výjimkou je škála Nemohu říci (?), tedy počet vynechaných položek v MMPI-2, která i přes vysokou vnitřní konzistenci jen velmi slabě souvisí se škálami SPARO. Lze tedy soudit, že sklon k vynechávání odpovědí je sice stabilní napříč celým dotazníkem, avšak příliš nesouvisí se skutečně volenými odpověďmi. Celkově se pak zdá, že nízká vnitřní konzistence škál MMPI-2 a SPARO opravdu limituje interpretační možnosti obou metod.

Stejně jako v případě dotazníku SPARO (Cígler & Rudá, 2021) rovněž ani většina škál MMPI-2 nevykazuje jednodimenzionální strukturu. Výjimkou je jen část obsahových a několik doplňkových škál. To však samo o sobě nemusí snižovat validitu MMPI-2 – většina skóru byla konstruována za účelem maximalizace prediktivní validity bez ohledu na faktorovou strukturu. Navíc se většina standardních klinických škál sestává z dílčích, Harrisových-Lingeosových subškál, identifikovaných

mj. s využitím faktorové analýzy (Butcher et al., 2013), a nelze tedy předpokládat jejich jednodimenzionalitu. Je však nutné upozornit, že ani tyto subškály zpravidla nejsou jednodimenzionální (s výjimkou části subškál deprese, D). Nelze proto říci, že by vícedimenzionální struktura klinických škál byla definovaná jejich subškálami (např. že by šest subškál schizofrenie Sc1–Sc6 představovalo šest faktorů schizofrenie, Sc). V takovém případě by totiž subškály měly být samy jednodimenzionální, což zpravidla nejsou.

Zásadnějším problémem je však faktorová struktura v případě restrukturalizovaných klinických škál (RC), které byly konstruovány právě s ohledem na dobré psychometrické parametry a „čistou“ faktorovou strukturu (A. Tellegen et al., 2003). Podobné závěry se však objevují i jinde (Hoelzle & Meyer, 2008), ačkoli např. Cheng (2006) pozoroval v kanadském vzorku výrazně lepší shodu modelů s daty než my v naší studii. Stejně tak Hajný (2018) na velmi podobném vzorku českých policistů pozoroval přiměřeně jednodimenzionální strukturu nejen u RC škál (vyjma RC2), ale též u validizační F škály a standardních klinických Pa a Pt škál.

Ačkoli jím zvolený dvouparametrový IRT model je matematicky ekvivalentní námi zvolené ordinální faktorové analýze, liší se použité metody odhadu. Zatímco Hajný (2018) použil metodu maximální věrohodnosti (ML), my jsme pracovali s diagonálně-váženou sumou čtverců (WLSMV). V našem případě byl kromě *TLI* (kde příčinou jsou celkově slabé korelace položek) nejčastěji problematický index *SRMR*, který je založen na rozdílu volně odhadnuté tetrachorické korelační matice a modelem implikované matice. A právě volně odhadované tetrachorické korelace mohly být v důsledku již zmíněného velkého množství nulových hodnot v bivariačních frekvenčních tabulkách odhadnuty nepřesně, s velkou chybou. *SRMR* mohlo být špatné (vysoké), protože pracuje s absolutní velikostí rozdílu; naopak jemu podobné *RMSEA* (root mean-square error of approximation) zvažuje i chybu odhadu, a vycházelo tak lépe. V případě IRT a ML estimátoru je podobný index *SRMSR* počítán jinak, bez využití tetrachorické korelační matice, a tento efekt se tak nemusel projevit. Uvažované nedostatky indexu *SRMR* však nebyly v literatuře dosud popsány a bylo by proto velmi užitečné naše závěry ověřit: jak v případě dotazníku MMPI-2, tak i metody SPARO, kde jsou závěry velmi podobné (Cígler & Rudá, 2021).

Souběžná validita

Hlavním cílem studie bylo ověřit sérii hypotéz o souběžné validitě dotazníku SPARO vůči metodě MMPI-2. Ačkoli z předchozího textu je zřejmé, že česká adaptace MMPI-2 není zcela spolehlivým kritériem, domníváme se, že jej lze využít. Na druhou stranu však nelze z případně nízké souběžné validity vinit výhradně dotazníku SPARO, příčinou mohou být i horší psychometrické parametry některých škál MMPI-2.

