

SNIJDERS-OOMEN NEVERBÁLNÍ INTELIGENČNÍ TEST

Recenze metody

AUTOR RECENZE: Edita Chvojková¹, Štěpán Postulka¹, Tereza Horáková²
Katedra psychologie, Fakulta sociálních studií Masarykovy univerzity

datum vzniku recenze:	30. 11. 2017
1.1 název nástroje:	Snijders-Oomen neverbální inteligenční test
zkrácený název:	SON-R 2½-7
1.2 původní název:	Snijders-Oomen-Nonverbal-Intelligence Test – Revised 2½ - 7
1.4 autoři původního testu:	Peter J. Tellegen, Jacob A. Laros
1.3 autoři lokální adaptace:	Dana Krejčířová, Dana Černochová, Jana Zapletalová, Helena Pacnerová, Hana Lukášová
1.7 lokální distributor:	Hogrefe - Testcentrum
1.9.1/1.9.2 datum vydání:	2008

Obecné informace

SON-R 2½-7³ je neverbální testová baterie pro měření intelektu u dětí ve věku od 24 měsíců do 7 let určená k individuální administraci. Jedná se o v pořadí třetí revizi inteligenčního testu SON, sestaveného Nan Snijders-Oomenovou a Janem Snijdersem v roce 1943 (Tellegen, Winkel, Wijnberg-Williams, & Laros, 1998) původně navrženého pro testování neslyšících dětí. Později došlo k rozšíření metody i pro děti slyšící a poslední revize už obsahuje normy obecné. Uplatnění nachází především v klinické a poradenské praxi.

Baterie sestává z 6 subtestů: kategorií, analogií a situací, které dohromady tvoří úsudkovou škálu, a mozaik, skládanek a vzorů, tvořících škálu performační. Jednotlivé subtesty obsahují průměrně 15 položek seřazených podle obtížnosti. Každý z nich, je tvořen dvěma rozdílnými typy úloh. Jednotlivé subtesty jsou v manuálu popsány společně s příklady konkrétních položek. Hrubé skóry v jednotlivých subtestech se převádějí na vážené skóry (dále VS; M = 10, SD = 3) podle instrukcí uvedených v manuálu v závislosti

¹ Katedra psychologie, Fakulta sociálních studií MU, Joštova 10, 602 00 Brno

² Mateřská škola, Speciální základní škola a Praktická škola, Hradecká 1231/11b, 500 03 Hradec Králové

³ Tento článek byl vytvořen za finanční podpory Národního ústavu pro vzdělávání (NÚV)

na věku dítěte. Součet VS každých tří subtestů vytváří dva IQ skóre ($M = 100$, $SD = 15$) – SON-PŠ (IQ skór škály performační) a SON-ŠÚ (IQ skór škály úsudkové), které jsou poměrně složitě interpretovatelné. Interpretovatelným je skór SON-IQ, který je tvořen součtem VS všech subtestů a vypovídá přímo o fluidní inteligenci dítěte. Výkon dítěte je porovnáván s normami vytvořenými na reprezentativním vzorku 450 dětí rozdělených do devíti skupin po 50. Uživatelům je k dispozici srovnání výkonu vyšetřovaného dítěte s ostatními pomocí percentilů.

Specifikem SON-R 2½-7 je fakt, že administrátor poskytuje dítěti zpětnou vazbu po každé položce a vede jej k opravě chybných odpovědí. Položky jsou nicméně skórovány dichotomicky (0-1), korigovaná odpověď není skórována jako správná a její zařazení má spíše motivační charakter. Zároveň však umožňuje sledovat, jak u dítěte probíhá proces učení, čímž interpretaci obohacuje o prvky dynamické diagnostiky. Podrobnější informace o výběru položek a jejich fungování jsou uvedeny v manuálu k holandské standardizaci testu (Tellegen et al., 1998), český uživatel k těmto informacím nemá přístup. Metoda je administrována formou tužka-papír, vyhodnocena může být jak ručně, tak s využitím počítačového programu. Testovací materiál je atraktivní, a jeho přehledné uspořádání umožňuje pohodlnou administraci. Některé obrázkové položky (např. pokrývky hlavy v subtestu Kategorie) však mohou být pro dítě obtížné z důvodu jejich zastaralosti.

Test má podle autorů měřit fluidní inteligenci v rámci modelu Cattella (1971). Fluidní inteligence má na *g-faktoru* nejsilnější náboj z ostatních sekundárních CHC faktorů. Není však založen na žádném specifickém modelu inteligence, což si o testu myslí i jiní autoři (např. Hülür, Wilhelm & Robitzsch, 2011). Ve většině strukturně založených teorií (Cattell, 1991; Horn & Cattell, 1966) bychom subtesty SON-R 2½-7 mohli opravdu klasifikovat jako měřící fluidní inteligenci (v CHC jako *Gf*) nebo širší vizualizační schopnosti (v CHC jako *Gv*; McGrew, 2009). Škála performační potom tvoří jeden faktor druhého řádu, který sytí položky tří subtestů, v nichž jde o prokázání prostorového vhledu (Carroll, 1993). Škála úsudková tvoří druhý faktor druhého řádu, který sytí položky tří subtestů, v nichž jde o prokázání schopností konkrétního a abstraktního myšlení (Carroll, 1993).

Finanční náklady na pořízení metody jsou následující: úplný testový soubor lze zakoupit za 45 200 Kč, samotný testovací materiál za 39 700 Kč, manuál za 1 500 Kč, 25 ks záznamových archů vyjde na 800 Kč a 50 ks pracovního sešitu pro subtest Vzory na 4500 Kč.

Psychometrické parametry

Reliabilita

V české a slovenské standardizaci byla reliabilita celkového skóru zjišťována na subtestových vážených skórech jako míra zobecnitelnosti (v manuálu i záznamovém archu označována jako "g"). To je v podstatě klasická Cronbachova α , počítaná na skórech subtestů namísto skórů položek. Domníváme se, že tento postup je motivován snahou autorů o očištění celkového skóru o specifický rozptyl subtestů. Pokoušejí se tak zjistit reliabilitu měření *g-faktoru* inteligence. Uvedený postup nicméně nepovažujeme za vhodný. Uživatel testu se z manuálu nedozví, jakým způsobem byl *g-index* spočítán.

Domníváme se, že měla být pro stejný účel spíše použita například McDonaldova hierarchická omega. Reportovaná reliabilita celkového skóru v testu tak může být (a pravděpodobně je) podhodnocena z důvodu nedodržení poměrně přísných předpokladů na Cronbachovu α , například tau-ekvivalence subtestů a unidimenzionality škál. Kromě indexů zobecnitelnosti nacházíme v manuálu i vnitřní konzistence škály performační a škály úsudkové (počítány nad položkami příslušných subtestů), stejně jako celkového IQ skóru (taktéž počítáno nad všemi položkami). Průměrná Cronbachova α pro Performační škálu je 0,88, pro Škálu usuzování je 0,86 a pro celkové SON-IQ je 0,92. Hodnoty pro reliabilitu celkového skóru jsou velmi vysoké. Ukazatele vnitřní konzistence nicméně nejsou pro další výpočty v rámci norem a standardních chyb odhadu směrodatné.

Autoři manuálu uvádějí dva různé typy chyb. Jednak standardní chybu měření (počítanou v jednotkách vážených skóru), jednak standardní chybu odhadu, pro jejíž výpočet je použit právě zmíněný koeficient zobecnitelnosti g . Intervaly spolehlivosti v normách jsou počítány právě s využitím standardní chyby odhadu. Hodnota g se ve věkových skupinách pohybuje mezi 0,78 až 0,81, což je pro hierarchický koeficient v pořádku. Oceňujeme fakt, že uživatel má v tabulkách norem k dispozici odhady 80procentních intervalů spolehlivosti.

Součástí ověření souběžné validity je i korelace při retestu, kterou bychom uvedli spíše v rámci posouzení reliability. Ta dosahuje hodnoty $r = 0,79$. Vezmeme-li v úvahu fakt, že každá nesprávná odpověď je korigována, což významně přispívá k efektu zácviku, zdá se tato hodnota být menší, než bychom očekávali. Na druhou stranu nevíme, po jaké době se opakované testování odehrálo, a tedy nelze výsledky jednoznačně interpretovat. V úvahu je třeba vzít i poměrně rychle se zvyšující se reliabilitu IQ v dětství (Schuerger & Witt, 1989).

Validita

Důkazy o validitě můžeme rozdělit do několika skupin. Jako první uvádějí autoři důkazy pro kriteriální validitu, a to prostřednictvím testování rozdílů v IQ mezi skupinami dětí s rodiči různého vzdělání. Autoři dále srovnávají IQ na základě bydliště dětí (město, nebo vesnice). Dalšími kritérii jsou úplnost rodiny, SES a pohlaví. Výsledky provedených testů jsou ve shodě s předpoklady.

Na konstruktovou, respektive faktorovou validitu autoři testu usuzují z analýzy hlavních komponent s ortogonální rotací Varimax. Tento postup není adekvátní. Za prvé, analýza hlavních komponent není faktorovou analýzou a nepracuje s výběrovou chybou, výsledky tedy mohou být zkreslené. Za druhé, použití ortogonální rotace nedává příliš smysl v případě, že jsou faktory zjevně korelované. Třetí, nejzásadnější problém spočívá v tom, že byla použita explorační metoda pro „konfirmasi“ určité faktorové struktury. Myslíme si, že použití konfirmační faktorové analýzy by pro dané účely bylo vhodnější a že stávající analýza hlavních komponent je neadekvátně interpretována. Naše znalosti o fungování testu jsou z tohoto důvodu omezené. Argumentace pro zařazení konkrétních subtestů a položek je podložena pouze odvoláním se na fakt, že takovéto typy úloh jsou tradiční součástí IQ testů. Na druhou stranu se jedná o dlouho používaný a několikrát revidovaný test. Z kapitoly o reliabilitě můžeme tušit, že autoři se snaží o extrakci g -faktoru, takový

model by ale spíše vyžadoval konstrukci bifaktorového řešení než volbu v manuálu popsanou.

