

UNIVERZITA KARLOVA

Filozofická fakulta

Katedra psychologie

Diplomová práce



Bc. Nikola Doubravová

**Paralelnost verzí A a B Neuropsychologické baterie pro
děti v doméně Exekutivní funkce**

**Parallelism of Versions A and B of the Neuropsychological
Battery for Children in the Executive Function Domain**

Vedoucí diplomové práce: PhDr. Alice Maulisová, Ph.D.

Konzultantka diplomové práce: Mgr. Kateřina Bukačová

2023/2024

Poděkování

Na tomto místě bych ráda poděkovala své vedoucí práce PhDr. Alici Maulisové, PhD. a konzultantce práce Mgr. Kateřině Bukačové za možnost administrace nově vznikající Neuropsychologické baterie pro děti a za možnost využití získaných dat v rámci své diplomové práce. Také chci oběma poděkovat za trpělivost, ochotu a podnětné komentáře a rady jak při sběru dat, tak při zpracování literárně-přehledové a empirické části práce. Velký dík patří i všem zúčastněným probandům a jejich rodinám za ochotu dobrovolně se studii účastnit.

Prohlášení

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 10. dubna 2024

Bc. Nikola Doubravová

Abstrakt

Cílem diplomové práce s názvem *Paralelnost verzí A a B Neuropsychologické baterie pro děti v doméně Exekutivní funkce* bylo ověřit paralelnost dvou verzí nově vznikající Neuropsychologické baterie pro děti, a to v doméně exekutivních funkcí. Tato doména zahrnuje subtesty Třídění, Inhibice a přesun pozornosti, Pracovní paměť, Verbální fluence a Prospektivní paměť.

Část literárně-přehledová v první řadě shrnuje dosavadní poznatky a teoretické modely zaměřené na exekutivní funkce. Dále se zabývá různými způsoby odhadu reliability a v poslední kapitole se zaměřuje právě na odhad reliability paralelních forem a jeho užití v testech exekutivních funkcí.

Empirická část se pak zabývá samotnou výzkumnou studií. Obě verze testové baterie byly probandům, dětem ve věku od šesti do devatenácti let, administrovány v rozmezí dvou až čtyřech týdnů. Výzkumný soubor ($N = 31$) byl náhodně rozdělen na dvě skupiny, kterým byly verze zadány v opačném pořadí. Dále byl všem probandům administrován anamnestický dotazník a Zkrácená Wechslerova inteligenční škála, druhé vydání (Wechsler, 2011). Statistická analýza dat zahrnovala korelační analýzu (Pearsonův nebo Spearmanův korelační koeficient) a rozdílové statistiky (t -test nebo Wilcoxonův test pro párové vzorky). Na základě statistických analýz se jako ekvivalentní jeví *Třídění*, *Pracovní paměť* a *Prospektivní paměť*. Dále se ekvivalence potvrdila pro verze *Fonemické fluence* pro hlásky V (ve verzi A) a M (ve verzi B), a měřený čas v části *Inhibice* subtestu *Inhibice a přesun pozornosti*.

Práce je spojena se standardizací nově vznikající Neuropsychologické baterie pro děti a jako první se zabývá potřebou ověření paralelnosti obou verzí této testové baterie. K umožnění praktického využití těchto poznatků je nutné zabývat se paralelností forem i v ostatních doménách baterie a také vyrovnáváním těchto alternativních forem, což by vedlo k zajištění co nejvyšší srovnatelnosti testových skóre.

Klíčová slova: Reliabilita; Exekutivní funkce; Neuropsychologická diagnostika; Paralelní formy

Abstract

The aim of the thesis entitled "Parallelism of Versions A and B of the Neuropsychological Battery for Children in the Executive Function Domain" was to verify the parallelism of the two versions of the newly developed Neuropsychological Battery for Children in the domain of executive functions. This domain includes subtests *Sorting*, *Inhibition and Switching*, *Working Memory*, *Verbal Fluency*, and *Prospective Memory*.

The literature review section first summarizes previous findings and theoretical models focused on executive functions. It then discusses various methods of estimating reliability and, in the final chapter, focuses on estimating the reliability of parallel forms and its use in tests of executive functions.

The empirical part deals with the research study itself. Both versions of the test battery were administered to participants, children aged six to nineteen, within a time range of two to four weeks. The research sample ($N = 31$) was randomly divided into two groups, each receiving the versions in opposite order. Additionally, all participants completed an anamnestic questionnaire and the Wechsler Abbreviated Scale of Intelligence, Second Edition (Wechsler, 2011). Data analysis included correlation analysis (Pearson's or Spearman's correlation coefficient) and differential statistics (t -test or Wilcoxon test for paired samples). Based on the statistical analyses, *Sorting*, *Working Memory*, and *Prospective Memory* appear to be equivalent. Equivalence was then confirmed for the *Phonemic Fluency* versions for the letters V (in version A) and M (in version B), and the measured time in the *Inhibition* part of the *Inhibition and Attention Shift* subtest.

The thesis is associated with the standardization of the newly developed Neuropsychological Battery for Children and is the first to address the need to verify the parallelism of both versions of this test battery. To enable practical application of these findings, it is necessary to investigate the parallelism of forms in other domains of the battery and to use an equating method for these alternative forms to ensure the highest possible comparability of test scores.

Key words: Reliability; Executive Functions; Neuropsychological Assessment; Parallel Forms

Obsah

Seznam zkratk	8
Úvod	10
I. Teoretická část	11
1. Exekutivní funkce	11
1.1. Modely exekutivních funkcí	12
1.1.1. Duncanův model	12
1.1.2. Model B. Millera & J. Cummingse	13
1.1.3. Model Lezakové	13
1.1.4. Model Stusse a Bensona	14
1.1.5. Unity and diversity model	14
1.2. Biologický podklad exekutivních funkcí	18
1.3. Vývoj exekutivních funkcí	21
1.3.1. Senzitivní období	23
1.4. Reflexe dosavadních modelů exekutivních funkcí	24
2. Reliabilita	27
2.1. Definice reliability	27
2.1.1. Standardní chyba měření	28
2.1.2. Domain sampling model	30
2.1.3. Item response Theory	31
2.2. Typy reliability	31
2.2.1. Test-retest reliabilita	32
2.2.2. Split-half reliabilita	33
2.2.3. Item-total korelace	34
2.2.4. Reliabilita jako vnitřní konzistence	34
2.2.5. Kuderova-Richardsonova reliabilita	35
2.2.6. Reliabilita jako shoda posuzovatelů	36
2.2.7. Reliabilita rozdílového skóru	36
2.2.8. Zvyšování reliability nástroje	37
3. Reliabilita paralelních forem a měření exekutivních funkcí	39
3.1. Odhad reliability paralelních forem	39
3.1.1. Vyrovnávání alternativních forem testu	40
3.2. Reliabilita paralelních forem ve vybraných testech exekutivních funkcí	41
3.2.1. Verbální fluence	43
3.2.2. Wisconsinský test třídění karet	45
3.2.3. Test cesty	46
3.2.4. Stroopův test	48
3.2.5. Rey-Osterriethova komplexní figura a komplexní figura Taylorové	50
3.2.6. Limity alternativních forem testů exekutivních funkcí	51
II. Empirická část	53
4. Cíl výzkumu	54

4.1. Výzkumné otázky a hypotézy	54
5. Metodika	55
5.1. Výzkumný soubor	55
5.2. Měřicí nástroje	57
5.2.1. Anamnestický dotazník	57
5.2.2. Zkrácená Wechslerova inteligenční škála, druhé vydání (WASI-II)	58
5.2.3. Neuropsychologická baterie pro děti (NB-D)	58
5.3. Proces sběru dat	60
5.4. Statistická analýza	61
5.5. Etika výzkumu	61
6. Výsledky	62
6.1. Deskriptivní a inferenční statistika	62
6.1.1. Třídění	62
6.1.2. Inhibice a přesun pozornosti	62
6.1.3. Pracovní paměť	64
6.1.4. Verbální fluence	64
6.1.5. Prospektivní paměť	66
6.2. Statistická analýza dat	66
6.2.1. Třídění	67
6.2.2. Inhibice a přesun pozornosti	67
6.2.3. Pracovní paměť	68
6.2.4. Verbální fluence	69
6.2.5. Prospektivní paměť	70
7. Diskuse	71
7.1. Limity studie	74
Závěr	77
Reference	79
Seznam obrázků	96
Seznam tabulek	97

Seznam zkratek

ADHD	Porucha aktivity a pozornosti
APA	Americká psychologická asociace
ARAS	Ascendentní retikulární aktivační systém
BADS-C	The Behavioural Assessment of the Dysexecutive Syndrome in Children
BRIEF	Behavior Rating Inventory of Executive Function (Škála hodnocení exekutivních funkcí u dětí)
BTC	Barevný test cesty
CANTAB	The Cambridge Neuropsychological Test Automated Battery
CFA	Konfirmatorní faktorová analýza
COWAT/FAS	Controlled Oral Word Association Test
CST	Clevelandský test třídění karet
CoT	Connections Task
CT	Počítačová tomografie
CTMT	Comprehensive Trail Making Test
DCCS	Dimensional Change Card Sort Test
D-KEFS	Delis-Kaplan Executive Function System
DK-TMT	Delis-Kaplan Executive Function System, Test cesty
DSM-V	Diagnostický a statistický manuál duševních poruch, páté vydání
DTPS	Difúzní thalamický projekční systém
EEG	Elektroencefalogram
fMRI	Funkční magnetická rezonance
GABA	Gamma-aminomáselná kyselina
ID	Identifikační číslo
IRT	Item Response Theory
LSPAN	Listening Span Task
MRI	Magnetická rezonance
MTCF	Modifikovaná komplexní figura Taylorové
NB-D	Neuropsychologická baterie pro děti
NEPSY II	A Developmental Neuropsychological Assessment, Second Edition (Vývojová neuropsychologická baterie, druhé vydání)
PET	Pozitronová emisní tomografie

ROCF	Rey-Osterriethova figura
SD	Směrodatná odchylka
SEM	Standardní chyba měření
SOPT	Self-Ordered Pointing task
TCF	Komplexní figura Taylorové
TEA-Ch	The Test of Everyday Attention for Children
TMT	Trail Making Test (Test cesty)
WASI-II	Wechsler Abbreviated Scale of Intelligence, Second Edition (Zkrácená Wechslerova inteligenční škála, druhé vydání)
WCST	Wisconsinový test třídění karet
WEIRD	Western, Educated, Industrialized, Rich, and Democratic

Úvod

Tato práce je součástí vývoje a normativní studie Neuropsychologické baterie pro děti (NB-D). NB-D představuje komplexní soubor testů vyvinutý pod hlavičkou Fakultní nemocnice Motol, určený pro neuropsychologickou diagnostiku kognitivních funkcí u dětí. Hlavním cílem práce je posoudit, zda verze A a B NB-D v doméně exekutivních funkcí poskytují srovnatelná data a zda mohou být považovány za ekvivalentní. To by umožnilo jejich využití pro retestování exekutivních funkcí v případě potřeby a práce má tedy potenciál přispět klinické praxi v oblasti dětské neuropsychologie.

První, literárně přehledová, část práce zahrnuje nejprve přehled teoretických modelů exekutivních funkcí a zabývá se jejich neurobiologickým podkladem a vývojem. Exekutivní funkce jsou souborem kognitivních procesů, které zajišťují řízení ostatních kognitivních funkcí. Mezi aspekty exekutivních funkcí se řadí pracovní paměť, inhibice a kognitivní flexibilita, dále také schopnost volit cíle, plánovat nebo udržovat a měnit činnost (Diamond, 2013). V další kapitole se práce zabývá metodami odhadu reliability v psychometrii. Reliabilita paralelních forem je pak jednou z těchto metod a právě tou se zabývá i třetí kapitola literárně přehledové části. Reliabilita paralelních forem používá dvě ekvivalentní formy testu, které měří stejný konstrukt. Tím je zajištěna vysoká míra srovnatelnosti výsledků a eliminace vlivu paměti na opakované testování (Helmstadter, 1964). Reliabilita paralelních forem je důležitá pro vytvoření alternativní testové metody s minimálním rizikem efektu učení (McDonald, 1999; Rust et al., 2021; Urbánek et al., 2011).

Druhá, empirická část, této diplomové práce se zabývá samotnou studií paralelnosti verzí A a B NB-D. Shrnuje cíle výzkumu a ověřované hypotézy, metodologii výzkumu, včetně použitých metod, a výzkumný soubor. Dále se zabývá provedenou statistickou analýzou a výsledky výzkumu a na závěr diskuzí výsledků v kontextu dosavadních poznatků.

Diplomová práce byla citována dle manuálu Americké psychologické asociace (APA; American Psychological Association, 2020).

I. Teoretická část

Teoretická část práce je přehledem současných poznatků v oblasti exekutivních funkcí a psychometrického zkoumání kognice. V první části se zabývá exekutivními funkcemi z pohledu různých teoretických modelů a také jejich biologickým podkladem a vývojem. V další kapitole se práce soustředí na odhad reliability v psychologii, zabývá se jednotlivými typy reliability i standardní chybou měření. V poslední, třetí kapitole, teoretické části se zaměřuje na odhad reliability paralelních forem a jeho užití v měření psychometrických charakteristik testů exekutivních funkcí v České republice i zahraničí.

1. Exekutivní funkce

Exekutivní funkce jsou souborem kognitivních funkcí, které zajišťují řízení ostatních mentálních procesů. Zapojují se v situacích, kdy není možné nebo efektivní spoléhat se na automatické a intuitivní reakce. Umožňují nám vědomě volit cíle a směřovat k jejich plnění, plánovat a udržovat, zastavovat nebo měnit činnost (Diamond, 2013).

Existují různé modely, které se snaží teoreticky zachytit exekutivní funkce. Tradičně se tyto modely dělí na unitární modely, jehož příkladem může být Duncanův model a které vysvětlují exekutivní funkce jako fluidní inteligenci, tedy obecný „g“ faktor inteligence (Duncan et al., 1996), a non-unitární modely, mezi které patří model Millera & Cummingse a které vykládají exekutivní funkce jako samostatné specifické procesy (Miller & Cummings, 2007).

Studie z posledních let (Diamond, 2013; Karr et al., 2018; Lehto et al., 2003; Miyake et al., 2000; Miyake & Friedman, 2017) zmiňují tři vzájemně korelované, ale oddělitelné složky exekutivních funkcí, a to inhibici, pracovní paměť a kognitivní flexibilitu. Z těch pak vychází vyšší exekutivní funkce, jako je deduktivní a induktivní myšlení, řešení problémů a plánování. Model založený na těchto studiích bývá označován jako Unity and diversity model exekutivních funkcí (Miyake et al., 2000). Po jeho uvedení ho Miyake i jiní autoři (Karr et al., 2018; Lehto et al., 2003; Miyake & Friedman, 2017) dále propracovávali.

Mezi nejčastěji citované modely exekutivních funkcí patří vedle modelu Unity and Diversity ještě model Lezakové (2004), dle které jsou základními komponentami exekutivních funkcí vůle, plánování, účelné jednání a úspěšný výkon, a model Stusse a Bensona (1986), kteří identifikují tři systémy interagující mezi sebou, jež jsou zodpovědné za monitoring pozornosti a exekutivních funkcí jedince. Mezi tyto systémy patří ascendentní retikulární aktivační systém (ARAS), difúzní thalamický projekční systém (DTPS) a fronto-thalamický gating systém (Dias et al., 2023).

1.1. Modely exekutivních funkcí

1.1.1. Duncanův model

Duncanova teorie exekutivních funkcí reprezentuje unitární modely (Duncan et al., 1996), tedy chápe exekutivu jako jednotnou funkci. Tento pohled vychází z konceptu obecné inteligence formulovaného Spearmanem, která předpovídá úspěšné zvládnání různých kognitivních úkolů. K tomuto obecnému faktoru jsou Duncanem přirovnávány exekutivní funkce (Duncan et al., 1996).

Ve chvíli, kdy lidské chování směřuje k dosažení cíle, dochází k aktivaci frontálních laloků. Duncan et al. (1996) prezentovali případy pacientů s lézemi ve frontálních lalocích, u kterých pozorovali jev nazvaný *goal neglect* – tito pacienti nesplnili zadaný úkol i přesto, že ho správně chápali i si ho zapamatovali. Vedle toho jedinci s poškozením v dané oblasti často dosahovali nižších výsledků v testech fluidní inteligence. Na základě těchto poznatků Duncan prosazuje právě unitární model a fluidní inteligenci považuje za nejlepšího prediktora úrovně exekutivních funkcí.

Unitární model je v současnosti považován za příliš zjednodušující. Jedním z důvodů je nesoulad s existujícími poznatky, které naznačují, že exekutivní funkce jsou spíše tvořeny vícero odlišnými schopnostmi. Například studie s pacienty s lézí v různých oblastech mozku ukázala, že poškození může vést k problémům ve specifické oblasti exekutivních funkcí (Tsuchida & Fellows, 2009, 2013). Nezabývá se navíc ani tím, jak se tyto funkce uplatňují v různých kognitivních úkolech. Různé funkce totiž mohou být zapotřebí v odlišných úkolových situacích (Miyake et al., 2000; Miller & Cummings, 2007).

1.1.2. Model B. Millera & J. Cummingse

Bruce L. Miller a Jeffrey L. Cummings (2007) prezentují svůj model exekutivních funkcí, který předpokládá, že jsou tvořeny vícero odlišnými schopnostmi. Mezi ty zahrnují:

- selektivní pozornost, tedy schopnost soustředit se na důležité informace a ignorovat rozptýlení;
- plánování jako schopnost vymyslet a realizovat postup k dosažení cíle;
- organizaci, která nám umožňuje udržet si přehled o informacích a úkolech;
- behaviorální kontrolu, zahrnující schopnost ovládat své emoce a chování.

Uvádějí, že tyto schopnosti jsou nezávislé na sobě a mohou být poškozeny nebo oslabeny v důsledku různých faktorů, včetně zranění, nemoci nebo vývojových poruch. Zmiňují také ontogenezi exekutivních funkcí, přičemž vysvětlují, jak se s věkem exekutivní funkce zlepšují a schopnost dětí plnit úkoly, které vyžadují plánování, organizaci a kontrolu chování, se zvyšuje.

1.1.3. Model Lezakové

Lezaková v rámci svého konceptu exekutivních funkcí definuje tento komplex mentálních procesů jako prvek, který umožňuje jednotlivcům cílevědomé a plánované chování. Tato schopnost řídit lidské aktivity tak, aby směřovaly k dosažení stanovených cílů, výrazně odlišuje exekutivní funkce od kognitivních, které jsou spíše zaměřeny na zpracování informací. Zatímco kognitivní funkce podle ní odpovídají na otázky „co“ nebo „kolik“, exekutivní funkce se zaměřují spíše na otázky typu „zda“ a „jak“ (Lezak et al., 2004).

V rámci svého modelu pak identifikuje čtyři základní složky exekutivních funkcí. První z nich, vůle, je manifestována schopností jedince motivovat se k činnostem, jež směřují k dosažení stanoveného cíle. Druhá složka, plánování, zahrnuje schopnost organizovat jednotlivé kroky k dosažení zamýšleného výsledku. Třetí složka, účelné jednání, se týká schopnosti iniciovat, strukturovat a řídit aktivitu tak, aby dosáhla úspěšného dokončení. Poslední, čtvrtou složkou je úspěšný výkon, který demonstruje schopnost dotáhnout činnost do konce a dosáhnout očekávaného výsledku (Lezak et al., 2004).

1.1.4. Model Stusse a Bensona

Podle Stusse a Bensona (1986) existují tři systémy, které vzájemně interagují při monitorování pozornosti a exekutivních funkcí jednatelce.

Prvním systémem je ascendentní retikulární aktivační systém (ARAS), který je zodpovědný za udržování obecné úrovně bdělosti jednatelce. Poškození systému ARAS tedy může vést ke ztrátě vědomí.

Druhým systémem je difúzní thalamický projekční systém (DTPS), který se podílí na krátkodobém udržování jednatelce zaměření na vnější podněty, jde o fázičké změny pozornosti. Narušení tohoto systému může způsobit nadměrné rozptylování pozornosti externími podněty.

Třetím systémem je fronto-thalamický gating systém, který je zodpovědný za vyšší kognitivní funkce, jako jsou plánování, výběr stimulů a odpovědí a monitorování každodenního výkonu, a podílí se tedy na exekutivní kontrole pozornosti. Poškození tohoto systému může ovlivnit vyšší úrovně kortikálních funkcí.

Zatímco první dva systémy se starají o udržení úrovně bdělosti, třetí systém se zaměřuje na exekutivní kontrolu pozornosti, konkrétně na plánování, výběr stimulů a odpovědí, a monitorování každodenního výkonu. Později se Stuss et al. (1995; 2005; 2011) zabýval propojením této teorie s empirickými důkazy z neuroanatomických, neuropsychologických a fyziologických studií.

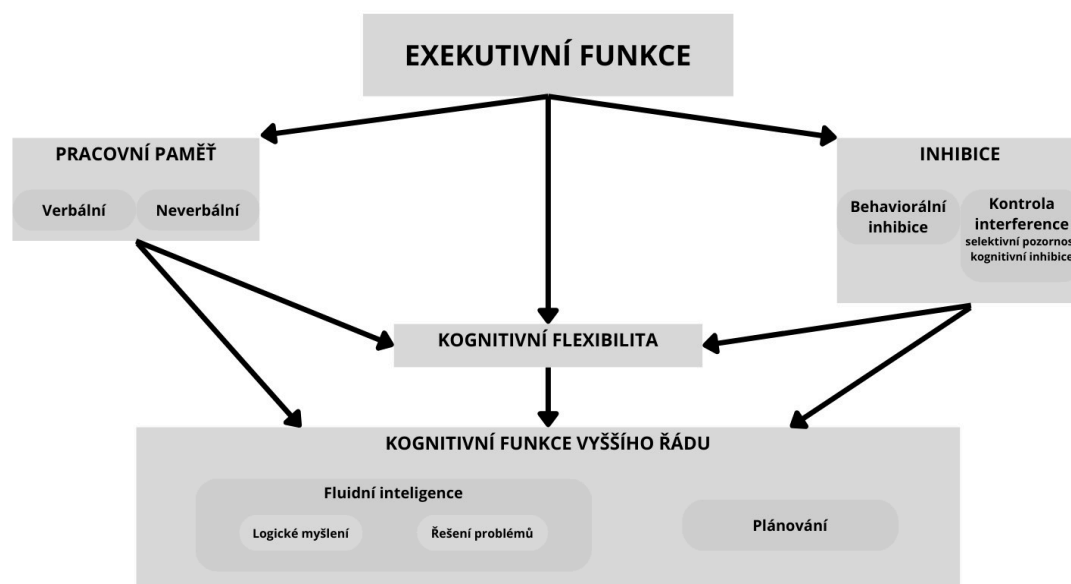
1.1.5. Unity and diversity model

Miyake et al. (2000) na základě poznatků ze svých studií naznačují, že exekutivní funkce nejsou pouze samostatné jednotky, ale jsou na sobě také závislé a jsou vzájemně propojené. Tento nový pohled na exekutivní funkce, který je založený na statistickém přístupu, integruje unitární a non-unitární modely. K závěru, že jde o samostatné složky exekutivy, došli výzkumníci na základě faktorové analýzy, která odhalila tři nezávislé faktory odpovídající inhibici, pracovní paměti a kognitivní flexibilitě (Obrázek 1). Z nich pak vychází exekutivní funkce vyššího řádu, tedy logické myšlení, řešení problémů a plánování. Z analýzy latentních struktur, která odhalila vazby mezi jednotlivými složkami, pak vyvozují, že jde o vzájemně propojené funkce. (Miyake et al., 2000; Miyake & Friedman, 2017).

Studie založené na tomto konceptu se zároveň zabývaly i souvislostí exekutivních funkcí a obecné inteligence. Došli k závěru, že jsou tyto koncepty úzce spojeny a že lidé s vyšší úrovní exekutivních funkcí mají také obvykle vyšší intelektovou výkonnost (Friedman et al., 2008; 2011). Nejde ale o zcela totožné koncepty. Exekutivní funkce sdílí jen asi 25 % variance s obecným g-faktorem, mají vlastní genetický podklad odlišný od genetického podkladu inteligence a poskytují informace o schopnostech a potenciálu jednotlivce, které nejsou obsaženy v inteligenčních testech (Miyake & Friedman, 2017). Vyvrací tak i unitární model, který považuje g-faktor za nejlepší prediktor úrovně exekutivních funkcí.

Obrázek 1

Unity and Diversity model exekutivních funkcí



Inhibice zahrnuje sebekontrolu (behaviorální inhibici) a kontrolu interference (selektivní pozornost a kognitivní inhibici). Inhibiční kontrola znamená schopnost ovládat svou pozornost, chování, myšlenky nebo emoce, a to navzdory vnitřním tendencím či vnějším podnětům. Místo toho nám umožňuje jednat v souladu s tím, co je vhodnější nebo potřebné. Namísto přenechání řízení našeho chování návykům a impulzům nám tedy nabízí možnost vybrat si způsob, jakým budeme reagovat. Pokud selektivně zaměřujeme pozornost na to, co si vědomě zvolíme, případně pokud potlačujeme pozornost na jiné podněty, využíváme inhibiční kontrolu pozornosti. Když je ale naše pozornost přitahována nějakým výrazným podnět, aniž bychom si zaměření na tento podnět vědomě volili (např. hlasitý zvuk), exekutivní funkce do tohoto procesu nejsou zapojeny. Jedná se o automatickou, podnětově řízenou pozornost, která směřuje zdola nahoru, oproti tomu

inhibiční kontrola směřuje shora dolů. Vedle inhibiční kontroly pozornosti je důležitým procesem také kognitivní inhibice, díky které můžeme potlačit mentální reprezentace, například nepodstatné a nechtěné myšlenky a vzpomínky. To znamená, že můžeme odolávat proaktivní nebo retroaktivní interferenci. Dalším aspektem inhibiční kontroly, který zahrnuje ovládání chování a emocí, je sebekontrola, která spočívá v odolávání pokušením a nejednání impulzivně nebo ve schopnosti zůstat u úkolu navzdory rušivým elementům. Projevem sebekontroly je pak schopnost oddálení okamžitého uspokojení, tedy schopnost odříct si okamžité potěšení ve prospěch většího potěšení v budoucnu (Diamond, 2013).

Další exekutivní funkcí dle tohoto modelu je *pracovní paměť*, která zahrnuje udržování informací v paměti a mentální manipulaci s nimi. Miyake (2000) tuto funkci nazývá updating nebo updating a monitoring pracovní paměti. Tato funkce vyžaduje sledování a kódování příchozích informací podle relevance pro daný úkol a následné přiměřené revidování položek uložených v pracovní paměti. Jde o nahrazení starých, již nepodstatných informací, novými a relevantnějšími. Pracovní paměť je nezbytná pro porozumění všemu, co se odvíjí v čase, protože toto porozumění vždy vyžaduje udržování minulých událostí v paměti a jejich vztahování k tomu, co přichází později. Je také zásadní pro porozumění psanému nebo mluvenému jazyku, provádění matematických úvah a výpočtů, mentální přeorganizování položek (například přeorganizování seznamu úkolů), převádění instrukcí na akční plány, aktualizace myšlenkových nebo akčních plánů, zkoumání alternativ a mentální vztahování informací k odvození obecného principu nebo k vidění vztahů mezi položkami nebo myšlenkami. Důležitá je i pro naši schopnost vidět spojení mezi zdánlivě nesouvisejícími věcmi a rozkládat prvky z integrovaného celku. Tím je také podkladem pro kreativitu, která v tomto případě znamená rozkládání a znovu kombinování prvků originálním způsobem. Pracovní paměť nám dále umožňuje uplatňovat konceptuální znalosti při rozhodování a zohledňovat naše zkušenosti a budoucí cíle při plánování a rozhodování. Odlišuje se od krátkodobé paměti, protože není pouhým udržováním informací v paměti, ale její hlavní vlastností je právě manipulace s těmito informacemi a reprezentacemi. Pracovní a krátkodobá paměť se seskupují do samostatných faktorů při faktorových analýzách jak u dětí, tak u adolescentů a dospělých (Alloway et al. 2004; Gathercole et al. 2004). Dle obsahu se dá rozdělit na dva typy – verbální a neverbální (zrakově-prostorová) pracovní paměť. (Diamond, 2013).

Známý model pracovní paměti vytvořili A. Baddeley a G. Hitch (1974). V jejich původním modelu byly popsány tři komponenty, ze kterých se pracovní paměť skládá. Pracovní paměť považují za dynamickou a proměnlivou funkci. První složkou je centrální exekutiva, která je odpovědná za kontrolu a koordinaci ostatních dvou komponent: fonologické smyčky, která umožňuje uchovávání a zpracování zvukových informací, a vizuoprostorového náčrtníku, jež odpovídá za uchovávání a zpracování vizuálních a prostorových informací. V průběhu let byl model dále rozvíjen a modifikován a byla přidána další komponenta zvaná epizodický buffer, který uchovává epizodické informace (Baddeley, 2000).

Třetí exekutivní funkcí dle modelu Unity and Diversity je *kognitivní flexibilita*, nebo také shifting, mentální flexibilita nebo mentální set shifting (změna nastavení), ta staví na předchozích dvou funkcích a objevuje se také mnohem později ve vývoji. Zajišťuje schopnost změny perspektivy, ať už prostorové nebo interpersonální, schopnost přecházet mezi různými úkoly a provádět novou operaci v konfrontaci s proaktivní interferencí nebo negativním primingem. Pro změnu perspektivy potřebujeme potlačit nebo deaktivovat naši předchozí interpretaci situace a nahrát do pracovní paměti nebo aktivovat nový úhel pohledu. V tomto smyslu kognitivní flexibilita staví na inhibiční kontrole a pracovní paměti. Kognitivní flexibilita také zahrnuje dostatečnou pružnost k přizpůsobení se změněným požadavkům nebo prioritám, přiznání si chyby a využívání náhlých, neočekávaných příležitostí. Zajišťuje nám, že naše kognitivní procesy nejsou rigidní (Diamond, 2013; Miyake et al., 2000).

Mezi *exekutivní funkce vyššího řádu* můžeme zařadit schopnost logického myšlení, plánování a řešení problémů (Diamond, 2013). Jde o komplexní schopnosti, které nám umožňují efektivní fungování v různorodých situacích běžného života, profesního uplatnění nebo studia.

Logické myšlení zahrnuje induktivní myšlení, díky kterému můžeme zobecňovat na základě konkrétního pozorování, a deduktivní myšlení, které naopak poskytuje konkrétní závěry na základě obecných pravidel. Řešení problémů znamená identifikaci problému, generování možných řešení, jejich vyhodnocení a výběr toho nejlepšího. Plánování pak zahrnuje schopnost definovat cíle, rozdělit je na jednotlivé kroky, prioritizovat, vytvořit časový plán a předvídat různá úskalí a možné překážky (Cristofori et al., 2019). Schopnost

logického myšlení a řešení problémů se překrývají s fluidní inteligencí, což pravděpodobně také vysvětluje vysokou korelaci g-faktoru a exekutivních funkcí (Diamond, 2013).

1.2. Biologický podklad exekutivních funkcí

S rozvojem neurovizuálních metod se více studií v posledních letech zabývá i biologickým podkladem kognitivních funkcí. To platí i o výzkumu exekutivy, které jsou částečně v souladu s modelem Unity and Diversity. Ty tvrdí, že exekutivní funkce jsou založeny na spolupráci různých mozkových oblastí, které se mohou překrývat a spolupracovat, ale zároveň mají každá své specifické funkce. Využívají přitom různé metody, včetně funkčních zobrazovacích metod jako je funkční magnetická rezonance (fMRI; Aron et al., 2014; Banich & Depue, 2015; Kim et al., 2012; Miyake & Friedman, 2017; Novák, 2009; Rottschy et al., 2012; Sylvester et al., 2003; Wager et al., 2005), pozitronová emisní tomografie (PET; Kim et al., 2012; Novák, 2009; Rottschy et al., 2012) nebo elektroencefalogram (EEG; Novák, 2009), které umožňují sledovat aktivitu mozku v reálném čase, nebo neuroanatomických metod jako je magnetická rezonance (MRI; Miyake & Friedman, 2017; Novák, 2009; Tsuchida & Fellows, 2013) nebo počítačová tomografie (CT; Novák, 2009; Tsuchida & Fellows, 2013), které umožňují sledovat anatomickou strukturu mozku. Obvykle je tento typ výzkumu založen na korelaci a zjišťuje, které oblasti jsou signifikantně aktivní při plnění úloh zaměřených na specifické exekutivní funkce. To se liší od původních studií zaměřených na léze ve frontální oblasti. Ty identifikují anatomické oblasti, které jsou pro jednotlivé úkony nezbytné. Vyplývá z nich, že frontální léze způsobují rozsáhlé až globální poškození exekutivních funkcí, což je konzistentní s obecným faktorem exekutivních funkcí; existují ale disociace mezi jednotlivými úkolovými situacemi, což poukazuje na diverzitu funkcí (Tsuchida & Fellows, 2009; 2013).

Z posledních výzkumů tohoto typu vyplývá, že jsou exekutivní funkce založeny na spolupráci frontálních a parietálních oblastí mozku. Jde o oblasti, které jsou součástí frontoparietální sítě, označované také jako „multiple-demand system”. Tento systém je aktivovaný při různorodých komplexních úkolech, zahrnujících úlohy zaměřené na inhibici, pracovní paměť a kognitivní flexibilitu. Na druhé straně ale výsledky studií také indikují existenci odlišných oblastí asociovaných s konkrétními exekutivními funkcemi (Miyake & Friedman, 2017).

Inhibice je schopnost potlačovat nežádoucí myšlenky, pocity a chování. Ukazuje se, že je spojena s aktivitou pravého prefrontálního kortexu (Banich & Depue, 2015) a pravého inferiorního frontálního kortexu (Aron et al., 2014). Wager et al. (2005) zjistili, že úlohy vyžadující inhibici aktivovali oblasti anteriorní insuly, anteriorního prefrontálního kortexu, pravého dorzolaterálního prefrontálního kortexu a dále premotorický kortex, putamen, anteriorní cingulum a parietální kortex. Pouze aktivita v oblasti anteriorní insuly ale korelovala s výkonem ve všech zadávaných úlohách. Ze zobrazovacích studií (Tsuchida & Fellows, 2013) dále vyplývá, že schopnost inhibice byla narušena u pacientů s lézí v oblasti levého ventrolaterálního prefrontálního kortexu.

Pracovní paměť je schopnost udržet si informace v paměti po krátkou dobu, manipulovat jimi a využívat je k řešení problémů. Dle recentních poznatků je spojena s bilaterální aktivitou ve fronto-parietální oblasti. Brocova oblast byla pak selektivně aktivní při verbálních úlohách, ventrální premotorický kortex byl aktivní při paměti na objekty a dorzální premotorický kortex při úlohách vyžadujících paměť na lokaci (Rottschy et al., 2012). Smolker et al. (2015) zjistili, že lepší výkon v oblasti pracovní paměti koreluje s objemem šedé hmoty v dorsolaterálním prefrontálním kortexu. V předchozí studii Tsuchida a Fellows (2009) identifikovali vztah pracovní paměti s lézí v levém a mediálním prefrontálním kortexu.

Kognitivní flexibilita, tedy schopnost měnit své chování v reakci na nové informace nebo požadavky, je spojena s aktivitou parietálního kortexu, levého dorzolaterálního prefrontálního kortexu, premotorického kortexu a mediálního frontálního kortexu (Sylvester et al., 2003). Kim et al. (2012) kognitivní flexibilitu asociuje s inferiorní frontální juncí a posteriorním parietálním kortexem. Percepční flexibilita byla spojena s dorsálním premotorickým kortexem a kontextová flexibilita s frontopolárním kortexem. Výzkum s pacienty s lézí (Tsuchida & Fellows, 2013) pak ukazuje, že zhoršení kognitivní flexibility souvisí s poškozením v levém ventrolaterálním prefrontálním kortexu.

Výzkumy tedy ukazují, že jsou exekutivní funkce založeny na spolupráci různých mozkových oblastí. Tyto oblasti se mohou překrývat a spolupracovat, ale zároveň mají každá své specifické funkce, přesto se studie zaměřené na identifikaci konkrétních funkcí jednotlivých oblastí plně neshodují. To naznačuje, že diverzita exekutivních funkcí může být interpretována řadou způsobů a že model Unity and Diversity nemusí být pro

vysvětlení exekutivních funkcí dostatečně komplexní. Tento model tedy není v úplném souladu s nejnovějšími poznatky z neuroanatomických studií (Cisek & Kalaska, 2010).

Ukazuje se například, že rozhodování je lokalizováno ve stejných senzomotorických oblastech mozku jako plánování a provedení akcí, je tedy obtížné kortikálním asociačním oblastem jednoznačně přiřadit jednotlivé funkce (Cisek & Kalaska, 2010), různé oblasti paralelně zpracovávají různé aspekty rozhodování (Pezzulo & Cisek, 2016).

Skrze modely exekutivních funkcí jsou tedy identifikovány určité zásadní složky exekutivních funkcí, není ale jisté, že jsou jedinými relevantními faktory. Studie, které se zaměřují na exekutivní funkce skrze pohled modelu Unity and Diversity, tento model často potvrzují, a to zejména u dospělých populací (Karr et al., 2018). V některých studiích se přesto objevovaly nové faktory, jako je verbální fluence nebo rozhodování (Tirapu-Ustárrroz et al., 2017). Je možné, že tato předchozí zjištění vedla k neúměrnému důrazu na tři funkce dle modelu Unity and Diversity, a to na úkor jiných relevantních funkcí nebo více obsáhlých klinických modelů. To, že byly tyto funkce identifikovány a studiem opakovaně potvrzeny, sice naznačuje, že se jedná o dobře definované a relevantní složky, nicméně neznamena to, že jsou nutně jedinými (Dias et al., 2023). Struktura faktorů a dokonce i konkrétní rozměry nebo dovednosti považované za exekutivní funkce se liší napříč studiem (Tirapu-Ustárrroz et al., 2018). Příkladem dalších, v literatuře nejčastěji zmiňovaných, složek na základě uvažování různých autorů jsou plánování, iniciace, teorie mysli nebo rozhodování (Tirapu-Ustárrroz et al., 2017). Tirapu-Ustárrroz et al. (2017) proto navrhli integrační model založený na zjištěních ze studií používajících různé metody (faktorové modely, studie úrazů a neurovizuální metody) a navrhli osm složek exekutivních funkcí: pracovní paměť, inhibici, flexibilitu, rychlost zpracování, verbální fluenci, duální provedení (souběžnost), plánování a rozhodování. Kromě toho navrhli novou složku, multitasking (neboli větvení), která by měla být v budoucnosti dále zkoumána (Dias et al., 2023).

Ukazuje se také, že existují i individuální rozdíly v neuroanatomické struktuře, které mají vliv na úroveň exekutivních funkcí. Jedinci s vyšší úrovní těchto funkcí mají snížený objem šedé hmoty ve frontoparietální síti a také vyšší hustotu šedé hmoty v těchto oblastech, což naznačuje hustší propojení neuronů. Individuální rozdíly v exekutivních funkcích jsou tedy spojeny s rozdíly v neuroanatomické struktuře a funkčních odlišnostech mozku (Miyake & Friedman, 2017). Smolker et al. (2015) zjistili na základě MRI, že lepší

výkon ve všech komponentách exekutivních funkcí souvisí s objemem šedé hmoty, konkrétně ve ventromediální oblasti pro obecný faktor exekutivních funkcí, v dorzolaterální oblasti pro pracovní paměť a ve ventrolaterální oblasti pro kognitivní flexibilitu. Naopak při studii založené na funkčním zobrazení pomocí fMRI došli výzkumníci k závěru, že obecný faktor exekutivních funkcí je spojen se silnějším propojením pravé frontoparietální sítě s mozečkem a vyšší úroveň kognitivní flexibility je spojena s rozšířením somatomotorické a ventrální pozornostní sítě o levý angulární gyrus (Reineberg et al. 2015).

Smolker et al. (2018) dále zjistili, že vysoká úroveň obecného faktoru byla spojena s větším objemem a povrchovou plochou oblastí v pravém středním frontálním gyru, pravém dolním temporálním gyru a s frakční anizotropií v částech pravého superiorního podélného fasciculu a levé anteriorní thalamické dráhy. Lepší schopnost pracovní paměti byla spojena s větší tloušťkou kortexu v oblasti levého cunea a precunea a sníženou tloušťkou kortexu v oblasti pravého superiorního frontálního gyru a pravého mediálního a superiorního temporálního gyru. Kognitivní flexibilita byla spojena se zvýšenou mediální difuzivitou a sníženou radiální difuzivitou v převážné části mozku a se s níženou axiální difuzivitou v odlišných klastrech levého superiorního podélného fasciculu, corporu callosum a pravé optické dráhy.

1.3. Vývoj exekutivních funkcí

Úroveň exekutivních funkcí v dětství predikuje nejen akademickou úspěšnost, ale i zdraví a well-being jedince. Ve věku 24 měsíců jsou exekutivní funkce již relativně stabilní a prediktivní hodnota budoucí úrovně těchto funkcí je vysoká. Do této doby takovou stabilitu nevykazují (Broomell & Bell, 2022).

Vývoj exekutivních funkcí je ovlivněn několika faktory. Prvním jsou genetické faktory, které se v raném věku mohou projevat reakcí na různé podněty. Tyto reakce mohou nabývat různé rychlosti a intenzity. Dalším faktorem je zrání, funkční diferenciací frontální kůry a propojení frontální kůry s dalšími oblastmi mozku. Na vývoj exekutivních funkcí mají ale vliv i environmentální faktory, jako je výchova nebo dítětem získané zkušenosti. Pokud dítě vyrůstá v podnětném prostředí, má také více příležitostí získat zkušenosti, které pro rozvoj těchto funkcí potřebuje (Doebel, 2020; Lynch et al., 2023; Miller et al., 2023).

V dětství je složité jednotlivé exekutivní funkce (inhibici, pracovní paměť a kognitivní flexibilitu) kvantifikovat, což vede k názoru, že jde v dětství o unitární faktor, který se separuje do diskrétních faktorů až později během vývoje. Existují už ale i empirické důkazy, které podporují v dětství dvoufaktorový model, složený z pracovní paměti a inhibiční kontroly (Miller et al., 2012). Přestože panuje nejistota ve struktuře exekutivních funkcí v raném dětství, zdá se, že jejich vývoj začíná v prvním roce života a pokračuje během vývoje až do adolescence (Blankenship et al., 2019; Broomell & Bell, 2022).

Prvotní projevy *inhibice* se objevují v dětství, když dítě začíná být schopné oddálit uspokojení. Raná schopnost inhibice se projevuje již kolem desátého měsíce (Cuevas et al., 2012). Tyto počátky exekutivy se od prvního roku zkvalitňují, přičemž nejvýraznější vývoj probíhá mezi třetím a pátým rokem (Diamond, 2013). Ve čtyřech letech jsou děti úspěšné jak v jednoduchých, tak komplexních úkolech zaměřených na inhibiční kontrolu. Inhibice se dále rozvíjí, a to především ve věku od pěti do osmi let (Romine & Reynolds, 2005), obzvláště v úlohách které kombinují inhibici a pracovní paměť (Carlson et al., 2005). Kognitivní, behaviorální i neurální vývoj inhibice ukazuje rychlé zlepšení v nízkém věku, které následuje pomalejším zkvalitňováním inhibiční kontroly až do adolescence. To je doprovázeno větší lokalizací této funkce v mozku. Tento vývoj se může dít skrze maturaci mozku, zvýšenou schopnost zvládnout komplexní úkoly, vyšší schopnost řídit se pravidly a rozvoj metakognice (Best & Miller, 2010).

Již děti ve věku osmi měsíců udrží mentální reprezentaci po krátkou dobu, což lze považovat za první měřítko *pracovní paměti* (Bell, 2012). Dle Gathercole et al. (2004) mají děti ve věku sedmi let pracovní paměť vyvinutou dostatečně, aby zvládly komplexní úkoly. Zjistili také, že jednoduché a komplexní úkoly zaměřené na pracovní paměť měly podobnou trajektorii vývoje, tedy lineární nárůst od čtyř do čtrnácti let a jejich vyrovnání mezi čtrnáctým a patnáctým rokem. Rozvoj pracovní paměti tedy záleží na komplexitě, to znamená, že jednodušší úlohy budou zvládnuty v nižším věku, zatímco s narůstajícím věkem dítě zvládne i narůstající komplexitu úlohy. Vývoj pracovní paměti tedy nezávisí na modalitě, ale na komplexitě, a probíhá až do adolescence. I v neurální aktivitě dochází k vývoji, především se projevuje lokalizovanou aktivitou ve frontoparietálním laloku, včetně dorzolaterálního prefrontálního kortexu (Best & Miller, 2010).

První projevy připomínající *kognitivní flexibilitu* se objevují u předškolních dětí ve věku od tří do čtyř let, pokud jsou pravidla v jednotlivých úkolech jednoduchá a jsou kladeny nízké nároky na inhibici. Zdá se ale, že jde o funkci, která není u předškolních dětí diferencovaná od pracovní paměti a inhibiční kontroly. To odpovídá předpokladu, že jsou pracovní paměť a inhibice prerekvizitou kognitivní flexibility (Best & Miller, 2010). K nejvýraznějšímu zlepšení dochází na počátku školní docházky, tedy kolem šesti let. U starších dětí a adolescentů dochází k jejímu dalšímu vývoji, který se projevuje v komplexnějších úlohách a v adolescenci dosahuje dospělé úrovně (Best & Miller, 2010).

V současné době nabývá výzkum v oblasti exekutivních funkcí u dětí na významu. Tradiční přístup k tomuto tématu se opíral o představu o exekutivních funkcích jako o separabilních komponentách, jako jsou pracovní paměť, inhibiční kontrola a kognitivní flexibilita. Nové empirické i konceptuální poznatky však vedou k pochybnostem o tomto tradičním paradigmatu. Nový pohled na vývoj exekutivních funkcí představuje Doebelová (2020). Zdůrazňuje, že vývoj exekutivních funkcí je spíše spojen s vytvářením dovedností v užívání kontroly ve prospěch konkrétních cílů, které jsou ovlivněny mentálním obsahem, jako jsou znalost, přesvědčení, normy, hodnoty a preference. Zatímco tradiční modely se soustředily na identifikaci a izolaci jednotlivých komponent exekutivních funkcí (Bell, 2012; Carlson et al., 2005; Cuevas et al., 2012; Gathercole et al., 2004; Miller et al., 2012; Romine & Reynolds, 2005), tento nový pohled tvrdí, že exekutivní funkce nelze oddělit od konkrétních cílů a kontextů, ve kterých se uplatňují. Důraz na mentální obsah, který je aktivován v rámci specifických úkolů, nám umožňuje lépe porozumět individuálním rozdílům v exekutivních funkcích a jejich vývoji. Pro praxi pak doporučuje, aby měření a intervence v oblasti exekutivních funkcí lépe reflektovaly skutečné cíle a kontexty, ve kterých se děti nacházejí. Namísto abstraktních laboratorních úkolů by se měly používat ekologicky validní metody, které lépe odráží skutečné situace, ve kterých se exekutivní funkce uplatňují. Navíc by intervence měly být navrženy s ohledem na konkrétní cíle a potřeby dětí a měly by brát v úvahu jejich mentální obsahy (Doebel, 2020).

1.3.1. Senzitivní období

Pro rozvoj exekutivních funkcí je nejdůležitější období raného dětství. To vychází z výzkumů zaměřených na senzitivní období vývoje (Andersen, 2018; Chen et al., 2022; Fiske & Holmboe, 2019; Gilman et al., 2019; Sharma et al., 2002; Sonuga-Barke et al., 2017; Thompson & Steinbeis, 2020). Senzitivní období je období vývoje, kdy je zvýšená

neurální senzitivita pro specifické environmentální stimuly. Při vystavení těmto stimulům dochází k typickému vývoji a po uzavření této periody je vývoj v dané oblasti omezen zkušenostmi ze senzitivního období. Exekutivní funkce jsou závislé na senzoričtých informacích, tedy i vývoj exekutivních funkcí je závislý na vývoji nižších systémů, které tyto informace sdružují, například na zrakové a sluchové percepci. Senzitivní období je spuštěno maturací GABAergního systému, tedy mediátorového systému syntetizujícího gamma-aminomáselnou kyselinu (GABA), hlavní inhibiční neurotransmiter centrálního nervového systému (Thompson & Steinbeis, 2020).

Poznatky o tomto období jsou často získávány ze studií zaměřených na deprivaci v určitém věku dítěte (Andersen, 2018; Sonuga-Barke et al., 2017; Thompson & Steinbeis, 2020). Deprivaci může být v takovém případě chudoba, zanedbání nebo například institucionální péče. V raném věku mají tyto události dlouhodobý efekt na vývoj exekutivních funkcí, a to především pokud se odehrávají v prvním půl roce života. Děti, které byly v institucionální péči a adoptovány do uplynutí prvních šesti měsíců, sice vykazovaly nižší úroveň exekutivních funkcí, ale nenabývaly klinické závažnosti, kdežto děti adoptované po uplynutí tohoto období častěji projevovali poruchy exekutivních funkcí, jako je například porucha aktivity a pozornosti (ADHD). Čím dříve tedy dochází ke kompenzaci deprivujících okolností, tím nižší jsou dopady na budoucí vývoj exekutivních funkcí. Přesto i relativně krátké působení těchto okolností má na jejich vývoj významný vliv (Sonuga-Barke et al., 2017).

Vývoj exekutivních funkcí odráží dozrávání prefrontálního kortexu, který v porovnání s ostatními mozkovými oblastmi dozrává pomalu a plně zralý je až v rané dospělosti. Šedá hmota nabývá na objemu v raném dětství přibližně do čtyř let, kdy dochází k synaptickému prořezávání, což vede k opětovnému snížení objemu šedé hmoty. Strukturální a funkční změny prefrontálního kortexu jsou zásadní pro rozvoj exekutivních funkcí a mají vliv na období, kdy může dojít ke stabilizaci těchto funkcí (Fiske & Holmboe, 2019). Individuální odlišnosti v dozrávání prefrontálního kortexu byly také spojeny s pozorovanými odlišnostmi ve výkonu v úkolových situacích zaměřených na exekutivní funkce (Broomell & Bell, 2022; McKenna et al., 2017).

1.4. Reflexe dosavadních modelů exekutivních funkcí

Mezi nejčastěji citované modely exekutivních funkcí patří kromě modelu Unity and Diversity dle Miyakeho (2000) také model Diamondové (2013), Lezakové (2004) nebo

model dle Stusse (2011). V předchozích kapitolách byly podrobně analyzovány teoretické modely exekutivních funkcí, s důrazem na nejčastěji citovaný model Unity and Diversity založený na faktorové analýze. Tento model identifikuje tři základní exekutivní funkce: pracovní paměť, kognitivní flexibilitu a inhibiční kontrolu.

Avšak i přesto, že Unity and Diversity model poskytuje strukturovaný rámec pro porozumění exekutivním funkcím, výzkum a diskuse v této oblasti se neustále vyvíjejí. Při prozkoumávání literatury a článků vznikají pochybnosti ohledně validity a úplnosti modelu, zvláště v souvislosti s výzvami spojenými s výzkumem.

Analýza literatury posledních let (Dias et al., 2023) odhaluje, že mnohé studie vycházejí z populací zvaných WEIRD (Western, Educated, Industrialized, Rich, and Democratic), což omezuje reprezentativnost a generalizovatelnost výsledků o lidské kognici a chování. Vzhledem k tomu, že kulturní vlivy mohou významně ovlivnit vývoj exekutivních funkcí, zdůrazňuje se nutnost získání empirických dat, která by objasnila tento vliv. Výzkum exekutivních funkcí by tedy neměl být omezen pouze na WIERD populaci, aby byla zajištěna vyšší validita.

Pochybnosti o validitě modelů exekutivních funkcí dále vyplývají z identifikace dalších exekutivních funkcí v rámci různých výzkumů, celkově jich bylo, mimo zmiňované tři vycházející z modelu Unity and Diversity, objeveno 76. Mezi nejčastější patří plánování, verbální fluence, abstraktní myšlení, řešení problémů, monitoring, konceptualizace, emocionální kontrola, organizace a plánování, rychlost zpracování, iniciace, selektivní pozornost a rozhodování. Dané oblasti zůstávají mimo dosavadní rámec modelů, včetně modelu Unity and Diversity, což naznačuje, že jeho současná podoba nemusí zcela pokrýt komplexní spektrum exekutivních funkcí (Dias et al., 2023).

Ani neuroanatomické studie nejsou plně v souladu s tímto modelem. Potvrzují sice existenci oblastí specifikovaných na jednotlivé funkce, které se ale dále překrývají a spolupracují, neshodují se však na identifikaci jednotlivých oblastí a zároveň poukazují i na další funkce kromě těch definovaných modelem Unity and Diversity. Tento model tedy není v úplném souladu s nejnovějšími poznatky z neuroanatomických studií (Cisek & Kalaska, 2010).

Unity and Diversity model, přestože poskytuje užitečný teoretický rámec pro studium exekutivních funkcí, čelí oprávněným pochybnostem ohledně své validity a

komplexnosti. Model zdánlivě omezuje svou schopnost zahrnout celé spektrum exekutivních funkcí a nedostatečně reflektuje kulturní rozmanitost v těchto funkcích. Otázky týkající se kulturních rozdílů, reprezentativnosti vzorků a kompletnosti modelů vedou k hledání nových dimenzí a přístupů k operacionalizaci a studiu exekutivních funkcí (Dias et al., 2023).

2. Reliabilita

Reliabilita je důležitý koncept v oblasti psychodiagnostiky a měření a hraje významnou roli při posuzování kvality a přesnosti výsledků získaných prostřednictvím různých testových metod. Umožňuje nám porozumět, do jaké míry jsou naměřené hodnoty konzistentní a svědčí o reálných rozdílech ve sledovaných jevech, přičemž zohledňuje vliv chyb měření. Veškerá variabilita v naměřených hodnotách vychází ze dvou základních zdrojů – skutečných rozdílů v měřené proměnné a chyb měření. Je nutné rozlišovat mezi manifestními (měřenými) a latentními (skutečnými) proměnnými, přičemž chyby měření jsou považovány za nepozorovatelné složky. Předpokladem pro správnou interpretaci reliability je také představa o nezávislosti chyb měření na hodnotách měřených atributů (Urbánek et al., 2011).

Z hlediska matematické statistiky je reliabilita definována jako podíl variability skutečných hodnot měřeného atributu k celkové variabilitě naměřených hodnot. Tato definice reflektuje snahu maximalizovat podíl latentní složky a minimalizovat šum v naměřených datech, což je nutné pro zajištění vysoké reliability měření. Přestože teoretická definice reliability ukazuje na podstatu této charakteristiky, je důležité zdůraznit, že v praxi nelze přímo měřit variability pravých skóre a náhodných chyb. Reliabilita je většinou odhadovaná pomocí různých metod, které zohledňují specifické charakteristiky chyb měření, a výsledky těchto odhadů jsou poté důležité pro interpretaci výsledků měření v různých oblastech vědeckého výzkumu (Urbánek et al., 2011).

2.1. Definice reliability

Reliabilita, neboli spolehlivost psychodiagnostické metody, představuje zásadní koncept ve vědeckém výzkumu a měření. Nicméně různí autoři přistupují k definici reliability s různými důrazy a perspektivami. Urbánek et al. (2011, s. 95) uvádí dvě možné definice:

„Reliabilita je charakteristika psychodiagnostické metody, která uvádí relativní nepřítomnost proměnných chyb v měření.“

„Reliabilita je v podstatě jiné označení pro spolehlivost nebo přesnost metody měření.“

Sürücü a Maslakçı (2020) zdůrazňují, že reliabilita je ukazatelem stability naměřených hodnot při opakovaných měřeních za stejných podmínek. Reliabilita není pouze vlastností měřicího nástroje, ale i výsledků tohoto nástroje. Hammersley (1987) pak zdůrazňuje, že mnoho definic reliability se orientuje na skóry opakovaných pokusů měřit stejný atribut stejným nástrojem. Uvádí také několik definic reliability od různých autorů, přičemž tyto definice se zaměřují na různé aspekty měření a jeho opakovatelnosti. Campbell a Fiske (1959) definují reliabilitu jako shodu mezi dvěma snahami o měření stejného atributu pomocí maximálně podobných metod, zatímco Lehner (1996) klade důraz na reprodukovatelnost měření a schopnost opakování stejných dat nejen samotným pozorovatelem, ale i dalšími, zdůrazňuje tedy reprodukovatelnost výsledků měření v různých situacích a prostřednictvím různých pozorovatelů. Jiné definice reliability, jako uvádí Johnston et al. (1981), se zaměřují na stabilitu měřicího nástroje pod stejnými podmínkami, tedy konzistenci měřicího nástroje a jeho schopnost poskytovat stejné hodnoty při opakovaném kontaktu se stejným jevem či stavem (Hammersley, 1987).

Z uvedených definic je patrné, že přístupy k reliabilitě se mohou lišit v závislosti na interpretaci, ať už se jedná o zaměření na reprodukovatelnost měření, stabilitu měřicího nástroje nebo schopnost konzistentně měřit požadovaný fenomén.

2.1.1. Standardní chyba měření

Podle teorie pravého skóru, vycházející z klasické teorie testů, je pozorovaný skór v testu součtem pravého skóru, který je neznámý, a chyby spojené s jeho měřením. Standardní chyba měření (SEM) poskytuje informace o přesnosti, s jakou pozorovaný skór odráží pravý skór. Znalost SEM a reliability testu nám umožňuje stanovit intervaly spolehlivosti a poskytuje nám informace pro interpretaci pozorovaných skórů v testech (Rust et al., 2021).

Chyby při odhadu reliability mohou mít různé zdroje a typy. Hammersley (1987) uvádí tři základní typy chyb: náhodné, konstantní a korelované chyby. Náhodné chyby se chovají tak, jako by jejich velikost a směr byly určeny náhodným výběrem, přičemž průměr těchto čísel je nula. Konstantní chyby nastávají, když průměr hodnot chyby není nula a vzniká nadhodnocení (nebo podhodnocení) o stejnou hodnotu. Korelované chyby jsou chyby související s různými skupinami, u kterých daný jev měříme, tvořenými na základě nějakého společného kritéria, například na základě pohlaví nebo socioekonomického statutu.

Porovnání skóre získaných za různých podmínek nám může pomoci posoudit účinky různých zdrojů a typů chyb. Například porovnáním skóre pro stejné objekty získané stejným pozorovatelem s použitím stejného nástroje v různých situacích můžeme odhadnout úroveň chyby způsobené variacemi mezi pozorováními. Stejně tak srovnáním skóre pro stejné objekty získaného různými pozorovateli s použitím stejného nástroje při stejné události můžeme odhadnout chybu odvozenou od nástroje. Kaplan a Saccuzzo (2018) zdůrazňují, že pozorované skóre se mohou lišit od pravého z mnoha důvodů, včetně situačních faktorů a nedostatečné reprezentativnosti testových položek.

Chyby měření jsou tedy běžným jevem v rámci testování a mohou významně ovlivnit přesnost získaných skóre. Reliabilita testu pak představuje relativní schopnost skóre odolávat těmto chybám měření. Mezi klíčové faktory, které mohou ovlivnit přesnost skóre testu, patří výběr obsahu (content sampling) a časový výběr (time sampling). Chyby způsobené výběrem obsahu (content sampling error) nastávají, když test měří pouze určitou část oblasti, kterou má zkoumat. Čím více položek obsahuje test, tím menší je pravděpodobnost, že tyto chyby budou mít významný dopad na konečný skóre. Vedle toho chyby způsobené časovým výběrem (time sampling error) vznikají, když test měří dovednosti nebo schopnosti, které se mohou v průběhu času měnit. Čím delší je interval mezi dvěma měřeními, tím vyšší je pravděpodobnost, že chyby způsobené časovým výběrem budou mít vliv na konečný skóre (Reynolds et al., 2021).

Content sampling error, chyba výběru obsahu, se definuje jako rozdíl mezi vzorkem úloh, které test obsahuje, a oblastí všech možných úloh, které by mohly být zahrnuty pro obsazení celé zkoumané oblasti. V praxi testy zřídka mohou zahrnovat všechny potenciální otázky nebo vyhodnocovat všechna relevantní chování, což otevírá prostor pro vznik této chyby. Content sampling error je obvykle považován za největší zdroj chyby v testových skórech. Jeho odhalení a správné zhodnocení mohou zásadně ovlivnit reliabilitu a platnost testových výsledků. Množství chyby měření způsobené výběrem obsahu je především dáno tím, jak efektivně a reprezentativně jsme vybrali vzorek úloh nebo otázek. Metody odhadu chyby výběru obsahu zahrnují analýzu míry podobnosti mezi jednotlivými úlohami. Je důležité, aby úlohy pokrývaly různé úrovně obtížnosti a reprezentovaly známé aspekty měřeného konstruktů, a proto je třeba tuto chybu pečlivě zvážit při konstrukci a interpretaci testů. Testy, které mají dobré vzorkování úloh, mají tendenci k nižší míře chyby výběru obsahu a jsou tak považovány za spolehlivější nástroje pro měření schopností a dovedností (Reynolds et al., 2021).

Time sampling error, chyba časového výběru, vzniká v důsledku náhodných změn u zkoušené osoby nebo vlivem prostředí. Změny u zkoušeného může způsobit například únava, nemoc nebo úzkost. Vliv prostředí při testování může souviset například s rušivými vlivy okolí. Tyto faktory mohou výrazně ovlivnit výkon zkoušené osoby a tím i výsledky testu. *Time sampling error* omezuje schopnost generalizovat výsledky testu napříč různými situacemi, což může ohrozit validitu testových zjištění. Je důležité, aby byly testy administrativně v kontrolovaných podmínkách a minimalizovala se tak chyba měření. To znamená, že prostředí zkoušení by mělo být co nejstabilnější a nejneutrálnější, aby se minimalizoval vliv vnějších faktorů na výkon zkoušené osoby. Takové testování má tendenci vést k validnějším a spolehlivějším výsledkům (Reynolds et al., 2021).

Chyby měření mohou být způsobeny i několika dalšími faktory, včetně chyb při sčítání skóre nebo při administraci testu. Chyby při sčítání skóre se mohou vyskytnout především při manuálním zpracování skóre. Chyby při administraci vznikají například chybným zadáním testového materiálu nebo nesprávným vysvětlením instrukcí. Rozdíly mezi hodnotiteli mohou také přispět k variabilitě v interpretaci výkonů zkoušených osob, například při hodnocení esejí (Reynolds et al., 2021).

2.1.2. Domain sampling model

Domain sampling model, základní koncept klasické testové teorie, se zabývá problémy vznikajícími při použití omezeného počtu položek k reprezentaci rozsáhlejšího a složitějšího konstruktů. Je užitečný především při snaze porozumět, jak počet položek v testu ovlivňuje jeho reliabilitu.

Tento model konceptualizuje reliabilitu jako poměr variance pozorovaného skóre na kratším testu k varianci pravého skóre. Obecně platí, že větší počet položek v testu povede k vyšší reliabilitě. Při konstrukci testu je důležité zajistit, aby každá položka adekvátně reprezentovala zkoumanou proměnnou. Díky náhodnému výběru položek by měl každý test nebo skupina položek poskytnout nezkreslený odhad pravého skóre, kvůli chybě způsobené výběrem vzorku ale mohou různé náhodné vzorky položek poskytnout odlišné odhady pravého skóre. Kdybychom vytvořili větší množství náhodných vzorků z jedné domény, odhady pravých skóre by měly mít normální rozložení.

Reliabilitu pak lze odhadnout výpočtem korelace mezi pozorovaným testovým skórem a pravým skórem. Pokud pravý skór není znám, lze jej odhadnout vytvořením několika paralelních testů a průměrováním korelací mezi nimi (Kaplan & Saccuzzo, 2018).

2.1.3. Item response Theory

Item Response Theory (IRT) reprezentuje inovativní metodologický přístup v psychometrii, který přináší efektivnější a spolehlivější mechanismy pro měření schopností jednotlivců. Oproti tradiční klasické testové teorii, která operuje s fixní sadou položek pro každého zkoumaného účastníka, IRT využívá algoritmy a adaptivní strategie pro selekci položek na základě úrovně schopností jedince. Tento adaptivní proces umožňuje testu přizpůsobit se individuálnímu schopnostem zkoumaného subjektu, což vede k preciznějším kvantifikování schopností (De Ayala & Santiago, 2016; DeMars, 2010; Templin, 2016). Díky IRT je možné získat důkladnější a podrobnější poznatky o schopnostech jednotlivců prostřednictvím kratších testových forem, což má význam zejména v oblastech jako je hodnocení inteligence či komplexních psychometrických konstruktů. Tento metodologický přístup vytváří nové možnosti a perspektivy pro výzkum v oblasti psychometrie a nabízí vědeckému i praktickému prostředí nástroje pro analýzu a interpretaci lidských schopností a behaviorálních projevů (Kaplan & Saccuzzo, 2018).

2.2. Typy reliability

Problematika teoretické definice reliability přináší do popředí výzvy spojené s určením podílu pravých skóru a náhodných chyb v naměřených hodnotách. Obě tyto proměnné jsou považovány za nepozorovatelné, latentní veličiny, jejichž přesné určení je často nemožné, a proto se spoléhá na jejich odhad. K odhadu reliability existuje několik přístupů, které vycházejí z různých definic náhodných chyb měření (Urbánek et al., 2011).

Jedním z nejčastěji používaných přístupů je zkoumání náhodných chyb z hlediska časových fluktuací naměřených hodnot, kdy se reliability chápe jako stabilita v čase, měřená test-retestovou metodou. Alternativní pohled na odhad reliability se zaměřuje na konzistenci v odpovědích testovaných osob, což vede k pojetí reliability jako vnitřní konzistence. Tuto konzistenci lze odhadovat různými metodami, včetně metody split-half nebo výpočtu Cronbachova koeficientu alfa. Vedle toho k výpočtu reliability můžeme přistoupit jako ke shodě mezi různými posuzovateli v rámci určitého výsledku. Toto pojetí

se využívá například ve shodě posouzení v testu tvořivosti nebo v zařazování odpovědí do kategorií projektivního testu (Urbánek et al., 2011).

2.2.1. Test-retest reliability

Test-retest reliability odkazuje na konzistenci výsledků získaných při opakované administraci daného nástroje na stejnou skupinu jedinců v různých časových obdobích. Zdrojem chyby je v tomto případě to, co způsobuje změnu skóru atributu měřeného opakovaně v odlišných časových bodech. Základním předpokladem této koncepce reliability je, že jsou měřené atributy stabilní v čase, proto se mluví o reliabilitě ve smyslu stability v čase. Probandům je při zjišťování test-retestové reliability zadán stejný dotazník nebo test dvakrát nebo vícekrát po sobě a odhad reliability spočívá v takovém případě v měření vztahu mezi jednotlivými výsledky měření, často vyjádřeném korelací (Urbánek et al., 2011). Vysoká korelace mezi srovnatelnými daty získanými v různých časových obdobích pak naznačuje vysokou test-retestovou reliabilitu nástroje.

Při určování reliability pomocí metody test-retest se obvykle využívá Pearsonův korelační koeficient, případně je možné data porovnávat pomocí *t*-testu. Je ale důležité mít na paměti, že pro provedení těchto analýz musí mít data normální rozložení (Oluwatayo, 2012). Pro interpretaci existují odlišné názory, přesto obecný konsenzus je, že korelace $r = 0,80$ a vyšší značí dobrou test-retest reliabilitu daného nástroje (Whiston, 2012).

Zásadní problém spočívá v tomto případě v předpokladu stálosti měřených proměnných, protože takový předpoklad obvykle není realistický, zejména pro měření v psychologii. Test-retest reliability se dá tedy využít pouze v případě měření atributů, které jsou alespoň relativně stabilní v časovém horizontu administrace všech pokusů. Při dlouhém časovém odstupu se i tyto atributy mohou měnit vlivem vnějších i vnitřních faktorů, což by způsobilo podhodnocení odhadu reliability. Zejména u dětí je nutné toto brát v potaz, jelikož právě u nich může docházet k významným změnám i v krátkém časovém rozmezí. Problém ale představuje i příliš krátké období mezi jednotlivými pokusy. Vlivem zapamatování odpovědí totiž může docházet k nadhodnocování reliability (Urbánek et al., 2011). Na doporučeném časovém odstupu jednotlivých měření ale není shoda. Někteří autoři se přiklání k delšímu rozestupu, například tři měsíce (Kline, 1993), dle Sürücüho a Maslakçihho (2020) je naopak vhodný kratší rozestup a uvádí interval 1–2 týdny.

Urbánek et al. (2011) argumentuje pro užívání test-retestové reliability navzdory nevýhodám, které se s ní pojí. Uvádí, že test-retest model je kromě odhadu reliability užitečný i pro posuzování samotné míry změny nebo stálosti proměnných. Dále tvrdí, že pouze odhad reliability na základě vnitřní konzistence je nedostatečný a zmiňuje, že u některých např. výkonových testů s rychlostní složkou ani není možné vnitřní konzistenci posuzovat.

2.2.2. Split-half reliability

Tento způsob odhadu reliability spočívá v rozdělení daného testu nebo dotazníku na dvě poloviny, ty se pak porovnávají a výsledná reliability je vyjádřena korelací těchto dvou polovin. V tomto případě je ale důležité přemýšlet nad tím, jak zajistit, aby byly obě poloviny ekvivalentní. Způsobů, jak test rozdělit, je několik. Nejlepší, ale zároveň nejsložitější postup, je skrze položkovou analýzu. Díky ní je možné nalézt položky podobné obsahově, obtížností i diskriminační účinností. Namísto tohoto komplikovaného postupu se ale často používá jednoduché rozdělení na první a druhou polovinu položek, tedy dle pořadí. Někdy se využívá i rozdělení na liché a sudé položky, což je vhodnější například u testů s rychlostní složkou nebo narůstající obtížností, kde by první a druhá polovina testu korelovala jen velmi málo (Urbánek et al., 2011).

Tento přístup k odhadu reliability řeší problémy spojené s variantou test-retest i paralelních forem. Vyhýbá se totiž změně měřeného atributu v čase, která může při těchto variantách nastat a způsobila by podhodnocení skutečné reliability. Zároveň se se split-half metodou nepojí riziko zapamatování odpovědí při příliš malém rozestupu mezi dvěma testováními a následného nadhodnocení reliability. Kromě toho také není nutné vytvářet paralelní formu, což je jednak složitý proces, jednak není tato forma nikdy dokonale paralelní (Urbánek et al., 2011).

Přináší s sebou ale také nový problém. Jelikož reliability testu roste s vyšším počtem položek, při snížení položek na polovinu se hodnota odhadu reliability snižuje (Helmstadter, 1964; Kline, 1993) a může tak docházet k jejímu podhodnocení. Ke kompenzaci tohoto snížení se využívá Spearmanův-Brownův vzorec. Kromě toho, že dokáže odhadnout reliability v případě split-half metody, umožňuje také odhadnout potenciální zvýšení nebo snížení reliability při změně počtu položek. Předpokládá ale, že přidané položky mají srovnatelné vlastnosti jako položky, ze kterých byl vypočítán odhad reliability testu. Spearmanův-Brownův vzorec nabývá hodnot od 0 do 1 a za reliabilní se

považuje test dosahující hodnoty 0,70 a vyšší. Škály a testy o méně než deseti položkách by tímto vzorcem neměly být měřeny (Kline, 1993). Předpokládá rovnost rozptylů obou polovin testu a také to, že jednotlivé položky měří stejný atribut. Někdy není možné předpoklad shodnosti rozptylů naplnit a v takovém případě se dá využít Guttmanův (Helmstadter, 1964) vzorec, který to nevyžaduje (Urbánek et al., 2011).

2.2.3. Item-total korelace

Item-total korelace je další způsob, jakým se dá k odhadu reliability přistoupit. Měří, jak moc souvisí jednotlivé položky s celkovým skórem daného testu. Obecně se za vhodné položky považují takové, které nabývají hodnoty od 0,30 do 0,80. Položky s nižšími hodnotami neměří daný koncept a položky s vyšší hodnotou se naopak vztahují pouze ke specifické oblasti. Pro tento způsob odhadu je vhodné mít mezi 100 a 200 respondentů (Sürücü & Maslakçı, 2020).

2.2.4. Reliabilita jako vnitřní konzistence

Přístup k odhadu reliability jako vnitřní konzistence reaguje na nedostatky split-half metody, přičemž neměří pouze korelaci dvou polovin testu, ale spíše korelaci všech možných dvojic položek mezi sebou. Cílem je zajištění vysoké reliability, což předpokládá, že probandi by měli na jednotlivé otázky odpovídat podobně. To vede k pozitivní a vysoké korelaci všech dvojic položek a tím k vysoké reliabilitě celého testu (Urbánek et al., 2011). Nejčastěji se k odhadu reliability jako vnitřní konzistence používá Cronbachův koeficient alfa (Cronbach, 1951; Urbánek et al., 2011).

Cronbachův alfa koeficient, vyjádřený statistickým ukazatelem v rozmezí 0 až 1, patří mezi nejefektivnější metody měření interní konzistence. Čím vyšší hodnota koeficientu, tím silnější je interní konzistence nástroje. Interpretace hodnot Cronbachova alfa koeficientu je následující:

- $\alpha \geq 0,90$ – Vysoká interní konzistence: Měřicí nástroj je velmi spolehlivý a jeho výsledky jsou konzistentní.
- $0,70 \leq \alpha \leq 0,89$ – Přijatelná interní konzistence: Měřicí nástroj je spolehlivý, i když výsledky mohou vykazovat menší variabilitu.
- $0,50 \leq \alpha \leq 0,69$ – Slabá interní konzistence: Měřicí nástroj je méně spolehlivý a výsledky mohou být značně proměnlivé.

- $\alpha \leq 0,49$ – Chybí interní konzistence: Měřicí nástroj není spolehlivý a výsledky jsou nepředvídatelné.

Vyšší interní konzistence indikuje větší reliabilitu měřicího nástroje a konzistentnější výsledky (Sürücü & Maslakçı, 2020).

2.2.5. Kuderova-Richardsonova reliabilita

Speciální případ Cronbachova koeficientu alfa představuje Kuderův-Richardsonův koeficient KR20, který se využívá zejména u dichotomických položek. Tento koeficient byl publikován již v roce 1937, před samotným Cronbachem (Cronbach, 1951), a později Cronbach ve své analýze koeficientů homogenity provedl jeho obecnější formulaci (Urbánek et al., 2011).

Primárním účelem Kuderova-Richardsonova koeficientu KR20 bylo měření reliability výkonových testů s rychlostní složkou, kde se kladl důraz na to, aby byly položky řazeny podle rostoucí obtížnosti. Tato vlastnost testu byla označována jako homogenita, což podle Helmstadtera znamená konzistenci výkonů v sadě testovacích položek. Helmstadter zdůraznil, že pokud jsou položky vhodně uspořádány podle obtížnosti, měl by proband řešit jednotlivé položky postupně, až dosáhne svého osobního „stropu“, kde už není schopen řešit žádnou další položku. V ideálním případě by výsledek všech probandů vykazoval podobný vzor, kde každý dosáhne bodu, před nímž vyřeší všechny položky a za nímž nedokáže vyřešit žádnou další, každý proband by tento bod měl na jiném místě podle své schopnosti. Takový test by podle Kudara a Richardsona (Helmstadter, 1964) projevoval dokonalou reliabilitu.

Přestože Helmstadter hovoří o homogenitě ve smyslu jednodimenzionality testu, ve skutečnosti se při popisu reliability výkonového testu pravděpodobně odvolává na škálovatelnost ve smyslu Guttmanovy škály, a ne na homogenitu, která je v podstatě totožná s vnitřní konzistencí. Kuderův-Richardsonův koeficient KR20, pracující s dichotomizovanými daty, však stále poskytuje užitečnou metodu pro výpočet reliability. V případě potřeby odhadu konzistence nebo homogenity pro položky konstruované ve variantě s pravdivou a nepravdivou odpovědí je KR20 vhodnějším odhadem (Urbánek et al., 2011).

2.2.6. Reliabilita jako shoda posuzovatelů

Předchozí analýzy reliability se zaměřovaly na situace, kdy bylo nutné zhodnotit spolehlivost měřícího nástroje s danou rozlišovací škálou, ať už se jednalo o škálu dichotomickou, ordinální, intervalovou nebo poměrovou. Nicméně v praxi se často setkáváme i se situacemi, kdy měření nebo přesněji zhodnocení odpovědí či výsledků aktivity vyžaduje účast interpreta. Příkladem může být hodnocení esejí, kde každý hodnotitel upřednostňuje různé aspekty. Dokonce i při hodnocení podle strukturovaných směrnic bude každý interpret vnášet do výsledků svou osobní perspektivu. Je evidentní, že při zapojení více interpretů do hodnocení téhož materiálu nemusí docházet vždy k dosažení shody ve výsledcích (Urbánek et al., 2011).

V nejjednodušším scénáři, kdy hodnotící stupnice odpovídá alespoň ordinální proměnné a kdy je k dispozici pouze dvojice hodnotitelů, lze pro určení shody mezi hodnotiteli použít korelační koeficient (Rust & Golombok, 1999). V případě, že je k dispozici větší počet hodnotitelů, kteří hodnotí stejný materiál, lze k posouzení shody využít koeficient konkordance nebo případně vnitrotřídní korelaci. V situacích, kdy jsou interpreti schopni klasifikovat jevy na nominální úrovni, může být odhad shody proveden pomocí koeficientů kappa (Urbánek et al., 2011).

2.2.7. Reliabilita rozdílového skóru

Rozdílový skór představuje numerickou hodnotu, která vychází z rozdílu mezi dvěma měřeními hodnotami. Jeho hlavním účelem je srovnání výsledků nebo schopností jedince v různých kontextech či časových obdobích. Například může sloužit k vyhodnocení změn v individuálním výkonu před a po absolvování specifického tréninkového programu či k porovnání úspěšnosti v různých oblastech, jako je matematické a čtenářské dovednosti. Rozdílový skór tak poskytuje analytický nástroj pro zhodnocení vývoje či změn výkonů, avšak vyžaduje obezřetný přístup při interpretaci a analýze dat vzhledem k jeho specifickým charakteristikám a omezením (Kaplan & Saccuzzo, 2018).

Reliabilita rozdílového skóru se týká míry, do jaké je skór konzistentní a spolehlivý při opakovaném použití nebo při aplikaci na stejnou populaci či stejné jedince. Jedná se o ukazatel, který vyjadřuje míru, do jaké můžeme důvěřovat rozdílu mezi dvěma měřeními hodnotami a jejich interpretaci. Vyšší reliabilita rozdílového skóru naznačuje, že změny

nebo rozdíly mezi dvěma měřeními jsou pravděpodobněji odrazem skutečného vývoje než pouhých náhodných variací nebo chyb měření. Reliabilitu rozdílového skóru lze obvykle vyjádřit pomocí korelačních koeficientů nebo jiných statistických metod, které berou v úvahu reliabilitu jednotlivých měření a jejich vzájemnou korelaci (Kaplan & Saccuzzo, 2018).

Při vytváření rozdílového skóru, které zachycuje rozdíl mezi dvěma měřenými hodnotami, čelíme několika výzvám. Jednou z nich je zvýšená chyba měření, protože tato chyba zahrnuje nedokonalosti z obou původních měření. Dále je reliabilita rozdílového skóru často nižší než u původních měření, což omezuje jeho využitelnost a přesnost při interpretaci výsledků. Tento pokles reliability může být dále ovlivněn závislostí na korelaci mezi dvěma původními měřeními; v případě nízké korelace může dojít k dalšímu snížení reliability rozdílového skóru. Tyto výzvy s reliabilitou rozdílového skóru mohou komplikovat jeho výklad a aplikaci v praxi, zejména v oblasti diagnostiky nebo pedagogiky. Odhady reliability jsou často založeny na pozorování populací, což může být problematické při individuálním použití. Je proto nezbytné přistupovat k interpretaci výsledků s opatrností a zvážit další faktory ovlivňující reliabilitu měření (Kaplan & Saccuzzo, 2018).

2.2.8. Zvyšování reliability nástroje

Zlepšení reliability testu je často žádoucí při jeho konstrukci a interpretaci. Existuje několik strategií, které mohou být aplikovány k zajištění toho, aby test poskytoval konzistentní a spolehlivé výsledky (Reynolds et al., 2021).

Nízká hodnota Cronbachova alfa koeficientu může být způsobena několika faktory. Jedním z nich je nedostatečná interní konzistence měřícího nástroje, což může být způsobeno nedostatečným počtem položek (Sürücü & Maslakçı, 2020). V takovém případě je možností zvýšení počtu položek v testu. Tímto způsobem lze zvýšit reliabilitu testu, zejména pokud jsou přidané položky stejně kvalitní jako ty původní. V případě, že je limitován počet položek, které lze do testu zahrnout, je možné reliabilitu zvýšit použitím více měření, která jsou poté kombinována do průměrného nebo kompozitního skóru (Reynolds et al., 2021).

Další možností je získání adekvátnějšího vzorku položek ze zkoumané domény. Tím se zajistí, že test obsahuje rozmanité položky, které pokrývají širokou škálu úrovní

obtížnosti a měřených aspektů domény, a lze tak zvýšit reliabilitu měření. Je také důležité vybírat a vyvíjet položky s dobrými měřicími charakteristikami. K tomu je užitečná položková analýza, která identifikuje reliabilní a validní položky dané testové metody. Dále je důležité dbát na správné postupy při administraci a skórování testu (Reynolds et al., 2021). Důvodem nízké hodnoty Cronbachova alfa může být i neznalost nebo opomenutí reverzních položek použité škály (Sürücü & Maslakçı, 2020).

V případě, že alfa koeficient dosahuje hodnoty 0,40 a nižší, může to signalizovat, že většina účastníků vyplnila dotazník bez přečtení nebo poskytla náhodné odpovědi. V takovém případě je vhodné provést důkladnou analýzu dat a vyloučit systematicky vyplněná nebo náhodně vyplněná data ze zkoumaného rozsahu. Toho se dá docílit například odstraněním outlierů (Sürücü & Maslakçı, 2020).

Není však vždy vhodné usilovat o co nejvyšší hodnotu Cronbachova alfa koeficientu. Kritika této snahy směřuje k tomu, že vyšší interní konzistence může vést ke snížení validity měřicího nástroje. Pokud jsou položky testu nebo dotazníku interně konzistentní, avšak obsahově specifitější, mohou mít nižší validitu. Proto je důležité zohlednit i další kritéria při výběru položek, jako je například diskriminační účinnost. Úzkým zaměřením na interní konzistenci bychom mohli ztratit důležité informace o testovaném jevu a omezit tak vysvětlující schopnost měřicího nástroje (Urbánek et al., 2011).

3. Reliabilita paralelních forem a měření exekutivních funkcí

Reliabilita paralelních forem je klíčovým prvkem při posuzování konzistence výsledků psychometrických testů, což je zásadní pro validní měření kognitivních funkcí, včetně exekutivních funkcí (Urbánek et al., 2011). Exekutivní funkce jako inhibice, flexibilita, pracovní paměť a plánování jsou nezbytné v situacích, kdy nestačí automatické nebo intuitivní reakce, a umožňují nám aktivně volit cíle, plánovat strategie a adaptivně reagovat na nové podněty (Diamond, 2013). Přesná a spolehlivá měřítka těchto funkcí jsou proto nezbytná pro pochopení jejich role v každodenním fungování.

3.1. Odhad reliability paralelních forem

Reliabilita paralelních forem je metoda odhadu reliability testu, která byla původně navržena Helmstadterem (1964) s cílem eliminovat vliv paměti při test-retestovém způsobu odhadu reliability. Používá dvě odlišné ale ekvivalentní verze testu, které měří stejnou vlastnost (Rust et al., 2021). Díky tomu zajišťuje v porovnání s jinými metodami odhadu reliability vyšší úroveň srovnatelnosti. Doložení reliability paralelních forem je také důležité pro vytvoření alternativní testové metody s minimálním rizikem efektu zácviků. Odhad reliability paralelních forem předpokládá, že chybou měření je rozdíl ve skórech obou forem testu (Urbánek et al., 2011).

Posouzení shody mezi paralelními formami lze provést pomocí expertního posudku, detailní analýzy položek (McDonald, 1999) nebo strukturního modelování (Urbánek, 2000). V praxi je důležité empirické posouzení reliability paralelních forem, které je založeno na statistické analýze průměrů, rozptylů a korelací mezi položkami obou forem. Přestože se často využívá korelace pro odhad reliability, v případě paralelních forem není tato statistika vždy postačující, ačkoliv je často jedinou uváděnou (Urbánek et al., 2011).

Tato metoda k posouzení reliability je často považována za nejlepší formu jejího odhadu, protože je méně ovlivněna faktory jako je učení nebo motivace (Kaplan & Saccuzzo, 2018; Rust et al., 2021). Zdá se, že technika paralelních forem je podobná metodě test-retest reliability, s výjimkou použití dvou odlišných měřících nástrojů k měření stejného chování nebo kvality. Proto některá omezení metody odhadu test-retestové reliability platí také pro metodu paralelních forem (Sürücü & Maslakçı, 2020). Nejzásadnějším problémem je předpoklad stálosti měřených proměnných, což je zejména v psychologii nereálné a zjednodušující. Zkoumané atributy se mohou vlivem vnějších i

vnitřních faktorů měnit, a to zejména u dětí i v krátkém časovém rozmezí. Jde tedy o odhad reliability vhodný v případě měření atributů, které jsou alespoň relativně stabilní v časovém horizontu administrace obou pokusů (Urbánek et al., 2011). Oproti test-retestové reliabilitě je reliabilita paralelních forem navíc náročnější na konstrukci, protože vyžaduje vytvoření dvou srovnatelných verzí testu a je tak méně často používána než metoda test-retest (Kaplan & Saccuzzo, 2018; Rust et al., 2021).

Při tvorbě paralelních forem u testů exekutivních funkcí se objevuje ještě další problém, související s efektem učení. V testech zaměřených na exekutivu mají participanti za úkol generovat nové strategie nebo rozpoznat abstraktní koncepty. Při řešení paralelní formy testu, i když jednotlivé úlohy vypadají jinak, může docházet k rozvzpomenutí na strategie, které užili v minulém řešení ekvivalentních úloh, a tak mohou dosahovat v druhém testování lepšího výsledku. Tvorba paralelních forem testů exekutivních funkcí je tak problematická a většina takto zaměřených testových metod paralelní formu nemá (Kurtz et al., 2004).

3.1.1. Vyrovnávání alternativních forem testu

Obtížnost vytváření skutečně paralelních forem vedla k vývoji postupů vyrovnávání (equating), které se snaží maximalizovat srovnatelnost skóre a tím umožňuje tvorbu alternativních forem testu pro praktické využití (Urbánek et al., 2011).

Proces vyrovnání testů zahrnuje úpravu skóru tak, aby byly kompenzovány rozdíly v obtížnosti mezi různými verzemi testů. Vyrovnané skóry umožňují porovnání výkonů jednotlivců napříč různými verzemi testů. Jde tedy o statistické metody, které se používají k dosažení srovnatelnosti skóre mezi dvěma nebo více verzemi testu. Skóry na těchto různých testových formách, které mají měřit stejný atribut, pak mohou být použity vzájemně zaměnitelně (González & Wiberg, 2017).

Podle Livingstona (2004) jsou testy, které měří stejný atribut, považovány za ekvivalentní, pokud umisťují jednotlivce do stejné relativní pozice ve skupině. Existuje několik metod vyrovnávání, zahrnující průměrné, lineární a equipercenilní vyrovnávání, které se liší ve způsobu, jakým definují relativní pozice. Každá z těchto metod vyžaduje specifické úpravy skóru, aby se dosáhlo srovnatelnosti výsledků mezi různými testovacími formuláři, přičemž každá má své vlastní matematické vyjádření a postup (Gross et al., 2012).

Equipercenilní vyrovnávání je jedním ze způsobů vyrovnávání, který identifikuje skóry na prvním testovém formuláři, které mají stejné percentilové pořadí jako skóry na druhém formuláři testu (González & Wiberg, 2017). Relativní pozici definuje pomocí percentilového pořadí skórů ve skupině (Gross et al., 2012).

Dalším typem je lineární vyrovnávání, které se používá k vyrovnání skórů pomocí lineární funkce a opírá se o standardní odchylky od průměru skupiny (González & Wiberg, 2017; Gross et al., 2012). Tato funkce přiřazuje každému skóru na jednom formuláři odpovídající skór na druhém formuláři tak, aby byla zachována lineární vztahová struktura mezi nimi. Lineární vyrovnávání se často používá v případě, kdy je potřeba skóry vyrovnat rychle a efektivně (González & Wiberg, 2017).

Průměrné vyrovnávání je další způsob vyrovnávání, který předpokládá, že formuláře testů se liší pouze v průměrech skórů a zaměřuje se na absolutní rozdíl od průměru. Tato metoda je často používána pouze pro ilustrativní účely, protože může být příliš zjednodušující (González & Wiberg, 2017; Gross et al., 2012).

Jinou populární metodou je Kernelovo vyrovnávání, které kombinuje parametrické a neparametrické techniky, protože se na odhadu podílejí jak konečné parametry, tak distribuční funkce. Tato metoda odhaduje parametry rozdělení pravděpodobnosti skórů parametricky, zatímco funkce rozdělení skórů jsou odhadovány neparametricky. To umožňuje flexibilní a robustní vyrovnání skórů, což je užitečné především u dat, pro která neznáme rozdělení skórů (González & Wiberg, 2017).

Equipercenilní vyrovnávání se zdá jako nejvhodnější pro neuropsychologické testování jednotlivce, protože nejlépe zohledňuje rozdíl v obtížnosti jednotlivých forem testu. Jde o metodu s širokou možností aplikace v klinickém i výzkumném prostředí a umožňuje využití dvou neekvivalentních forem testu. Přestože se ukázala tato metoda jako ideální, nemusí být vhodná ve všech případech (Gross et al., 2012).

3.2. Reliabilita paralelních forem ve vybraných testech exekutivních funkcí

Pro měření exekutivních funkcí existuje široká škála metod zaměřených na jednotlivé funkce, ale i testových baterií, které se snaží pojmout exekutivní funkce jako komplex. Některé metody vychází z různých teoretických modelů, jiné z klinické praxe.

Diamondová (2013) uvádí metody měření jednotlivých exekutivních funkcí podle modelu Unity and Diversity. Pro měření inhibice navrhuje například Stroopův test (Krivá, 2013; Stroop, 1935), Simonův test (Simon, 1969), Flanker test (Eriksen & Eriksen, 1974), antisakádové testy, testy oddáleného uspokojení, go/no-go testy a stop-signal testy. Mezi zkouškami pracovní paměti zmiňuje opakování čísel v opačném pořadí, reorganizaci slyšených položek, Corsiho test (Corsi, 1972), SOPT (Self-Ordered Pointing task; Petrides & Milner, 1982) nebo n-back úlohy. A pro kognitivní flexibilitu se užívají testy neobvyklého použití, testy verbální a kategoričké fluence, dále WCST (Wisconsinický test třídění karet; Grant & Berg, 1948; Telecká, 2013) nebo DCCS (Dimensional Change Card Sort Test; Zelazo, 2006).

V měření ostatních exekutivních funkcí, mimo inhibici, pracovní paměť a kognitivní flexibilitu, jsou podle Diasové et al. (2023) mezi deseti nejčastěji používanými metodami testy verbální fluence, Test cesty (Bezdiček et al. 2012; Tombaugh, 2004), Stroopův test (Krivá, 2013; Stroop, 1935), WCST (Grant & Berg, 1948; Telecká, 2013), BRIEF (Škála hodnocení exekutivních funkcí u dětí, Behavior Rating Inventory of Executive Function; Gioia et al., 2015; Ptáček, 2011), Londýnská věž (Bezdiček et al., 2018; Shallice, 1982), opakování čísel, test neobvyklého použití, Rey-Osterriethova komplexní figura (Košč & Novák, 1997; Osterrieth, 1944; Rey, 1941) a D-KEFS (Delis-Kaplan Executive Function System; Delis et al., 2001). Kulišťák (2017) ještě kromě zmíněných uvádí test přísloví, Podobnosti ve Wechslerově inteligenčním testu (Černochová et al., 2010; Wechsler, 1997), Perceptual Maze test (Elithorn, 1955) a Hanojskou nebo Torontskou věž.

Pro měření exekutivních funkcí u dětí jsou vhodné baterie D-KEFS (Delis-Kaplan Executive Function System; Delis et al., 2001), CANTAB (The Cambridge Neuropsychological Test Automated Battery; Wild & Musser, 2013), TEA-Ch (The Test of Everyday Attention for Children; Manly et al., 1998), BADS-C (The Behavioural Assessment of the Dysexecutive Syndrome in Children; Emslie et al., 2013) nebo NEPSY II (A Developmental Neuropsychological Assessment, Second Edition; Kemp & Korkman, 2010). Dále se dají využít například testy LSPAN (Listening Span Task), opakování čísel v opačném pořadí, odd one out test, WCST (Grant & Berg, 1948; Telecká, 2013) nebo bludiště (Henry & Bettenay, 2010).

Metod měření exekutivních funkcí je pro výzkum i klinické využití mnoho. Následující kapitola se bude zabývat některými nejčastěji používanými z nich, a to v kontextu paralelních forem v měření exekutivy.

3.2.1. Verbální fluence

Verbální fluence je produkce plynulé řeči, která vyžaduje schopnost flexibilně přemýšlet a organizovanost myšlení. Je také nepřímým indikátorem krátkodobé paměti, kvůli nutnosti udržovat povědomí o tom, která slova již byla řečena. Spočívá ve vymýšlení co nejvyššího počtu slov dle určitého kritéria za omezený čas.

Při vymýšlení slov pro test verbální fluence se lidé často zaměřují na tvorbu klastrového uspořádání slov. Existují dva hlavní typy klastrování: fonologické klastry a sémantické klastry. Fonologické klastry spojují slova podle jejich prvního zvuku nebo homonyma, zatímco sémantické klastry spojují slova podle jejich významu nebo kategorie. Po vyčerpání jednoho klastru musí jedinec efektivně přejít na další, což vyžaduje flexibilitu myšlení (Lezak, 2012).

První test verbální fluence, původně psaný, publikovali Thurstone a Thurstonová (1962). Později Benton (1968) vytvořil COWAT/FAS (Controlled Oral Word Association Test), což je mluvená verze testu se třemi hláskami (F, A, S). Jeho další verze (Benton et al. 1994) byla vyvinuta v rámci Multilingvální examinační afázie pro písmena C, F, L a P, R, W. Písmena byla vybírána na základě frekvence slov, které na ně v angličtině začínají: první písmena C a P mají vysokou frekvenci, zatímco F a R nižší a L a W nejnižší. Úkolem je pak vymyslet co nejvíce slov začínajících na dané písmeno během jedné minuty, jde tedy o tři série slov po jedné minutě.

Fonemická verbální fluence, jako je COWAT/FAS, se ukázala být náročnější a mít menší variabilitu výkonů než kategorická fluence, která je jednodušší a zahrnuje vymýšlení co nejvíce slov v dané kategorii, například zvířata nebo ovoce a zelenina (Lezak, 2012).

V prostředí České republiky je test oblíbený a často využívaný v pedagogicko-psychologické diagnostice i klinické psychologii. Nejčastěji se používá u dospělé populace především pro diagnostiku mírné kognitivní poruchy nebo demence (Cígler & Durmeková, 2018). Pro tu existují například normy pro osoby nad 59 let (Nikolai et al., 2015). Chybí však aktuální normy pro celou dětskou populaci, i pro ni však může být test verbální fluence užitečný pro diagnostiku neurovývojových a

neurodegenerativních onemocnění, psychiatrických onemocnění nebo poškození mozku (Villalobos et al., 2022). V současnosti existují normy jen pro věkově omezené populace, jako jsou studenti druhých ročníků středních škol (Cimlerová et al., 2007), děti z Prahy ve věku 9–14 let (Preiss et al., 2012) nebo děti ve věku 5–12 let (Cígler & Durmeková, 2018).

Ekvivalenci hlásek ve fonemické fluenci se na české populaci zabývali Kopeček a Kuncová (2006). Dobrovolníkům zadali test verbální fluence pro hlásky N, K, P, B, D, L, M, T, R, S. Mezi nimi hledali takové páry hlásek, které budou signifikantně korelovat a zároveň mezi nimi nebude signifikantní rozdíl ve skórech při retestu. Na základě těchto kritérií identifikovali dvojice hlásek NB, KP a TL. Efekt učení byl signifikantní u obou verzí, celkový skóre se mezi měřeními nelišil signifikantně a pro obě měření obě verze signifikantně korelovaly při prvním i druhém měření. Studii uzavírají s tím, že jsou verze NKT a BPL alternativní. Již předtím se Štorková et al. (2004) pokoušela najít ekvivalentní verzi k české verbální fluenci pro hlásky NKP. Na základě frekvence výskytu slov s danými počátečními písmeny jako ekvivalentní formu ověřovali test verbální fluence pro hlásky VRS, zjistili ale, že je tato verze signifikantně složitější a není tak jako alternativa pro retestování vhodná.

Novější českou studií alternativní verze testu verbální fluence je studie Pastrnáka et al. (2018). Autoři použili jako alternativní test verbální fluence pro písmena NKP test s písmeny BTL. Pro kategorickou fluenci použili pro kategorie zvířata a zelenina alternativní kategorie jména a profese. Alternativní forma pro fonemickou verbální fluenci se ukázala jako složitější a není tak považována za ekvivalentní. Duální sady v kategorické fluenci ekvivalentní také nebyly, ale při hodnocení kategorií samostatně se jako ekvivalentní ukázaly kategorie zvířata a jména.

I původní Bentonova verze v rámci Multilingvální examinace afázie obsahovala dvě paralelní verze fonemické fluence, a to pro hlásky C, F, L a paralelně pro P, R, W (Benton et al., 1994). Normy pro obě trojice hlásek existují pro dospělé různých věkových kategorií a vzdělání (Lezak, 2012). Ekvivalence byla zjišťována u studentů ($N = 66$). Nebyly zjištěny žádné rozdíly průměru pro žádný index a pouze jeden (průměrná velikost klastru) se lišil při zkoumání standardních odchylek. V druhé části studie se zjistilo, že použití obou verzí generuje srovnatelná normativní data a korelace mezi skóry. Tyto studie naznačují, že verze C, F, L a P, R, W mohou být vnímány jako alternativní formy a že pro většinu účelů produkují srovnatelné výsledky (Ross et al., 2006).

3.2.2. *Wisconsinský test třídění karet*

Wisconsinský test třídění karet (WCST) byl vyvinut k odhadu schopnosti abstrakce a přesunu kognitivních strategií (Grant & Berg, 1948). V osmdesátých letech byl WCST standardizován a skórován Heatonem, čímž se stal vhodným pro klinickou praxi (Heaton, 1981). Další revidovaný manuál WCST z roku 1993 obsahuje upřesnění pravidel skórování a normy pro široký věkový rozsah (Heaton et al., 1993). V prvním formálním vydání testové příručky je WCST označován jako test měřící schopnost strategického plánování a účelného jednání s využitím zpětné vazby (Heaton, 1981).

WCST je úkolem párování karet, kde účastníci párují odpovědní kartu s jednou ze čtyř podnětových karet bez specifické instrukce od administrátora. Test končí buď po dokončení všech šesti kategorií nebo po 128 pokusech. Původně byl prováděn manuálně pomocí fyzických karet a později byl adaptován i do počítačového formátu (Heaton & PAR Staff, 2008). V České republice existují normy pro dětskou populaci od školního věku a pro dospělou populaci (Telecká, 2013).

Na vzorku dospělé populace ($N = 100$) se Steinmetz et al. (2010) pokoušel o odhad ekvivalence počítačové a manuální verze WCST, přičemž zkoumali čtyři skóry: celkový počet správných odpovědí, procento chyby, perseverativní chyby a chyby nedodržení sady. Výsledky ukázaly značné rozdíly ve variancích, malé až střední koeficienty reliability paralelních forem a malé až střední koeficienty časové stability. Poukazují, že skóry na manuální verzi a počítačové verzi WCST vykazují neúplnou psychometrickou ekvivalenci.

Jako alternativní forma WCST byl vytvořen Clevelandský test třídění karet (CST; Poreh et al., 2007), autoři však měřili pouze korelaci obou testů a více se paralelností forem nezabývali. Pozdější studie toto téma otevírá, v první části porovnávala normativní data získaná pro test CST s demograficky upravenými normami pro test WCST. CST produkoval ekvivalentní skóry pro některé, ale ne všechny věkové skupiny. Druhá část studie se zabývala právě paralelností obou verzí při opakovaném měření. Probandi byli rozděleni do tří skupin, z nichž první skupině byl administrován nejprve CST, poté WCST, u druhé skupiny tomu bylo naopak a třetí skupině byl administrován CST dvakrát. Všechny tři testové skupiny vykazaly srovnatelné skóry a korelace neprokázaly statisticky významné rozdíly mezi různými skupinami. Zdá se, že jsou oba testy srovnatelné, ovšem nepodařilo se zamezit efektu učení. Jakmile se proband naučí princip v testu, v následujícím testování dosahuje lepších výsledků (Poreh et al., 2012).

3.2.3. *Test cesty*

Metoda známá jako Test cesty, nebo anglicky Trail Making Test (TMT), byla vytvořena Reitanem a Wolfsonem, primárně sloužila jako součást armádních zkoušek pro diagnostiku neuropsychologických poruch a byla veřejně publikována v roce 1944 (Motýl, 2015). Tento test obsahuje dvě části, označené jako TMT-A a TMT-B. Účastníci v první části spojují 25 čísel ve vzestupném pořadí, zatímco ve druhé části střídají mezi čísly a písmeny (Bezdíček et al., 2012).

TMT vyžaduje postupné pohyby rukou, které zapojují jemnou motoriku a zároveň kombinují kognitivní, sensorické a motorické schopnosti (Klaming & Vlaskamp, 2018). Zatímco TMT-A měří rychlost psychomotorických reakcí a vizuální pozornost, TMT-B poskytuje informace o schopnosti exekutivní kontroly, jako je kognitivní flexibilita a přepínání mezi úkoly (Reitan & Wolfson, 2004). Kromě toho TMT hodnotí i jemnou motorickou manuální výkonnost (Park & Schott, 2022) a je citlivý na různé neuropsychologické poruchy u dospělých i dětí (Reitan & Wolfson, 2004).

Pro českou populaci jsou k dispozici normativní data pro TMT pro mladé dospělé a osoby nad 75 let (Bezdíček et al., 2012). Normy existují i pro dětskou populaci od 8 do 14 let (Preiss & Preiss, 2006) a pro studenty 9. tříd základních škol až 4. ročníků středních škol (Sokolová & Cígler, 2018).

Kromě standardní verze Testu cesty byl vyvinut Barevný test cesty (BTC) pro děti předškolního a mladšího školního věku, tedy ve věku od pěti do sedmi let. Tato modifikace testu je zaměřena na detekci neurologických poruch a zkoumá kognitivní funkce a exekutivní procesy v tomto specifickém věku. Pro úspěšné absolvování testu je zapotřebí selektivity i inhibice. Dítě se musí koncentrovat na různé podněty a flexibilně vybírat ty, které jsou aktuálně relevantní, zatímco zanedbává rušivé vlivy (Sedláčková & Galbavá, 2021).

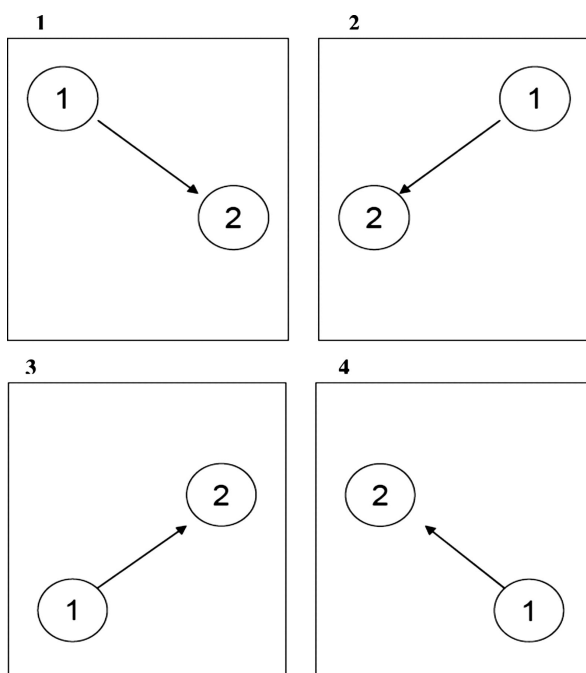
Právě ve věku před nástupem do školy dochází u dětí k významnému rozvoji kognitivních funkcí a vývoj exekutivní funkcí zde hraje zásadní roli. Dítě v tomto období rozvíjí schopnost selektivně zaměřovat pozornost a efektivně ji rozdělovat mezi různé podněty (Šnobllová & Krejčová, 2017). Díky důležitosti této fáze ve vývoji je výhodné, že je Barevný test cesty (BTC) přizpůsoben dětem od pěti let věku. V České republice, kde je povinné předškolní vzdělávání, může být tato forma užitečná pro částečné srovnání

kognitivního vývoje (Šnoblová & Krejčová, 2017). Včasná identifikace obtíží prostřednictvím tohoto testu může vést k intervencím, které předcházejí souvisejícím problémům ve školním výkonu (Sedláčková & Galbavá, 2021).

Alternativní formou testu cesty se zabývala Wagner et al. (2011) ve snaze zajistit test pro měření efektu intervence u pacientů s depresí. Před nimi Franzen et al. (1990) vyvinuli alternativní verze TMT C a D, ale ty neumožňovaly dostatečné sledování změn v průběhu léčby, jen změnu mezi výkonem před a po léčbě. McCracken a Franzen (1992) prokázali vysokou spolehlivost a validitu těchto alternativních forem. Alternativní verze TMT A a B vytvořili Wagner et al. (2011) tak, že zrcadlově zobrazili původní verzi testu. Tři nové verze vznikly otočením původní verze kolem horizontálních a vertikálních os (Obrázek 2).

Obrázek 2

Tvorba alternativní verze TMT



(Wagner et al., 2011)

Tyto verze pak administrovali svému vzorku v jednom sezení, bez určeného časového rozestupu. Probandi byli rozděleni do čtyř skupiny, každé skupině byly jednotlivé verze zadány v jiném pořadí, aby tak kontrolovali efekt učení. Cílem bylo zjistit, jestli se výkon (tedy průměrný čas potřebný k vypracování jedné verze testu) liší mezi jednotlivými verzemi. Nejprve byla použita ANCOVA k porovnání průměrných časů zpracování mezi čtyřmi verzemi TMT A a B. Výsledky ukázaly, že mezi verzemi nebyly významné rozdíly. Byly také zjištěny významné efekty věku a inteligence na výkonnost v TMT. Dále byla provedena regresní analýza ke zkoumání vztahu mezi výkonnostmi v první

verzi TMT a v poslední verzi. Výsledky pro TMT A ukázaly, že výkonnost v první verzi, věk a inteligence signifikantně předpovídali výkonnost v poslední verzi. Pro TMT B bylo zjištěno, že pouze výkonnost v první verzi signifikantně předpovídala výkonnost v poslední verzi. Studie zjistila vysokou spolehlivost a ekvivalenci nových verzí (Wagner et al., 2011).

Další studie (Atkinson et al., 2011) zkoumala tři různé testy, které mají podobné vlastnosti jako Test cesty (TMT) A a B, ale liší se v některých detailech. Tyto testy zahrnují Testy cesty v Delis-Kaplan Executive Function System (DK-TMT), Comprehensive Trail Making Test (CTMT) a Connections Task (CoT). Každý z těchto testů obsahuje podobné úlohy jako TMT A a B, ale také přidává další podmínky, jako je například zkoušení rychlosti motorických dovedností. Účastníci studie byli náhodně rozděleni do skupin a každému z nich byly testy zadány v průběhu tří týdnů. Interval mezi jednotlivými testy byl týden a pořadí, v jakém byly testy zadány, bylo pro každou skupinu jiné (výzkumníci využili všech šest možných pořadí). Statistické analýzy zahrnovaly konfirmatorní faktorovou analýzu (CFA) a strukturální multi-group analýzu. Ukázalo se, že pořadí testů nemělo vliv na výsledky testů a že skóre v testech jsou vzájemně ekvivalentní.

3.2.4. Stroopův test

Stroopův test byl původně publikován k vyšetření potenciální interference čtení slov na pojmenování barev a naopak (Stroop, 1935), jde v současnosti o bohatě využívaný nástroj pro měření selhání selektivní pozornosti a funkcí spojených s kognitivní kontrolou, jako je inhibice (Henik et al., 2018).

V tomto testu mají účastníci za úkol reagovat na barvu, kterou jsou vytištěna jednotlivá slova, a zároveň opomíjet význam těchto slov. Barva, kterou jsou slova zobrazena, a význam slov mohou být kongruentní, inkongruentní nebo neutrální. Pokud budeme mít červeně napsané slovo červená, jsou význam slova a barva kongruentní, zeleně napsané slovo červená by znamenalo inkongruenci a neutrální vztah by byl například mezi červeně napsaným řetězcem písmen X (Obrázek 3). Tím je ovlivněna úspěšnost při plnění testu, přičemž bylo zjištěno, že nejrychlejší reakce účastníci mají, pokud jsou slovo a barva kongruentní, naopak nejpomalejší jsou reakce při inkongruenci. Při tomto testu je měřen reakční čas, z nějž je pak možné získat informaci o míře facilitace a interference. Za facilitaci se považuje rozdíl v reakčním čase mezi kongruentní a neutrální verzí, jako

interference se uvádí rozdíl reakčního času mezi neutrální a inkongruentní verzí (Henik et al., 2018).

Obrázek 3

Kongruentní, inkongruentní a neutrální vztah



Od původní publikace (Stroop, 1935) byl test mnohokrát použit a modifikován do různých verzí, MacLeodova přehledová práce (1991) uvádí více než tisíc publikací. V minulosti se pozornost v této oblasti věnovala především příčinám interference. V posledních letech se výzkumy zabývají exekutivním procesům, které se s interferencí pojí, jako je schopnost interference ovládat (Henik et al., 2018).

V České republice byl Stroopův test poprvé publikován jako příručka (Krivá, 2013) vycházející z diplomové práce (Krivá, 2010), která obsahuje normy pro dospělé populaci. Nejnovější českou verzí tohoto testu je Pražská verze Stroopova testu (Bezdíček et al., 2021), která je zaměřena na populaci osob vyššího věku.

Alternativní formou tohoto testu se zabýval již Stroop (1935). Vytvořil ji obrácením položek v každém subtestu. Nikdy však psychometrickou ekvivalenci nezjišťoval. Alternativními verzemi se pak zabývali i další výzkumníci v průběhu let (Beglinger et al., 2005; Shenoy et al., 2020).

Studie z roku 2020 (Shenoy et al., 2020) se zabývala efektem učení při použití alternativních forem Stroopova testu. Jejím cílem bylo určit počet pokusů potřebných k oslabení efektu učení, mimo to výzkumníci zjišťovali také ekvivalenci těchto alternativních forem. Jednotlivé verze testu se lišily v pořadí barev a slov. Celkem bylo připraveno pět alternativních verzí testových formulářů. Participanti byli náhodně rozděleni do pěti různých skupin na základě pořadí administrace různých verzí testu. Každý účastník absolvoval všech pět forem testu, a to v pěti po sobě jdoucích dnech v dopoledních hodinách. Zjistili, že efekt učení se při opakovaném testování projevil, ale jeho velikost se snížila po čtyřech opakováních. Nebyl zjištěn žádný efekt pořadí administrace jednotlivých

forem testu, proto považují jednotlivé formy testu za ekvivalentní. Pro omezení efektu učení a zajištění co nejstabilnějších výsledků doporučují administraci čtyř zkušebních testů (Shenoy et al., 2020).

3.2.5. Rey-Osterriethova komplexní figura a komplexní figura Taylorové

Rey-Osterriethova figura (ROCF; Osterrieth, 1944; Rey, 1941) a figura Taylorové (TCF; Taylor, 1969) jsou testy, které se běžně užívají v klinickém i výzkumném kontextu (Tremblay et al., 2015) a zjišťují vizuoprostorové dovednosti, paměť, pozornost, organizační dovednosti a plánování. V obou testech musí účastníci překreslit figuru a bez předchozího upozornění ji s odstupem tří minut nakreslit znovu, z paměti. Po půl hodině se administruje ještě oddálené vybavení (Meyers & Meyers, 1995). Podnětovým materiálem je figura v obou případech složená z 18 geometrických segmentů. Účastníci si figuru, resp. její základní charakteristiky, pamatují i po roce od první administrace (Yamashita, 2009), proto je pro retest vhodná alternativní figura.

V České republice je vydaná příručka z roku 1997 (Košč & Novák, 1997). Její využití má ale několik nedostatků, včetně nedostatečných a zastaralých norem. Pro děti od 7,6 do 15,5 let uvádí příručka normy založené na datech z roku 1981. Pro nižší věk a pro dospělé používá normy z původního vydání Osterrietha z roku 1945 a pro věk 15,5 až 17,5 let jsou z roku 1993 (Kölesová, 1993), tedy mírně novější (Krčová, 2014). Nově také existují normy pro osoby nad 60 let (Drozdová et al., 2015).

Protože ROCF i TCF obsahují stejné množství elementů, které jsou považovány za srovnatelně komplexní, byly v minulosti i obě figury často považovány za ekvivalentní a bývají interpretovány za použití norem pro ROCF (Strauss et al., 2006). Více studií již ale ukázalo, že figura Taylorové je méně obtížná a dochází tak k nadhodnocování výsledků (Delaney et al., 1992; Hamby et al., 1993; Paštrnák et al., 2018; Strauss & Spreen, 1990; Tombaugh et al., 1992). TCF totiž sice replikuje vizuopercepční charakteristiky ROCF, ale ne její organizační kvality (Hamby et al., 1993), proto může TCF poskytovat srovnatelné výsledky v prvotní kopii, ale ne v oddáleném vybavení figury (Lezak et al., 2004).

Proto Hubleyová (1996) vyvinula Modifikovanou komplexní figuru Taylorové (MTCF), u které se potvrdilo, že poskytuje srovnatelné výsledky pro vizuokonstrukci i vizuoprostorovou paměť. Pozdější studie potvrdily srovnatelný výkon mezi ROCF a MTCF při prvotním překreslení s předlohou, při okamžitém vybavení i při oddáleném

vybavení po 20 minutách. Participantům byl test zadán s odstupem jednoho týdne, byli také rozděleni na dvě skupiny podle pořadí administrace testu, čímž výzkumníci kontrolovali efekt učení. Narozdíl od komplexní figury Taylorové je tak její modifikovaná verze srovnatelná s původní Rey-Osterriethovou figurou (Hubley & Jassal, 2005).

V České republice se ekvivalencí těchto dvou komplexních figur zabýval Paštrnák et al. (2018). Přestože se TCF ukazuje být jednodušší než ROCF, je v Česku pořád hojně užívaná jako její alternativa. Proto se rozhodli zkoumat její psychometrické charakteristiky jako alternativní formy testu. Zjistili signifikantní rozdíly průměrů mezi verzemi testu v několika měřících. Rozdíl průměru pro prvotní kopii byl signifikantní u dvou ze tří hodnotitelů, rozdíl hrubých skóre byl ale minimální. V klinické praxi považují výzkumníci rozdíl hrubého skóre o jeden bod jako marginální. Rozdíly průměru u třetího hodnotitele signifikantní nebyly. Pro vybavení po třech i třiceti minutách byl rozdíl průměru signifikantní u všech tří hodnotitelů. Na základě těchto zjištění doporučují použití TCF kopie s předlohou pro výzkumné užití, ale s obezřetností a přihlédnutím k dosavadním zjištěním o ekvivalenci obou verzí testu (Paštrnák et al., 2018).

3.2.6. Limity alternativních forem testů exekutivních funkcí

S tvorbou alternativních forem testů exekutivních funkcí se pojí několik úskalí. Prvním je předpoklad stálosti těchto atributů. Stejně jako u odhadu test-retestové reliability, u odhadu reliability paralelních forem také předpokládáme stabilitu měřených funkcí, alespoň v období mezi jednotlivými měřeními. Tento předpoklad nemusí být správný, především u dětí může v souvislosti s vývojem docházet ke změnám i v jinak relativně stálých funkcích. Je tedy nutné správně zvolit časový rozestup mezi jednotlivými měřeními (Urbánek et al., 2011).

Tato metoda je navíc velmi náročná na konstrukci, vzhledem k nutnosti vytvořit dvě (nebo více) srovnatelné verze testu (Kaplan & Saccuzzo, 2018; Rust et al., 2021). Taková snaha navíc nemusí být úspěšná. Například při snaze vytvořit alternativní verzi pro ROCF byla vytvořena TCF, u té se ale zjistilo, že je jednodušší a může tak docházet k nadhodnocování výsledků při opakovaném testování (Delaney et al., 1992; Hamby et al., 1993; Paštrnák et al., 2018; Strauss & Spreen, 1990; Tombaugh et al., 1992). Až její modifikovaná verze MTCF (Hubley, 1996) splňuje dostatečně podmínky ekvivalence (Hubley & Jassal, 2005).

Největším problémem se však zdá být efekt učení. Testy exekutivních funkcí totiž často vyžadují rozpoznávání abstraktních konceptů nebo generování strategií. I když tedy jednotlivé úkoly alternativní verze testu mohou vypadat jinak, participant si může pamatovat strategii, kterou použil u předchozího testování, a dochází tak znovu k nadhodnocování výsledků z druhého měření (Kurtz et al., 2004). Jako kontrola efektu učení se obvykle používá strategie rozdělení probandů do několika skupin, přičemž každé skupině jsou pak alternativní testy administrovány v různém pořadí (Atkinson et al., 2011; Hubley & Jassal, 2005; Poreh et al., 2012; Shenoy et al., 2020; Wagner et al., 2011). Ovšem takové studie někdy dochází k závěru, že není možné efektu učení zamezit, protože se probandi princip testu naučí a v dalším testování se vyvarují chyb (Poreh et al., 2012). Možnou strategií je také zjistit, po kolika pokusech se efekt učení stabilizuje. Tím se například zabývala studie alternativních forem Stroopova testu (Shenoy et al., 2020). Zjistili, že až po čtvrté administraci různých forem Stroopova testu dochází ke snížení efektu učení, proto pro dosažení co nejstabilnějších výsledků doporučují administraci čtyř cvičných úloh.

II. Empirická část

Pro posouzení exekutivních funkcí existuje mnoho neuropsychologických metod, a to i pro děti. Dostupnost těchto testů je však v České republice velmi omezená, zejména pak chybí normativní studie. Pokud pak máme potřebu dítě testovat opakovaně v kratším časovém úseku, tak k tomu nemáme obvykle žádný nástroj. Přitom mohou být metody pro posouzení exekutivních funkcí důležité při posouzení kognitivní výkonnosti u dětí s neurovývojovým onemocněním, získaným onemocněním (např. po kraniotraumatu) nebo u dětí s poškozením centrální nervové soustavy, neméně důležité je to i u dětí s neurodegenerativním onemocněním nebo psychiatrickým onemocněním, jako jsou například poruchy nálady či schizofrenie (Villalobos et al., 2022). Alternativní verze testů, které by umožnily opakované testování v krátkém časovém intervalu, jsou důležité pro posouzení dynamiky vývoje či změny v naměřené úrovni exekutivních funkcí, obvykle například po nějaké intervenci (po nasazení farmakoterapie, po chirurgickém zákroku, ad.). K tomu není většinou vhodné využívat stejnou variantu testu, a to zejména kvůli efektu učení. Efekt učení je problém, který právě i tvorbu alternativní verze velmi komplikuje a mnoho výzkumníků se o to v měření exekutivních funkcí nepokouší (Kurtz et al., 2004).

Právě efektu učení u měření exekutivních funkcí je specifický tím, že si probandi pamatují nejen podobu prvního zadaného testu, ale mohou si zapamatovat i strategie, které k testu vymysleli, a ty použít u alternativní formy. Dochází tak k nadhodnocování výsledků i přes to, že se od sebe na první pohled alternativní formy liší (Kurtz et al., 2004).

V návaznosti na uvedené problémy v praxi je realizován projekt zaměřený na vývoj a standardizaci Neuropsychologické baterie pro děti (NB-D). Baterie by měla být souborem zkoušek, které pomohou komplexně zhodnotit kognitivní schopnosti v celé jejich šíři u dětí od 6 do 19 let. Baterie má několik domén, z čehož jedna je zaměřena právě na exekutivní funkce, přičemž některé subtesty jsou zaměřeny pouze na exekutivu, některé jsou komplexnější, a zahrnují více kognitivních domén. NB-D by měla splňovat psychometrická kritéria pro kvalitní psychodiagnostickou pomůcku, včetně standardizace, vysoké spolehlivosti metody (reliability) a validity. V rámci naplnění kritérií probíhá i tato práce, která se zaměřuje na odhad reliability dvou paralelních forem (A a B) u domény exekutivních funkcí NB-D.

4. Cíl výzkumu

Cílem této práce je posouzení ekvivalence dvou verzí Neuropsychologické baterie pro děti (NB-D) v oblasti exekutivních funkcí u dětí ve věku od 6 do 19 let. Tato analýza zahrnuje srovnání výkonu dětí v jednotlivých subtestech zaměřených na exekutivní funkce v první a druhé verzi testové baterie a zabývá se odhadem reliability paralelních forem. Studie vznikla jako součást projektu 2. LF UK s č. TL03000328, který byl spolufinancován se státní podporou Technologické agentury ČR v rámci Programu ÉTA. Práce je tedy samostatnou součástí procesu standardizace NB-D a klade si za cíl poskytnout ucelené poznatky o použitelnosti obou verzí testu v klinickém a diagnostickém prostředí. Studie vznikla jako součást projektu 2. LF UK s č. TL03000328, který byl spolufinancován se státní podporou Technologické agentury ČR v rámci Programu ÉTA.

4.1. Výzkumné otázky a hypotézy

Otázky směřují k posouzení ekvivalence mezi verzemi testové baterie v oblasti exekutivních funkcí u dětí.

Otázka č. 1: Jaká je korelace mezi paralelními formami A a B NB-D v oblasti exekutivních funkcí?

Otázka č. 2: Bude mezi průměrnými hrubými skóry pro verzi A a B NB-D v oblasti exekutivních funkcí významný rozdíl?

Hypotézy dále zpřesňují výzkumné otázky a zmiňují, v jakých situacích bude nebo nebude daná hypotéza přijata.

Hypotéza č. 1: Skóry dětí ve věku od 6 do 19 let, kterým byly administrovány subtesty cílené na posouzení exekutivních funkcí z NB-D, verze A budou dosahovat alespoň středně silné korelace se skóry, kterých tyto děti dosáhly při administraci stejnojmenných subtestů z NB-D, verze B.

Hypotéza č. 2: Ve skupině dětí ve věku od 6 do 19 let, kterým byly administrovány subtesty posuzující exekutivní funkce NB-D, verze A i verze B, bude pozorován statisticky nevýznamný rozdíl ($p \leq 0,05$) při porovnání jejich průměrných výkonů mezi verzemi.

5. Metodika

5.1. Výzkumný soubor

Kritéria pro účast ve výzkumu byla zjišťována předem pomocí anamnestického dotazníku. Zájemci, kteří je nespĺňovali byli tudíž vyloučeni z výzkumu ještě před sběrem dat. Vylučovací kritéria byla následující:

- kognitivní deficit
- onemocnění (aktuálně nebo v anamnéze):
 - neurologické, včetně poranění hlavy s bezvědomím delším než 5 minut;
 - psychiatrické, včetně léčeného a neurovývojového;
 - jiné závažné somatické s vlivem na CNS;
- závažná prenatální nebo perinatální zátěž;
- historie užívání psychoaktivních látek;
- nekorigované senzorní postižení;
- specifické vývojové poruchy nebo poruchy řeči.

Power analýza byla provedena v programu GPower 3.1 (Faul et al., 2007; 2009) s cílem určit požadovanou velikost vzorku pro účely empirické části této práce. Zadané parametry zahrnovaly korelaci o hodnotě $r = 0,45$ a hladinu významnosti stanovenou na $p \leq 0,05$. Cílem bylo dosáhnout síly testu o velikosti minimálně 0,8, což znamená, že pravděpodobnost detekce skutečného efektu, pokud existuje, je 80 %. Výsledkem power analýzy byla doporučená velikost vzorku $N = 29$. S ohledem na možné vyloučení některých účastníků a další faktory, byl nakonec nasbírán vzorek o velikosti $N = 31$, což by mělo zajistit dostatečnou robustnost a spolehlivost výsledků v této studii.

Výzkumu se tedy celkem zúčastnilo 31 dětí, z toho 15 dívek a 16 chlapců. Průměrný věk byl při prvním měření 11,9 roku ($SD = 3.45$), přičemž nejnižší věk byl 6,6 roku a nejvyšší 18,4 roku. Většina dětí docházela na běžnou základní školu ($N = 25$), menší část na střední odborné učiliště zakončené výučním listem ($N = 2$) a střední odbornou školu zakončenou maturitní zkouškou ($N = 1$). Zbytek dětí docházel na gymnázium ($N = 3$). U 3 dětí se objevila obtížnější adaptace při nástupu do mateřské školy a 7 dětí mělo odklad školní docházky (avšak bez vážných příčin). Žádný z probandů neopakoval ani jeden ročník na základní či střední škole, dokonce dvě třetiny

děti ($N = 20$) měly na posledním vysvědčení vyznamenání. Většina dětí ($N = 28$) jsou praváci.

U 1 dítěte rodiče reportovali potíže v těhotenství (konkrétně hlídaný porod z důvodu předchozího předčasného porodu a potratu) a u 1 dítěte se objevily blíže nespecifikované nezávažné potíže při porodu. Rodinná anamnéza je u všech probandů bez jednoznačné psychiatrické či neurologické zátěže. U 1 dítěte se projevil pomalejší vývoj řeči, bez diagnózy opožděného vývoje řeči nebo vývojové dysfázie, jinak nikdo žádné potíže v psychomotorickém vývoji dítěte nereportoval.

Větší část rodin ($N = 17$) žije ve městě, zbytek na vesnici, a většina dětí žije v úplné rodině ($N = 19$). Přibližně polovina matek měla dokončenou vysokou školu ($N = 15$), ostatní měly dokončenou střední školu s maturitní zkouškou ($N = 13$) nebo střední odborné učiliště ($N = 3$). I mezi otci měla asi polovina dokončenou vysokou školu ($N = 15$), méně jich ale mělo maturitní zkoušku ($N = 8$) a 1 z nich měl základní vzdělání. Zaměstnání rodičů našich probandů je různorodé jak v oborech, tak v míře kvalifikace (Tabulka 1).

Tabulka 1

Četnost povolání rodičů

Matka		Otec	
Povolání	Četnost	Povolání	Četnost
Administrativa - celní správa	1	Analytik, logistik	2
Asistent pedagoga	4	Datový a finanční analytik	2
Fyzioterapeut	1	Dělník	3
Laborant - chemická výroba	1	Elektrikář	2
Lékař	2	Letecký pozemní technik	1
Nepracuje - pečující osoba	2	Lékař	2
OSVČ - prodej oděvů	2	Manažer	5
Pomocná kuchařka	1	Masér	2
Pracovník České pošty	1	Muzikant	2
Prodáváč	1	Opravář	1
Projektová manažerka	1	Pastorační asistent	1
Psycholog	2	Revizní technik	2
Správce dokumentace	2	Řidič kamionu	1
Telefonista	1	Technik	2
Účetní/ekonom	2	Zahradník	1
Učitelka	5	Zedník	1
Vychovatelka	1		
Zdravotní sestra	1		

Všem dětem byla také zadána Zkrácená Wechslerova inteligenční škála, druhé vydání (WASI-II; Wechsler, 2011). Výsledky byly vyhodnoceny na základě norem získaných na americké populaci. Průměrný intelekt výzkumného souboru byl 108,1 ($SD = 11,68$), přičemž minimální hodnota byla 88, maximální 137. Žádné z dětí tedy nemělo intelekt, který by dle mezinárodní klasifikace nemocí (MKN-10; Ústav zdravotnických informací a statistiky ČR, 2021) spadal do pásma významného podprůměru (tedy žádné z dětí nesplňovalo kritéria pro diagnózu mentální retardace).

5.2. Měřicí nástroje

Respondentům studie byla administrována celá nově vznikající Neuropsychologická baterie pro děti (NB-D), která obsahuje subtesty pro sedm domén: Paměť a učení, Pozornost, Exekutivní funkce a pracovní paměť, Řeč, Zrakově-percepční funkce, Motorické funkce a Sociální kognice (Bukačová et al., 2021). Kromě toho byl zadán všem respondentům i test WASI-II (Zkrácená Wechslerova Inteligenční Škála, druhé vydání; Wechsler Abbreviated Scale of Intelligence, Second Edition; Wechsler, 2011) pro zjištění úrovně intelektových předpokladů, a anamnestický dotazník. Všichni probandi, resp. jejich zákonní zástupci, také vyplnili informovaný souhlas s účastí ve studii. Celá studie byla schválena etickou komisí 2. LF UK.

5.2.1. Anamnestický dotazník

Anamnestický dotazník byl součástí záznamového archu NB-D a zjišťuje především základní a demografické údaje probanda. Zjišťuje datum narození (a datum administrace), pohlaví a studovanou školu (vybírání se z možností běžná ZŠ, ZŠ speciální, ZŠ praktická, střední odborné učiliště s výučním listem, střední odborná škola s maturitou, gymnázium, vyšší odborná škola, vysoká škola, mateřská škola), dále studovaný ročník. Ptá se, zda mělo dítě odklad školní docházky, zda opakovalo nějakou třídu a zda má přiděleného asistenta pedagoga, zjišťuje také, zda nemá diagnostikované specifické poruchy učení a jaká je školní úspěšnost dítěte (zda má vyznamenání a jaké má známky z českého jazyka, matematiky a cizího jazyka). Další část se zaměřuje na raný vývoj a zdravotní stav dítěte, zjišťuje rizikové faktory v prenatální a perinatální anamnéze a problémy v psychomotorickém vývoji nebo v adaptaci na školní prostředí, ptá se také na sensorické poruchy, úrazy hlavy, neurologická a psychiatrická onemocnění nebo jiná chronická onemocnění a užívání psychoaktivních látek. Poslední sekce se zaměřuje na rodiče, jejich

vzdělání, povolání, bydliště (město nebo vesnice) a rizikové faktory v jejich zdravotní anamnéze.

5.2.2. Zkrácená Wechslerova inteligenční škála, druhé vydání (WASI-II)

Zkrácená Wechslerova inteligenční škála, druhé vydání (WASI-II; Wechsler Abbreviated Scale of Intelligence, Second Edition; Wechsler, 2011; užito v podmínkách Licenční smlouvy č. LSR-191984), je krátký test posuzující intelektové předpoklady a je určený pro jedince od 6 do 90 let. Ve studii byla škála použita za smluvně stanovených podmínek s psychodiagnostickým nakladatelstvím. Skládá se ze čtyř subtestů.

V subtestu *Kostky* má proband za úkol skládat konstrukce v časovém limitu podle návodu. Návod je buď složená konstrukce nebo obrázek v podnětové knize. Kostky jsou všechny stejné, mají vždy dvě bílé stěny, dvě červené a dvě červenobílé. Subtest je ukončen po dvou po sobě jdoucích neúspěších. Body získává proband za správné řešení, přičemž při rychlém splnění úkolu dostává více bodů.

Dalším subtestem je *Slovník*. V první části jsou probandovi předkládány obrázky a jeho úkolem je obrázky zobrazené v podnětové knize pojmenovat. Následují verbální položky, ve kterých administrátor nahlas přečte slovo, které má proband definovat. Odpovědi jsou bodovány dle kvality od 0 do 2 bodů a subtest se ukončuje, pokud proband dosáhne ve třech po sobě jdoucích položkách skóru 0.

Jako třetí je administrován subtest *Matrice*, ve kterém je cílem vybrat z možností tu, která doplní sérii nebo neúplnou matici v předložené podnětové knize. Na odpovědi mají probandi časový limit a odpovědi jsou hodnoceny dle správnosti 0 nebo 1 bodem. Ukončuje se po třech nesprávných odpovědích.

Posledním subtestem WASI-II jsou *Podobnosti*. Zde jsou představovány vždy dvě slova, vyjadřující běžné objekty nebo koncepty. Úkolem je vysvětlit, v čem jsou si slova podobná. Odpovědi jsou opět hodnoceny dle kvality a subtest je ukončen po třech po sobě jdoucích neúspěšných odpovědích, tedy odpovědích za 0 bodů.

5.2.3. Neuropsychologická baterie pro děti (NB-D)

NB-D je nově vznikající testová baterie pro děti od 6 do 19 let. Hodnotí kognitivní profil v celkem sedmi doménách (Bukačová et al., 2021).

První doménou je *Paměť a učení*. Ta obsahuje subtest Verbální paměť a učení, Paměť na příběhy, Neverbální paměť a Vybavení s nápovědou. Zaměřuje se tedy na verbální i neverbální paměť a zjišťuje křivku učení, vybavení po interferenci, oddálené vybavení nebo rekognici. Druhou doménou je *Pozornost*. Zde baterie obsahuje subtesty Sluchová pozornost a Zraková pozornost. Na ni navazuje doména *Exekutivní funkce a pracovní paměť*, která je pro tuto práci ústřední a budu se jí zabývat níže. Čtvrtou doménou je *Řeč*, která obsahuje subtesty Konfrontační pojmenování a Porozumění pokynům. Pro doménu *Zrakově-percepční funkce*, byly vybrány subtesty Zrakové vnímání, Zrakově prostorové vnímání a Orientace v prostoru. Pro předposlední doménu *Motorické funkce* pak obsahuje subtest Motorická koordinace, Zrakově-motorická přesnost a Kopie tvarů, a pro doménu *Sociální kognice* subtesty Teorie mysli a Rozpoznávání emocí. Toto dělení na domény volně odráží dělení kognitivních funkcí dle DSM-V (Diagnostický a statistický manuál duševních poruch, páté vydání), kde jsou kognitivní funkce rozděleny do kategorií Komplexní pozornost, Exekutivní funkce, Učení a paměť, Řeč, Percepčně-motorické funkce a Sociální kognice (American Psychiatric Association, 2013).

Doména, kterou se zabývá tato práce, tedy *Exekutivní funkce a pracovní paměť*, je doménou třetí. Obsahuje celkem pět subtestů – Třídění, Inhibice a přesun pozornosti, Pracovní paměť, Verbální fluence a Prospektivní paměť. Tím prvním je tedy *Třídění*, ve kterém se před dítě rozloží karty s obrázky, celkem osm karet, a v časovém limitu max. 420 sekund má za úkol vytvořit co nejvíce možných dělení na dvě kategorie. Kategorie musí být podle smysluplného kritéria, které karty rozdělí na dva díly po čtyřech kartách. Příkladem takového kritéria jsou třeba malé a velké karty, nebo dívky a chlapci. Každý způsob dělení karet se zapisuje a hodnotí se od 0 do 2 bodů dle kvality odpovědi.

Dalším subtestem v této doméně je *Inhibice a přesun pozornosti*. Tato úloha má celkem tři části. Před probandy je položena podnětová kniha s několika řádky obrázků slunce a měsíce (v alternativní verzi jsou šipky směřující nahoru a dolů). V první části, *Pojmenování*, je má proband pouze co nejrychleji všechny přečíst. V další fázi, *Inhibice*, má číst obrázek opačný, než který vidí, tedy pokud vidí slunce, přečte měsíc a naopak. V poslední části, *Přesun pozornosti*, má za úkol přečíst opačný obrázek, pokud je vybarven bíle, a obrázek, který vidí, pokud je vybarven černě. U všech částí se zaznamenává doba potřebná pro přečtení všech obrázků a počet chyb.

Následuje subtest *Pracovní paměť*. V podnětové knize jsou probandovi jeden po druhém ukazovány obrázky, posléze sada všech možných devíti obrázků, na kterých má říct viděné obrázky v opačném pořadí, než ve kterém mu byly předkládány. Začíná se na dvou obrázcích a nejvyšším počtem je devět, každý počet obrázků proband zkouší dvakrát. Po neúspěchu u obou pokusů se stejným počtem obrázků je subtest ukončen. Hodnotí se jak správnost zapamatovaných obrázků, tak správnost opačného pořadí.

Čtvrtým subtestem v této doméně je *Verbální fluence*. Zde je probandovým úkolem vymýšlet co nejvíce slov za jednu minutu dle daného kritéria. Nejprve je zadána fonemická fluence pro písmena K a V (pro verzi B písmena T a M), poté kategorická fluence pro kategorii zvířata (pro verzi B kategorie činnosti) a nakonec kategorická fluence s nutností střídání mezi kategoriemi oblečení a nádobí (pro verzi B kategorie jídlo a pohádkové postavy). Celkově jde tedy o čtyři série slov po jedné minutě. Pro jednotlivá kritéria se hodnotí počet unikátních slov a u poslední se navíc hodnotí počet switchů, tedy kolikrát proband vystřídal kategorie.

Prospektivní paměť spočívá v plnění různých úkolů po určité aktivitě nebo v určitý čas. Celkem obsahuje osm položek, které jsou probandovi přečteny. Čtyři z nich zadávají úkoly, které má udělat po nějaké aktivitě, další čtyři jsou úkoly, které má splnit v určitý čas (za 10 minut). Po přečtení se pokračuje v dalších subtestech, dokud neuplyne čas pro splnění poslední úlohy. Hodnotí se každý úkol maximálně dvěma body, přičemž jeden bod je možné získat za splnění úkolu ve správný čas a jeden bod za správné splnění úkolu. Na úlohu navazuje rekognice s nuceným výběrem.

5.3. Proces sběru dat

Probandi byli vybíráni na základě dostupnosti a dobrovolnosti. Možnost účasti ve výzkumu byla včetně informace o finanční odměně sdílena na sociálních sítích Facebook a Instagram. První kontakt s rodiči dětí probíhal obvykle přes dané sociální sítě nebo e-mail, následně jim byly po telefonu vysvětleny podmínky účasti, účastnit se mohly pouze zdravé děti ve věku od 6 do 19 let, které neprodělaly závažné poranění hlavy a neměli žádné neurologické, psychiatrické nebo jiné chronické onemocnění. Zároveň musely mít možnost účastnit se obou částí výzkumu, tedy dvou vyšetření s rozestupem dvou až čtyř týdnů. Sběr dat probíhal od května 2023 do listopadu 2023.

Při prvním setkání byl s rodiči podepsán informovaný souhlas a měli zároveň prostor doptat se na doplňující informace. Následně byl s rodiči vyplněn anamnestický dotazník a s dětmi bylo zahájeno první vyšetření. Děti byly náhodně rozděleny do dvou skupin, přičemž první skupině byla NB-D administrována v pořadí verze A, pak verze B, druhé skupině naopak. Během prvního setkání byl také administrován WASI-II. Celkově první vyšetření trvalo asi tři hodiny. Druhé vyšetření proběhlo s rozestupem dvou až čtyřech týdnů a dětem byla administrována již jen druhá verze NB-D. Průměrná doba administrace NB-D pro obě vyšetření byla dvě hodiny a patnáct minut. Po absolvování obou částí byla odeslána finanční kompenzace na bankovní účet uvedený v informovaném souhlasu.

5.4. Statistická analýza

Všechny použité testy byly vyhodnoceny dle manuálů a hrubé skóry včetně času plnění zaneseny do matice v MS Excel (Microsoft Corporation, 2018). Pomocí Shapiro-Wilkova testu byla následně posouzena normalita rozložení dat. Pro data s normálním rozložením byly použity parametrické, pro ostatní neparametrické statistické analýzy. Pro odhad reliability paralelních forem byla použita korelace, v případě normálního rozložení Pearsonův korelační koeficient, v případě nenormálního rozložení Spearmanův koeficient. Dále pro zjišťování rozdílů mezi hrubými skóry byl použit *t*-test pro data s normální distribucí a pro nenormální distribuci Wilcoxonův test pro párové vzorky. Pro všechny statistické analýzy byla stanovena hladině signifikance $p \leq 0,05$.

5.5. Etika výzkumu

Rodičům dětí, které se účastnily výzkumu, byl předán písemný informovaný souhlas, který obsahoval informace o účelu studie, průběhu vyšetření a jeho délce a možných rizicích pro dítě (mírná únava). Všechny náležitosti byly rodičům vysvětleny i ústně a měli možnost doptat se na doplňující informace. Za účast ve výzkumu probandům náležela finanční kompenzace ve výši 500,- Kč. Data jsou dále anonymizovaná po přepisu dat ze záznamových archů, každému dítěti je přiřazeno náhodně identifikační číslo (ID). Záznamové archy s osobními údaji respondentů jsou uloženy ve Fakultní nemocnici Motol dle pravidel pro uskladňování dokumentů obsahujících citlivé údaje. Celý výzkum byl schválen Etickou komisí 2. Lékařské fakulty Univerzity Karlovy.

6. Výsledky

6.1. Deskriptivní a inferenční statistika

V této kapitole jsou zaznamenány výkonové charakteristiky pro obě verze NB-D v jednotlivých zkoumaných subtestech zaměřených na exekutivní funkce. Jsou uvedeny jejich průměry, směrodatné odchylky (*SD*), mediány a minimální a maximální hodnoty. Dále je ověřována normalita rozložení dat, k tomu byl využit Shapirův-Wilkův test, pro který byla zvolena hladina signifikance $p \leq 0,05$. Všechny analýzy v této kapitole byly provedeny v programu R 4.3.1 (R Core Team, 2023).

6.1.1. Třídění

Tabulka 2 shrnuje deskriptivní statistiku pro subtest *Třídění*, kde byla hodnocena proměnná *Správná pravidla*. Na základě Shapirova-Wilkova testu a stanovené hladiny signifikance $p \leq 0,05$ byl potvrzen předpoklad normality rozložení dat v tomto subtestu.

Tabulka 2

Třídění – deskriptivní statistika a Shapirův-Wilkův test

	Verze testu	Správná pravidla
Průměr	A	10,97
	B	11,81
Medián	A	11
	B	12
SD	A	3,67
	B	4,31
Minimum	A	4
	B	3
Maximum	A	17
	B	21
Shapiro-Wilk <i>W</i>	A	0,96
	B	0,97
Shapiro-Wilk <i>p</i>	A	0,257
	B	0,532

Pozn. SD – směrodatná odchylka.

6.1.2. Inhibice a přesun pozornosti

Tento subtest má tři části – Pojmenování, Inhibice a Přesun pozornosti. Ve všech třech částech je hodnocen celkový čas, počet neopravených chyb, opravených chyb a celkový počet chyb. V následující tabulce (Tabulka 3) jsou uvedeny průměry a *SD*, mediány, minimální a maximální hodnoty a výsledky Shapirova-Wilkova testu včetně

hladiny signifikance pro část *Pojmenování* tohoto subtestu. Z výsledků je zřejmé, že předpoklad normality můžeme pro zvolenou hladinu signifikance zamítnout pro všechna hodnocená kritéria.

Tabulka 3

Pojmenování – deskriptivní statistika a Shapirův-Wilkův test

	Verze testu	Čas	Neopravené chyby	Opravené chyby	Chyby celkem
Průměr	A	21,66	0,06	0,42	0,48
	B	21,58	0,03	0,29	0,32
Medián	A	20,05	0	0	0
	B	21,00	0	0	0
SD	A	6,56	0,25	0,67	0,72
	B	7,86	0,18	0,46	0,48
Minimum	A	13,75	0	0	0
	B	12,26	0	0	0
Maximum	A	45,00	1	2	2
	B	53,00	1	1	1
Shapiro-Wilk <i>W</i>	A	0,86	0,27	0,65	0,67
	B	0,81	0,18	0,57	0,59
Shapiro-Wilk <i>p</i>	A	< 0,001	< 0,001	< 0,001	< 0,001
	B	< 0,001	< 0,001	< 0,001	< 0,001

Pozn. SD – směrodatná odchylka.

Tabulka 4 se zabývá částí *Inhibice* tohoto subtestu. Na základě Shapirova-Wilkova testu můžeme i u této části pro zvolenou hladinu signifikance zamítnout předpoklad normality pro měřená kritéria.

Tabulka 4

Inhibice – deskriptivní statistika a Shapirův-Wilkův test

	Verze testu	Čas	Neopravené chyby	Opravené chyby	Chyby celkem
Průměr	A	28,48	0,26	0,61	0,87
	B	33,89	0,39	1,10	1,48
Medián	A	26,00	0	0	0
	B	30,00	0	1	1
SD	A	11,32	0,82	0,95	1,15
	B	20,13	0,99	1,08	1,55
Minimum	A	16,12	0	0	0
	B	14,22	0	0	0
Maximum	A	73,00	4	3	4
	B	130,00	4	3	7
Shapiro-Wilk <i>W</i>	A	0,82	0,37	0,67	0,76
	B	0,61	0,46	0,82	0,81
Shapiro-Wilk <i>p</i>	A	< 0,001	< 0,001	< 0,001	< 0,001
	B	< 0,001	< 0,001	< 0,001	< 0,001

Pozn. SD – směrodatná odchylka.

Na poslední část, *Přesun pozornosti*, se zaměřuje Tabulka 5. Na stanovené hladině signifikance můžeme v této části zamítnout předpoklad normality rozložení dat pro všechny měřené charakteristiky, kromě času pro verzi testu A.

Tabulka 5

Přesun pozornosti – deskriptivní statistika a Shapirův-Wilkův test

	Verze testu	Čas	Neopravené chyby	Opravené chyby	Chyby celkem
Průměr	A	46,71	0,87	1,71	2,58
	B	43,98	0,35	1,90	2,26
Medián	A	42,12	0	2	2
	B	38,00	0	2	2
SD	A	14,49	1,88	1,72	2,09
	B	22,68	0,75	1,70	1,65
Minimum	A	26,58	0	0	0
	B	23,43	0	0	0
Maximum	A	78,00	9	6	9
	B	151,00	3	7	7
Shapiro-Wilk <i>W</i>	A	0,94	0,53	0,86	0,91
	B	0,62	0,54	0,88	0,92
Shapiro-Wilk <i>p</i>	A	0,075	< 0,001	< 0,001	0,013
	B	< 0,001	< 0,001	0,002	0,019

Pozn. SD – směrodatná odchylka.

6.1.3. Pracovní paměť

Pro subtest Pracovní paměť je hodnotícím kritériem získaný hrubý skór. Deskriptivní statistiky shrnuje Tabulka 6. Normalitu rozložení můžeme na stanovené hladině signifikance zamítnout pro verzi A.

Tabulka 6

Pracovní paměť – deskriptivní statistika a Shapirův-Wilkův test

	Verze testu	Hrubý skór
Průměr	A	10,77
	B	11,10
Medián	A	9
	B	12
SD	A	4,62
	B	3,88
Minimum	A	6
	B	5
Maximum	A	28
	B	20
Shapiro-Wilk <i>W</i>	A	0,79
	B	0,96
Shapiro-Wilk <i>p</i>	A	< 0,001
	B	0,270

Pozn. SD – směrodatná odchylka.

6.1.4. Verbální fluence

V subtestu *Verbální fluence* jsou čtyři části – fonemická fluence pro dvě různá písmena, kategoričká fluence a kategoričká fluence s nutností střídání kategorič. Hodnotí se pak počet slov pro obě fonemické fluence zvlášť, počet slov pro fonemickou fluenci

celkem, počet slov v kategorické fluenci, počet slov v kategorické fluenci s nutností střídání a počet switchů, tedy změny kategorie, v kategorické fluenci s nutností střídání. Deskriptivní statistiku pro tyto hodnoty shrnují následující tabulky (Tabulka 7, Tabulka 8). Pro fonemickou fluenci mají všechna hodnocená kritéria normální rozložení dat na základě Shapirova-Wilkova testu na stanovené hladině signifikance $p \leq 0,05$ (Tabulka 7).

Tabulka 7

Fonemická fluence – deskriptivní statistika a Shapirův-Wilkův test

	Verze testu	Fonemická pro K/T	Fonemická pro V/M	Fonemická celkem
Průměr	A	11,52	9,26	20,77
	B	9,45	8,55	18,00
Medián	A	13	10	22
	B	9	7	19
SD	A	6,10	5,33	10,86
	B	4,46	3,93	7,91
Minimum	A	3	1	4
	B	2	2	4
Maximum	A	27	19	46
	B	19	16	35
Shapiro-Wilk W	A	0,93	0,95	0,96
	B	0,97	0,96	0,97
Shapiro-Wilk p	A	0,051	0,163	0,244
	B	0,525	0,220	0,534

Pozn. SD – směrodatná odchylka.

Stejně tak i pro kategorickou fluenci mají data ve všech hodnocených kritériích normální rozložení (Tabulka 8).

Tabulka 8

Kategorická fluence – deskriptivní statistika a Shapirův-Wilkův test

	Verze testu	Kategorická	Kategorická se střídáním	Počet switchů
Průměr	A	21,00	13,32	12,06
	B	16,16	11,71	10,65
Medián	A	20	13	12
	B	17	11	10
SD	A	7,27	3,94	3,93
	B	4,90	4,26	4,27
Minimum	A	10	6	5
	B	6	5	4
Maximum	A	37	23	22
	B	27	23	22
Shapiro-Wilk W	A	0,96	0,97	0,96
	B	0,99	0,95	0,95
Shapiro-Wilk p	A	0,267	0,399	0,374
	B	0,985	0,173	0,128

Pozn. SD – směrodatná odchylka.

6.1.5. Prospektivní paměť

V subtestu *Prospektivní paměť* jsou hodnotícím kritériem *Hrubý skór* v první části a *hrubý skór* v *Rekognici*. Deskriptivní statistiku k obojímu shrnuje Tabulka 9. Na hladině významnosti $p \leq 0,05$ můžeme zamítnout předpoklad normality rozložení dat pro obě hodnocená kritéria.

Tabulka 9

Prospektivní paměť – deskriptivní statistika a Shapirov-Wilkův test

	Verze testu	Hrubý skór	Rekognice
Průměr	A	12,03	7,97
	B	11,00	7,90
Medián	A	12	8
	B	12	8
SD	A	3,08	0,18
	B	4,02	0,30
Minimum	A	2	7
	B	3	7
Maximum	A	16	8
	B	16	8
Shapiro-Wilk W	A	0,84	0,18
	B	0,89	0,34
Shapiro-Wilk p	A	< 0,001	< 0,001
	B	0,004	< 0,001

Pozn. SD – směrodatná odchylka.

6.2. Statistická analýza dat

K testování stanovených hypotéz byl využit Spearmanův korelační koeficient pro ověření Hypotézy č. 1 u subtestů, pro které jsme v předchozí části zamítli normalitu rozložení dat na základě Shapirova-Wilkova testu, a Pearsonův korelační koeficient u subtestů, které předpoklad normálního rozložení dat splňují. Síla korelace byla interpretována následujícím způsobem:

- Pearsonovo r /Spearmanovo $\rho \leq 0,39$ – slabá korelace;
- $0,40 \leq$ Pearsonovo r /Spearmanovo $\rho \leq 0,69$ – středně silná korelace;
- Pearsonovo r /Spearmanovo $\rho \geq 0,70$ – silná korelace.

Dále pro ověření Hypotézy č. 2 byl použit t -test nebo Wilcoxonův test pro párové vzorky v závislosti na normalitě rozložení daných dat. Pro t -test je uvedena také velikost efektu (Cohenovo d). Velikost efektu byla interpretována takto:

- $d \leq 0,49$ – malý efekt;
- $0,50 \leq d \leq 0,79$ – střední efekt;

- $d \geq 0,80$ – velký efekt.

U Wilcoxonova testu pro párové vzorky byla velikost efektu vyjádřena pomocí r hodnoty a interpretace této hodnoty se řídila následujícím pravidlem:

- $r \leq 0,29$ – malý efekt;
- $0,30 \leq r \leq 0,49$ – střední efekt;
- $r \geq 0,50$ – velký efekt.

Statistické analýzy byly provedeny v programu R 4.3.1 (R Core Team, 2023).

6.2.1. Třídění

Data v tomto subtestu měla normální rozložení, proto byla splněna podmínka pro použití Pearsonova korelačního koeficientu. Tento koeficient je statisticky signifikantní a dosahuje silné korelace (Tabulka 10).

Tabulka 10

Třídění – Pearsonův korelační koeficient

	Pearsonovo r	p -hodnota
Správná pravidla	0,76	< 0,001

T -testem nebyl zjištěn statisticky signifikantní rozdíl hrubých skóre (Tabulka 11).

Tabulka 11

Třídění – t -test

	t -statistika	p hodnota	Cohenovo d
Správná pravidla	-1,66	0,106	-0,30

6.2.2. Inhibice a přesun pozornosti

Protože většina položek nesplňovala podle Shapirova-Wilkova testu předpoklad normálního rozložení, byl pro ověření korelace hrubých skóre mezi verzí A a B použit Spearmanův korelační koeficient. Pro *Čas* ve všech částech tohoto subtestu ukazují výsledky (Tabulka 12) silné korelace, statisticky významné na stanovené úrovni $p \leq 0,05$. Pro chybové skóre jsou korelace v části *Pojmenování* záporné, ve většině skóre nejsou dostatečně silné, kromě hodnot pro *Opravené chyby* a *Chyby celkem* v části *Přesun pozornosti* a *Neopravené chyby* a *Opravené chyby* v části *Inhibice*, kde korelace dosahují střední síly a jsou statisticky signifikantní. Tyto výsledky lze vysvětlit nízkou variancí v těchto proměnných.

Tabulka 12

Inhibice a přesun pozornosti – Spearmanův korelační koeficient

		Spearmanovo ρ	p hodnota
Pojmenování	Čas	0,71	<0,001
	Neopravené chyby	-0,05	0,798
	Opravené chyby	-0,16	0,378
	Chyby celkem	-0,25	0,166
Inhibice	Čas	0,81	<0,001
	Neopravené chyby	0,40	0,028
	Opravené chyby	0,40	0,028
	Chyby celkem	0,36	0,047
Přesun pozornosti	Čas	0,79	<0,001
	Neopravené chyby	0,15	0,417
	Opravené chyby	0,45	0,012
	Chyby celkem	0,43	0,015

Po provedení Wilcoxonova testu pro párové vzorky byl odhalen statisticky signifikantní rozdíl v části *Pojmenování* pro proměnnou *Čas*, r hodnota naznačuje velkou sílu efektu. Dále byl významný rozdíl v části *Přesun pozornosti* pro *Čas* se střední silou efektu.

Tabulka 13

Inhibice a přesun pozornosti – Wilcoxonův test pro párové vzorky

		p hodnota (Wilcoxonův test)	r hodnota
Pojmenování	Čas	0,004	0,52
	Neopravené chyby	0,590	0,02
	Opravené chyby	0,061	0,44
	Chyby celkem	0,056	0,36
Inhibice	Čas	0,629	0,10
	Neopravené chyby	0,773	0,10
	Opravené chyby	0,397	0,09
	Chyby celkem	0,317	0,10
Přesun pozornosti	Čas	0,019	0,43
	Neopravené chyby	0,151	0,22
	Opravené chyby	0,610	0,14
	Chyby celkem	0,562	0,12

6.2.3. Pracovní paměť

I pro pracovní paměť byl využit neparametrický Spearmanův korelační koeficient, protože data ve verzi A nesplňují podmínku normality rozložení dat. Pro hrubý skór v tomto subtestu dosahoval Spearmanův korelační koeficient $\rho = 0,55$ ($p = 0,001$), tedy středně silné korelace, výsledek je statisticky signifikantní.

Pro zjištění rozdílu mezi hrubými skóry v obou verzích testu byl použit Wilcoxonův test pro párové vzorky, rozdíl nebyl statisticky signifikantní ($p = 0,585$).

6.2.4. Verbální fluence

Pro *Fonemickou fluenci* výsledky naznačují na hladině významnosti $p \leq 0,05$ statisticky signifikantní a silné korelace. Stejně tak pro *Kategorickou fluenci s nutností střídání* mezi kategoriemi i pro *Počet switchů* dosahuje Pearsonův korelační koeficient silné korelace a je statisticky signifikantní. To neplatí pro *Kategorickou fluenci*, kde se ukazuje sice signifikantní korelace pro $p \leq 0,05$, ale korelace je pouze středně silná (Tabulka 14).

Tabulka 14

Verbální fluence – Pearsonův korelační koeficient

	Pearsonovo r	p hodnota
Fonemická pro K/T	0,81	< 0,001
Fonemická pro V/M	0,79	< 0,001
Fonemická celkem	0,90	< 0,001
Kategorická	0,40	0,025
Kategorická se střídání	0,59	< 0,001
Počet switchů	0,60	< 0,001

Kromě *Fonemické fluence V/M* byl dle výsledků zaznamenán statisticky signifikantní rozdíl hrubých skóru u všech subtestů verbální fluence. Tento efekt dosahuje na základě Cohenova d střední velikosti pro hodnoty *Fonemická fluence pro K/T*, *Fonemická fluence celkem* a *Kategorická fluence*. Pro ostatní měřené proměnné je velikost efektu malá (Tabulka 15).

Tabulka 15

Verbální fluence – t-test

	t -statistika	p hodnota	Cohenovo d
Fonemická pro K/T	3,205	0,003	0,58
Fonemická pro V/M	1,194	0,242	0,21
Fonemická celkem	1,194	0,005	0,54
Kategorická	3,878	< 0,001	0,70
Kategorická se střídání	2,425	0,021	0,44
Počet switchů	2,139	0,041	0,38

6.2.5. *Prospektivní paměť*

V tomto subtestu byly měřeny dvě proměnné, v obou proměnných ukazují výsledky středně silnou korelaci, která je statisticky signifikantní pro $p \leq 0,05$ (Tabulka 16).

Tabulka 16

Prospektivní paměť – Spearmanův korelační koeficient

	Spearmanovo ρ	p -hodnota
Hrubý skór	0,41	0,022
Rekognice	0,56	0,001

Wilcoxonův test pro párové vzorky neodhalil statisticky signifikantní rozdíl v hrubých skórech mezi verzemi NB-D pro žádnou z proměnných (Tabulka 17).

Tabulka 17

Prospektivní paměť – Wilcoxonův test pro párové vzorky

	p -hodnota (Wilcoxonův test)	r hodnota
Hrubý skór	0,180	0,24
Rekognice	0,346	0,25

7. Diskuse

Cílem diplomové práce bylo provést studii zaměřenou na posouzení reliability, konkrétně pro paralelnost verzí A a B Neuropsychologické baterie pro děti, a to pro subtesty v doméně Exekutivní funkce. Zkoumanými subtesty byly subtesty *Třídění*, *Inhibice a přesun pozornosti*, *Pracovní paměť*, *Verbální fluence* a *Prospektivní paměť*. Vzhledem k tomu, že vývoj této baterie společně s normativní studií na dětské populaci teprve probíhá, jde o první příspěvek na dané téma.

První otázkou této studie bylo, zda spolu budou subtesty verze A a verze B korelovat. Výsledky předkládané studie ukazují signifikantní korelace pro subtesty *Třídění*, *Pracovní paměť*, *Prospektivní paměť* a *Verbální fluence*, dále pak pro měřené časy v subtestu *Inhibice a přesun pozornosti* a pro chybové skóry *Opravené chyby a Chyby celkem* v části *Přesun pozornosti a Neopravené chyby* a *Opravené chyby* v části *Inhibice* tohoto subtestu.

Druhou otázkou bylo, zda budou rozdílové statistiky skóru jednotlivých subtestů verze A a verze B statisticky signifikantní, což by svědčilo proti ekvivalenci verzí. Signifikantní rozdílové statistiky se objevily u měřených časů subtestu *Inhibice a přesun pozornosti*, kromě času u části *Inhibice*, a také pro všechny proměnné ve *Verbální fluenci*, kromě *Fonemické fluence pro V/M*.

Obě hypotézy můžeme přijmout jako platné u subtestů *Třídění*, *Pracovní paměť* a *Prospektivní paměť* a tyto subtesty je tedy možné považovat za ekvivalentní. Subtest *Třídění* dosahoval mezi jednotlivými verzemi silné korelace ($r = 0,76$), která byla zároveň statisticky signifikantní ($p < 0,001$). Rozdíl hrubých skóru statisticky signifikantní nebyl ($p = 0,106$). V subtestu *Pracovní paměť* byla korelace statisticky významná a dosahovala středně silné korelace ($\rho = 0,55$; $p = 0,001$) a rozdílová statistika taktéž nebyla signifikantní ($p = 0,585$). V subtestu *Prospektivní paměť* byla pro proměnnou *Hrubý skór* zjištěna středně silná korelace ($\rho = 0,41$; $p = 0,022$) a nesignifikantní rozdílová statistika ($p = 0,180$). I pro proměnnou *Rekognice* pak byla zjištěna středně silná korelace ($\rho = 0,56$; $p = 0,001$) a nesignifikantní rozdíl skóru ($p = 0,346$). Statisticky signifikantní středně silné až silné korelace a nevýznamné rozdíly v průměrných výsledcích obou verzí testu odráží konzistenci v měřených exekutivních funkcích u zkoumané populace a naznačují ekvivalenci obou verzí NB-D.

Dobrá psychometrická ekvivalence subtestů ale ještě není dostačující pro jejich doporučení pro praktické využití při opakovaném měření. Například k WCST byla vytvořena alternativní verze CST, která se sice na základě statistických analýz jeví jako ekvivalentní, nebrání ale efektu učení a není tak k opakovanému měření v kratším časovém rozmezí vhodná (Poreh et al., 2012). Podobně tomu bylo při tvorbě alternativní formy Testu cesty (Wagner et al., 2011), ani v této studii se nepodařilo zamezit efektu učení i přes dobrou ekvivalenci všech verzí testu. Shenoyová et al. (2020) na tento problém reagovali snahou zjistit, po kolikáté administraci alternativních verzí Stroopova testu se efekt učení stabilizuje, a zjistili, že jsou k tomu nutná čtyři testování. Doporučují proto administraci čtyřech cvičných úloh. Pro využití ekvivalentních verzí v klinické praxi je tedy nutné zabývat se i dalšími psychometrickými charakteristikami těchto testů, jako je test-retest reliabilita.

V subtestu *Inhibice a přesun pozornosti* se projeví silné korelace pro všechny tři měřené časy, ovšem korelace chybových skóre nebyla dostatečně silná nebo signifikantní ve všech měřených proměnných. Hypotézu zaměřenou na korelaci není možné přijmout pro chybové skóre v části *Pojmenování* (nesignifikantní korelace), pro proměnnou *Chyby celkem* části *Inhibice* ($p = 0,36$; $p = 0,047$) a proměnnou *Neopravené chyby* v části *Přesun pozornosti* ($p = 0,15$; $p = 0,417$). Tyto výsledky je možné připsat nedostatečné varianci rozdílových skóre a neměly by mít vliv na běžnou využitelnost testu. Statisticky signifikantní rozdíly mezi hrubými skóre se pak projeví v proměnné *Čas* v části *Pojmenování*, kde se projevil silný efekt ($p = 0,004$; $r = 0,52$) a části *Přesun pozornosti*, kde se projevil středně silný efekt ($p = 0,019$; $r = 0,43$). Výsledky tedy poukazují na neúplnou psychometrickou ekvivalenci v částech *Pojmenování* a *Přesun pozornosti* pro měřené časy a ve všech chybových skórech.

Vzhledem k charakteru tohoto testu je nutná obezřetnost k efektu učení. V tomto subtestu se mohou probandi naučit strategii k lepšímu řešení úlohy a tu pak využít k při pozdějším testování alternativní formou testu (Kurtz et al., 2004). Proto je vhodné zabývat se nejen ekvivalencí forem, ale také test-retestovou reliabilitou nástroje. Je také možné zabývat se strategií, která by efekt učení snížila. Tou by mohla být například administrace více pokusů, což je způsob, který navrhuje Shenoyová et al. (2020) na základě své studie efektu učení u Stroopova testu. Podobně v testové baterii NEPSY II se v *Inhibici* administrují dva pokusy za sebou (Kemp & Korkman, 2010). Předpokladem je, že by měl efekt učení dosáhnout stropu, po němž se již výsledky

učením nezlepšují. Na druhé straně tomu tak nemusí být u klinické populace, a proto by bylo vhodné výsledky jednotlivých testů i na těchto populacích zkoumat v budoucích studiích.

Posledním zkoumaným subtestem byla *Verbální fluence*. Korelace byla ve všech proměnných signifikantní a pro všechny proměnné kromě *Kategorické fluence* ($r = 0,40$; $p = 0,025$) také dosahovala silné korelace. Nesignifikantní rozdíl mezi hrubými skóry se pak ukázal v proměnné *Fonemická fluence pro V/M* ($p = 0,242$), v ostatních proměnných byly rozdíly signifikantní a velikost efektů byla malá až střední. I v tomto subtestu je psychometrická ekvivalence na základě těchto výsledků neúplná a za statisticky plně ekvivalentní lze považovat pouze *Fonemickou fluenci* pro hlásky V a M.

Ekvivalencí hlásek v testu fonemické fluence se zabývalo již více autorů (Benton et al., 1994; Kopeček & Kuncová, 2006; Paštrnáka et al., 2018; Štorková et al., 2004), ovšem zajištění takové ekvivalence se na české populaci povedlo zatím jen pro hlásky NKT v jedné verzi a BPL ve verzi druhé. Narozdíl od studie předkládané v této práci šlo ale o studii opírající se o dospělé populaci (Kopeček & Kuncová, 2006). Neekvivalence různých hlásek může mít ve verbální fluenci několik důvodů. Mezi ně patří frekvence slov začínajících na určené hlásky v daném jazyce. Pokud je frekvence slov vyšší, může být i jejich výbavnost vyšší. Také ale záleží na složitost těchto slov, protože slova začínající na některé hlásky mohou být kognitivně náročnější na generování než jiné i přes to, že jsou frekventovanější ve zkoumaném jazyce (Kopeček & Kuncová, 2006). V NB-D je pro každou verzi stanovena jedna tvrdá a jedna obojetná souhláska (tedy taková, která se objevuje ve vyjmenovaných slovech). Právě dvojice hlásek obojetných vykazuje dostatečnou psychometrickou ekvivalenci, narozdíl od dvojice souhlásek tvrdých. Bylo by vhodné se v budoucích studiích zabývat tím, zda druh souhlásky v českém jazyce nemá vliv na výbavnost slov, které na ni začínají. Ovšem frekvenci slov v jazyce je obtížné zjistit, navíc je slovní zásoba u dospělých jiná než u dětí. U dětské populace se zpravidla nevyvíjí lineárně a frekvence užívání různých slov je v každém věku odlišná.

Z měřených subtestů je možné považovat za ekvivalentní subtesty *Třídění*, *Pracovní paměť* a *Prospektivní paměť*, jsou tedy vhodné pro praktické využití. Dále se jako ekvivalentní jeví *Fonemická fluence* pro hlásky V a M, a měřený čas v části *Inhibice* subtestu *Inhibice a přesun pozornosti*. I ty je možné považovat za vhodné pro

praktické užití. U většiny subtestů se tedy podařilo prokázat plnou statistickou ekvivalenci obou verzí NB-D. Jednoznačné ekvivalence v subtestech *Inhibice a přesun pozornosti* a *Verbální fluence* bude možné dosáhnout vyrovnáním norem korekcí bodového hodnocení výkonu.

7.1. Limity studie

Velikost výzkumného souboru byla zvolena na základě a priori power analýzy a podařilo se získat dostatečně velký soubor pro stanovenou sílu efektů. Soubor je takové heterogenní na základě věku i pohlaví. Z informací o rodinném zázemí (např. povolání a vzdělání rodičů) se dá usuzovat také na rozmanitý socioekonomický status rodin, ze kterých děti pochází. Přesto že byla velikost vzorku stanovena na základě a priori power analýzy, jde stále o relativně malý soubor, což může být limitem této studie a ovlivnit generalizovatelnost výsledků.

Probandi byli náhodně rozděleni do dvou skupin, kde každé byly verze NB-D administrovány v jiném pořadí, tedy první skupině byla nejprve administrována verze A a poté verze B, u druhé skupiny tomu bylo naopak. Obě verze byly administrovány v rozmezí dvou až čtyřech týdnů. Takto nastavenou metodologií se studie snaží předcházet nejčastějším úskalím při tvorbě alternativních forem testových metod. Tím je jednoznačně předpoklad stálosti měřených atributů, přičemž tento předpoklad je problematický u většiny psychologických fenoménů, a to především u dětí (Urbánek et al., 2011). Tím, že není rozestup mezi jednotlivými testováními příliš dlouhý, neměl by se ve výsledcích testů projevit přirozený vývoj těchto atributů. Ovšem i přes tuto snahu mohlo k vývoji i v tomto intervalu dojít, což by mělo potenciálně vliv na výsledky výzkumu.

Významným problémem je také efekt učení. Při testování exekutivních funkcí musí proband obvykle generovat nové strategie. Přestože je pak alternativní forma testu odlišná, může být zapamatována již použitá strategie a mohlo by tak docházet k nadhodnocování výsledků (Kurtz et al., 2004). Po vzoru předchozích studií byl proto zvolen postup, při kterém došlo k náhodnému rozdělení probandů na dvě skupiny a zadání verzí testu těmto skupinám v odlišném pořadí (Atkinson et al., 2011; Hubley & Jassal, 2005; Poreh et al., 2012; Shenoy et al., 2020; Wagner et al., 2011), aby se předešlo vlivu efektu učení na výsledky statistických analýz. To ale nemusí znamenat, že by alternativní verze efektu učení dostatečně předcházely, protože by mohlo

docházet k zapamatování použité strategie. Pro budoucí užití těchto verzí je nutné zabývat se také test-retestovou reliabilitou (Kurtz et al., 2004).

Mezi další limity studie patří subjektivní faktory na straně probanda, které také mohly mít vliv na výsledky testování. Mezi tyto faktory můžeme zařadit motivaci a subjektivní vnímání testové situace, únavu nebo náladu.

Hlavním přínosem předkládané studie je příspěvek ke studii zaměřené na standardizaci Neuropsychologické baterie pro děti, jíž je tato práce součástí. Jde o první příspěvek na téma paralelnosti forem dvou verzí této testové baterie a přináší tak významné poznatky pro praktické využití testové metody.

Dalším krokem, který umožní aplikaci zjištěných výsledků, je vyrovnávání alternativních forem testu. Přestože byla na základě statistických analýz u většiny subtestů prokázána jejich ekvivalence, nejde o dokonalou paralelnost, té je v psychologickém testování v podstatě nemožné dosáhnout. Proces vyrovnávání se snaží zvýšit srovnatelnost výsledků obou verzí testu a umožnit tak jejich praktické využití (Urbánek et al., 2011). Jako nejvhodnější forma vyrovnávání v testování kognitivních funkcí se jeví equipercentilní vyrovnávání (Gross et al., 2012). To identifikuje skóry na prvním testovém formuláři, které mají stejné percentilové pořadí jako skóry na druhém formuláři testu (González & Wiberg, 2017).

Dále je pro využitelnost celé testové baterie nutné provést analýzu ekvivalence obou verzí testu i v ostatních doménách. Těmi jsou kromě exekutivních funkcí *Paměť a učení*, *Pozornost*, *Řeč*, *Zrakově-percepční funkce*, *Motorické funkce* a *Sociální kognice*. Zároveň je pro užití alternativních forem pro opakované testování v klinické praxi nutné znát i test-retest reliabilitu nástroje. I pokud se podaří zajistit dobrou psychometrickou ekvivalenci dvou verzí testu, nemusí se podařit předejít efektu učení, tak jako se to stalo například při ověřování ekvivalence WCST a CST (Poreh et al., 2012), alternativních forem Testu cesty (Wagner et al., 2011) nebo Stroopova testu (Shenoy et al., 2020). Po prvním testování se probandi naučili princip testu a tuto naučenou strategii pak využili i v dalším testování alternativní verzí (Kurtz et al., 2004; Poreh et al., 2012; Shenoy et al., 2020; Wagner et al., 2011).

Důležitým aspektem je také výkonnostní profil v testové baterii u různých klinických populací. Bylo by tedy vhodné, aby se v budoucnu realizovaly také klinické

studie zaměřené na různé populace. Diagnostická pomůcka by tedy měla poskytovat komplexní psychometrické údaje a uživatel bude volit podle nich, například při opakovaném testování bude zvažovat i alternativu retestu stejným podnětovým materiálem.

Závěr

Cílem práce bylo ověřit paralelnost dvou verzí Neuropsychologické baterie pro děti, a to v doméně exekutivních funkcí. Tato doména zahrnuje subtesty *Třídění*, *Inhibice a přesun pozornosti*, *Pracovní paměť*, *Verbální fluence* a *Prospektivní paměť*.

Obě verze testové baterie byly probandům, dětem ve věku od šesti do devatenácti let, administrovány v rozmezí dvou až čtyřech týdnů. Výzkumný soubor (N = 31) byl náhodně rozdělen na dvě skupiny, kterým byly verze zadány v opačném pořadí. Dále byl všem probandům administrován anamnestický dotazník a Zkrácená Wechslerova inteligenční škála, druhé vydání. Statistická analýza dat zahrnovala korelační analýzu (Pearsonův nebo Spearmanův korelační koeficient) a rozdílové statistiky (t-test nebo Wilcoxonův test pro párové vzorky).

Na základě statistických analýz mohou být alternativní verze NB-D považovány za ekvivalentní v subtestu *Třídění*, kde se projevila silná korelace a statisticky nesignifikantní rozdíl hrubých skóre. Středně silné korelace pak byly zjištěny v subtestech *Pracovní paměť* a *Prospektivní paměť*, rozdílové statistiky nebyly signifikantní. I v těchto subtestech lze na základě této studie předpokládat ekvivalenci.

V subtestech *Inhibice a přesun pozornosti* a *Verbální fluence* byla zjištěna neúplná psychometrická ekvivalence verzí testové baterie. Pro první zmíněný subtest lze za ekvivalentní považovat pouze měřený čas v části *Inhibice*. Signifikantní a dostatečně silné korelace se projevily pouze v měřených časech u všech tří částí subtestu. Rozdílové statistiky pak byly signifikantní v měřených časech v částech *Pojmenování* a *Přesun pozornosti*. V subtestu *Verbální fluence* dosahovaly korelační analýzy dostatečné statistické signifikance i síly, ale signifikantní rozdíly hrubých skóre byly ve všech proměnných kromě *Fonemické fluence* pro hlásky V a M, ekvivalenci se tedy podařilo potvrdit pouze u této proměnné.

Předkládaná studie potvrzuje, že podnětový materiál zmíněných subtestů Neuropsychologické baterie pro děti ve verzi A i ve verzi B vybízí testované ke srovnatelné výkonnosti, výstupy mohou být mezi verzemi považovány za konzistentní. U většiny sledovaných subtestů se podařilo prokázat plnou paralelnost mezi verzemi A a B, k dosažení jednoznačnější ekvivalence u subtestů *Inhibice a přesun pozornosti* a

Verbální fluence bude možné dojít přes analýzy směřující k vyrovnání norem korekcí bodového hodnocení výkonu, tedy dosaženého testového skóru.

Diplomová práce je spojena se standardizací nově vznikající Neuropsychologické baterie pro děti a jako první se zabývá potřebou ověření paralelnosti obou verzí této testové baterie. Jedná se však pouze o první krok, obdobným procesem jako doména *Exekutivní funkce* budou muset projít i ostatní testové domény.

Reference

- Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, 87(2), 85–106. <https://doi.org/10.1016/j.jecp.2003.10.002>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*. <https://doi.org/10.1176/appi.books.9780890425596>
- American Psychological Association (APA). (2020). *Publication Manual of the American Psychological Association (7th ed.)*. American Psychological Association.
- Andersen S. L. (2018). Stress, sensitive periods, and substance abuse. *Neurobiology of Stress*, 10, 100140, 1–14. <https://doi.org/10.1016/j.ynstr.2018.100140>
- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences*, 18(4), 177–185. <https://doi.org/10.1016/j.tics.2013.12.003>
- Atkinson, T. M., Ryan, J. P., Kryza, M., & Charette, L. M. (2011). Using versions of the trail making test as alternate forms. *The Clinical Neuropsychologist*, 25(7), 1193–1206. <https://doi.org/10.1080/13854046.2011.589410>
- Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, 8, 47–89. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- Banich, M. T., & Depue, B. E. (2015). Recent advances in understanding neural systems that support inhibitory control. *Current Opinion in Behavioral Sciences*, 1, 17–22. <https://doi.org/10.1016/j.cobeha.2014.07.006>
- Beglinger, L. J., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D. A., Crawford, J., Fastenau, P. S., & Siemers, E. R. (2005). Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical*

Neuropsychology : The Official Journal of the National Academy of Neuropsychologists, 20(4), 517–529. <https://doi.org/10.1016/j.acn.2004.12.003>

Bell, M. A. (2012). A psychobiological perspective on working memory performance at 8 months of age. *Child Development*, 83(1), 251–265. <https://doi.org/10.1111/j.1467-8624.2011.01684.x>

Benton, A. L. (1968). Differential behavioral effects in frontal lobe disease. *Neuropsychologia*, 6, 53–60. [https://doi.org/10.1016/0028-3932\(68\)90038-9](https://doi.org/10.1016/0028-3932(68)90038-9)

Benton, A. L., Hamsher, K. S., & Sivan, A. B. (1994). *Multilingual Aphasia Examination (3rd ed.)*. Psychological Assessment Resources.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, 81(6), 1641–1660. <https://doi.org/10.1111/j.1467-8624.2010.01499.x>

Bezdiček, O., Georgi, H., Nikolai, T., & Kopeček, M. (2021). *Pražská verze Stroopova testu*. Karolinum.

Bezdiček, O., Michalec, J., & Shallice, T. (2018). *Londýnská věž (ToL), Schalliceova verze ToL*. Karolinum.

Bezdicek, O., Motak, L., Axelrod, B. N., Preiss, M., Nikolai, T., Vyhnálek, M., Poreh, A., & Růžička, E. (2012). Czech version of the trail making test. *Archives of Clinical Neuropsychology*, 27, 906–914. <https://doi.org/10.1093/arclin/acs084>

Blankenship, T. L., Slough, M. A., Calkins, S. D., Deater-Deckard, K., Kim-Spoon, J., & Bell, M. A. (2019). Attention and executive functioning in infancy: Links to childhood executive function and reading achievement. *Developmental Science*, 22(6), 1–9. <https://doi.org/10.1111/desc.12824>

Broomell, A. P. R., & Bell, M. A. (2022). Longitudinal development of executive function from infancy to late childhood. *Cognitive Development*, 63, 101229. <https://doi.org/10.1016/j.cogdev.2022.101229>

Bukačová, K., Lhotová, P., & Maulisová, A. (2021). Neuropsychologická testová baterie pro děti. *E-psychologie*, 15(1), 90–91. <https://doi.org/10.29364/epsy.394>

- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. <https://doi.org/10.1037/h0046016>
- Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more: Executive function and symbolic representation in preschool children. *Psychological Science*, *16*(8), 609–616. <https://doi.org/10.1111/j.1467-9280.2005.01583.x>
- Chen, J., Scheller, M., Wu, C., Hu, B., Peng, R., Liu, C., Liu, S., Zhu, L., & Chen, J. (2022). The relationship between early musical training and executive functions: Validation of effects of the sensitive period. *Psychology of Music*, *50*(1), 86–99. <https://doi.org/10.1177/0305735620978690>
- 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. <http://dx.doi.org/10.29364/epsy.329>
- Cimlerová, P., Pokorná, D., & Chalupová, E. (2007). *Diagnostika specifických poruch učení u adolescentů a dospělých osob*. Institut pedagogicko-psychologického poradenství ČR.
- Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience*, *33*(1), 269–298. <https://doi.org/10.1146/annurev.neuro.051508.135409>
- Corsi, P. M. (1972). *Human Memory and the Medial Temporal Region of the Brain*. [Thesis, McGill University, Montreal]. <https://escholarship.mcgill.ca/concern/theses/05741s554>
- Cristofori, I., Cohen-Zimmerman, S., & Grafman, J. (2019). Executive functions. *Handbook of Clinical Neurology*, *163*, 197–219. <https://doi.org/10.1016/B978-0-12-804281-6.00011-2>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334. <https://doi.org/10.1007/BF02310555>
- Cuevas, K., Swingler, M. M., Bell, M. A., Marcovitch, S., & Calkins, S. D. (2012). Measures of frontal functioning and the emergence of inhibitory control processes at 10

months of age. *Developmental Cognitive Neuroscience*, 2(2), 235–243.
<https://doi.org/10.1016/j.dcn.2012.01.002>

Černochová, D., Goldmann, P., Král, P., Soukupová, T., Šnorek, V., & Havlůj, V. (2010). *WAIS-III - Wechslerova inteligenční škála pro dospělé*. Hogrefe–Testcentrum.

De Ayala, R., & Santiago, S. (2016). An introduction to mixture item response theory models. *Journal of School Psychology*, 60, 25–40.
<https://doi.org/10.1016/j.jsp.2016.01.002>

Delaney, R. C., Prevey, M. L., Cramer, J., Mattson, R. H., & VA Epilepsy Cooperative Study #264 Research Group. (1992). Test-retest comparability and control subject data for the rey-auditory verbal learning test and rey-osterrieth/taylor complex figures. *Archives of Clinical Neuropsychology*, 7(6), 523–528.
<https://doi.org/10.1093/arclin/7.6.523>

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS)*. Pearson.

DeMars, C. (2010). *Item Response Theory*. Oxford University Press.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.
<https://doi.org/10.1146/annurev-psych-113011-143750>

Dias, N. M., Helsdingen, I. E., Lins, E. K. R. M., Etcheverria, C. E., Dechen, V. A., Steffen, L., Cardoso, C. O., & Lopes, F. M. (2023). Executive functions beyond the “Holy trinity”: A scoping review. *Neuropsychology*, 38(2), 107–125.
<https://dx.doi.org/10.1037/neu0000922>

Doebel, S. (2020). Rethinking executive function and its development. *Perspectives on Psychological Science*, 15(4), 942–956. <https://doi.org/10.1177/1745691620904771>

Drozdová, K., Štěpánková, H., Lukavský, J., Bezdíček, O., & Kopeček, M. (2015). Normativní studie testu Reyovy-Osterriethovy komplexní figury v populaci českých seniorů. *Česká a slovenská neurologie a neurochirurgie*, 78/III(5), 542–549.
<https://www.csn.eu/casopisy/ceska-slovenska-neurologie/2015-5-3/normativni-studie-testu-reyovy-osterriethovy-komplexni-figury-v-populaci-ceskych-senioru-55997>

- Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cognitive Psychology*, *30*(3), 257–303. <https://doi.org/10.1006/cogp.1996.0008>
- Elithorn, A. (1955). A preliminary report on a perceptual maze test sensitive to brain damage. *Journal of Neurology, Neurosurgery, and Psychiatry*, *18*(4), 287–292. <https://10.1136/jnnp.18.4.287>
- Emslie, H., Wilson, F. C., Burden, V., Nimmo–Smith, I., & Wilson, B. A. (2013). *Behavioural Assessment of Dysexecutive Syndrome for Children (BADS-C)*. Pearson.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. <https://doi.org/10.3758/BF03203267>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160.
- Fiske, A., & Holmboe, K. (2019). Neural substrates of early executive function development. *Developmental Review*, *52*, 42–62. <https://doi.org/10.1016/j.dr.2019.100866>
- Franzen, M. D., Paul, D. S., & Price, G. (1990). Alternate form reliability of Trails A, B, C, and D. In *Ninth Annual Convention of the National Academy of Neuropsychology*. Reno.
- Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology*, *47*(5), 1410–1430. <https://doi.org/10.1037/a0023750>
- Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in

origin. *Journal of Experimental Psychology: General*, 137(2), 201–225.
<https://doi.org/10.1037/0096-3445.137.2.201>

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: evidence from national curriculum assessments at 7 and 14 years of age. *Cognitive Psychology*, 18(1), 1–16.
<https://doi.org/10.1002/acp.934>

Gilman, S. E., Huang, Y. T., Jimenez, M. P., Agha, G., Chu, S. H., Eaton, C. B., Goldstein, R. B., Kelsey, K. T., Buka, S. L., & Loucks, E. B. (2019). Early life disadvantage and adult adiposity: tests of sensitive periods during childhood and behavioural mediation in adulthood. *International Journal of Epidemiology*, 48(1), 98–107. <https://doi.org/10.1093/ije/dyy199>

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). *Behavior Rating Inventory of Executive Function, Second Edition*. Psychological Assessment Resources.

González, J., & Wiberg, M. (2017). *Applying Test Equating Methods Using R*. Springer.

Grant, D. A., & Berg, E. A. (1948). *Wisconsin Card Sorting Test*. APA PsycTests.

Gross, A. L., Inouye, S. K., Rebok, G. W., Brandt, J., Crane, P. K., Parisi, J. M., Tommet, D., Bandeen-Roche, K., Carlson, M. C., & Jones, R. N. (2012). Parallel but not equivalent: Challenges and solutions for repeated assessment of cognition over time. *Journal of Clinical and Experimental Neuropsychology*, 34(7), 758–772.
<https://doi.org/10.1080/13803395.2012.681628>

Hamby, S., Wilkins, J. W., & Barry, N. S. (1993). Organizational quality on the Rey-Osterrieth and Taylor complex figure tests: A new scoring system. *Psychological Assessment*, 5(1), 27–33. <https://doi.org/10.1037/1040-3590.5.1.27>

Hammersley, M. (1987). Some notes on the terms “Validity” and “Reliability.” *British Educational Research Journal*, 13(1), 73–82.
<https://doi.org/10.1080/0141192870130107>

Heaton, R. K. (1981). *Wisconsin card sorting test manual*. Psychological Assessment Resources.

- Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and Expanded*. Psychological Assessment Resources.
- Heaton, R. K., & PAR Staff. (2008). *Wisconsin Card Sorting Test: Computer Version 4—Research Edition*. PAR.
- Helmstadter, G. C. (1964). *Principles of Psychological Measurement*. Appleton-Century-Crofts.
- Henry, L. A., & Bettenay, C. (2010). The assessment of executive functioning in children. *Child and Adolescent Mental Health, 15*(2), 110–119. <https://doi.org/10.1111/j.1475-3588.2010.00557.x>
- Henik, A., Bugg, J. M., & Goldfarb, L. (2018). Inspired by the past and looking to the future of the Stroop effect. *Acta Psychologica, 189*, 1–3. <https://doi.org/10.1016/j.actpsy.2018.06.007>
- Huble, A. M. (1996). Modification of the Taylor complex figure: A comparable figure to the Rey-Osterrieth figure? *Edgeworth Series in Quantitative Behavioural Science, 96–97*.
- Huble, A. M., & Jassal, S. (2005). Comparability of the Rey-Osterrieth and Modified Taylor complex figures using total scores, completion times, and construct validation. *Journal of Clinical and Experimental Neuropsychology, 28*(8), 1482–1497. <https://doi.org/10.1080/13803390500434441>
- Johnston, J. M., Pennypacker, H. S. & Deitz, S. M. (1981). Strategies and tactics of human behavioral research. *The Behavior Analyst 4*, 61–68. <https://doi.org/10.1007/BF03391853>
- Kaplan, R. M. & Saccuzzo, D. P. (2018). *Psychological Testing: Principles, Applications, & Issues (9th ed.)*. Cengage Learning.
- Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin, 144*(11), 1147–1185. <https://doi.org/10.1037/bul0000160>

Kemp, S. L., & Korkman, M. (2010). *Essentials of NEPSY-II Assessment*. John Wiley & Sons.

Kim, C., Cilles, S. E., Johnson, N. F., & Gold, B. T. (2012). Domain general and domain preferential brain regions associated with different types of task switching: A meta-analysis. *Human Brain Mapping, 33*(1), 130–142. <https://doi.org/10.1002/hbm.21199>

Klaming, L., & Vlaskamp, B. N. S. (2018). Non-dominant hand use increases completion time on part B of the Trail Making Test but not on part A. *Behavior Research Methods, 50*(3), 1074–1087. <https://doi.org/10.3758/s13428-017-0927-1>

Kline, P. (1993). *The Handbook of Psychological Testing*. Routledge.

Kölesová, E. (1993). *Rey-Osterriethov test komplexnej figúry ako pomôcka diagnostiky priestorových schopností u žiakov stredných škôl – Štandardizačná skúška*. [Diplomová práca, Masarykova univerzita v Brně]. https://is.muni.cz/th/kubwf/Samankova_Disertacni_prace.pdf

Kopeček, M., & Kuncová, A. (2006). Efekt nácviku testu generování slov a testování alternativní verze. Pilotní studie. *Psychiatrie, 10*(4), 211–215. https://www.researchgate.net/publication/283795475_Practice_Effect_in_Verbal_Fluency_Test_and_Assessment_of_an_Alternative_Version_Pilot_Study_Efekt_nacviku_testu_generovani_slov_a_testovani_alternivni_verze_Pilotni_studie

Košč, M. , Novák, J. (1997). *Rey-Osterriethova komplexní figura - příručka k testu*. Psychodiagnostika.

Krčová, V. (2014). Rey-Osterriethova komplexní figura – recenze metody. *Testforum, 4*, 22–26. <http://testforum.cz/domains/testforum.cz/index.php/testforum/article/view/24#.VapGEreIq70>

Krivá, E. (2010). *Validizace Stroopova testu na české populaci*. [Diplomová práce, Univerzita Karlova v Praze]. https://dspace.cuni.cz/bitstream/handle/20.500.11956/21392/DPTX_2009_2_11210_AS_ZK10001_126943_0_86274.pdf?sequence=1&isAllowed=y

- Krivá, E. (2013). *Stroopův test*. Testcentrum.
- Kulišťák, P. (2017). *Klinická neuropsychologie v praxi*. Karolinum.
- Kurtz, M. M., Ragland, J. D., Moberg, P. J., & Gur, R. C. (2004). The Penn Conditional Exclusion Test: A new measure of executive-function with alternate forms for repeat administration. *Archives of Clinical Neuropsychology*, 19(2), 191–201. [https://doi.org/10.1016/S0887-6177\(03\)00003-9](https://doi.org/10.1016/S0887-6177(03)00003-9)
- Lehner P. N. (1996). *Handbook of Ethological Methods (2nd ed.)*. Cambridge University Press.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21(1), 59–80. <https://doi.org/10.1348/026151003321164627>
- Lezak, M. D. (2012). *Neuropsychological Assessment (5th ed.)*. Oxford University Press.
- Lezak, M. D., Howieson, D. B. & Loring, D. W. (2004). *Neuropsychological Assessment (4th ed.)*. Oxford University Press.
- Livingston, S. A. (2004). *Equating Test Scores (Without IRT)*. Educational Testing Service. <http://www.ets.org/Media/Research/pdf/LIVINGSTON.pdf>
- Lynch, J. D., Xu, Yolton, K., Khoury, J. C., Chen, A., Lanphear, B. P., Cecil, K. M., Braun, J. M., & Epstein, J. N. (2023). Environmental predictors of children's executive functioning development. *Child Neuropsychology*, 1–21. <https://doi.org/10.1080/09297049.2023.2247603>
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163–203. <https://doi.org/10.1037/0033-2909.109.2.163>
- Manly, T., Robertson, I. H., Anderson, V., & Nimmo-Smith, I. (1998). *Test of Everyday Attention for Children (TEA-Ch)*. Pearson.

McCracken, L. M., & Franzen, M. D. (1992). Principal-components analysis of the equivalence of alternate forms of the Trail Making Test. *Psychological Assessment*, 4(2), 235–238. <https://doi.org/10.1037/1040-3590.4.2.235>

McDonald, R. P. (1999). *Test Theory: A Unified Treatment*. Lawrence Erlbaum Associates.

McKenna, R., Rushe, T., & Woodcock, K. A. (2017). Informing the structure of executive function in children: A meta-analysis. *Frontiers in Human Neuroscience*, 11. <https://doi.org/10.3389/fnhum.2017.00154>

Meyers, J.E. and Meyers, K.R. (1995) *Rey Complex Figure Test and Recognition Trial: Professional Manual*. Psychological Assessment Resources.

Microsoft Corporation. (2018). *Microsoft Excel*. <https://office.microsoft.com/excel>

Miller, B. L., & Cummings, J. L. (2007). *The Human Frontal Lobes: Functions and Disorders (2nd ed.)*. The Guilford Press.

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, 13(3), 395–423. <https://doi.org/10.1080/15248372.2011.585478>

Miller, S. E., Galvagno, L. G., & Elgier, Á. (2023). Universality and context-specificity in early executive function development. *Infant Behavior & Development*, 71, 1–22. <https://doi.org/10.1016/j.infbeh.2023.101841>

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>

Miyake, A., & Friedman, N. P. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, 86, 186–204. <https://doi.org/10.1016/j.cortex.2016.04.023>

Motýl, J. (2015). Test cesty. Recenze metody. *Testforum*, 4(6), 47–52. <https://doi.org/10.5817/TF2015-6-85>

- Nikolai, T., Štěpánková, H., Michalec, J., Bezdíček, O., Horáková, K., Marková, H., Růžička, E., & Kopeček, M. (2015). Testy verbální fluence, česká normativní studie pro osoby vyššího věku. *Česká a slovenská neurologie a neurochirurgie*, 78/111(3), 292–299. <https://doi.org/10.14735/amcsnn2015292>
- Novák, T. (2009). Zobrazovací metody v psychiatrické praxi. *Psychiatrie pro praxi*, 10(1), 12–16. <https://www.solen.sk/storage/file/article/73491a06bcb49ac2c66cc341a14033e3.pdf>
- Oluwatayo, J. A. (2012). Validity and reliability issues in educational research. *Journal of Educational and Social Research*, 2(2), 391–400. <https://ideas.repec.org/a/bjz/jesrjr/135.html>
- Osterrieth, P. A. (1944). Le test de copie d'une figure complexe; contribution à l'étude de la perception et de la mémoire. *Archives de Psychologie*, 30, 206–356.
- Paštrnák, M., Šulcová, K., Dorazilová, A., & Rodriguez, M. (2018). Alternative forms parallel to the Czech versions of Rey Auditory Verbal Learning Test, Complex Figure Test and Verbal Fluency. *Czech and Slovak Neurology and Neurosurgery*, 81(1), 73–80. <https://doi.org/10.14735/amcsnn201873>
- Park, S. Y., & Schott, N. (2022). The trail-making-test: Comparison between paper-and-pencil and computerized versions in young and healthy older adults. *Applied neuropsychology. Adult*, 29(5), 1208–1220. <https://doi.org/10.1080/23279095.2020.1864374>
- Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal-and temporal-lobe lesions in man. *Neuropsychologia*, 20(3), 249–262. [https://doi.org/10.1016/0028-3932\(82\)90100-2](https://doi.org/10.1016/0028-3932(82)90100-2)
- Pezzulo, G., & Cisek, P. (2016). Navigating the affordance landscape: Feedback control as a process model of behavior and cognition. *Trends in Cognitive Sciences*, 20(6), 414–424. <https://doi.org/10.1016/j.tics.2016.03.013>
- Poreh, A. M., Fishman, D. M., & Yocum, A. A. (2007). The Cleveland Card Sorting Test. *Archives of clinical neuropsychology*, 22(7), 880.

- Poreh, A. M., Pastel, D., Miller, A., & Levin, J. (2012). The Cleveland Sorting Test: A preliminary study of an alternate form of the Wisconsin Card-Sorting Test. *Applied Neuropsychology: Adult*, *19*(2), 147–152. <https://doi.org/10.1080/09084282.2011.643952>
- Preiss, M. et al. (2012). *Neuropsychologická baterie Psychiatrického centra Praha: klinické vyšetření základních kognitivních funkcí, 3. přepracované vydání*. Psychiatrické centrum.
- Preiss, M., & Preiss, J. (2006). *Test cesty*. Psychodiagnostika s.r.o.
- Ptáček, R. (2011). *BRIEF – Škála hodnocení exekutivních funkcí u dětí, 1. české vydání*. Hogrefe – Testcentrum.
- R Core Team (2023). *R (4.3.1): A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org>
- Reineberg, A. E., Andrews-Hanna, J. R., Depue, B. E., Friedman, N. P., & Banich, M. T. (2015). Resting-state networks predict individual differences in common and specific aspects of executive function. *NeuroImage*, *104*, 69–78. <https://doi.org/10.1016/j.neuroimage.2014.09.045>
- Reitan, R. M., & Wolfson, M. (2004). The Trail Making Test as an initial screening procedure for neuropsychological impairment in older children. *Archives of Clinical Neuropsychology*, *19*(2), 281–288. [https://doi.org/10.1016/S0887-6177\(03\)00042-8](https://doi.org/10.1016/S0887-6177(03)00042-8)
- Rey, A. (1941). The psychological examination in cases of traumatic encephalopathy. Problems. *Archives de Psychologie*, *28*, 215–285.
- Reynolds, C. R., Altmann, R. A., & Allen, D. N. (2021). *Mastering Modern Psychological Testing: Theory and Methods*. Springer Nature.
- Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Applied Neuropsychology*, *12*(4), 190–201. https://doi.org/10.1207/s15324826an1204_2
- Ross, T.P., Furr, A. E., Carter, S. E., & Weinberg, M. (2006). The psychometric equivalence of two alternate forms of the Controlled Oral Word Association Test. *The*

Clinical Neuropsychologist, 20(3), 414–431.
<https://doi.org/10.1080/13854040590967153>

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., Fox, P. T., & Eickhoff, S. B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *NeuroImage*, 60(1), 830–846.
<https://doi.org/10.1016/j.neuroimage.2011.11.050>

Rust, J., Golombok, S. (1999). *Modern Psychometrics. The Science of Psychological Assessment*. Routledge.

Rust, J., Kosinski, M., & Stillwell, D. (2021). *Modern Psychometrics. The Science of Psychological Assessment. (4th ed.)*. Routledge.

Sedláčková, P., & Galbavá, S. (2021). Barevný test cesty pro děti: Recenze metody. *Testforum*, 14, 1–7. <https://doi.org/10.5817/TF2021-14-14205>

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London, Biology*, 298, 199–209. <https://doi.org/10.1098/rstb.1982.0082>

Sharma, A., Dorman, M. F., & Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. *Ear and Hearing*, 23(6), 532–539.
<https://doi.org/10.1097/00003446-200212000-00004>

Shenoy, S., Sathe, P. K., & Sathe, A. (2020). Practice and order effect of 5 alternate forms of Colour Word Stroop Test: A pilot study. *Medical Science*, 24(104), 2760–2766.

https://www.researchgate.net/publication/342877889_Practice_and_order_effect_of_5_alternate_forms_of_Colour_Word_Stroop_Test_A_Pilot_Study

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, 81(1), 174–176. <https://doi.org/10.1037/h0027448>

Smolker, H. R., Depue, B. E., Reineberg, A. E., Orr, J. M., & Banich, M. T. (2015). Individual differences in regional prefrontal gray matter morphometry and fractional anisotropy are associated with different constructs of executive function. *Brain Structure and Function* 220, 1291–1306. <https://doi.org/10.1007/s00429-014-0723-y>

Smolker, H. R., Friedman, N. P., Hewitt, J. K., & Banich, M. T. (2018). Neuroanatomical correlates of the unity and diversity model of executive function in young adults. *Frontiers in Human Neuroscience*, *12*. <https://doi.org/10.3389/fnhum.2018.00283>

Sokolová, H., & Cígler, H. (2018). Test cesty: Normativní česká data pro žáky 9. tříd ZŠ až 4. ročníků SŠ. *Testforum*, *11*, 1–20. <http://dx.doi.org/10.5817/TF2018-11-197>

Sonuga-Barke, E. J. S., Kennedy, M., Kumsta, R., Knights, N., Golm, D., Rutter, M., Maughan, B., Schlotz, W., & Kreppner, J. (2017). Child-to-adult neurodevelopmental and mental health trajectories after early life deprivation: the young adult follow-up of the longitudinal English and Romanian adoptees study. *Lancet*, *389*(10078), 1539–1548. [https://doi.org/10.1016/S0140-6736\(17\)30045-4](https://doi.org/10.1016/S0140-6736(17)30045-4)

Steinmetz, J.-P., