

**VERBÁLNÍ FLUENCE U DĚTÍ VE VĚKU 5–12 LET: ČESKÉ NORMY
A VYBRANÉ PSYCHOMETRICKÉ UKAZATELE¹**

Hynek Cígler, Světlana Durmeková

Abstrakt

Test verbální fluence je často využívanou zkouškou nejen v pedagogicko-psychologické praxi. Předložená metodická studie obsahuje základní psychometrické ukazatele, zejména odhady reliabilit, a hlavně pak normy pro počty slov vyjmenovaných za jednu minutu začínajících na hlásky N, K a P (a celkem) na vzorku 411 dětí ve věku 5–12 let.

Klíčová slova: verbální fluence; psychomotorické tempo; reliabilita; standardizace

***VERBAL FLUENCY IN CHILDREN AGED 5 TO 12: CZECH NORMS
AND PSYCHOMETRIC PROPERTIES***

Abstract

The Verbal Fluency Test is the common task not only in the school psychological services. This methodological study provides the main psychometrical parameters of the test, particularly the reliability estimates, and norms for the number of words named within one minute beginning on the sounds N, K and P (and total) on the sample of 411 pupils aged 5–12 years.

Keywords: *verbal fluency; psychomotor speed; reliability; standardization*

Došlo: 11. 6. 2018

Schváleno: 3. 12. 2018

¹ Příprava publikace byla podpořena grantem GAČR GA17-09797S.

Úvod

Ačkoliv² test verbální fluence není explicitně uveden v Urbánkové (2012) přehledu nejpoužívanějších českých diagnostických metod, jedná se o velmi často používanou zkoušku jak v pedagogicko-psychologické diagnostice, tak i v klinické neuropsychologii. V českém prostředí jsou k dispozici např. normy pro starší osoby nad 59 let (Nikolai a kol., 2015), nespécifický dospělý vzorek ve věku 18–60 let (Janásová, 2015), adolescenty z druhého ročníku středních škol (Cimlerová, Pokorná a Chalupová, 2007) či pražské děti ve věku 9–14 let a dospělé 20–85 let (Preiss a kol., 2012). Aktuální normy pro děti z celé české populace však nebyly dosud dostupné.

Testy verbální fluence jsou založeny na principu vyjmenování co největšího počtu slov splňujících předem dané kritérium za omezenou časovou jednotku, zpravidla 60 s. Nejčastěji jsou používány tzv. sémantická slovní fluence, kde má proband jmenovat slova dané lingvistické kategorie (např. zvířata, dopravní prostředky atp.) a fonemická slovní fluence, kde slova musí začínat vždy na určitou hlásku (např. P, N atp.). Méně často se užívá syntaktické slovní fluence, kdy slova musí splňovat např. předepsaný počet slabik.

Verbální fluence je používána zejména u dospělých a starších osob pro diagnostiku mírné kognitivní poruchy (Green a kol., 1999) či demence (Levy a kol., 2002), a to i pro účely diferenciální diagnostiky (Rinehardt, 2014).

Využití u dětí je méně časté. Metoda by podle některých studií měla být užitečná pro identifikaci dyslexie (Landerl, Fussenegger, Moll, a Willburger, 2009), ADHD (Frazier, Demaree a Youngstrom, 2004), diferenciální diagnostiku dyslexie oproti vizuo-prostorovým obtížím či ADHD, kde by děti s dyslexií měly dosahovat horších skóre (Cohen, Morgan, Vaughn, Riccio a Hall, 1999), a jiné. Přestože pozorované efekty jsou spíše slabší, vzhledem k extrémně krátkému času zkoušky může být test verbální fluence použit pro rychlý screening, který je v případě potřeby následně doplněn podrobnějším vyšetřením.

Rovněž v českém prostředí proběhlo několik studií zahrnujících testy verbální fluence; souhrnné informace o fungování testu pro české čtenáře předložil zejména Marek Preiss (1996a, 1996b, cit. dle Preiss, 1997; 1997; Preiss a kol., 2012). Mezera (1988; cit. podle Preiss, 1997) použil hlásky CDFHKNRT s limitem jedné minuty na hlásku. Poměrně nečekaně přitom nepozoroval žádný vztah s inteligencí a kromě hlásek K a T ani rozdíl mezi dětmi s „LMD“ (lehkou mozkovou dysfunkcí) a zdravým souborem. Štorková, Preiss a Kopeček (2004) ověřovali reliabilitu paralelních forem (hlásky KNP vs. VRS) při dvou měřeních s odstupem jednoho týdne na vzorku 20 zdravých dospělých respondentů; reliabilita typu test-retest bohužel není reportována. Reliabilita paralelních forem (odhadnutá pomocí Spearmanovy pořadové korelace) byla $\rho = 0,63$ při prvním a $\rho = 0,83$ při druhém měření.³ Při retestu též došlo ke statisticky významnému zlepšení o 13 %; souhrnný skóre pro hlásky KNP byl statisticky významně vyšší než u hlásek VRS, průměrně o 5,2 slov. Velmi podobné výsledky s celou řadou hlásek předkládají i Kopeček a Kuncová (2006) a navrhují alternativní verze k nejvíce užívané trojici hlásek NKP. Konečně pak Preiss a kol. (2012) předkládají normy na pražském vzorku dětí ve věku 9–14 let pro hlásky KT a dále pak normy pro dospělé (20–85) let s hláskami NKP.

² Na tomto místě chceme poděkovat anonymnímu recenzentovi tohoto textu, který nás (1.) upozornil na chybějící české zdroje, a (2.) uvedl na pravou míru naši chybnou domněnku o příčině rozdílů průměrného výkonu dětí v naší a Preissové a kol. (2012) studii. Tyto zásahy podstatným způsobem zvýšily výslednou kvalitu předloženého článku.

³ Podle našich vlastních výpočtů se tyto odhady statisticky významně nelišily, $z(20) = 1,3$, $p = 0,19$, a průměrná hodnota (po Fisherově transformaci a zpětné transformaci) je $\rho_M = 0,75$.

Primárním cílem této studie není hodnotit kritériální validitu testu, ale poskytnout kvalitní české normy těm uživatelům, kteří již test v praxi využívají. Zvolili jsme test fonemické verbální fluence a pracovali jsme s počty slov začínajících postupně na hlásku N, K a P vyřčených za jednu minutu – a to jak zvlášť, tak i pro všechny hlásky dohromady. Dále se pokusíme odhadnout reliabilitu metody, a to rovněž jak pro celkový skór, tak i pro jednotlivé dílčí zkoušky. Závěrem pak ověříme souvislost získaných výsledků s demografickými proměnnými.

Nad rámec této studie dáváme čtenářům k dispozici rovněž online supplement (viz odkaz na konci studie) v podobě frekvenční analýzy nejčastějších slov generovaných během testu podle úvodních hlásek, a to jak pro celý výzkumný soubor dohromady, tak i pro jednotlivé věkové kategorie. Ačkoliv se tento materiál nabízí pro kvalitativní analýzu výsledků testu, nejsme si vědomi žádných empirických zdrojů, které by takovéto využití podporovaly. Přesto věříme, že může jít o zajímavý soubor informací.