Souběžná validita (jedná se o kapitolu plně převzatou ze zahraničního manuálu) je demonstrována na řadě zahraničních studií prostřednictvím korelací skóre s jinými intelektovými, vývojovými, řečovými a dalšími specifickými testy. Velikost těchto korelací (viz Tabulka 1) považujeme s přihlédnutím k typu použitých testů za uspokojivou. Některé další studie mimo ty uvedené v manuálu (např. Moore, O'Keefe & Lawhon, 1998) hovoří o souběžné validitě s nástrojem WPPSI (Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition, 2012), která dosahuje pro celkové IQ hodnoty $r = 0,87$ (u populace zdravých dětí). Za pozitivní považujeme vysokou korelaci SON-IQ s inteligenčními testy i u postižených dětí (s WISC-R, $r = 0,62$, a s WPPSI-R, $r = 0,82$) – což implikuje fakt, že by SON-R 2½-7 mohl být dobrou alternativou v jejich případě.

Tabulka 1: Testy použité pro validizaci SON-R 2½-7 (převzato ze zahraničních studií)

Intelektové testy	r	Vývojové testy	r	Řečové testy	r	Další specifické testy	r
RAKIT	0,60	BOS 2-30	0,59	TvK	0,59	TOMAL	0,45
TONI-2	0,51	GOS 2½-4½	0,65	SSV	0,29	DTVP-2	0,73
WPPSI-R	0,59			PVT-R	0,29		
BAS (všech 6 subtestů)	0,83			PLS-3	0,47		
K-ABC (neverbální škála)	0,67			Reynell-Schlichting	0,48		

Autoři manuálu tvrdí, že SON-R se odlišuje od konvenčních intelektových testů „jak po stránce obsahové, tak i způsobem administrace“, je tedy vhodné ho administrovat jako druhý test v případech, kdy je nutné učinit závažná rozhodnutí. Dle našeho názoru může SON-R skutečně sloužit jako vhodný diagnostický nástroj v situacích, kdy je evidentně výkon dítěte v jiném testu snižován například jeho úzkostností či komunikační nedostatečností. Čekali bychom tedy, že inkrementální validita testu bude v manuálu více komentována.

Normy a interpretace výsledků

Česká standardizační studie SON-R byla provedena na vzorku 450 dětí, které byly rozděleny do devíti skupin po šesti měsících věku od 2;6 do 7;0 let, zajišťovaném pravděpodobně klastrovou metodou. Reprezentativnost vzorku se autoři snažili zajistit podchycením důležitých sociodemografických kritérií, jejichž rozložení bylo stanoveno pomocí údajů ČSÚ. Mezi kritéria řadí následující: pohlaví, trvalé bydliště, úplnost rodiny, vzdělání matky, vzdělání otce, SES, somatické postižení, imigranti/odlišná etnická či národnostní příslušnost. Očekávaná procenta výskytu v populaci a reálné zastoupení dětí v kategoriích podle stanovených kvót se podařilo dodržet. Jednotlivá kritéria se zdají být relevantní vůči měřenému rysu. Nicméně se domníváme, že je důležité upozornit uživatele testů na některé postupy. Všechna data byla sbírána v rámci vyšetření v PPP, kterých proběhlo celkově 600, a z těchto vyšetření bylo záměrně (tak, aby výsledný

soubor kopíroval předem stanovená standardizační kritéria) vybráno 450 dětí. Domníváme se, že vhodnější by bylo použít váženého výpočtu statistik na všech dětech bez selekce. Vzhledem k reprezentativnosti je dobré mít na mysli případnou specifickou takto vybraných dětí. Zastoupení kritérií “somatické postižení” a “odlišná etnická/národnostní příslušnost” není komentováno na úrovni jednotlivých věkových skupin. Čtenář tedy podle manuálu nedokáže rozpoznat, zda byly všechny děti splňující zmiňovaná kritéria zařazeny do jedné věkové skupiny. To by mohlo vést ke zkreslení výsledků. Je zajímavé, že na stránkách distributora testu je uvedeno, že standardizační studie proběhla na 500 dětech – to je o celou 1 věkovou skupinu více než je uvedeno v manuálu (Hogrefe, n. d.). V manuálu není k dispozici informace o tom, které děti byly ze vzorku explicitně vyloučeny.

Poněkud nesrozumitelně může na uživatele testu působit nejednotnost mezi věkovým rozpětím vzorku uváděným v sekci “3. Česká standardizační studie” (zmiňovaných 2;6 až 7;0 let) a tím, co uživatel nachází v samotné Tabulce norem, ve které se objevují i normy pro děti ve věku 2;4, 2;5 a 7;1 až 7;6. Na základě jakého vzorku byly normy skutečně zkonstruovány, se nám nepodařilo dohledat.

Interpretace výsledku dítěte na základě norem dostupných v papírové verzi manuálu je pro uživatele přívětivá. Každá věková skupina má svou vlastní dvojstranu, tudíž není nutno nepřehledně a složitě otáčet listy. Na této dvojstraně se také nacházejí tabulky s převody HS na VS, což celý proces usnadňuje. Pro přesnější diagnostiku nám schází intervaly spolehlivosti pro percentily.

Uživatel testu by měl mít na paměti to, že na obou koncích věkového rozložení nastávají při interpretaci problémy. U nejnižších věkových skupin je znatelný efekt podlahy, který v praxi znamená, že dítě prakticky nemůže dosáhnout nižšího součtového skóre, než je 35 a tím pádem ani menšího IQ skóre než 68, který je roven 2. percentilu. To je pro způsob, jakým se percentily interpretují, docela paradoxní. Tento nedostatek autoři v manuálu reflektují a u nejmladších dětí doporučují test pouze pro výzkumné účely. Naopak u nejstarších dětí nacházíme efekt stropu, který v praxi znamená, že žádné z takto starých dětí nemůže dosáhnout vyššího IQ, než je 129 (97. percentil).

Z hlediska interpretace je dobré zvažovat i případnou diskrepanci mezi škálou usuzování a škálou performační, která se v praxi objevuje zřídka. Může však poukázat na řečový deficit (ten se promítá i do oblastí usuzování, zobecňování a kategorizace) nebo naopak na výrazný deficit prostorového vnímání či zrakovou vadu.

Závěr

Při práci se SON-R 2½–7 uživatel musí mít na paměti určité nedostatky, kterými test trpí a které jsme již zmínili. Jde především o nedostatečnost norem v nejmladší a nejstarší věkové skupině způsobenou efektem podlahy a stropu, nedostatečné ověření konstruktové a faktorové validity a poněkud kuriózní přístup k odhadu reliability celkového skóre, která zamlžuje skutečnou přesnost měření celého testu. Poněkud matoucí je také uspořádání manuálu, který mísí poznatky z české standardizační studie a ze studie německé či holandské, což může uživateli znesnadnit práci.

Žádný z těchto problémů však není tak závažný, abychom mohli test nedoporučit k užívání, nebo jeho užívání jinak omezovali. Z vícera zdrojů je zřejmé, že SON-R 2½–7 intelekt měří (např. Jenkinson, Roberts, Dennehy & Tellegen, 1996; Moore, O'Keefe & Lawhon, 1998). Nicméně je na místě, aby se autoři české adaptace snažili uživateli testu podat vysvětlení, proč se tomu tak děje, a v návaznosti na to dokázali vytvořit smysluplný strukturní model metody. Jak by takový model mohl vypadat, jsme nastínili již v kapitole Validita. I když se může zdát, že tato vylepšení nebudou mít dopad na využití metody v praxi, není tomu tak. Domníváme se, že lepší uchopení dat může vést k přesnějším odhadům reliability i lepšímu pochopení toho, v čem dané dítě exceluje, nebo naopak zaostává, což může praktickému využití značně pomoci. Bez tohoto totiž máme pouze šest subtestů, které v nevalně provedené explorativní faktorové analýze vytvoří dva faktory po třech subtestech.

Z pohledu reliability je metoda adekvátní. Autoři u odhadu reliability neskončili pouze u vnitřní konzistence, ale snažili se uživatelům testu poskytnout index zobecnitelnosti (g). Pokus o výpočet hierarchického indexu je při měření g-faktoru inteligence chvályhodný, způsob jeho výpočtu ale není na základě informací z manuálu pochopitelný. Pokud by byl index odhadnut neadekvátním způsobem, mohla by být standardní chyba odhadu zkreslena oběma směry, a v tom důsledku i širě 80procentního intervalu spolehlivosti, který je autory doporučován jako hlavní interpretační nástroj.

O nedostatečnosti konstruktové validity se zmiňujeme v odstavci výše. Další důkazy o validitě nám autoři poskytují skrze validitu kriteriální a souběžnou. Důkazy o kriteriální validitě jsou uspokojivé, i přes některá, hlavně statistická, pochybení v rámci analýz. Kapitulu o souběžné validitě potom tvoří překlad zahraničního manuálu a v rámci něho provedených analýz. Předpokládáme, že tento způsob lze považovat za dostatečný s přihlédnutím k neverbální povaze testu a faktu, že děti z Německa, Nizozemí i USA žijí v podobné kultuře, používají stejné hračky, setkávají se stejnými podněty apod.

Vzhledem ke své neverbální povaze je SON R 2½-7 považován za tzv. *culture fair* test. Toto pojetí je nicméně problematické a neznamená, že je možné test administrovat komukoli bez dalšího zvážení. Například u dětí z odlišného socio-kulturního prostředí, které na implicitní úrovni uvažují v odlišných jazykových kódech, je třeba k jejich specifikům přihlížet a poněkud modifikovat způsob administrace, tak, aby dítě nebylo znevýhodněno pouze z důvodu atypického přístupu k řešení úloh, které však nemusí vypovídat nic o jeho adaptaci. Podrobná doporučení pro práci s takovými dětmi uvádějí ve svém textu Pekárková, Lábusová, Rendl a Nikolai (2010). Považujeme za vhodné se tímto textem seznámit.

Vyzdvihnout musíme velice kvalitní zpracování testového materiálu, který je pro děti přitažlivý. Podobně tak i způsob administrace, který zohledňuje to, že se pracuje s opravdu malými dětmi, které potřebují motivaci, aby vydržely pracovat. Samotný manuál je, až na pár překlepů, kvalitně zpracován. Vyplnění záznamového archu je otázkou několika minut a není třeba vykonávat speciální výpočetní či jiné úkony, neboť autoři vše potřebné poskytují na dané dvojstraně v sekci Normy.