V případě skóru věrohodnosti podávaných výpovědí (L) metody SPARO výsledky nepodpořily jeho souběžnou validitu s validizačními škálami L a K dotazníku MMPI-2. Zatímco obě validizační škály z MMPI-2 středně silně souvisejí s klinickými a dalšími škálami téhož dotazníku, to stejné neplatí pro L škálu dotazníku SPARO. Navíc L škála dotazníku SPARO současně vykazuje i neuspokojivou reliabilitu a zdá se být tedy vůbec nevhodná pro jakékoli diagnostické závěry.

Jedinou hlavní komponentou nebo obecným faktorem dotazníku SPARO, jehož validita byla bez výhrad podpořena, je škála emocionální variabilnosti (EM) díky silným korelacím se škálami psychastenie (Pt) a úzkosti (ANX) a rovněž i uspokojivé vnitřní konzistenci. I zde však musíme mít na paměti silné korelace s jinými hlavními komponentami dotazníku, které jsou teoreticky

nepodložené. Částečně pak byla podpořena i validita škály regulační variabilnosti (RE), její korelace se škálami hypomanie (Ma) a zloby (ANG) jsou však jako důkaz konvergentní validity spíše nízké. Její reliabilita je rovněž uspokojivá.

Námi pozorovaná validita EM (a nikoliv ostatních škál) je rovněž v souladu s předchozími studiemi. Emocionální variabilnost (EM) silně korelovala s Eysenckovou škálou neuroticismu (Mikšík, 2004), který je podobný námi zkoumaným škálám MMPI-2 Psychastenien (Pt) a Úzkost (ANX). Smysluplné, avšak slabé korelace škály EM se schopností se přizpůsobit a celkovou psychosociální výkonností dotazníku WPSI pozorovali rovněž Preiss a Haas (1997). Škála emocionální variabilnosti se tak zdá být jedinou validní škálou mezi hlavními skóry dotazníku SPARO. Příčinou může být velmi silné zešíkmení škály, kdy většina respondentů odpovídá na většinu položek nesouhlasně, a dosahuje tak celkově nízkého skóre – je možné, že respondenti s vysokým skóre mají tendenci odpovídat jiným způsobem i v jiných dotaznících. Chceme navíc upozornit, že EM škála se skládá z pouhých 20 položek z celkových 300, zahrnuje tedy méně než 7 % dotazníku. Zároveň není jisté, jakým způsobem by fungovalo těchto 20 položek administrovaných samostatně.

Ostatní klíčové skóry dotazníku SPARO nelze označit za konstruktově-validní. Příčinou je zpravidla nízká konvergentní validita, v případě komponenty adjustační variabilnosti (AD) jsme dokonce pozorovali opačný vztah se škálou hysterie (Hy), než jsme předpokládali. Neuspokojivě vychází i divergentní validita, např. obecná hladina psychické vzrušivosti (PV) dokonce s obsahovými škálami MMPI-2 souvisí středně silně až silně, ačkoli jsme předpokládali žádný nebo zanedbatelný vztah. Zde je však nutné podotknout, že konkrétně škála PV má zároveň uspokojivou souběžnou validitu s psychastenii (Pt) a hypomanií (Ma). Zdá se tedy, že ačkoli její obsahová validita je uspokojivá, rozptyl skórů je kontaminovaný obsahovými charakteristikami psychopatologie tak, jak je měří dotazník MMPI-2. Validita ostatních škál nebyla podpořena ani zčásti.

Limity

Většina limitů naší studie byla zmíněna přímo v textu. Zde bychom chtěli zdůraznit, že naše výsledky jsou založeny výhradně na vzorku ověřování profesní způsobilosti policistů a uchazečů o práci u policie. Je možné, že v případě klinického vzorku by MMPI-2 fungovalo lépe. Na druhou stranu, korelační struktura dotazníku SPARO reportovaná Cíglerem a Rudou (2021) byla shodná s originální strukturou (Mikšík, 2009, pp. 116–118).