Metody

Výzkumný vzorek a sběr dat

Sběr dat proběhl u dětí pěti až dvanáctiletých v rámci standardizační studie k testu Vytváření příběhů (Cígler a Durmeková, 2017), kde je popsán podrobněji. Každému dítěti, kterému byl zadán test Vytváření příběhů, byl rovněž bezprostředně poté administrován test verbální fluence. Výkon v testu verbální fluence tak mohl být ovlivněn z hlediska sémantických významů jednotlivých vyřčených slov, ale nepatrně (pozitivně) možná i co do jejich počtu, jak podnětovým materiálem testu Vytváření příběhů, tak vlastními příběhy, které dítě vytvořilo. Data v případě předškolních dětí sbírali převážně psychologové z pedagogicko-psychologických poraden, kteří oslovovali rodiče dětí zpravidla s pomocí učitelů, a to včetně učitelů přípravných tříd základních škol. V případě dětí v základním vzdělávání většinu dat zajistili školní psychologové. Samotný sběr probíhal mezi 2. listopadem 2015 a 13. červnem 2016. Studie se zúčastnilo celkem 17 různých administrátorů ze 12 krajů České republiky, rodiče všech dětí podepsali informovaný souhlas. Data byla sbírána kvótním způsobem se zohledněním velikosti sídla a vzděláním rodičů. Všechny děti byly české národnosti a jejich mateřským jazykem byl český jazyk.

Celkem se studie zúčastnilo 420 dětí, 9 z nich však bylo vyřazeno z důvodu nevalidního protokolu. Pracovali jsme tedy s výsledným vzorkem $N = 411$, z toho bylo 200 chlapců (49 %) a 211 dívek (51 %). Rozložení podle věku obsahuje tabulka 1. Vzorek se rovněž lišil od rozložení populace dle velikosti sídla, $\chi^2(4) = 246$, $p < 0,001$, a to s vysokou velikostí účinku, Cramerovo $V = 0,40$; ve vzorku byla převaha dětí z malých měst a málo respondentů z menších obcí, jak je patrné z tabulky 2.

Tabulka 1: Počet respondentů podle věku a pohlaví

	věk (v celých dosažených letech)							
	5	6	7	8	9	10	11	12
dívky	26	28	31	22	26	27	26	25
chlapci	27	23	26	25	23	27	25	24

Tabulka 2: Počet respondentů dle velikosti sídla

	N (rel.)	v populaci
do 2 000	42 (10 %)	27 %
2 001 – 15 000	233 (57 %)	25 %
15 001 – 80 000	88 (21 %)	22 %
80 001 – 1 000 000	33 (8 %)	14 %
přes 1 000 000	15 (4 %)	12 %

Test verbální fluence

Administrován byl standardní test fonemické verbální fluence. Úkolem probanda bylo vyjmenovat za 1 minutu co nejvíce slov začínajících na hlásky N, K a P (v tomto pořadí). Výhodiskem pro stávající adaptaci testu byla instrukce uvedená v příručce k testové baterii Diagnostika specifických poruch učení u adolescentů a dospělých osob (Cimlerová, Pokorná a Chalupová, 2007). Tuto administraci jsme jen mírně upravili s ohledem na nižší věk probandů našeho souboru. Výsledná instrukce zněla:

„Řeknu Ti určitou hlásku (písmeno) a Ty budeš mít jednu minutu na to, abys vymyslel/a co nejvíce slov, která na danou hlásku začínají (např. „B“ – bagr, batoh, budík). Nepoužívej slova začínající velkými písmeny (například Barbora, Bohdan, Brno) a slova příbuzná (bydlení, bydlet) – dovysvětlit, pokud tomu dítě nerozumí.“

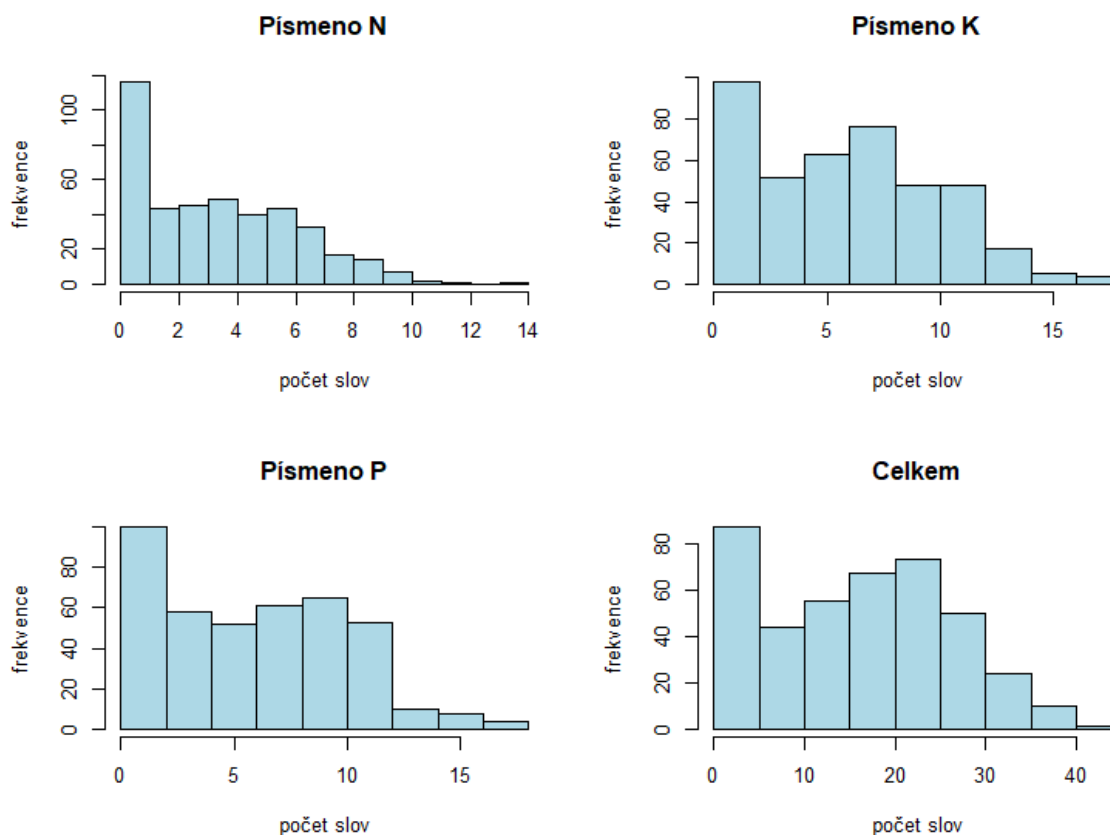
Statistická analýza dat

Veškeré analýzy byly provedeny v prostředí R s příslušnými balíčky, zejm. *psych* (Revelle, 2017). Dále byly využity knihovny *ks* (Duong, 2016) pro vyhlazení percentilových norem, *ppcor* (Kim, 2015) pro výpočet parciálních korelací a *sjstats* (Lüdecke, 2018) pro odhad velikostí efektů lineárních modelů. Pro reportování souhrnných výsledků lineárních modelů byla použita výhradně suma čtverců typu III, odhadnutá v balíčku *car* (Fox a Weisberg, 2011).