Shrnutí

Dle našeho názoru je metoda použitelná v oblasti klinické, školní a poradenské psychologie. Omezená je svou menší diferenciační schopností na obou koncích věkového rozložení. Významně limitováno je využití testu u dětí s jakýmkoli narušením zrakového vývoje, především u těch s centrálními zrakovými vadami (CZV), jež v předškolním věku často ještě nebývají diagnostikovány. CZV se v tomto případě promítne zpravidla do všech subtestů a výrazně zkresluje interpretaci výsledků.

Kromě PPP a dalších specializovaných center najde nástroj využití i ve výzkumu (např. Udholm, Aaberg, Bloch, Sandahl & Ovesen, 2017; Kiechl-Kohlendorfer, Ralser, Peglow, Pehboeck-Walser & Fussenegger, 2013). V případě vyšetření nadaných dětí předškolního věku je administrace problematická v případě pomalého pracovního tempa či vyšší úzkostnosti dítěte. Test pak může trvat i více než 60 minut, což je u takto malého dítěte s ohledem na schopnost udržení pozornosti doba příliš dlouhá. Velice vhodné je SON-R 2½–7 používat jako potvrzení výsledků předešlých metod či v situacích, kdy nelze administrovat běžnější inteligenční test.

V poradenské praxi lze test využít v rámci mimoúrovňového testování i u starších dětí s mentálním postižením pro stanovení přibližné vývojové úrovně. Kvůli nedostatku standardizovaných metod v našem prostředí je možné SON R využít jako metodu první volby pro stanovení mentální úrovně dětí mladších pěti let. Metodu považujeme za kvalitní a je podle nás škoda, že česká adaptace nedosahuje kvalit odpovídajících testu samotnému, čímž jeho interpretační možnosti spíše narušuje.

Zdroje

- Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Cattell, R.B. (1971). *Abilities: Their structure, growth and action*. Boston, MD: Houghton Mifflin.
- Hogrefe. (n.d.). *SON-R 2½ – 7*. Staženo z:
<http://www.testcentrum.com/testy/son%E2%80%932-%E2%80%937>
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 57, 253–270.
- Hülür, G., Wilhelm, O., & Robitzsch, A. (2011). Intelligence differentiation in early childhood. *Journal Of Individual Differences*, 32(3), 170-179. doi:10.1027/1614-0001/a000049
- Jenkinson, J., Roberts, S., Dennehy, S., & Tellegen, P. (1996). Validation of the Snijders-Oomen Nonverbal Intelligence Test—Revised 21/2–7 for Australian children with disabilities. *Journal Of Psychoeducational Assessment*, 14(3), 276-286. doi:10.1177/073428299601400307
- Kiechl-Kohlendorfer, U., Ralser, E., Peglow, U. P., Pehboeck-Walser, N., & Fussenegger, B. (2013). Early risk predictors for impaired numerical skills in 5-year-old children born before 32 weeks of gestation. *Acta Paediatrica*, 102(1), 66-71. doi:10.1111/apa.12036

- McGrew, K.S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37(1), 1-10. doi:10.1016/j.intell.2008.08.004
- Moore, C., O'Keefe, S.L., Lawhon, D., & Tellegen, P. (1998). Concurrent validity of the Snijders-Oomen Nonverbal Intelligence Test 2 ½-7—Revised with the Wechsler Preschool and Primary Scale of Intelligence—Revised. *Psychological Reports*, 82(2), 619-625. doi:10.2466/pr0.1998.82.2.619
- Pekárková, S., Lábusová, A., Rendl M., & Nikolai T. (2010). Nemoc bezmocných: Lehká mentální retardace. Analýza inteligenčního testu SON-R. Praha: Člověk v tísni. Staženo z: <http://osf.cz/wp-content/uploads/2015/04/analyza-nemoc-bezmocnych.pdf>
- Schuerger, J., & Witt, A. (1989). The temporal stability of individually tested intelligence. *Journal of Clinical Psychology*, 45(2), 294-302.
- Tellegen, P.J., Laros, J.A., & Heider, D. (2008). *Neverbální inteligenční test SON-R 2 ½-7 - uživatelská příručka*. Praha: Testcentrum – Hogrefe.
- Tellegen, P.J., Winkel, M., Wijnberg-Williams, B.J. & Laros, J.A. (1998). *Snijders-Oomen Nonverbal Intelligence Test. SON-R 2 ½ -7 Manual and Research Report*. Lisse: Swets & Zeitlinger B.V.
- Udholm, N., Aaberg, K., Bloch, C., Sandahl, M., & Ovesen, T. (2017). Cognitive and outcome measures seem suboptimal in children with cochlear implants - a cross-sectional study. *Clinical Otolaryngology*, 42(2), 315-321. doi:10.1111/coa.12723
- Wechsler, D. (2012). Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition. *Technical and interpretive manual*. San Antonio, TX: NCS Pearson, Inc.

MODEL RECENZE PODLE EFPA PRO POPIS A HODNOCENÍ PSYCHOLOGICKÝCH TESTŮ *FORMULÁŘ RECENZE TESTU A POZNÁMKY PRO RECENZENTY¹*

Toto je lokální úprava dokumentu pro účely publikace v časopise Test Fórum.
Originální český překlad je k dispozici na stránkách EFPA
(www.efpa.eu/download/505cd9db4144ecb16174087909c9cd6d).

Původní verzi sestavil a uspořádal Dave Bartram
Doplnili a revidovali Patricia Lindley, Dave Bartram a Natalie Kennedy v dubnu 2004²
Současná verze 3.42: květen 2005
Český překlad: Tomáš Urbánek

Od uživatelů tohoto dokumentu a jeho obsahu žádá EFPA, aby uznali tento zdroj prostřednictvím následujícího textu:

“Kritéria pro recenzi testu podle EFPA do značné míry vychází z formy a obsahu kritérií pro recenze testů Britské psychologické společnosti (BPS) a kritérií vytvořených Komisí pro testové záležitosti (COTAN) Holandské asociace psychologů (NIP). Dave Bartram a Patricia Lindley původně vyvinuli kritéria BPS a recenzní procedury pro UK Employment Service a později rozšířili jejich používání pro celou BPS. Arne Evers připravil k vydání nizozemský systém posuzování kvality testů.

EFPA je vděčná BPS a NIP za svolení použít jejich kritéria jako základ pro vytvoření evropského modelu. EFPA je také vděčná Davu Bartramovi, Arnu Eversovi a Patricii Lindley za jejich příspěvek k vývoji tohoto modelu. Veškerá intelektuální vlastnická práva původních kritérií podle BPS a NIP jsou nadále uznávána a náleží těmto orgánům.”

¹ Tento dokument byl vytvořen z několika zdrojů, včetně Hodnotícího formuláře pro recenzi testu používaného v BPS (NPAL a Řídící komise pro testové standardy při BPS – Steering Committee on Test Standards), Španělského dotazníku pro hodnocení psychometrických testů (Španělská psychologická asociace) a Systému pro posuzování kvality testu (Komise pro testování Holandské asociace psychologů). Některé části byly adaptovány se svolením z dokumentu: BPS Books Reviews of Level B Assessment Instruments for use in Occupational Assessment, Notes for Reviewers: Version 3.1. December 1998: Copyright © NPAL, 1989, 1993, 1998.

² Současná verze je spojením dvou oddělených dokumentů (Formuláře recenze a Poznámek pro recenzenty). Obsah byl navíc uspořádán a doplněn na základě jeho používání recenzenty online testů v BPS.

.....
Část 1:

Popis nástroje: Obecné informace a klasifikace
.....

EFPA 3.2 reference

	Recenzent 1:	Bc. Edita Chvojková
	Recenzent 2:	Bc. Štěpán Postulka
	Recenzent 3:	Mgr. Tereza Horáková
	Konzultující editor:	Mgr. Hynek Cígler, Ph.D.
	Vedoucí editor:	Mgr. Hynek Cígler, Ph.D.
	Vedoucí editor aktualizace: (pouze v případě aktualizací)	
	Editor aktualizace: (pouze v případě aktualizací)	
	Datum vzniku této recenze:	30. 11. 2017
1.1	Název nástroje (lokální verze):	Snijders-Oomen neverbální inteligenční test
	Zkrácená verze názvu testu:	SON-R 2½-7
1.2	Původní název testu (pokud je lokální verze adaptací):	Snijders-Oomen-Nonverbal-Intelligence Test – Revised 2½ - 7
1.4	Autoři původního testu:	Peter J. Tellegen & Jacob A. Laros
1.3	Autoři lokální adaptace:	PhDr. Dana Krejčířová, Mgr. Dana Černochová, PhDr. Jana Zapletalová, Mgr. Helena Pacnerová & Mgr. Hana Lukášová
1.7	Lokální distributor/vydavatel testu:	Testcentrum - Hogrefe
1.8	Vydavatel původní verze testu (pokud je jiný než současný distributor/vydavatel):	Hogrefe, Göttingen
1.9.1	Datum vydání současné revize/vydání:	1998
1.9.2	Datum vydání adaptace pro lokální užívání:	2008
1.9.3	Datum vydání původního testu:	1975

Obecný popis nástroje

SON-R 2½-7 je baterie pro měření fluidní inteligence určena pro děti ve věku 2;4 až 7;7 let. Jedná se o třetí revizi neverbálního inteligenčního testu původně určeného neslyšícím dětem. Později došlo k rozšíření metody i pro děti slyšící a poslední revize už obsahuje normy obecné. Svými autory je test považován za kulturně univerzální právě díky neverbální povaze, která na testované dítě neklade specifické nároky v podobě jazykových schopností.

Baterie obsahuje celkem šest subtestů rozčleněných do dvou škál – performační a úsudkové. Do performační škály náleží subtesty Mozaiky, Skládanky a Vzory. Do škály úsudkové potom náleží subtesty Kategorie, Analogie a Situace. Každý ze subtestů má dvě části, které jsou od sebe odděleny zácvikovou položkou. Položky jsou skórovány dichotomicky – nula bodů je dítěti uděleno v případě, kdy položku nezvládne, nebo mu administrátor musí pomoci, a jeden bod je udělen vždy, když je položka zodpovězena správně. Hrubé skóry v jednotlivých subtestech se převádí na vážené skóry ($M = 10$, $SD = 3$) podle instrukcí uvedených v manuálu v závislosti na věku dítěte. Součet VS každých tří subtestů vytváří dva IQ skóry ($M = 100$, $SD = 15$) – SON-PŠ (IQ skór škály performační) a SON-ŠÚ (IQ skór škály úsudkové), které jsou poměrně složitě interpretovatelné. Interpretovatelným je skór SON-IQ, který je tvořen součtem VS všech subtestů a vypovídá přímo o fluidní inteligenci dítěte.