Dalším limitem je fakt, že u šesti škál MMPI-2 (FP, MA3, HEA1, DO, PSYC, ?) neodpovídá námi vytvořený hrubý skór skóru exportovanému z programu. Zdá se, že se náš a originální postup liší vždy o jednu položku (vyjma škály PSYC, kde jde o dvě položky) – není však jisté, zda je chybně uvedeno zařazení položek v nám dostupných materiálech (Netík, 2003), nebo zda je chyba ve vyhodnocovacím programu. Dopady na validitu našich výsledků a fungování metody MMPI-2 jsou však zanedbatelné. A konečně jsme pracovali pouze s hrubými skóry, odvozené standardní skóry jsme neměli k dispozici. To by však rovněž nemělo mít vliv na výsledky naší studie.

Významný limit při ověřování souběžné validity dotazníku SPARO vzniká na straně dotazníku MMPI-2. Při návrhu výzkumného designu jsme nepočítali s tak nízkou reliabilitou škál MMPI-2 a jejich neuspokojivou faktorovou strukturou. Je možné, že nízká souběžná validita SPARO je přinejmenším zčásti způsobena MMPI-2 jako nevhodně zvoleným kritériem. To je však podpořeno rozsáhlým zahraničním výzkumem a i přes určité nedostatky lze předpokládat přiměřenou úroveň validity.

A konečně, určitým omezením je i samotná konstrukce výzkumných hypotéz. I přes usilovnou snahu bylo velmi obtížné porovnat obsahovou validitu škál MMPI-2 a SPARO. Je proto možné, že některé potenciálně důležité vztahy jsme přehlédli, zatímco smysluplnost některých hypotéz je omezena. Z toho důvodu odkazujeme čtenáře do on-line supplementu, kde má k dispozici plnou korelační matici všech škál, a může je sám ověřit.

Shrnutí obou studií

Výsledky podporují předchozí závěry (Cígler & Rudá, 2021), že neexistuje dostatek empirické podpory pro validní použití dotazníku SPARO v kontextu profesní diagnostiky. Naopak, získaná data svědčí proti používání dotazníku, a to s využitím tří zdrojů důkazů.

Reliabilita. V první řadě příliš velké množství škál (včetně těch klíčových) má neuspokojivou reliabilitu. Navíc jsme ukázali, že škály s nízkou reliabilitou slaběji souvisí s kritériem, a to bez ohledu na obsah. Lze očekávat, že nereliabilní škály budou hůře predikovat i budoucí pracovní výkon, selhání či schopnost vykonávat zaměstnání (Cígler & Rudá, 2021).

Faktorová struktura. Interpretace dotazníku SPARO spočívá na předpokladu, že hlavní komponenty kognitivní (KO), emocionální (EM), regulační (RE) a adjustační (AD) variabilnosti představují čtyři relativně nezávislé psychické rysy, a že existují obecnější a rovněž nezávislé faktory psychické vzrušivosti (PV) a motorické hybnosti (MH) (Mikšík, 1992, 2001, 2004, 2009). To není pravda. Korelace komponent jsou extrémně vysoké, v případě škál AD a EM dosahují dokonce 0,83. Obecnější faktory PV a MH jsme v našich datech nepozorovali vůbec. Tyto obtíže jsou navíc patrné již z originálních, Mikšíkem (2009) reportovaných informací; Mikšík totiž nevzal v potaz nedostatečnou reliabilitu škál, která korelace zásadním způsobem podhodnotila. To vše invaliduje nejen dotazník SPARO, ale i celou Mikšíkovu (1992, 2001, 2004, 2009) teorii bazální psychické autoregulace osobnosti, která je na těchto předpokladech postavena (Cígler & Rudá, 2021).

Souběžná validita. S výjimkou škály emocionální variabilnosti (EM) nebyla prokázána souběžná validita hlavních komponent, obecnějších faktorů a škály věrohodnosti podávaných výpovědí (L) dotazníku SPARO vůči skórum MMPI-2. Nepozorovali jsme na základě teorie očekávané korelace a vybrané skóry oproti předpokladům naopak souvisely s obsahovými škálami MMPI-2.