Výsledky

Deskriptivní statistiky

Bez zohlednění věku dětí byly všechny skóry silně zprava zešikmené, jak ukazuje obrázek 1. Zajímavý je vysoký počet dětí, které selhaly při generování slov začínajících na hlásku N. Toto selhání patrně souviselo s jejich delší adaptací na zadanou zkoušku – hlásku N totiž byla v rámci testu verbální fluence zadávána jako první. Základní deskriptivní statistiky výkonu v jednotlivých subtestech i dohromady (průměry a směrodatné odchylky) jsou níže v tabulce 5.



Obrázek 1

Rozložení počtů slov bez zohlednění věku dítěte

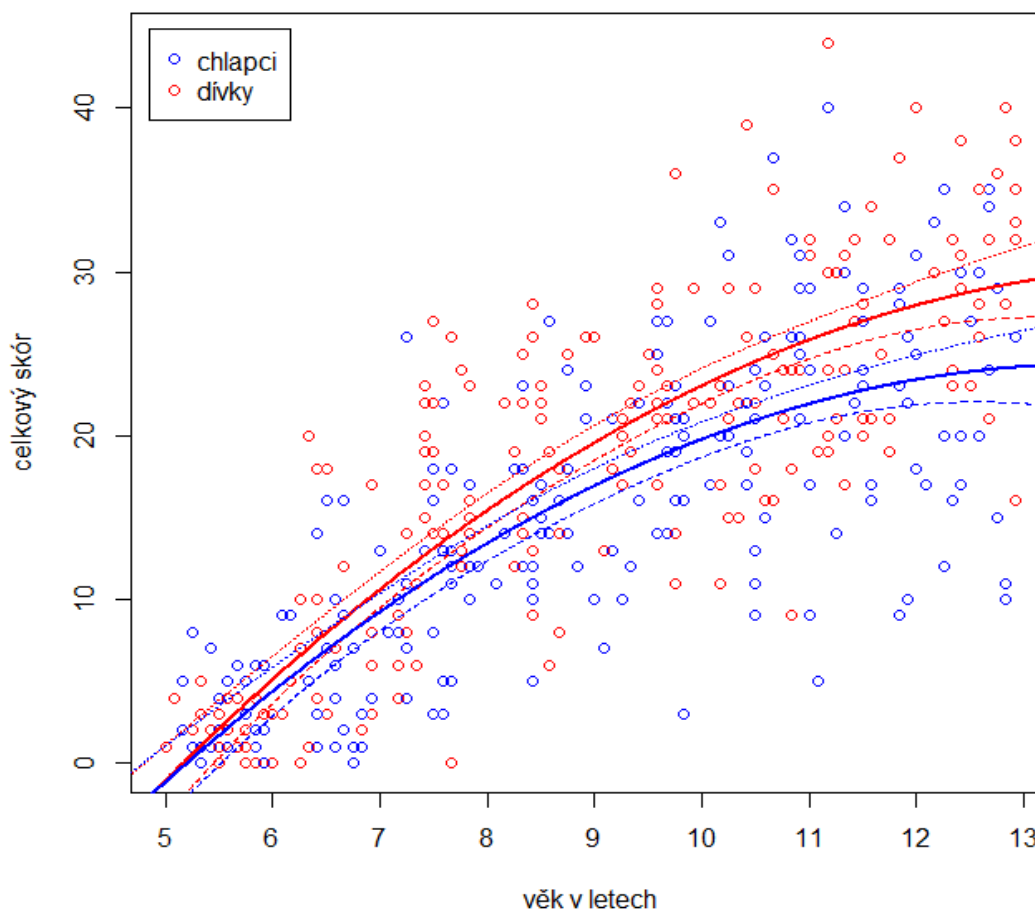
Tabulka 3 obsahuje výsledky lineárního modelu s různými prediktory v podobě demografických proměnných a celkovým počtem slov ve všech třech subtestech jako závislou proměnnou. Z důvodu vysoké kolinearit jsme použili věk dítěte (v měsících, centrovaný), nikoliv už školní ročník (protože normy budou konstruovány právě podle věku). Přidání interakce věku dítěte a pohlaví zvýšilo rozptyl vysvětlený modelem, $\Delta F(1, 386) = 6,14$, $p = 0,013$; následné přidání interakce čtverce věku a pohlaví již nikoliv, $\Delta F(1, 385) = 0,02$, $p = 0,868$. Reportujeme tedy pouze druhý model s prostou interakcí věku a pohlaví. Je nicméně nutné počítat s podhodnocením skutečných velikostí efektu a drobným zkreslením modelu z důvodu heteroskedacity dat, Breusch-Paganův test $\chi^2(1) = 20,14$, $p < 0,001$.

Nejsilnějším prediktorem byl věk dítěte, lineární efekt $\eta_p^2 = 0,495$, kvadratický pak $\eta_p^2 = 0,036$. Děti vzdělanějších matek dosahovaly vyšších skóre, $\eta_p^2 = 0,068$, vzdělání otce však nemělo vliv (velikost efektu navíc byla slabší, $\eta_p^2 = 0,023$). Signifikantním prediktorem byla i velikost sídla, $\eta_p^2 = 0,031$. Děti z nejmenších obcí do 2 000 obyvatel společně s dětmi z Prahy měly oproti ostatním dětem skóre vyšší o 3,8 bodu, $p = 0,002$, $\beta = 0,11$; ostatní rozdíly již statisticky významné nebyly. Chlapci pak dosahovali v průměru o 2,4 bodu horšího skóre než dívky a vliv pohlaví narůstal s věkem. U pětiletých dětí rozdíl nebyl pozorovaný, v kohortě nejstarších 12letých dětí však dívky vyprodukovaly v průměru o 6,4 slov více než chlapci – viz obrázek 2.

Tabulka 3: Vztah celkového skóre a demografických charakteristik

	df	F	p	ω^2	η_p^2
věk	1	377,742	<0,001	0,432	0,495
věk ²	1	14,238	<0,001	0,015	0,036
pohlaví	1	16,813	<0,001	0,018	0,042
vzdělání matky	7	4,019	<0,001	0,024	0,068
vzdělání otce	9	0,992	0,446	0,000	0,023
velikost sídla	4	3,068	0,017	0,009	0,031
věk*pohlaví	1	6,144	0,014	0,006	0,016

Pozn.: Celý model byl signifikantní $F(24, 386) = 37,8$, $p < 0,001$, a vysvětlil většinu pozorovaného rozptylu testu, $\text{adj. } R^2 = 0,683$. Věk dítěte v měsících byl centrováný (odečtením průměrného věku dětí $M = 107,53$).



Obrázek 2

Vztah věku a celkového skóre podle pohlaví včetně 95% intervalu spolehlivosti

Konvergentní a kriteriální validita

Celkový skór souvisel se školními známkami: parciální Spearmanova korelace po kontrole školního ročníku s hodnocením z matematiky byla $\rho_p = -0,271$, $p < 0,001$, a z českého jazyka $\rho_p = -0,202$, $p < 0,001$.

Vztah s inteligencí měřenou testem WISC-III a kreativitou měřenou testem Vytváření příběhů byl ověřený jinde (Cígler a Durmeková, 2017), výsledky jsou v diskuzi.