Test má podle autorů měřit fluidní inteligenci v rámci modelu Cattella (1971). Fluidní inteligence zde slouží jako jakýsi g-faktor. Není však založen na žádném specifickém modelu inteligence, což si o testu myslí i jiní autoři (např. Hülür, Wilhelm & Robitzsch, 2011). Ve většině strukturně založených teorií (Cattell, 1991; Horn & Cattell, 1966) bychom subtesty SON-R 2½-7 mohli opravdu klasifikovat jako měřící fluidní inteligenci (v CHC jako Gf) nebo širší vizualizační schopnosti (v CHC jako Gv; McGrew, 2009). Škála performační potom tvoří jeden faktor druhého řádu, který sytí položky tří subtestů, v nichž jde o prokázání prostorového vhledu (Carroll, 1993). Škála úsudková tvoří druhý faktor druhého řádu, který sytí položky tří subtestů, v nichž jde o prokázání schopností konkrétního a abstraktního myšlení (Carroll, 1993).

Výkon dítěte je porovnávám s normami vytvořenými na reprezentativním vzorku 450 dětí rozdělených do devíti skupin po 50. Uživatelům je k dispozici srovnání výkonu vyšetřovaného dítěte s ostatními pomocí percentilů. Specifikem testu je možná korekce ze strany administrátora v situacích, ve kterých dítě odpověď neví či položku zodpoví nesprávně. Oproti jiným inteligenčním bateriím pro děti ve stejné věkové kategorii také administrátor podává dětem po každé položce zpětnou vazbu. Test je administrován individuálně formou tužka-papír a jedna administrace zabere přibližně 50 minut v závislosti na věku testovaného dítěte. Uživatel má možnost vyhodnotit výsledky administrace pomocí speciálního počítačového programu, který mu práci usnadní a také zpřesní, neboť nedochází k zaokrouhlování. Testování je adaptivním v tom smyslu, že se administrátor může rozhodnout, zda začne od položky jedna, tři či pět. Rozhodnutí se nejčastěji zakládá na věku dítěte, ale může být založeno i na odhadu kognitivní úrovně dítěte. Autoři se tak snaží zabránit demotivaci, která by mohla vzniknout řešením příliš jednoduchých položek. Naopak je testování ukončeno, pokud dítě v subtestech ÚS třikrát chybuje nebo pokud v subtestech PŠ chybuje dvakrát za sebou, čímž se autoři znovu snaží vyhnout demotivaci.

.....

Část 2:
Klasifikace

.....

1.10.1	Obsahová doména	<input checked="" type="checkbox"/> Školní schopnosti <input checked="" type="checkbox"/> Všeobecné schopnosti <input type="checkbox"/> Verbální schopnosti <input type="checkbox"/> Numerické schopnosti <input checked="" type="checkbox"/> Prostorové schopnosti <input checked="" type="checkbox"/> Neverbální schopnosti <input type="checkbox"/> Rychlost vnímání <input type="checkbox"/> Paměť <input type="checkbox"/> Manuální zručnost <input type="checkbox"/> Osobnost – Rys <input type="checkbox"/> Osobnost – Typ <input type="checkbox"/> Osobnost – Stav <input type="checkbox"/> Kognitivní styly <input type="checkbox"/> Motivace <input type="checkbox"/> Hodnoty <input type="checkbox"/> Zájmy <input type="checkbox"/> Přesvědčení <input type="checkbox"/> Poruchy a patologie <input type="checkbox"/> Skupinové procesy <input type="checkbox"/> Rodina <input type="checkbox"/> Organizace, její fungování, agregovaná měření, klima atd. <input type="checkbox"/> Školní nebo výchovné funkce <input type="checkbox"/> Jiné:
1.10.2	Zamýšlená(é) nebo hlavní oblast(i) použití.	<input checked="" type="checkbox"/> Klinická psychologie <input type="checkbox"/> Neuropsychologie <input type="checkbox"/> Forenzní psychologie <input checked="" type="checkbox"/> Psychologie výchovy a vzdělávání <input type="checkbox"/> Psychologie práce a personalistika <input checked="" type="checkbox"/> Poradenství, doporučení, vedení a volba povolání <input type="checkbox"/> Psychologie zdraví, životní styl a životní spokojenost <input type="checkbox"/> Sporty a volný čas <input type="checkbox"/> Jiné:

1.10.3	Zamýšlený způsob použití (podmínky, za jakých byl nástroj standardizován a validizován)	<input type="checkbox"/> Nesupervidovaná administrace bez kontroly nad identitou respondenta a bez úplné kontroly nad podmínkami administrace (např. volně přístupný test na internetu, test dostupný ke koupi v knihkupectví). <input type="checkbox"/> Kontrolovaný nesupervidovaná administrací. Kontrola nad podmínkami (čas atd.) a určitá kontrola nad identitou uživatele testu (např. testy administrované přes internet, ale pouze známým osobám – přístup omezený heslem). <input checked="" type="checkbox"/> Supervidovaná a kontrolovaná administrace. Administrace testu pod kontrolou kvalifikovaného administrátora nebo dohlázele. <input type="checkbox"/> Řízená administrace. Administrace testu prováděná pouze přes určená testovací centra (např. programy hodnocení licencí a certifikace).
1.10.4	Popis populací, pro které je test určen:	Děti ve věku 2;4 až 7;7 let.
1.10.5	Počet škál a krátký popis proměnné nebo proměnných měřených nástrojem	<p>2 škály po 3 subtestech, každá v průměru 15 položek:</p> <ul style="list-style-type: none"> - Škála performační: dítě řeší jednotlivé položky a souběžně s tím manipuluje s testovým materiálem - Škála úsudková: dělí se na testy abstraktního myšlení (pochopení principu a aplikace na nové materiály) a konkrétního myšlení (nastolení realistického vztahu mezi osobami či objekty) <p>Test má měřit fluidní inteligenci podle modelu Cattella (1971), která zde slouží jako jakýsi g-faktor. Škála performační potom tvoří jeden faktor druhého řádu, který sytí položky tří subtestů, v nichž jde o prokázání prostorového vhledu (Carroll, 1993). Škála úsudková tvoří druhý faktor druhého řádu, který sytí položky tří subtestů, v nichž jde o prokázání schopností konkrétního a abstraktního myšlení (Carroll, 1993).</p>

1.11	Formát položek	<input checked="" type="checkbox"/> Otevřený <input type="checkbox"/> Mnohonásobná volba, alternativy na stejné škále <input type="checkbox"/> Bipolární adjektiva <input type="checkbox"/> Likertovy ratingy (škály) <input type="checkbox"/> Nucená volba, alternativy na smíšených škálách (ipsativní) – vysvětlení viz Poznámky <input type="checkbox"/> Mnohonásobná volba, alternativy na smíšených škálách (ipsativní) – vysvětlení viz Poznámky <input type="checkbox"/> Sady párů adjektiv (sémantický diferenciál), smíšené škály (ipsativní) <input type="checkbox"/> Jiné: <p style="padding-left: 40px;">popište</p>
1.12	Počet položek testu:	Performační škála: subtesty Mozaiky (15 položek), Skládanky (14) a Vzory (16) Úsudková škála: subtesty Kategorie (15), Analogie (17) a Situace (14).
1.13	Způsob(y) administrace:	<input checked="" type="checkbox"/> Interaktivní individuální administrace <input type="checkbox"/> Supervidovaná skupinová administrace <input type="checkbox"/> Počítačová lokálně nainstalovaná aplikace – pod supervizí/dohledem <input type="checkbox"/> Počítačová aplikace na webu – pod supervizí/dohledem <input type="checkbox"/> Počítačová lokálně nainstalovaná aplikace – bez supervise/testování sebe <input type="checkbox"/> Počítačová aplikace na webu – bez supervise/testování sebe <input type="checkbox"/> Jiné:
1.14	Způsob odpovídání:	<input type="checkbox"/> Ústní rozhovor <input checked="" type="checkbox"/> Papír a tužka <input checked="" type="checkbox"/> Manuální operace <input type="checkbox"/> Na počítači <input type="checkbox"/> Jiné:

1.15	<p>Čas potřebný pro administraci nástroje:</p> <p>Čas na přípravu (čas, který zabere administrátorovi připravit a rozložit materiály pro diagnostické sezení).</p> <ul style="list-style-type: none"> • Čas na administraci na sezení: zahrnuje čas potřebný pro dokončení všech položek a odhad času potřebného pro podání instrukcí, projití zácvičných položek a nějaké doplňující komentáře na konci sezení. • Skórování: čas nutný pro získání hrubých skórů. • Analýza: čas strávený prováděním dalších prací s hrubými skóry, aby se z nich odvodily další míry a zformulovala rozumně úplná interpretace (za předpokladu, že znáte nástroj). • Zpětná vazba: čas potřebný k přípravě a poskytnutí zpětné vazby vyšetřované osobě. <p>Připouští se, že čas posledních dvou komponent se může značně lišit – v závislosti na kontextu, ve kterém se nástroj používá. Ale aspoň nějaké údaje a komentáře budou užitečné.</p>	<p>Příprava: 5 minut</p> <p>Administrace: přibližně 50 minut</p> <p>Skórování: 5 minut</p> <p>Analýza: 10 minut</p> <p>Zpětná vazba: 5 minut</p> <p>Vše v závislosti na tom, od jaké položky se test začne administrovat a s jak starým dítětem uživatel testu pracuje.</p>
1.16	Jsou k dispozici různé formy nástroje?	Nejsou.

.....

Část 3:

Měření a skórování

.....

1.17	Procedura skórování testu:	<input type="checkbox"/> Počítačové skórování s přímým vstupem odpovědí testovanou osobou <input checked="" type="checkbox"/> Počítačové skórování s ručním vstupem odpovědí z papírového záznamového archu <input type="checkbox"/> Počítačové skórování odpovědí z papírového záznamového archu pomocí jejich načtení pomocí skeneru <input checked="" type="checkbox"/> Jednoduchý ruční skórovací klíč – nutné pouze kancelářské dovednosti <input type="checkbox"/> Komplexní ruční skórování – vyžadující trénink ve skórování nástroje <input type="checkbox"/> Služby zpracování dat – např. skórování společností prodávající nástroj <input type="checkbox"/> Jiné:
1.18	Skóry:	Hrubé skóry z jednotlivých subtestů se podle přiloženého klíče převedou na skóry vážené. Tyto VS je poté možné sčítat a převádět na interpretovatelné a srovnatelné IQ skóry.