Bez dalších důkazů, zejm. v oblasti prediktivní validity, proto v žádném případě nelze doporučit použití dotazníku SPARO v praxi a celkově pro jiné než výzkumné účely. Jeho použití v praktické diagnostice může být v rozporu s běžnými etickými požadavky na psychologické vyšetření tak, jak je specifikují např. body 9.3 a 9.4 Standardů pro pedagogické a psychologické testování Americké psychologické asociace (AERA et al., 2014). Zejména není možné na základě výsledků dotazníku SPARO činit žádná rozhodnutí s důsledky pro vyšetřenou osobu, tedy například ověření profesní způsobilosti zaměstnanců či uchazečů o zaměstnání. Taková rozhodnutí nejsou založena na důkazech („evidence based“), což je dokonce například podle Stříteského (2018) v rozporu s českou legislativou.

Nedostatečné vnitřní konzistence jsme ovšem pozorovali i u škál dotazníku MMPI-II, které jsou navíc zřejmě nižší než v originální americké verzi; příčinou může být omezení variability v našem policejním vzorku policistů, případně jazykové nedostatky českého překladu (který však působí kvalitním dojmem). Nízkých reliabilit dosahují zejména komponenty obsahových škál, poruchy regulace chování, nové validizační škály a jednotlivé komponenty obsahových škál. Konstruktovou

validitu restrukturalizovaných klinických škál pak snižuje jejich nejednoznačná a vícedimenzionální faktorová struktura; přitom tyto škály byly vyvinuty právě s ohledem na lepší psychometrické parametry a jasnou faktorovou strukturu (A. Tellegen et al., 2003). Zde však může být na vině použitý analytický postup.

Je patrné, že ačkoli se MMPI-II může opřít o rozsáhlý zahraniční výzkum, bylo by velmi vhodné jeho psychometrické vlastnosti – zejména prediktivní validitu – ověřit i na českém vzorku. Je totiž potřeba vyloučit možnost, že překlad do češtiny a kulturní odlišnosti snižují validitu a možnosti interpretace dotazníkových výsledků. Interpretaci výsledků dotazníku MMPI-2 je s ohledem na tato zjištění nutné provádět obezřetně.

Je již zvykem české psychodiagnostické praxe, že není kladen dostatečný důraz na důkazy validity a reliability u používaných nástrojů. Ostatně mezi v nedávné době nejčastěji používané diagnostické metody (Svoboda et al., 2004; Urbánek, 2010) patřily testy, které jsou experty hodnoceny jako nevěrohodné či přímo zdiskreditované (Norcross et al., 2006) – příkladem může být Lüscherův test barev, hand test, kresba postavy či stromu. Stejně tak byly využívány zastaralé metody (Pražský dětský Wechsler, WAIS-R) či metody bez českých norem (Stanford-Binetova inteligenční škála).

Podobná situace bohužel nastala i u dotazníku SPARO, který i přes svou značnou oblibu postrádá empirické důkazy validity a reliability, které by jeho využití ospravedlňovaly. Naše studie takové důkazy nenalezla, a naopak její výsledky svědčí spíše proti jeho použití v běžné praxi. Rovněž i v případě MMPI-2 chybí důkazy o fungování české verze metody, zde je však možné se spoléhat na zobecnitelnost zahraničních výzkumů alespoň do té doby, než budou k dispozici místní validizační studie.

Poděkování

Tento článek vznikl za finanční podpory Grantové agentury České republiky v rámci projektu Metodické studie pro české verze měřítek psychologických charakteristik adolescentů a vynořujících se dospělých (GA17-09797S).

Chceme poděkovat Policejnímu prezidiu České republiky za poskytnutí dat a dalších materiálů, bez kterých by tento článek nevznikl. Dále děkujeme Stanislavu Ježkovi za pomoc při úpravě struktury tohoto článku, Michaele Borovanské za pomoc při přípravě podkladů a Tomáši Urbánkovi za tvorbu dílčího podkladu pro výsledný analytický skript, který přiřazuje položky MMPI-2 do škál.

Doplňující výsledky, rozšiřující materiály, analytické skripty a další podklady jsou uvedeny v on-line supplementu dostupném na <https://doi.org/10.17605/OSF.IO/CNU58>.

Literatura

- AERA, APA, & NCME. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.
- Bednářová, B. (2016). *Skresľovanie odpovedí v osobnostných dotazníkoch v kontexte výberového konania. Diplomová práca*. Masarykova univerzita. <https://is.muni.cz/th/o0fgs/>
- Benton, T. (2015). An empirical assessment of Guttman's lambda 4 reliability coefficient. In R. Millsap, D. Bolt, L. van der Ark, & W. C. Wang (Eds.), *Quantitative Psychology Research. Springer Proceedings in Mathematics & Statistics*, vol. 89 (pp. 301–310). Springer. <https://doi.org/10.1007/978-3-319-07503-7>
- Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., Kaemmer, B., & Netík, K. (2002). *MMPI 2: Minnesota Multiphasic Personality Inventory-2*.

- Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., Kaemmer, B., & Netík, K. (2013). *MMPI-2 - Minnesota Multiphasic Personality Inventory-2 : Druhé rozšířené vydání*. Testcentrum.
- Chalmers, R. P. (2018). On misconceptions and the limited usefulness of ordinal alpha. *Educational and Psychological Measurement*, 78(6). <https://doi.org/10.1177/0013164417727036>
- Cheng, M. K. (2006). *Structural equation modeling of the Restructured Clinical scales of the MMPI-2*. Thesis. University of Windsor. <https://scholar.uwindsor.ca/etd/7215>
- Cígler, H., & Rudá, A. (2021). Reliabilita a faktorová validita dotazníku SPARO. *E-psychologie*, 15(1), 16-39. <https://doi.org/10.29364/epsy.391>
- Čurdová, J., & Pourová, M. (2019). Minnesota Multiphasic Personality Inventory 2 – Recenze metody. *TESTFÓRUM*, 7(12), 6. <https://doi.org/10.5817/TF2019-12-12447>
- Farnikova, K., Obereigneru, R., Kanovsky, P., & Prasko, J. (2012). Comparison of personality characteristics in Parkinson disease patients with and without impulse control disorders and in healthy volunteers. *Cognitive And Behavioral Neurology*, 25(1), 25–33. <https://doi.org/10.1097/WNN.0b013e31824b4103>
- Friedman, A. F., Bolinsky, P. K., Levak, R. W., & Nichols, D. S. (2015). *Psychological assessment with the MMPI-2/MMPI-2-RF* (3rd ed.). Routledge.
- Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-Type and ordinal Item Response Data: A conceptual, empirical, and practical guide - practical assessment, research & evaluation. *Practical Assessment, Research, and Evaluation*, 17(3), 1–13. <https://doi.org/10.7275/n560-j767>
- Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika*, 74(1), 155–167. <https://doi.org/10.1007/s11336-008-9099-3>
- Greene, R. L. (2000). *The MMPI-2: An interpretive manual*. Allyn and Bacon.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10(4), 255–282. <https://doi.org/10.1007/BF02288892>
- Hajný, M. (2018). *Vybrané psychometrické analýzy dotazníku MMPI-2*. Diplomová práce. Masarykova univerzita. <https://is.muni.cz/th/b8prp/>
- Hansmanová, L., Obereigneru, R., & Šmahaj, J. (2007). Psychosomatické aspekty zhoubného nádoru těla děložního. *Česká gynekologie*, 72(4), 304–308.
- Hathaway, S. R., & McKinley, J. C. (1942). *Manual for the Minnesota Multiphasic Personality Inventory*. University of Minnesota Press.
- Hoelzle, J. B., & Meyer, G. J. (2008). The factor structure of the MMPI-2 restructured clinical (RC) scales. *Journal of Personality Assessment*, 90(5), 443–455. <https://doi.org/10.1080/00223890802248711>
- Jensovsky, J., Ruzicka, E., Spackova, N., & Hejdukova, B. (2002). Changes of event related potentials and cognitive processes in patients with subclinical hypothyroidism after thyroxine treatment. *Endocrine Regulations*, 36(3), 115–122.
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2020). *semTools: Useful tools for structural equation modeling* (0.5-3). <https://cran.r-project.org/package=semTools>
- Jurišová, E., & Sarmány-Schuller, I. (2013). Structure of basal psychical self-regulation and personality integration in relation to coping strategies in decision-making in paramedics. *Studia Psychologica*, 55(1), 3–18. <https://doi.org/10.21909/sp.2013.01.617>
- Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2014). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, 44(3), 486–507. <https://doi.org/10.1177/0049124114543236>
- Kudela, M., & Hansmanová, L. (2014). Psychosomatické aspekty a léčba psychofarmaky v etiopatogenezi karcinomu endometria. *Česká gynekologie*, 79(5), 378–381.

- Lakens, D. (2017). Equivalence tests. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>
- Líbalová, Z., Feyereisl, J., Martan, A., Cepický, P., Halaska, M., Krofta, L., Váchová, D., Balcarová, J., & Pecená, M. (2001). Psychologie inkontinence moce. I. Srovnání žen s urgentní a stresovou inkontinencí před zahájením terapie. *Česká gynekologie*, 66(3), 171–174.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley.
- Mangiafico, S. (2020). *rcompanion: Functions to Support Extension Education Program Evaluation* (R package version 2.3.21). <https://cran.r-project.org/package=rcompanion>
- Mikšík, O. (1992). *IHAVEZ – SPIDO – VAROS (příručka)*. Psychodiagnostika.
- Mikšík, O. (2001). *Zjišťování bazální struktury a dynamiky autoregulace, integrovanosti a psychické odolnosti osobnosti dotazníky SPARO, BAROM, IHATRANS (manuál)*. Diaros.
- Mikšík, O. (2004). *Dotazník SPARO (příručka)*. Psychodiagnostika.
- Mikšík, O. (2007). *Psychologická charakteristika osobnosti*. Karolinum.
- Mikšík, O. (2009). *Psychika osobnosti v období závažných životních a společenských změn*. Karolinum.
- Netík, K. (2003). *Příloha C: Složení škál*.
- Nichols, D. S. (2011). *Essentials of MMPI-2 Assessment*. John Wiley & Sons, Inc.
- Norcross, J. C., Koocher, G. P., & Garofalo, A. (2006). Discredited psychological treatments and tests: A Delphi poll. *Professional Psychology: Research and Practice*, 37(5), 515–522. <https://doi.org/10.1037/0735-7028.37.5.515>
- Obereignerů, K., Mareš, J., Obereignerů, R., & Kanovský, P. (2010). The Occurrence of Psychogenic Disorders in Neurology. *Česká a slovenská neurologie a neurochirurgie*, 73(5), 529–533.
- Phillips, N. (2017). *yarr: A Companion to the e-Book “YaRrr!: The Pirate’s Guide to R”* (R package version 0.1.5). <https://cran.r-project.org/package=yarr>
- Preiss, J., & Haas, T. (1997). Některé psychometrické charakteristiky české verze Washingtonského psychosociálního dotazníku pro záchvatová onemocnění (WPSI). *Československá psychologie*, 41(4), 334–346.
- Prokopová, P. (2011). *Analýza osobnostní struktury uchazeče o službu v profesionálním hasičském záchranném sboru. Diplomová práce*. Masarykova univerzita. <https://is.muni.cz/th/hny4h/>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Revelle, W. (2019). *psych: Procedures for Personality and Psychological Research* (1.9.12). Northwestern University. <https://cran.r-project.org/package=psych>
- Revelle, W., & Zinbarg, R. E. (2009). Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74(1), 145–154. <https://doi.org/10.1007/s11336-008-9102-z>
- Rose, E. M., Mathew, T., Coss, D. A., Lohr, B., & Omland, K. E. (2018). A new statistical method to test equivalence: an application in male and female eastern bluebird song. *Animal Behaviour*, 145, 77–85. <https://doi.org/10.1016/j.anbehav.2018.09.004>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36. <http://www.jstatsoft.org/v48/i02/>
- Sellbom, M., Fischler, G. L., & Ben-Porath, Y. S. (2007). Identifying MMPI-2 predictors of police officer integrity and misconduct. *Criminal Justice and Behavior*, 34(8), 985–1004. <https://doi.org/10.1177/0093854807301224>
- Signorell, A., & et al. (2020). *DescTools: Tools for descriptive statistics* (R package version 0.99.32). <https://cran.r-project.org/package=DescTools>
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach’s Alpha. *Psychometrika*, 74(1), 107–120. <https://doi.org/10.1007/s11336-008-9101-0>

- Spearman, C. (1904). The Proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72. <https://doi.org/10.2307/1412159>
- Stancak, A., Hubova, D., Skorodensky, M., Kollar, J., & Mrinak, J. (1986). Clinical symptomatology of the style of Type A behavior pattern according to the MMPI-100. *Studia Psychologica*, 28(1), 73–79.
- Stříteský, M. (2018). Právní aspekty pracovně psychologické diagnostiky. *Testforum*, 6(10), 23–31. <https://doi.org/10.5817/TF2018-10-179>
- Šucha, M., Šťastná, L., & Zámečník, P. (2017). Osobnostní vlastnosti řidičů řídících pod vlivem alkoholu. *Adiktologie*, 17(1), 34–44.
- Svoboda, M., Řehan, V., Vtípil, Z., Klimusová, H., Humpolíček, P., Urbánek, T., & Kouhoutek, T. (2004). *Aplikovaná psychodiagnostika v České republice*. MSD.
- Tellegen, A., Ben-Porath, Y. S., McNulty, J. L., Arbisi, P. A., Graham, J. R., & Kaemmer, B. (2003). *MMPI-2 Restructured Clinical (RC) Scales: Development, validation, and interpretation*. University of Minnesota Press.
- Tellegen, A., & Ben-Porath, Y. S. (1992). The new uniform T scores for the MMPI--2: Rationale, derivation, and appraisal. *Psychological Assessment*, 4(2), 145–155. <https://doi.org/10.1037//1040-3590.4.2.145>
- Torchiano, M. (2019). *effsize: Efficient Effect Size Computation* (R package version 0.7.6). <https://doi.org/10.5281/zenodo.1480624>
- Urbánek, T. (2010). Nejpoužívanější psychodiagnostické metody v České republice. *TESTFÓRUM*, 1(1), 6–9. <https://doi.org/10.5817/TF2010-1-3>
- Wickham, H., & Bryan, J. (2019). *readxl: Read Excel Files* (R package version 1.3.1). <https://cran.r-project.org/package=readxl>
- Wise, E. A., Streiner, D. L., & Walfish, S. (2010). A review and comparison of the reliabilities of the MMPI-2, MCMI-III, and PAI presented in their respective test manuals. *Measurement and Evaluation in Counseling and Development*, 42(4), 246–254. <https://doi.org/10.1177/0748175609354594>
- Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods*, 6(1), 21–29. <https://doi.org/10.22237/jmasm/1177992180>

Údaje o autorech

Mgr. Hynek Cígler, PhD. přednáší na katedře psychologie a je výzkumným pracovníkem v Institutu výzkumu dětí, mládeže a rodiny, Fakulta sociálních studií Masarykovy univerzity.

 <https://orcid.org/0000-0001-9959-6227>

Kontaktní údaje

Adresa: Katedra psychologie, Fakulta sociálních studií Masarykovy univerzity, Joštova 10, 602 00 Brno

E-mail: hynek.cigler@mail.muni.cz

Mgr. Adéla Rudá vystudovala psychologii na Filozofické fakultě Univerzity Karlovy v Praze. Pracovala jako psychologka na Policejním prezidiu ČR, nyní působí v soukromé praxi.

Kontaktní údaje

Email: adela.ruda@seznam.cz

Přílohy

Příloha 1: Seznam a stručný popis klinických škál MMPI-2

Tabulka 1.1: Standardní klinické škály MMPI-2

1: Hs – Hypochondrie	vágní a nespecifické stížnosti v oblasti tělesných funkcí, zvýšený zájem o tělo
2: D – Deprese	depresivní symptomatika (např. špatná nálada, obecná nespokojenost)
3: Hy – Hysterie	specifické somatické obtíže, pocit dobré socializace a adaptace
4: Pd – Psychopatie	napětí v sociálním prostředí
5: Mf – Maskulinita– feminita	sleduje zájmy a některé povahové rysy, neukazuje na psychopatologii
6: Pa – Paranoia	zvýšená sociální citlivost, tendence k ostražitosti, vztahovačnosti, persekčním myšlenkám
7: Pt – Psychastenie	úzkostnost, nerozhodnost, sebepochybnosti
8: Sc – Schizofrenie	bizarní myšlenkové procesy a neobvyklé percepce, sociální odcizení, útoky do fantazií
9: Ma – Hypomanie	mírnější stupně manické excitace (např. impulzivita, povrchnost)
0: Si – Sociální introverze	sociální staženost, plachost, nevztahuje se k psychopatologii

Pozn.: Pro přehled všech škál a jejich zkratk viz originální manuál 2. českého vydání, tab. 4–15 (Butcher et al., 2013). S kritickými položkami (tab. 16 v manuálu) v naší studii nepracujeme.

Příloha 2: Zařazení položek do škál

Správnost zařazení položek do škál byla ověřena pomocí vizuální inspekce bodových grafů, které srovnávaly skóry škál exportované z administracních systémů se skóry ručně vytvořenými z odpovědí na jednotlivé položky (viz on-line supplement). Kromě toho byly pro každou škálu oba druhy skórů vzájemně srovnány Pearsonovou korelací.

Korelace sedmi námi spočítaných skórů se skóry exportovanými z vyhodnocovacího programu byly menší než jedna, a tedy došlo k chybě na naší straně nebo na straně vyhodnocovacího programu. Rozdíly byly ale celkově spíše malé a ve všech případech se pohybovaly v rozmezí ± 1 bod hrubého skóru. Rovněž námi odhadnutý součtový skór, s výjimkou škály HEA1 ovlivněné malým počtem položek, vždy silně koreloval se skórem exportovaným ze systému. Konkrétně šlo o tyto škály: $r_{FP} = 0,999$; $r_{MA3} = 0,933$; $r_{HEA1} = 0,763$; $r_{DO} = 0,964$; $r_{PSYC} = 0,972$; $r_{VRIN} = 0,966$ a $r_{?} = 0,998$.

Naše přiřazení položek do škál nicméně plně odpovídalo originálnímu složení škál (Netík, 2003) a pečlivá kontrola neodhalila žádný systematický rozdíl. Pokusili jsme se proto identifikovat chyby prostřednictvím lineární regrese, kde závislou proměnnou byl skór spočítaný programem a prediktory jednotlivé položky³. Ve všech případech se podařilo identifikovat problematickou položku; její odstranění však nevedlo k perfektní shodě námi spočítaných a originálních skórů; zřejmě je položka zaměněna s jinou. Je však otázka, zda je chyba v reportované struktuře škál (Netík, 2003), nebo ve vyhodnocovacím softwaru. Dopad na validitu skórů je však v každém případě minimální.

V případě škály „Nedovedu říci“ (?), která je definovaná jako počet vynechaných položek, byl námi spočítaný skór u 21 respondentů o jeden bod vyšší oproti skóru exportovanému z programu. Z těchto 21 respondentů ani jeden neposkytl odpověď na 10 položek, z nichž však na jedinou zároveň pouze na jedinou poskytli odpověď všichni ostatní respondenti: šlo o položku 477. Je tedy evidentní, že počítačový program tuto položku chybně nezapočítal do hrubého skóre škály.

Drobná rozdílnost vznikla rovněž u škály TRIN, kde sice byl mezi naším a původním odhadem perfektně lineární vztah, avšak námi spočítané skóry ležely o devět bodů níže než skóry originální. Způsobilo to rozdílné zacházení s 9 páry položek, u nichž dle instrukcí (Netík, 2003) měl být odečten skór, pokud na ně respondent odpověděl shodně. Vyhodnocovací software naopak bod přičetl, pokud na tyto položky respondent odpověděl rozdílně. Je proto otázkou, zda tento jev nemá vliv na validitu skóru TRIN reportovaného programem.

³ FP: položka 487. MA3: položka 115. HEA1: položka 008. DO: položka 339. PSYC: položky 446 a 138. VRIN: páry položek 472 N – 533 S a 491 S – 509 N.