Nakonec jsme ověřovali souvislost s mimořádným intelektovým nadáním ($n = 5$) a různými druhy znevýhodnění (celkové $n = 55$).⁴ Do modelu jsme zařadili diagnózy, které se vyskytly alespoň u dvou probandů. Přidání těchto proměnných do dříve prezentovaného lineárního modelu (k věku, demografickým proměnným atp.) zvýšilo množství vysvětleného rozptylu o jedno procento, $\Delta R^2_{adj.} = 0,013$, $\Delta F(8, 369) = 2,88$, $p = 0,004$. Signifikantními prediktory byly mimořádné nadání, dyskalkulie, ADD a poruchy autistického spektra (kde děti s autismem skórovaly lépe než běžné děti). Efekty však byly spíše slabé, podrobné výsledky obsahuje tabulka 4, a zároveň všechny testy měly velmi malou sílu, jak je patrné ze šířky konfidenčních intervalů pro velikost efektu.

Jednotlivé subtesty (počty slov na jednotlivé hlásky) spolu středně silně korelovaly. Parciální Pearsonovy korelace počtů slov začínajících na jednotlivé hlásky po kontrole věku v měsících byly: N a K $r_p = 0,44$, N a P $r_p = 0,49$, K a P $r_p = 0,61$ (všechna $p < 0,001$).

Tabulka 4: Kriteriální validita

diagnóza	N	B [CI _{95%}]	df	F	p	ω^2	η_p^2
mimořádné nadání	5	6,58 [1,1; 12,0]	1	5,669	0,018	0,005	0,015
dyslexie	20	0,59 [-2,7; 3,9]	1	0,125	0,724	-0,001	0,000
dysortografie	13	1,02 [-3,1; 5,1]	1	0,238	0,626	-0,001	0,001
dysgrafie	13	-0,93 [-5,3; 3,4]	1	0,177	0,674	-0,001	0,000
dyskalkulie	6	-7,68 [-14,2; -1,2]	1	5,367	0,021	0,005	0,014
ADHD	14	1,47 [-1,8; 4,7]	1	0,793	0,374	0,000	0,002
ADD	11	-5,40 [-9,0; -1,8]	1	8,472	0,004	0,009	0,022
PAS	2	11,88 [1,4; 22,3]	1	5,023	0,026	0,005	0,013

Pozn.: Proměnné přidány do modelu z tab. 3. $\Delta R^2_{adj.} = 0,010$, $\Delta F(8, 347) = 2,35$, $p = 0,018$; $n = 402$.⁴

Reliabilita

Reliabilita ve smyslu vnitřní konzistence byla pro celkový počet slov ve všech třech subtestech odhadnuta pro každou věkovou kohortu zvlášť prostřednictvím McDonaldovy omegy, která na rozdíl od běžnější Cronbachovy alfy nepředpokládá tau-ekvivalenci položek (tedy shodné faktorové náboje) a poskytuje tak méně zkreslený odhad (Marko, 2016). Zároveň jsme odhadli reliabilitu počtu slov začínajících na každou hlásku zvlášť, a to pomocí komunality každého subtestu (tedy standardizovaného podílu rozptylu v daném subtestu vysvětleného příslušným faktorem) ze stejného faktorového modelu, který byl použit pro odhad koeficientu omega.

⁴ Vztah se specifickými poruchami učení (SPU) a mimořádným nadáním byl ověřován na zmenšeném vzorku $N = 402$ žáků, protože u 9 dětí chyběly informace o výskytu SPU.

Faktorový model použitý pro všechny tyto výpočty byl odhadnut funkcí omega z psych balíčku (Revelle, 2017) s extrakcí metodou minimálních reziduí (minres). Reliabilita celého vzorku byla velmi dobrá, pro celkový skór $\omega = 0,91$, komunality subtestů $\omega_N = 0,65$, $\omega_K = 0,78$, $\omega_P = 0,86$, čemuž by odpovídaly standardní chyby měření $SE_{sum} = 3,16$, $SE_N = 1,67$, $SE_K = 1,94$ a $SE_P = 1,59$. Výsledky včetně základních deskriptiv pro jednotlivé kohorty pak jsou v tabulce 5.

Tabulka 5: Reliabilita, standardní chyba měření a deskriptivní statistiky podle kohorty

věk	deskriptivy								reliabilita				standardní chyba měření (SE)				
	N		K		P		SUM		N	K	P	SUM	N	K	P	SUM	
5	53	0,5	(0,61)	0,9	(1)	1,0	(1,06)	2,4	(2)	0,34	0,41	0,27	0,61	0,38	0,63	0,66	1,26
6	51	1,6	(1,58)	2,4	(2,42)	2,0	(2,07)	6,0	(5,43)	0,62	0,72	0,73	0,87	0,57	0,87	0,74	1,95
7	57	3,1	(2,02)	5,2	(2,83)	5,3	(3,01)	13,7	(6,6)	0,49	0,54	0,63	0,79	0,94	1,31	1,39	3,06
8	47	3,4	(1,97)	7,1	(2,43)	6,2	(2,77)	16,7	(5,98)	0,39	0,55	0,68	0,78	0,93	1,15	1,31	2,83
9	49	4,4	(2,11)	7,4	(3,06)	7,9	(3,04)	19,8	(6,31)	0,12	0,32	0,91	0,68	1,20	1,73	1,72	3,57
10	54	5,1	(2,33)	8,5	(2,8)	8,7	(2,97)	22,3	(6,73)	0,34	0,43	0,91	0,79	1,08	1,30	1,38	3,12
11	51	5,8	(2,66)	8,9	(3,56)	9,7	(3,11)	24,4	(7,83)	0,43	0,54	0,70	0,79	1,22	1,64	1,43	3,60
12	49	6,5	(2,53)	10,2	(3,55)	10,2	(3,52)	26,8	(7,74)	0,25	0,69	0,51	0,73	1,32	1,85	1,83	4,03
<i>Md</i>										0,37	0,54	0,69	0,78	1,01	1,30	1,39	3,09

Pozn.: Reliabilita celkového skóre byla odhadnutá prostřednictvím McDonaldovy ω ; pro dílčí skóre v produkci slov na jednotlivá písmena pak pomocí komunality daného subtestu na celkovém faktoru. *Md* – medián.

Srovnání subtestů a konstruktová validita

Jednotlivé subtesty, tedy počty slov začínajících na hlásky N, K a P vyjmenované dítětem za jednu minutu jsme vzájemně srovnali s využitím strukturního modelování.

V první řadě jsme ověřili model s latentní proměnnou „fluence“, která sytila všechny tři subtesty. Fluenci predikoval věk v měsících (centrovaný) a jeho druhá mocnina (vzájemně nekorelované); intercepty těchto manifestních exogenních proměnných byly volně odhadovány. Tento model popsal data perfektně, $\chi^2(5) = 5,75$, $p = 0,331$, $TLI = 0,999$, $SRMR = 0,024$, $RMSEA = 0,021$ s $CI_{95\%} = [0,000; 0,081]$, a nelišil se tedy statisticky významně od dat. Věk tak působí shodným způsobem na počty vygenerovaných slov začínajících na jednotlivé hlásky skrze latentní proměnnou (verbální fluenci) a naopak se neliší v míře svého vlivu na produkci slov začínajících na různé hlásky.

Zafixování interceptů položek (subtestů) výrazně zhoršilo model, $\Delta\chi^2(2) = 360,2$, $p < 0,001$, a vedlo ke zcela neadekvátní shodě modelu s daty, $\chi^2(7) = 330,2$, $p < 0,001$, $TLI = 0,557$, $SRMR = 0,203$, $RMSEA = 0,362$ s $CI_{95\%} = [0,329; 0,396]$. Generovat slova začínající na hlásky N/K/P je tak rozdílně obtížné, viz tab. 6. Rovněž zafixování faktorových nábojů vedlo k výraznému zhoršení modelu, $\Delta\chi^2(2) = 65,9$, a neshodě s daty, $\chi^2(7) = 152,6$, $p < 0,001$, $TLI = 0,802$, $SRMR = 0,242$, $RMSEA = 0,242$ s $CI_{95\%} = [0,210; 0,276]$. Produkce slov začínajících na jednotlivé hlásky tak různým způsobem souvisí s celkovou verbální fluencí a vyjmenované počty slov v jednotlivých subtestech proto nejsou na stejné „škále“. Reálný efekt je spíše slabý, rozdíly ve standardizovaných faktorových nábojích byly spíše malé.

Hodnoty všech volně odhadnutých parametrů finálního modelu jsou prezentovány v tabulce 6. Při porovnání intervalů spolehlivosti je patrné, že slov začínajících na N řekli respondenti méně než slov začínajících na K a P; tento subtest rovněž slaběji souvisel s latentním rysem.

Tabulka 6: Parametry strukturního modelu (celý vzorek)

	faktorové náboje (regrese)			intercepty		reziduální rozptyl		
	λ	(CI _{95%})	std.	v	(CI _{95%})	ψ	(CI _{95%})	std.
N	1,00		0,814	4,18	(3,91; 4,45)	2,69	(2,17; 3,21)	0,34
K	1,60	(1,47; 1,73)	0,888	6,94	(6,56; 7,32)	3,64	(2,85; 4,42)	0,21
P	1,69	(1,56; 1,83)	0,919	7,02	(6,62; 7,42)	2,79	(2,08; 3,51)	0,16
věk	0,07	(0,06; 0,08)	0,813					
věk ²	-0,06	(-0,08; -0,03)	-0,158					

Pozn.: Model se statisticky významně nelišil od dat, $\chi^2(5) = 5,75$, $p = 0,331$, $TLI = 0,999$, $SRMR = 0,024$, $RMSEA = 0,021$ s $CI_{95\%} = [0,000; 0,081]$. Všechny koeficienty jsou signifikantní na $p < 0,001$. λ , v , ψ – nestandardizované faktorové náboje, intercepty, resp. reziduální rozptyly (λ v případě proměnných věk a věk² označuje regresní koeficient); std. – standardizované hodnoty; CI_{95%} – 95% interval spolehlivosti.

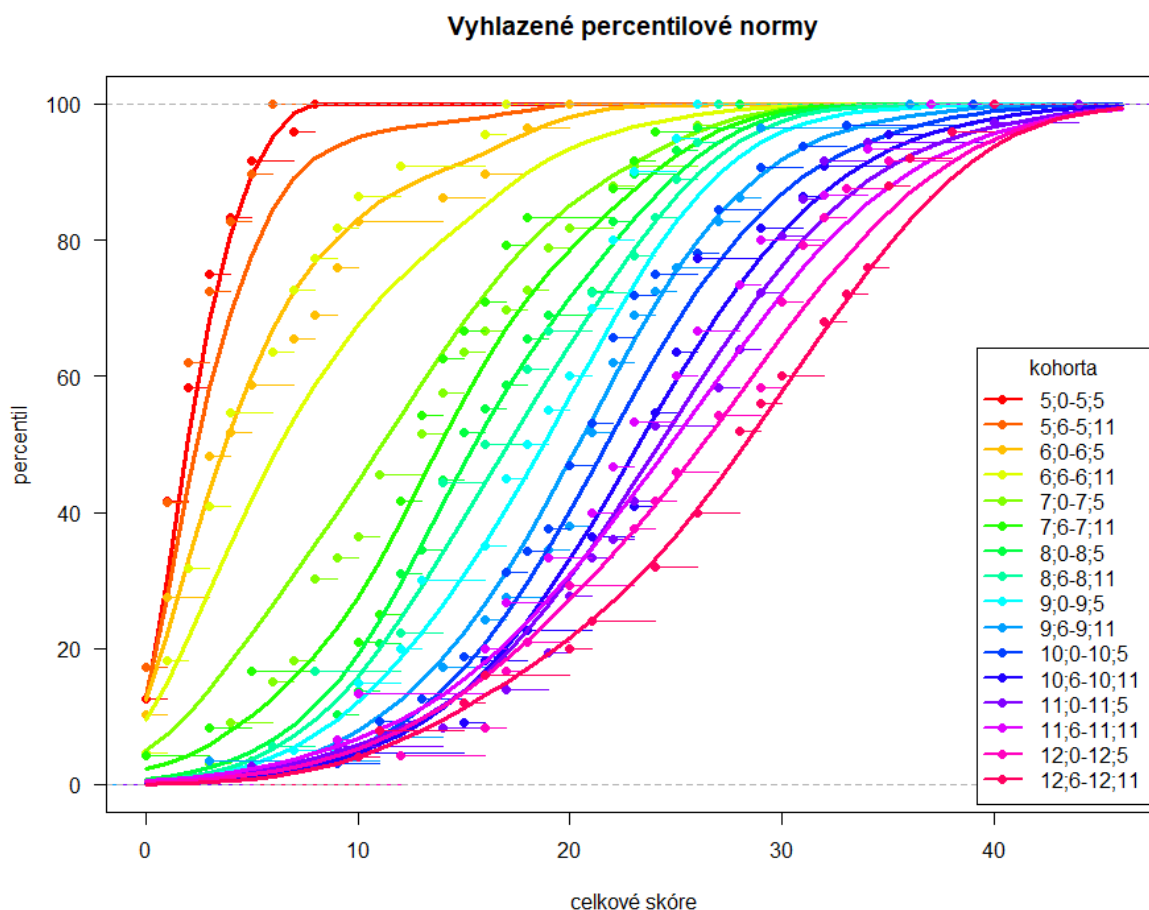
Normy

Vzhledem k velmi těsnému vztahu věku a skóru v testu verbální fluence jsou normy konstruovány pro kohorty po půl roce věku; malý počet respondentů však bohužel neumožnil tvorbu norem zvlášť podle pohlaví, prezentujeme tedy společné normy pro chlapce a dívky. Je ale nutné předpokládat, že zatímco průměrný výkon nejmladších chlapců a dívek se neliší, nejstarší dívky dosahují lepšího skóru než chlapci; rozdíl by však neměl být výrazný. Normy byly vytvořeny zvlášť pro celkový skór (součet všech vygenerovaných slov začínajících na N, K a P) a zvlášť pro počet slov začínajících na jednotlivé hlásky. U norem neuvádíme intervaly spolehlivosti; čtenář si je však může snadno dopočítat na základě chyby měření z tabulky 5 (95 % interval spolehlivosti je roven $\pm 1,96$ násobku chyby měření; 90% interval $\pm 1,64$ násobku; 80 % pak $\pm 1,28$ násobku).

Výpočet proběhl pro oba reportované skóry ve dvou krocích. V každé věkové kohortě po šesti měsících věku byla odhadnuta kumulativní distribuční funkce neparametrickou metodou kernel smoothing pomocí funkce *kcde* z R balíčku *ks* (Duong, 2016), následně jsme pak predikovali hodnotu této funkce pro hrubé skóry od nuly po maximální pozorovaný skór zvýšený o dvě slova. Tím byl do jisté míry snížen vliv výběrové chyby pro každou kohortu „separátně“.

Vizuální kontrola takto vzniklých forem však odhalila nepřesnosti – v některých případech shodný skór například znamenal vyšší percentil u vyšší věkové kohorty než u nižší. Ve druhém kroku proto byly navíc ještě hodnoty kumulativní distribuční funkce pro každý z uvažovaných hrubých skórů vyhlazeny „napříč“ věkovými kohortami kubickou spline regresí a následně omezeny do intervalu $<0; 1>$. Tím byl redukován vliv výběrové chyby „napříč“ věkovými kohortami. I přes tento postup nejsou normy ideální, v ojedinělých případech (zejména v případě počtu slov začínajících na jednotlivé hlásky) stejnému skóru nadále odpovídá vyšší percentil u starších dětí než u mladších. Protože jsou však tyto rozdíly celkově malé, dále už je nekorigujeme a případnou korekci necháváme na uživateli norem.

Výsledek pro celkový skór, tedy vyhlazené normy i původně pozorovaná kumulativní distribuce, je pro ilustraci na obrázku 2 (je patrné, že vyhlazení napříč kohortami redukovalo například výběrovou chybu pozorovaných percentilů u dětí ve věku 6;0–6;6, které podaly lepší výkon, než děti o půl roku starší). Normy jsou součástí příloh.



Obrázek 2

Plné křivky: Finální vyhlazené kumulativní distribuční funkce, s jejichž pomocí byly odhadovány percentily. Horizontální úsečky: pozorované kumulativní distribuční funkce.

Diskuze

Předpokládáme, že předložený článek bude v první řadě užitečný psychologům a (speciálním) pedagogům (např. z pedagogicko-psychologických poraden a školních poradenských pracovišť), kteří v praxi používají test verbální fluence, ale neměli pro něj dosud aktuální a reprezentativní normy. V přílohách tohoto textu jsou proto k dispozici normy pro počty slov začínajících na hlásku N/K/P vyřčených za jednu minutu a pro celkový počet slov za všechny tři hlásky N, K a P, a to pro děti ve věku 5;0–12;11. Vzhledem k velmi silnému vztahu s věkem jsou tyto normy konstruovány pro kohorty po půlrocích věku.

Ačkoliv vztah s věkem pozorujeme hlavně u mladších dětí a u starších pak slabně, schopnost verbální fluence se vyvíjí i po dvanáctém roce života. Celkový výkon studentů druhých ročníků středních škol (tedy cca 16 let) je podle Cimlerové, Pokorné a Chalupové (2007) $M = 36,79$. To je prakticky shodné s celkovým výkonem běžné, blíže nespecifikované dospělé populace ve věku 18–60 let, který by podle Janásově (2015) měl být $M = 36,8$. To odpovídá v našem vzorku u nejstarších, dvanáctiletých dětí 83.–86. percentilu. Obdobně ve studii Nikolaie a kol. (2015) byl průměrný celkový skóre pro hlásky K, P a S u 60–75letých respondentů 42–46 slov podle dosaženého vzdělání. To odpovídá asi 97.–99. percentilu našich dvanáctiletých dětí; naopak průměrný skóre dvanáctiletých dětí v naší studii by odpovídal T-skóre cca 35–40 v téže studii. Zde je však nutné vzít v úvahu použití odlišných hlásek, které sice autoři pečlivě zdůvodňují, ale které zároveň znesnadňuje možnost srovnání s naším výzkumem – z našich i cizích výsledků (např. Kopeček a Kuncová, 2006; Štorková, Preiss a Kopeček, 2004) je zřejmé, že výkon v testu fonemické slovní fluence se napříč jednotlivými hláskami liší. Celkově se nicméně zdá, že schopnost verbální fluence se vyvíjí zhruba do konce základní školy a dále pak setrvává na podobné úrovni.

Dále je třeba uvést, že naše výsledky se od výsledků uvedených v publikaci Preisse a kol. (2012) poněkud liší. V počtu slov za jednu minutu začínajících na hlásku K (jen ta byla použita v obou studiích) se pouze u devítiletých dětí jejich průměrný výkon napříč studii statisticky významně neliší, $t(72,9) = 0,82$, Cohenovo $d = 0,13$. Ve zbylých kohortách děti ze studie Preisse a kol. (2012) vyjmenovaly v průměru vyšší počet slov: v desíti letech $t(93,8) = 2,81$, $p = 0,008$, $d = 0,37$; v jedenácti letech $t(76,5) = 4,13$, $p < 0,001$, $d = 0,64$; a konečně ve dvanácti letech $t(64,1) = 3,19$, $p = 0,002$, $d = 0,74$. Průměrná velikost efektu byla Cohenovo $d_M = 0,47$. Příčina tohoto rozdílu není zřejmá, se způsobem skórování vzhledem k prakticky totožně vymezeným kritériím hodnocení v obou studiích patrně nesouvisí (anonymní recenzent, osobní komunikace [recenzní posudek], 2. 12. 2018). Preiss a kol. (2012) nicméně pracovali výhradně s pražskými dětmi, které v našem vlastním vzorku skórovaly (společně s dětmi z obcí do 2 000 obyvatel) statisticky významně lépe než ostatní děti, průměrně o 3,8 bodu celkového skóre. Počet dětí z Prahy v požadovaném věku však byl v našem vzorku příliš malý na podrobnější analýzu. Jiným vysvětlením může být tzv. negativní Flynnův efekt (Dutton, van der Linden a Lynn, 2016), případně změny průměrné verbální fluence v čase způsobené jinými příčinami. Pro tuto hypotézu však neexistuje žádná podpora v literatuře; jediná alespoň zčásti relevantní studie (Wicherts a kol., 2004) nepozorovala žádné statisticky významné změny v průměrné verbální fluenci u pětiletých holandských dětí mezi lety 1981 a 1992. Je nicméně pravdou, že od devadesátých let došlo k řadě kulturních a společenských změn a případný trend z 80. let (navíc z Nizozemska) nemusí být možné zobecnit na současnou českou společnost. Bylo by velmi vhodné získat nová data pražských dětí alespoň pro jednu z věkových kohort a zjištěné výsledky porovnat s výsledky uvedenými ve studii Preiss a kol. (2012).

Hlavním omezením našich norem je menší vzorek, který i přes pozorované rozdíly dle pohlaví neumožnil konstrukci samostatných norem pro chlapce a dívky. Uživatel však může rozdíl přibližně korigovat s pomocí obrázku 2 uvedeného výše. U nejmladších dětí rozdíl mezi chlapci a dívkami pozorován nebyl, s věkem se však zvětšoval. Jiné studie rozdíl ve výkonu mezi pohlavími buď nepozorovaly, nebo jen konstantní bez moderace věkem (Riva, Nichelli a Devoti, 2000; Hurks a kol., 2004). Dalším omezením předložených norem je i nereprezentativnost vzorku vzhledem ke vzdělání rodičů a velikosti sídla, přičemž tyto proměnné měly vliv na výsledný skóre – normy tak mohou být tímto faktem mírně zkreslené.

Na druhou stranu jsme se pokusili nevýhody způsobené menším vzorkem redukovat pokročilými statistickými postupy, které dle nás vedou k vyšší stabilitě odhadnutých norem vůči výběrové populaci.

Kromě norem reportujeme odhady reliability pro celkový skór i pro dílčí skóry jednotlivých hlásek. V praxi nicméně doporučujeme používat zejména celkový skór – tedy při administraci testu fonematické fluence dítěti postupně zadat hlásky N, K a P (v tomto pořadí). Vzhledem k pozorovaným rozdílům v počtu slov i reliabilitě napříč počty slov na různé hlásky není možné zhodnotit případný vliv jednotlivých hlásek a pořadí jejich administrace (které bylo konstantní pro všechny děti) na získaný hrubý skór – souhrnný vliv obou těchto proměnných je však na základě výsledků strukturního modelu shodný napříč celým zkoumaným věkovým rozpětím. Efekt zácviku je přitom z některých studií patrný (Kopeček a Kuncová, 2006; Štorková, Preiss a Kopeček, 2004). Při administraci všech tří subtestů bude nicméně efekt pořadí shodný jako v této standardizační studii a nebude tak mít vliv na pozorované skóry z hlediska interpretace výsledků. Zároveň pak skóry pro jednotlivé hlásky vykazují příliš nízkou reliability pro praktické použití; reliability je však poměrně nízká i u celkového skóre. Tyto odhady reliability jsou přesto cenné, protože studie s podobnými informacemi (navíc na vzorku dětí) prakticky neexistují. Cohen a kol. (1999) reportují Cronbachovu alfu $\alpha = 0,80$ pro čtyři půlminutové subtesty (písmena) dětí ve věku 6–12 let. Po upravení Spearman-Brownovým vzorcem je odhad reliability pro náš design $r^* = 0,86$, což je srovnatelné s naším odhadem na celém vzorku $\omega = 0,91$ (navíc jsme měli o něco větší rozptyl věku, což reliability v našem vzorku nutně navýšilo).⁵

Celkový skór po kontrole věku dále slabě souvisel se známkou z matematiky i českého jazyka. Rovněž děti intelektově nadané (a také děti s poruchami autistického spektra, PAS) dosahovaly vyšších skóre, a naopak děti s dyskalkulií a poruchami pozornosti (ADD) skóre horšího než běžné děti. Oproti očekávání nebyl pozorován rozdíl kontrolního vzorku a dětí s ADHD, který by měl být na základě metaanalýzy pro Cohenovo $d = 0,46$ s $CI_{95\%} = [0,33-0,59]$ (Frazier, Demaree a Youngstrom, 2004). Naopak byl pozorován rozdíl mezi dětmi s PAS a kontrolním vzorkem, což odporuje předchozím zjištěním (Corbett, Constantine, Hendren, Locke a Ozonoff, 2009). Podle výzkumu Cohena a kol. (1999) by měla verbální fluence diskriminovat děti s jazykově-dyslektickými obtížemi od dětí s vizuálně-prostorovými obtížemi či ADHD, což se rovněž nepotvrdilo. Dále ve studii Landerla a kol. (2009) byl pozorován vztah verbální fluence s dyslexií, ale nikoliv dyskalkulií; tedy přesně naopak, než v našem vzorku. Podstatným limitem všech našich analýz je však rozhodně velmi nízká velikost „klinických“ subvzorků (například jen dvě děti s PAS. Síla všech testů byla velmi malá, jak je vidět i z extrémně širokých intervalů spolehlivosti. Bez dalších výzkumných dat na českých respondentech proto doporučujeme tyto výsledky interpretovat s maximální obezřetností.

U dětí osmi- a devítiletých ($n = 93$) byl dále pomocí strukturního modelu ověřován v dřívější studii (Cígler a Durmeková, 2017) vztah verbální fluence s testy WISC-III (Wechsler, Krejčířová, Boschek a Dan, (2002) a Vytváření příběhů (Cígler a Durmeková, 2017). Ukázalo se, že na latentní úrovni (tedy po očištění o chybu měření) verbální fluence mediuje veškerý vztah verbálního IQ ve WISC-III s kreativitou zjišťovanou prostřednictvím testu Vytváření příběhů (přímý efekt nebyl statisticky významný). Verbální fluence byla predikována

⁵ Protože rovněž odhad vnitřní konzistence pomocí McDonaldovy omegy vede k vyššímu (a přesnějšímu) odhadu než Cronbachova alfa, pro lepší srovnání s Cohenem a kol. (1999) jsme dopočítali ještě i Cronbachovu alfu ze součtu všech tří subtestů. V tomto případě $\alpha = 0,892$, s intervalem spolehlivosti založeným na 1000 bootstrapových vzorcích $CI_{95\%} = [0,876; 0,906]$, což se od Cohena a kol. (1999) výsledku již prakticky neliší.

verbálním IQ, $\beta = 0,50$ (vztah s performačním IQ nebyl statisticky významný a byl věcně zanedbatelný), a zároveň predikovala celkový skóre ve Vytváření příběhů, $\beta = 0,45$. Zdá se tedy, že verbální fluence středně silně souvisí s verbálními složkami intelektu a nesouvisí s performačními. Tyto závěry jsou prakticky shodné se závěry Ardila, Pineda a Rosselli (2000) na předchozí verzi testu WISC-R. Závěrem je nutné zdůraznit, že tyto souvislosti byly odhadovány na latentní úrovni (po očištění o chybu měření v rámci strukturního modelu), vztahy pozorovaných skóre proto budou o něco nižší.

Závěr

Uživatelům testu fonematické verbální fluence předkládáme české normy pro děti ve věku 5–12 let, které mohou přímo využít ve své praxi. Reliabilita testu je však bohužel nižší, než by uživatel očekával a než jaká by byla žádoucí. Vzhledem k velmi nenáročné a rychlé administraci si nicméně myslíme, že test verbální fluence lze využít jako screeningovou metodu pro různé účely. V takovém případě doporučujeme volit součet slov začínajících na hlásky N, K a P (v tomto pořadí). Získané výsledky je každopádně nutné interpretovat pečlivě, vzít v úvahu chybu měření i nejasnou validitu testu, a v případě potřeby je doplnit podrobnějším vyšetřením jinými metodami.

Literatura

- Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, 15(1), 31–36. [http://doi.org/10.1016/S0887-6177\(98\)00159-0](http://doi.org/10.1016/S0887-6177(98)00159-0)
- Cígler, H., & Durmeková, S. (2017). *Vytváření příběhů: Psychometrická studie*. Praha: Národní ústav pro vzdělávání.
- Cimlerová, P., Pokorná, D., Chalupová, E. & kol. (2007). *Diagnostika specifických poruch učení u adolescentů a dospělých osob*. Praha: Institut pedagogicko-psychologického poradenství ČR.
- Cohen, M. J., Morgan, A. M., Vaughn, M., Riccio, C. A., & Hall, J. (1999). Verbal Fluency in Children: Developmental Issues and Differential Validity in Distinguishing Children with Attention-Deficit Hyperactivity Disorder and Two Subtypes of Dyslexia. *Archives of Clinical Neuropsychology*, 14(5), 433–443. [http://doi.org/10.1016/S0887-6177\(98\)00038-9](http://doi.org/10.1016/S0887-6177(98)00038-9)
- Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry Research*, 166(2–3), 210–22. <http://doi.org/10.1016/j.psychres.2008.02.005>
- Duong, T. (2016). Non-parametric smoothed estimation of multivariate cumulative distribution and survival functions, and receiver operating characteristic curves. *Journal of the Korean Statistical Society*, 45(1), 33–50. <http://doi.org/10.1016/j.jkss.2015.06.002>
- Dutton, E., van der Linden, D., & Lynn, R. (2016). The negative Flynn Effect: A systematic literature review. *Intelligence*, 59, 163–169. <https://doi.org/10.1016/j.intell.2016.10.002>

- Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression*. Thousand Oaks, CA: Sage. <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>
- Frazier, T. W., Demaree, H. A., & Youngstrom, E. A. (2004). Meta-analysis of intellectual and neuropsychological test performance in attention-deficit/hyperactivity disorder. *Neuropsychology*, 18(3), 543–555. <http://doi.org/10.1037/0894-4105.18.3.543>
- Hurks, P. P. M., Hendriksen, J. G. M., Vles, J. S. H., Kalff, A. C., Feron, F. J. M., Kroes, M., ... Jolles, J. (2004). Verbal fluency over time as a measure of automatic and controlled processing in children with ADHD. *Brain and Cognition*, 55(3), 535–544. <http://doi.org/10.1016/j.bandc.2004.03.003>
- Janásová, T. (2015). *Analýza výkonů v testech figurální a verbální fluence*. Univerzita Palackého v Olomouci: Nepublikovaná bakalářská diplomová práce.
- Kim, S. (2015). ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. *Communications for Statistical Applications and Methods*, 22(6), 665–674. <http://doi.org/10.5351/CSAM.2015.22.6.665>
- Kopeček, M., & Kuncová, A. (2006). Efekt nácviku testu generování slov a testování alternativní verze. *Psychiatrie*, 10(4), 211–215.
- Landerl, K., Fussenegger, B., Moll, K., & Willburger, E. (2009). Dyslexia and dyscalculia: Two learning disorders with different cognitive profiles. *Journal of Experimental Child Psychology*, 103(3), 309–324. <http://doi.org/10.1016/j.jecp.2009.03.006>
- Levy, G., Jacobs, D. M., Tang, M.-X., Côté, L. J., Louis, E. D., Alfaró, B., ... Marder, K. (2002). Memory and executive function impairment predict dementia in Parkinson's disease. *Movement Disorders*, 17(6), 1221–1226. <http://doi.org/10.1002/mds.10280>
- Lüdecke, D. (2018). *sjstats: Statistical Functions for Regression Models. R package version 0.14.1*. <https://cran.r-project.org/package=sjstats>
- Marko, M. (2016). Využitie a zneužitie Cronbachovej alfy pri hodnotení psychodiagnostických nástrojov. *TESTFÓRUM*, 5(7). <http://doi.org/10.5817/TF2016-7-90>
- Nikolai, T., Štěpánková, H., Michalec, J., Bezdíček, O., Horáková, K., Marková, H., ... Kopeček, M. (2015). Testy verbální fluence, česká normativní studie pro osoby vyššího věku. *Česká a slovenská neurologie a neurochirurgie*, 78/111(3), 292–299. <http://doi.org/10.14735/amcsnn2015292>
- Preiss, M. (1997). Verbální fluence, metoda vyšetřování poškození mozku u dětí a dospělých. *Československá psychologie*, 41(3), 244–249.
- Preiss, M., Bartoš, A., Čermáková, R., Nondek, M., Benešová, M., Rodriguez, M., ... Nikolai, T. (2012). *Neuropsychologická baterie Psychiatrického centra Praha*. Praha: Psychiatrické centrum Praha.
- Revelle, W. (2017). *psych: Procedures for Personality and Psychological Research*. Evanston: Northwestern University. <https://cran.r-project.org/package=psych>
- Rinehardt, E., Eichstaedt, K., Schinka, J. A., Loewenstein, D. A., Mattingly, M., Fils, J., ... Schoenberg, M. R. (2014). Verbal Fluency Patterns in Mild Cognitive Impairment and Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders*, 38(1–2), 1–9. <http://doi.org/10.1159/000355558>
- Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and confrontation naming in children. *Brain and Language*, 71(2), 267–284. <http://doi.org/10.1006/brln.1999.2166>

- Stern, Y., Albert, S., Tang, M. X., Tsai, W. Y., Freeman, A., Haber, M., ... DeLong, M. R. (1999). Rate of memory decline in AD is related to education and occupation: cognitive reserve? *Neurology*, 53(9), 1942–7. <http://doi.org/10.1212/01.wnl.0000031426.21683.e2>
- Štorková, P., Preiss, M., & Kopeček, M. (2004). Efekt nácviku testu verbální fluence a testování alternativní verze. Pilotní studie. *Psychiatrie*, 8(3), 186–189.
- Urbánek, T. (2012). Nejpoužívanější psychodiagnostické metody v České republice. *TESTFÓRUM*, 1(1). <http://doi.org/10.5817/TF2010-1-3>
- Wechsler, D., Krejčířová, D., Boschek, P., & Dan, J. (2002). *WISC-III - Wechslerova inteligenční škála pro děti*. Praha: Testcentrum.
- Wicherts, J. M., Dolan, C. V., Hessen, D. J., Oosterveld, P., van Baal, G. C. M., Boomsma, D. I., & Span, M. M. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence*, 32(5), 509–537. <https://doi.org/10.1016/j.intell.2004.07.002>

Online supplement (frekvenční analýza nejčastějších slov generovaných během testu) je dostupná online: https://e-psycholog.eu/docs/suppl_4_2018_frekvence.pdf

Přílohy jsou dostupné na adrese : https://e-psycholog.eu/docs/epsy.329_prilohy.pdf

Příloha 1: Percentilové normy pro celkový skóre (hlásky N, K, P)

Příloha 2: Percentilové normy pro počet slov začínajících na N

Příloha 3: Percentilové normy pro počet slov začínajících na K

Příloha 4: Percentilové normy pro počet slov začínajících na P

O autorech

Hynek Cígler, Ph.D., působí na Katedře psychologie & Institut pro výzkum dětí, mládeže a rodiny, Fakulta sociálních studií Masarykovy univerzity v Brně.

Kontaktní údaje:

Adresa: Fakulta sociálních studií Masarykovy univerzity v Brně, Joštova 10, Brno

Email: hynek.cigler@mail.muni.cz

Světlana Durmeková, Ph.D., působí v Národním ústavu pro vzdělávání, školské poradenské zařízení a zařízení pro další vzdělávání pedagogických pracovníků.

Kontaktní údaje:

Adresa: Weilova 1271/6, 102 00 Praha 10, detašované pracoviště Novoborská 372/8, 190 00 Praha 9

E-mail: svetlana.durmekova@nuv.cz

Cígler, H., & Durmeková, S. (2018). Verbální fluence u dětí ve věku 5–12 let: české normy a vybrané psychometrické ukazatele. *E-psychologie*, 12(4), 16–30. <https://doi.org/10.29364/epsy.329>