1.19	Transformace skóru na standardní skóry:	<input type="checkbox"/> Normalizovaná – skóry se získají použitím normalizační tabulky <input checked="" type="checkbox"/> Nenormalizovaná – skóry se získají lineární transformací
1.20	Použité škály	Skóry založené na percentilech <input checked="" type="checkbox"/> Centily <input type="checkbox"/> 5-stupňová klasifikace: centilové rozdělení 10:20:40:20:10 <input type="checkbox"/> Decily a další klasifikace založené na stejném počtu percentilů Standardní skóry <input type="checkbox"/> z-skóry <input checked="" type="checkbox"/> IQ deviační kvocienty IQ atd. (např. M=100, SD=15 pro Weschlerův test) <input type="checkbox"/> Přijímací testy na VŠ (např. test SAT M=500, SD=100; GRE atp.) <input type="checkbox"/> steny, staniny, C skóry <input type="checkbox"/> T-skóry <input checked="" type="checkbox"/> Jiné: Vážené subtestové skóry s M=10, SD=3

.....

Část 4:

Počítačově generované zprávy

.....

Toto je čistě *popisné*. Hodnocení zpráv bude součástí části recenze nazvané Hodnocení.

1.21	Jsou počítačově generované zprávy k dispozici s nástrojem?	<input type="checkbox"/> Ano <input checked="" type="checkbox"/> Ne
------	---	--

.....

Část 5:

Nabídka, podmínky a náklady

.....

Tato část definuje, co vydavatel poskytne, komu, za jakých podmínek a za jaké ceny. Definuje podmínky kladené dodavatelem a týkající se toho, kdo smí a kdo nesmí získat materiál nástroje. Pokud jedna z možností neodpovídá podmínkám nabídky, doplňte popis relevantních podmínek.

1.23	Dokumentace poskytovaná distributorem jako součást testového balíku	<input checked="" type="checkbox"/> Uživatelský manuál <input type="checkbox"/> Technický (psychometrický) manuál <input checked="" type="checkbox"/> Doplňkové technické informace a aktualizace (např. lokální normy, lokální validizační studie atd.) <input type="checkbox"/> Rozšiřující informace v podobě knih a článků k tématu <input type="checkbox"/> Kombinace výše uvedených (uved'te)
------	--	---

1.24	Metody publikace	<input checked="" type="checkbox"/> Papír <input type="checkbox"/> PC – Diskety <input type="checkbox"/> PC – CD/ROM <input type="checkbox"/> Download z internetu <input type="checkbox"/> Živý internet (nástroj pracuje v internetovém prohlížeči) <input type="checkbox"/> Jiné:
1.25.1	Počáteční náklady. Cena kompletní sady materiálů (všechny manuály a další materiál nutný k aspoň jedné zkušební administraci). Kolik uchazečů lze vyšetřovat pomocí materiálů získaných za počáteční náklady, kde tyto náklady zahrnují materiály pro opakované vyšetření.	45 200 Kč – úplný testový soubor - Příručka – 1500 Kč - Testovací materiál – 39 700 Kč - Záznamový list 25 ks – 800 Kč - Pracovní sešit Vzory 50 ks – 4500 Kč
1.25.2	Opakující se náklady:	- Záznamový list 25 ks – 800 Kč - Pracovní sešit Vzory 50 ks – 4500 Kč Po zaplacení jednorázových poplatků (Příručka + Testovací materiál) činí náklady na vyšetření jedno vyšetření dítěte 122 Kč.
1.26.1	Ceny za zprávy generované softwarem nainstalovaným uživatelem:	
1.26.2	Ceny za vyhotovení zprávy zaslané prostřednictvím pošty/faxu:	
1.26.3	Ceny za vyhotovení zprávy zaslané prostřednictvím internetové služby:	
1.27	Ceny za další služby a zpracování dat: opravy nebo vývoj automatických zpráv:	
1.28	Kvalifikační požadavky na práci s testem vyžadované dodavatelem testu 1.28 se týká kvalifikací uživatele vyžadovaných dodavatelem. V této části, pokud vydavatel stanovil informace o kvalifikaci uživatele, mělo by to být uvedeno pomocí uvedených kategorií. Tam, kde kvalifikační požadavky nejsou jasné, mělo by to být vyjádřeno pomocí "Jiné", <i>ne</i> "Žádné". "Žádné" znamená, že existuje explicitní výrok týkající se toho, že není potřeba kvalifikace.	<input type="checkbox"/> Žádné <input checked="" type="checkbox"/> Oprávnění (certifikát) pro specifický test <input type="checkbox"/> Oprávnění (certifikát) pro obecné výkonové testy: i. e. míry maximálního výkonu ve schopnostech <input type="checkbox"/> Potvrzení v testování obecných schopností a dovedností: míry maximálního výkonu ve vztahu k potenciálu k výkonu <input type="checkbox"/> Potvrzení v obecné diagnostice a diagnostice osobnosti: míry typického chování, postojů a preferencí <input type="checkbox"/> Jiné:
1.29	Profesionální kvalifikace vyžadovaná pro používání nástroje 1.29 se týká kvalifikací uživatele vyžadovanou dodavatelem. V této části, pokud vydavatel stanovil informace o kvalifikaci uživatele, mělo by to být	<input type="checkbox"/> Žádné <input checked="" type="checkbox"/> Praktický psycholog s kvalifikací v relevantní aplikační oblasti <input checked="" type="checkbox"/> Praktický psycholog <input checked="" type="checkbox"/> Výzkumný psycholog <input type="checkbox"/> Nepsychologický akademický výzkumník

	uvedeno pomocí uvedených kategorií. Kde požadavky na kvalifikaci nejsou jasné, mělo by to být vyjádřeno pomocí "jiné", ne "Žádné". "Žádné" znamená, že existuje explicitní výrok týkající se toho, že není potřeba kvalifikace.	<input type="checkbox"/> Praktik v relevantních příbuzných profesích (terapie, medicína, poradenství, vzdělání, lidské zdroje atd.) <input type="checkbox"/> Držitel Certifikátu způsobilosti pro testování v psychologii práce A BPS <input type="checkbox"/> Držitel Certifikátu způsobilosti pro testování v oblasti vzdělávacím A BPS <input type="checkbox"/> Držitel Certifikátu způsobilosti pro testování v psychologii práce B BPS <input type="checkbox"/> Jiné:
--	---	--

.....

Část 6:

Hodnocení testových materiálů

.....

Vysvětlení hodnocení

V následujících částech jsou celková posouzení adekvátnosti informací týkajících se validity, reliability a norem zobrazeny automaticky tučně.

Jakýkoli nástroj s jedním nebo více posouzeními 0 nebo 2 týkajícími se atributů považovaných za kritické pro bezpečné používání nástroje, by neměl být považován za nástroj, který splňuje minimální standardy.

Vstup na posuzovacím formuláři	Posouzení podle standardů EFPA	Reprezentace recenze v UK	Vysvětlení
[n/a]	[n/a]	[n/a]	Tento atribut není u tohoto nástroje použitelný
0	[-]	[None]	Není možné posoudit jako ne nebo nedostatek poskytnutých informací
1	[-1]	[*]	Neadekvátní
2		[**]	NYNÍ NEPOUŽÍVÁNO
3	[0]	[***]	Adekvátní nebo přiměřený
4	[1]	[****]	Dobrý
5	[2]	[*****]	Vynikající
		[N.r.i.o.r] * (pouze pro aktualizace)	Položka nebyla v původní recenzi posuzována

V této části má být provedeno více hodnocení různých aspektů nebo atributů dokumentace dodávané s nástrojem (nebo balíkem). Termín „dokumentace“ byl vybrán, aby pokrýval všechny ty materiály dodávané s nástrojem nebo snadno dostupné kvalifikovanému uživateli: např. manual administrátora; technické příručky; brožury s normami; dodatky k manuálu; aktualizace od vydavatelů/dodavatelů atd.

Položky mají být posuzovány n/a nebo 0 až 5 (poloviční rating je přijatelný)

Rating

Kvalita vysvětlení principů, prezentace a kvalita poskytnuté informace: (Tento celkový rating se získá použitím posouzení založeného na ratingích daných pro položky 2.1–2.8)		3,5
2.1	Celkový rating kvality vysvětlení principů: (Tento celkový rating se získá použitím posouzení založeného na hodnotách ratingů daných pro položky 2.1.1 – 2.1.5)	3
2.1.1	i) Teoretické základy konstruktů:	1,5
2.1.2	ii) Procedura vývoje testu:	4

2.1.3	iii) Důkladnost analýz položek a model analýzy položek:	3,5
2.1.4	iv) Vysvětlení obsahové validity:	1,5
2.1.5	v) Souhrn relevantního výzkumu:	1,5
2.2	Adekvátnost dokumentace dostupné uživateli (uživatelské a technické manuály, dodatky týkající se norem atd.): (Tento celkový rating se získá použitím posouzení založeného na hodnotách ratingů daných pro položky 2.2.1 – 2.2.6) <i>Pro část 2.2 jsou stanoveny následující „měřítka“ pro rating „vynikající“ (5). Pozornost je zde zaměřena na kvalitu pokrytí poskytnutého v dokumentaci dostupné kvalifikovaným uživatelům. Všimněte si, že část 2.2 se týká úplnosti a jasnosti dokumentace dostupné uživateli (uživatelské a technické manuály, dodatky k normám atd.) v pojmech pokrytí a vysvětlení. V pojmech kvality nástroje, jak ji dosvědčuje dokumentace, jsou rozpracovány oblasti v této části pod čísly: 2.1, 2.3, 2.9, 2.10 a 2.11.</i>	3
2.2.1	Principy: [viz 2.1] Dobře argumentovaný a jasně prezentovaný popis toho, co má podle návrhu měřit a proč byl zkonstruován tak, jak je.	1,5
2.2.2	Vývoj: Úplné detaily týkající se zdrojů položek, pilotáže, analýz položek, srovnávacích studií a změn prováděných v průběhu vývojových pokusů.	1
2.2.3	Standardizace: Jasná a detailní informace poskytnuté o velikostech a zdrojích standardizačního souboru a standardizační proceduře.	4
2.2.4	Normy: Jasná a detailní informace poskytnuté o velikostech a zdrojích normalizačních skupin, podmínkách vyšetření atd.	4
2.2.5	Reliabilita: Dobře vysvětlení reliability a široký rozsah měř vnitřní konsistence a retestu spolu s vysvětlením jejich relevance a zobecnitelnosti nástroje vyšetření.	1,5
2.2.6	Validita: Dobře vysvětlení validity spolu s širokou škálou studií jasně a poctivě popsaných.	3
2.3	Kvalita procedurálních instrukcí poskytnutých uživateli: (Tento celkový rating se získá s použitím posouzení na základě hodnot ratingů daných pro položky 2.3.1 – 2.3.7)	4
2.3.1	Pro administraci testu: Poskytnutá jasná a detailní vysvětlení a procedurální průvodce krok za krokem spolu s dobrými radami týkajícími se otázek uchazečů a problémových situací.	5
2.3.2	Pro skórování testu, normy atd.: Poskytnuté jasná a detailní informace spolu s popsanými kontrolami pro vyhnutí se možným chybám skórování.	5
2.3.3	Pro interpretaci a vytváření zpráv: Detailní doporučení týkající se interpretace různých skóru, chápání normativních měr a zacházení se vztahy mezi různými škálami, s množstvím ilustrativních příkladů a případových studií.	4
2.3.4	Pro poskytnutí zpětné vazby a debriefingu respondentům testu a dalším: Detailní doporučení, jak prezentovat zpětnou vazbu uchazečům.	1
2.3.5	Pro poskytování dobrých praktických témat týkajících se poctivosti a zkreslení: <i>Uvedení detailních informací o studiích sexuálního a etnického zkreslení s relevantními varováními týkajícími se používání a zobecňování validit.</i>	4
2.3.6	Omezení používání: Jasně popisy, kdo by měl a kdo by neměl být vyšetřován spolu s dobře vysvětlenými odůvodněními těchto omezení (např. typy nezpůsobilostí, požadované úrovně gramotnosti atd.).	1,5
2.3.7	Reference a podpůrné materiály: Detailní odkazy na relevantní podpůrnou akademickou literaturu a křížové odkazy na další příbuzné materiály týkající se diagnostických nástrojů.	3
Kvalita materiálů: (Tento celkový rating se získá použitím posouzení založeného na hodnotách ratingů pro položky 2.4 – 2.8)		4
2.4	Všeobecná kvalita materiálů testu (testové brožury, odpověďové archy, testové objekty, software atd.):	5
2.5	Kvalita lokální adaptace testu (pokud byl test přeložen a adaptován do místního jazyka):	[n/a]
2.6	Snadnost, s jakou může respondent testu porozumět úkolu:	4
2.7	Snadnost, s jakou mohou být respondentem testu tvořeny reakce nebo odpovědi:	4
2.8	Kvalita položek:	4

Recenzentovy komentáře týkající se dokumentace:

(komentáře principů, designu, vývoje testu a jeho přijatelnosti)

Na základě poskytnuté dokumentace lze test administrovat bez problémů. Uživatelům poskytuje dostatečnou podporu a jeho vyhodnocení je snadné a pro respondenty je test velice přívětivý, shlédávají jej zábavným. Za problematickou spatřujeme úvodní části manuálu, které by měly diskutovat teoretické pozadí, vývoj položek, proč a jak položky vznikly apod. Rozumíme, že test má za sebou dlouhou řadu revizí a historii, avšak dle našeho názoru by tady tato část měla být obsáhlejší a explanativnější. Zvláště bychom ocenili informace o tom, proč položky fungují tak, jak fungují.

.....

Část 7:

Hodnocení norem, reliability a validity

.....

Položky mají být posuzovány n/a nebo 0 až 5 (jsou přijatelné poloviční ratingy)

Rating**Hodnocení technických informací – celková adekvátnost:**

(Tento celkový rating se získá použitím posouzení založeného na hodnotách ratingů daných pro položky 2.9 – 2.11)

4

Informace o normách nebo referenční skupině

2.9	Celková adekvátnost:	3,5
2.9.1	Vhodnost pro lokální použití, ať už pro lokální nebo mezinárodní normy: [n/a] Nepoužitelné 0 Žádná informace nepodána. 1 Není lokálně relevantní (např. nevhodné zahraniční výběry). 3 Lokální výběr z obecné populace nebo nelokální normy, které lze použít s varováním. 4 Výběry lokální země nebo relevantní mezinárodní výběry s dobrou relevancí pro zamýšlenou aplikaci. 5 Výběry lokální země nebo relevantní mezinárodní výběry vybrané z dobře definovaných výběrů z relevantních aplikačních oblastí.	4
2.9.2	Vhodnost pro zamýšlené aplikace: [n/a] Nepoužitelné 0 Žádná informace nepodána. 1 Norma nebo normy nejsou adekvátní pro zamýšlené aplikace. 3 Adekvátní normy pro obecnou populaci a/nebo rozmezí normativních tabulek. 4 Dobré rozmezí normativních tabulek. 5 Vynikající rozmezí výběrově relevantních norem vztahujících se k věku a pohlaví, s informacemi o dalších rozdílech v rámci skupin (např. směr etnických skupin).	3
2.9.3	Velikosti výběrů: [n/a] Nepoužitelné 0 Žádná informace nepodána. 1 Neadekvátní výběry (např. méně než 150). 3 Adekvátní výběry (např. 150-300). 4 Velké výběry (např. 300-1000). 5 Velmi velké výběry (např. 1000+).	1

2.9.4	Procedury použité při výběru souboru: (vyberte jednu a ohodnoťte kvalitu použitého postupu) <ul style="list-style-type: none"> <input type="checkbox"/> Žádná informace neposkytnuta <input checked="" type="checkbox"/> Reprezentativní populaci [sumarizujte kritéria] <input type="checkbox"/> Nahodilá <input type="checkbox"/> Náhodná 	3
2.9.5	Kvalita informací poskytnutých o minoritní/chráněné skupině, rozdílech, vlivech věku, rodu atd.: [n/a] Nepoužitelné 0 Žádná informace nepodána. 1 Neadekvátní informace. 3 Adekvátní obecné informace s minimální analýzou. 4 Dobré popisy a analýzy skupin a rozdílů 5 Vynikající série analýz a diskuse o relevantních tématech vztahujících se k použití a interpretaci.	4
2.9.6 Komentáře recenzentů k normám: Stručná zpráva o normách a jejich historii, včetně informací o doporučených učiněných vydavatelem/autorem pro aktualizaci norem obvyklým způsobem.		
<p>Normy byly vytvořeny na základě reprezentativního vzorku 450 dětí. Data byla sbírána mezi léty 2005 a 2006 dvacítkou pověřených psychologů. Sběr dat probíhal v rámci vyšetření v PPP, nebo při vyšetřeních ve specializovaných centrech. Z celkového vzorku 600 dětí bylo tedy záměrným výběrem vybráno 450 tak, aby se naplnila předem stanové kvóty podle sociodemografických kritérií. Mezi kritéria řadí následující: pohlaví, trvalé bydliště, úplnost rodiny, vzdělání matky, vzdělání otce, SES, somatické postižení, imigranti/odlišná etnická či národnostní příslušnost. Očekávaná procenta výskytu v populaci a reálné zastoupení dětí v kategoriích podle stanových kvót se podařilo dodržet. Jednotlivá kritéria se zdají být relevantní vůči měřenému rysu. I přes některé nejanosti, diskutované v textu této recenze, se domníváme, že normy jsou použitelné pro hodnocení a porovnávání výkonu respondentů.</p>		

Validita

2.10	Celková adekvátnost: (Tento celkový rating se získá na základě posouzení hodnot ratingů daných v položkách 2.10.1 – 2.10.2.4. Neprůměrujte pouze čísla, abyste získali celkový rating. Obvykle bude roven buď konstruktové validitě nebo validitě vztahující se ke kritériu, podle toho, která z nich je vyšší.)	4
2.10.1	Konstruktová validita – celková adekvátnost (Tento celkový rating se získá na základě posouzení hodnot ratingů daných v položkách 2.10.1.2 – 2.10.1.6. Neprůměrujte pouze čísla, abyste tento celkový rating získali.)	3
2.10.1.1	Použité plány: (zatrhněte tolik, kolik je jich použitelných) <ul style="list-style-type: none"> <input type="checkbox"/> Žádná informace nepodána <input checked="" type="checkbox"/> Korelace s dalšími nástroji a výkonovými kritérii <input type="checkbox"/> Vnitroškálový (korelace položky se zbytkem) <input checked="" type="checkbox"/> Rozdíly mezi skupinami <input type="checkbox"/> Matice mnoha rysů a mnoha metod (MTMM) <input checked="" type="checkbox"/> Explorační faktorová analýza <input type="checkbox"/> Konfirmační faktorová analýza <input type="checkbox"/> Experimentální plány <input type="checkbox"/> Jiné: uveďte 	
2.10.1.2	Velikosti výběrů: <ul style="list-style-type: none"> 0 Žádná informace neposkytnuta. 1 Jedna neadekvátní studie (např. velikost výběru menší než 100). 3 Jedna adekvátní studie (např. velikost výběru 100-200). 4 Více než jedna adekvátní nebo velká studie. 5 Dobrá série adekvátních až rozsáhlých studií. 	4

2.10.1.3	Procedura výběru souboru: (vyberte jednu) <input type="checkbox"/> Žádná informace neposkytnuta <input checked="" type="checkbox"/> Reprezentativní vůči populaci [sumarizujte kritéria] <input type="checkbox"/> Nahodilá <input type="checkbox"/> Náhodná	
2.10.1.4	Medián a rozsah korelací mezi testem a dalšími podobnými testy: 0 Žádná informace neposkytnuta. 1 Neadekvátní ($r < 0.55$). 3 Adekvátní ($0.55 < r < 0.65$). 4 Dobrý ($0.65 < r < 0.75$). 5 Vynikající ($r > 0.75$)	3
2.10.1.5	Kvalita nástrojů jako kritérií nebo markerů: 0 Žádná informace neposkytnuta. 1 Poskytnuta neadekvátní informace. 3 Adekvátní kvalita. 4 Dobrá kvalita. 5 Vynikající kvalita s širokým rozsahem relevantních markerů pro konvergentní a divergentní validizaci.	4
2.10.1.6	Analýzy diferenciálního fungování položek (DIF): [N/A] Nepoužitelné 0-5 hodnocení kvality DIF analýzy	0
2.10.2	Validita vztahující se ke kritériu - celková adekvátnost (Tento celkový rating se získá na základě posouzení hodnot ratingů daných v položkách 2.11.1 - 2.10.2.4. Nepřůměrujte pouze čísla, abyste získali celkový rating.)	4
2.10.2.1	Popis použitých kritérií a charakteristik populací: (zatrhněte tolik, kolik je použitelných) <input checked="" type="checkbox"/> Souběžná <input type="checkbox"/> Prediktivní <input checked="" type="checkbox"/> Postdiktivní	
2.10.2.2	Velikosti výběrů: 0 Žádná informace neposkytnuta. 1 Jedna neadekvátní studie (např. velikost výběru menší než 100). 3 Jedna adekvátní studie (např. velikost výběru 100-200). 4 Jedna velká nebo více než jedna adekvátně rozsáhlá studie. 5 Dobrá série adekvátních až rozsáhlých studií.	4
2.10.2.3	Procedura výběru souboru: <i>(vyberte jednu)</i> <input type="checkbox"/> Žádná informace neposkytnuta <input checked="" type="checkbox"/> Účelná nebo reprezentativní <input type="checkbox"/> Nahodilá <input type="checkbox"/> Náhodná	
2.10.2.4	Medián a rozsah korelací mezi testem a kritérii: 0 Žádná informace neposkytnuta. 1 Neadekvátní (např. $r < 0.2$). 3 Adekvátní (např. $0.2 < r < 0.35$). 4 Dobrý (např. $0.35 < r < 0.50$). 5 Vynikající (např. $r > 0.50$)	5

2.10.3 Komentáře recenzenta týkající se validity:

V sekci Validita pracují autoři české adaptace s daty získanými z české standardizační studie, ale i s daty z holandských, německých i amerických studií. Předpokládáme, že společenské kontexty jsou si v tomto případě dost podobné.

Pro rating 2.10.2.4 je použito výhradně studií zahraničních, neboť autoři českého manuálu ve svých analýzách používají t-testy, u kterých neuvádějí velikost účinku, kterou bychom mohli nějakým způsobem srovnat s korelací. Autoři v manuálu uvádějí pouze hodnoty signifikance, z nichž nelze velikosti účinku vypočítat.

Důkazy o validitě můžeme rozdělit do několika skupin. Jako první uvádějí autoři důkazy pro kritériální validitu, a to prostřednictvím testování rozdílů v IQ mezi skupinami dětí s rodiči různého vzdělání. Autoři dále srovnávají IQ na základě bydliště dětí (město, nebo vesnice). Dalšími kritérii jsou úplnost rodiny, SES a pohlaví. Výsledky provedených t-testů jsou ve shodě s předpoklady.

Na konstruktovou, respektive faktorovou validitu autoři testu usuzují z analýzy hlavních komponent s ortogonální rotací Varimax. Tento postup není adekvátní. Za prvé, analýza hlavních komponent není faktorovou analýzou a nepracuje s výběrovou chybou, výsledky tedy mohou být zkreslené. Za druhé, použití ortogonální rotace nedává příliš smysl v případě, že jsou faktory zjevně korelované. Třetí, nejzásadnější problém spočívá v tom, že byla použita explorační metoda pro „konfirmasi“ určité faktorové struktury. Myslíme si, že použití konfirmační faktorové analýzy by pro dané účely bylo vhodnější, a že stávající analýza hlavních komponent je neadekvátně interpretována. Naše znalosti o fungování testu jsou z tohoto důvodu omezené. Argumentace pro zařazení konkrétních subtestů a položek je podložena pouze odvoláním se na fakt, že takovéto typy úloh jsou tradiční součástí IQ testů. Na druhou stranu se jedná o dlouho používaný a několikrát revidovaný test. Z kapitoly o reliabilitě můžeme tušit, že autoři se snaží o extrakci g-faktoru, takový model by ale spíše vyžadoval konstrukci bifaktorového řešení než volbu v manuálu popsanou.

Souběžná validita (jedná se o kapitolu plně převzatou ze zahraničního manuálu) je demonstrována na řadě zahraničních studií prostřednictvím korelací skóre s jinými intelektovými, vývojovými, řečovými a dalšími specifickými testy. Velikosti těchto korelací považujeme s přihlédnutím k typu použitých testů uspokojivou. Některé další studie mimo ty uvedené v manuálu (např. Moore, O'Keefe & Lawhon, 1998) hovoří o souběžné validitě s nástrojem WPPSI (Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition, 2012), která dosahuje pro celkové IQ hodnoty $r = 0,87$ (u populace zdravých dětí). Za pozitivní považujeme vysokou korelaci SON-IQ s inteligenčními testy i u postižených dětí (s WISC-R, $r = 0,62$, a s WPPSI-R, $r = 0,82$) – což implikuje fakt, že by SON-R 2½-7 mohl být dobrou alternativou v jejich případě.

Autoři manuálu tvrdí, že SON-R se odlišuje od konvenčních intelektových testů „jak po stránce obsahové, tak i způsobem administrace“, je tedy vhodné ho administrovat jako druhý test v případech, kdy je nutné učinit závažná rozhodnutí. Dle našeho názoru může SON-R skutečně sloužit jako vhodný diagnostický nástroj v situacích, kdy je evidentně výkon dítěte v jiném testu snižován například jeho úzkostností či komunikační nedostatečností. Čekali bychom tedy, že inkrementální validita testu bude v manuálu více komentována.

Reliabilita

2.11	Celková adekvátnost: <i>(Tento celkový rating se získá na základě posouzení hodnot ratingů daných v položkách 2.11.1–2.10.2.4. Neprůměrujte pouze čísla, abyste získali celkový rating.)</i>	3
2.11.1.	Poskytnutá data týkající se reliability: <i>(vyberte jednu možnost)</i> <input type="checkbox"/> Uveden pouze jeden koeficient reliability <input type="checkbox"/> Uveden pouze jeden odhad standardní chyby měření <input checked="" type="checkbox"/> Koeficienty reliability pro několik různých skupin <input checked="" type="checkbox"/> Standardní chyba měření uvedená pro několik různých skupin	
2.11.1	Vnitřní konzistence:	
2.11.1.1	Velikost výběru: 0 Neposkytnuta žádná informace. 1 Jedna neadekvátní studie (např. rozsah výběru menší než 100). 3 Jedna adekvátní studie (např. rozsah výběru 100-200). 4 Jedna rozsáhlá nebo více než jedna adekvátně rozsáhlá studie. 5 Dobrá série adekvátních až rozsáhlých studií. [N/A] Nepoužitelné.	3

2.11.1.2	Medián koeficientů: 0 Neposkytnuta žádná informace. 1 Neadekvátní (např. $r < 0.7$) 3 Adekvátní (např. $r = 0.7$ až 0.79) 4 Dobrý (např. $r = 0.8$ až 0.89) 5 Vynikající (např. $r > 0.9$) [N/A] Nepoužitelné.	3
2.11.2	<i>Testová-retestová stabilita:</i>	
2.11.2.1	Rozsah výběru: 0 Neposkytnuta žádná informace. 1 Jedna neadekvátní studie (např. rozsah výběru menší než 100). 3 Jedna adekvátní studie (např. rozsah výběru 100-200). 4 Jedna rozsáhlá nebo více než jedna adekvátně rozsáhlá studie. 5 Dobrá série adekvátních až rozsáhlých studií.	3
2.11.2.2	Medián koeficientů: 0 Neposkytnuta žádná informace. 1 Neadekvátní (např. $r < 0.6$) 3 Adekvátní (např. $r = 0.6$ až 0.69) 4 Dobrý (např. $r = 0.7$ až 0.79) 5 Vynikající (např. $r > 0.8$)	4,5
2.11.3	<i>Reliabilita jako ekvivalence:</i>	
2.11.3.1	Rozsah výběru: 0 Neposkytnuta žádná informace. 1 Jedna neadekvátní studie (např. rozsah výběru menší než 100). 3 Jedna adekvátní studie (např. rozsah výběru 100-200). 4 Jedna rozsáhlá nebo více než jedna adekvátně rozsáhlá studie. 5 Dobrá série adekvátních až rozsáhlých studií. [N/A] Nepoužitelné.	[n/a]
2.11.3.2	Medián koeficientů: 0 Neposkytnuta žádná informace. 1 Neadekvátní (např. $r < 0.6$) 3 Adekvátní (např. $r = 0.6$ až 0.69) 4 Dobrý (např. $r = 0.7$ až 0.79) 5 Vynikající (např. $r > 0.8$) [N/A] Nepoužitelné	[n/a]

2.11.4 Komentáře recenzentů k reliabilitě:

- Komentujte intervaly spolehlivosti pro koeficienty reliability
- Uveďte Spearmanovy-Brownovy ekvivalenty

V české a slovenské standardizaci byla reliabilita celkového skóru zjišťována na subtestových vážených skórech jako míra zobecnitelnosti (v manuálu i záznamovém archu označována jako "g"). To je v podstatě klasická Cronbachova α , počítaná na skórech subtestů namísto skóre položek. Domníváme se, že tento postup je motivován snahou autorů o očištění celkového skóru o specifický rozptyl subtestů. Pokoušejí se tak zjistit reliabilitu měření g-faktoru inteligence. Uvedený postup nicméně nepovažujeme za vhodný. Uživatel testu se z manuálu nedozví, jakým způsobem byl g index spočítán. Domníváme se, že měla být pro stejný účel spíše použita například McDonaldova hierarchická omega. Reportovaná reliabilita celkového skóru v testu tak může být (a pravděpodobně je) podhodnocena z důvodu nedodržení poměrně přísných předpokladů na Cronbachovu α , například tau-ekvivalence subtestů a unidimenzionality škál. Kromě indexů zobecnitelnosti nacházíme v manuálu i vnitřní konzistence škály performační a škály úsudkové (počítány nad položkami příslušných subtestů), stejně jako celkového IQ skóru (taktéž počítáno nad všemi položkami). Průměrná Cronbachova α pro Performační škálu je 0,88, pro Škálu usuzování je 0,86 a pro celkové SON-IQ je 0,92. Hodnoty pro reliabilitu celkového skóru jsou velmi vysoké. Ukazatele vnitřní konzistence nicméně nejsou pro další výpočty v rámci norem a standardních chyb odhadu směrodatné.

Autoři manuálu uvádějí dva různé typy chyb. Jednak standardní chybu měření (počítanou v jednotkách vážených skóre), jednak standardní chybu odhadu, pro jejíž výpočet je použit právě zmíněný koeficient zobecnitelnosti g. Interval spolehlivosti v normách jsou počítány právě s využitím standardní chyby odhadu. Hodnota g se ve věkových skupinách pohybuje mezi 0,78 až 0,81,

což je pro hierarchický koeficient v pořádku. Oceňujeme fakt, že uživatel má v tabulkách norem k dispozici odhady 80procentních intervalů spolehlivosti.

Součástí ověření souběžné validity je i korelace při retestu, kterou bychom uvedli spíše v rámci posouzení reliability. Ta dosahuje hodnoty $r = 0,79$. Vezmeme-li v úvahu fakt, že každá nesprávná odpověď je korigována, což významně přispívá k efektu zácvičku, zdá se tato hodnota být menší, než bychom očekávali. Na druhou stranu nevíme, po jaké době se opakované testování odehrálo, a tedy nelze výsledky jednoznačně interpretovat. V úvahu je třeba vzít i poměrně rychle se zvyšující se reliabilitu IQ v dětství (Schuerger & Witt, 1989).

Část 9:

Závěrečné hodnocení:

3.0 Hodnotící zpráva testu:

Tato část by měla obsahovat stručné, jasně obhájené posouzení nástroje/produktu. Mělo by popisovat jeho pro a proti a poskytnout určitá obecná doporučení týkající se toho, jak a kdy by se měl používat – spolu s varováními (kde jsou potřebná) týkajícími se případů, kde by se používat neměl.

Při práci se SON-R 2½ - 7 uživatel musí mít na paměti určité nedostatky, kterými test trpí a které jsme již zmínili. Jde především o nedostatečnost norem v nejmladší a nejstarší věkové skupině způsobenou efektem podlahy a stropu, nedostatečné ověření konstruktové a faktorové validity a poněkud kuriózní přístup k odhadu reliability celkového skóru, která zamlžuje skutečnou přesnost měření celého testu. Poněkud matoucí je také uspořádání manuálu, který mísí poznatky z české standardizační studie a ze studie německé či holandské, což může uživateli znesnadnit práci.

Žádný z těchto problémů však není tak závažný, abychom nemohli test doporučit k užívání, nebo jeho užívání jinak omezovali. Z vícera zdrojů je zřejmé, že SON-R 2½ - 7 intelekt měří (např. Jenkinson, Roberts, Dennehy & Tellegen, 1996; Moore, O'Keefe & Lawhon, 1998). Nicméně je na místě, aby se autoři české adaptace snažili uživateli testu podat vysvětlení, proč se tomu tak děje a v návaznosti na to dokázali vytvořit smysluplný strukturní model metody. Jak by takový model mohl vypadat, jsme nastínili již v kapitole Validita. I když se může zdát, že tato vylepšení nebudou mít dopad na využití metody v praxi, není tomu tak. Domníváme se, že lepší uchopení dat může vést k přesnějším odhadům reliability i lepšímu pochopení toho, v čem dané dítě exceluje nebo naopak zaostává, což může praktickému využití značně pomoci. Bez tohoto totiž máme pouze šest subtestů, které v nevalně provedené explorativní faktorové analýze vytvoří dva faktory po třech subtestech.

Z pohledu reliability je metoda adekvátní. Autoři u odhadu reliability neskončili pouze u vnitřní konzistence, ale snažili se uživatelům testu poskytnout index zobecnitelnosti (g). Pokus o výpočet hierarchického indexu je při měření g-faktoru inteligence chvályhodný, způsob jeho výpočtu ale není na základě informací z manuálu pochopitelný. Pokud by byl index odhadnut neadekvátním způsobem, mohla by být standardní chyba odhadu zkreslena oběma směry, a v tom důsledku i šíře 80procentního intervalu spolehlivosti, který je autory doporučován jako hlavní interpretační nástroj.

O nedostatečnosti konstruktové validity se zmiňujeme v odstavci výše. Další důkazy o validitě nám autoři poskytují skrze validitu kriteriální a souběžnou. Důkazy o kriteriální validitě jsou uspokojivé, i přes některá, hlavně statistická, pochybení v rámci analýz. Kapitulu o souběžné validitě potom tvoří překlad zahraničního manuálu a v rámci něho provedených analýz. Předpokládáme, že tento způsob lze považovat za dostatečný s přihlédnutím k neverbální

povaze testu a faktu, že děti z Německa, Nizozemí i USA žijí v podobné kultuře, používají stejné hračky, setkávají se stejnými podněty apod.

Vyzdvihnout musíme velice kvalitní zpracování testového materiálu, který je pro děti přitažlivý. Podobně tak i způsob administrace, který zohledňuje to, že se pracuje s opravdu mladými dětmi, které potřebují motivaci, aby vydržely pracovat. Samotný manuál je až na pár překlepů kvalitně zpracován. Vyplnění záznamového archu je otázkou několika minut a není třeba vykonávat speciální výpočetní či jiné úkony, neboť autoři vše potřebné poskytují na dané dvojstraně v sekci Normy.

3.1 Závěry:

Dle našeho názoru je metoda použitelná v oblasti klinické, školní a poradenské psychologie. Omezená je pouze svou menší diferenciační schopností na obou koncích věkového rozložení. Kromě PPP a dalších specializovaných center najde využití i ve výzkumu (např. Udholm, Aaberg, Bloch, Sandahl & Ovesen, 2017; Kiechl-Kohlendorfer, Ralser, Peglow, Pehboeck-Walser & Fussenegger, 2013). Vhodné je SON-R 2½ – 7 používat jako potvrzení výsledků předešlých metod či v situacích, kdy nelze administrovat běžnější inteligenční test. Metodu považujeme za kvalitní a je podle nás škoda, že česká adaptace nedosahuje kvalit odpovídajících testu samotnému, čímž jeho interpretační možnosti spíše narušuje

4.0 Doporučení (vyberte jedno)

Všechny následující charakteristiky uvedené níže by měly mít ratingy [n/a], [2], [4], [5], pokud by měl být nástroj „doporučen“ pro obecné použití (hodnocení 5 nebo 6):

[2.9] Normy a referenční skupiny
[2.10.1] Konstruktová validita
[2.10.2] Kriteriační validita
[2.11] Reliabilita - celková
[2.12] Počítačově generované zprávy

Pokud má kterýkoli z výše uvedených ratingů hodnotu [] nebo [1], nástroj by měl být klasifikován pod doporučením 1, 2, 3 nebo 4, nebo klasifikován pod doporučením 7 "jiné" s adekvátním vysvětlením.

- 1 Pouze výzkumný nástroj. Ne pro užití v praxi.
- 2 Vhodný pouze pro užití expertním uživatelem za pečlivě kontrolovaných podmínek nebo ve velmi omezených aplikačních oblastech
- 3 Vhodný pro použití pod supervizí v aplikační(ch) oblasti(ech) definovaných distributorem, libovolnými uživateli s obecnými kompetencemi pro používání a administraci testů
- 4 Vyžaduje další vývoj. Vhodný pouze pro použití ve výzkumu.
- 5 Vhodný pro používání v aplikační(ch) oblasti(ech) definovaných distributorem, uživateli testů, kteří splňují speciální kvalifikační požadavky distributora
- 6 Vhodný pro sebevyšetření bez supervise v aplikační(ch) oblasti(ech) definovaných distributorem
- 7 Jiné:

5 Odkazy k poznámkám a bibliografie

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor- analytic studies*. New York: Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth and action*. Boston, MD: Houghton Mifflin.

Hogrefe. (n.d.). SON-R 2½ – 7. Staženo z: <http://www.testcentrum.com/testy/son%E2%80%937-%E2%80%937>

- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 57, 253–270.
- Hülür, G., Wilhelm, O., & Robitzsch, A. (2011). Intelligence differentiation in early childhood. *Journal Of Individual Differences*, 32(3), 170-179. doi:10.1027/1614-0001/a000049
- Jenkinson, J., Roberts, S., Dennehy, S., & Tellegen, P. (1996). Validation of the Snijders-Oomen Nonverbal Intelligence Test—Revised 21/2–7 for Australian children with disabilities. *Journal Of Psychoeducational Assessment*, 14(3), 276-286. doi:10.1177/073428299601400307
- Kiechl-Kohlendorfer, U., Ralser, E., Peglow, U. P., Pehboeck-Walser, N., & Fussenegger, B. (2013). Early risk predictors for impaired numerical skills in 5-year-old children born before 32 weeks of gestation. *Acta Paediatrica*, 102(1), 66-71. doi:10.1111/apa.12036
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37(1), 1-10. doi:10.1016/j.intell.2008.08.004
- Moore, C., O'Keefe, S. L., Lawhon, D., & Tellegen, P. (1998). Concurrent validity of the Snijders-Oomen Nonverbal Intelligence Test 2 ½-7—Revised with the Wechsler Preschool and Primary Scale of Intelligence—Revised. *Psychological Reports*, 82(2), 619-625. doi:10.2466/pr0.1998.82.2.619
- Schuerger, J., & Witt, A. (1989). The temporal stability of individually tested intelligence. *Journal of Clinical Psychology*, 45(2), 294-302.
- Tellegen, P.J., Laros, J.A., & Heider, D. (2008). *Neverbální inteligenční test SON-R 2½-7 - uživatelská příručka*. Praha: Testcentrum – Hogrefe.
- Tellegen, P.J., Winkel, M., Wijnberg-Williams, B.J. & Laros, J.A. (1998). *Snijders-Oomen Nonverbal Intelligence Test. SON-R 21/2-7 Manual and Research Report*. Lisse: Swets & Zeitlinger B.V.
- Udholm, N., Aaberg, K., Bloch, C., Sandahl, M., & Ovesen, T. (2017). Cognitive and outcome measures seem suboptimal in children with cochlear implants - a cross-sectional study. *Clinical Otolaryngology*, 42(2), 315-321. doi:10.1111/coa.12723
- Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition. Technical and interpretive manual*. San Antonio, TX: NCS Pearson, Inc.

Měřené konstrukty:

Fluidní inteligence, abstraktní myšlení a konkrétní myšlení (obojí schopnost usuzování), prostorový vhled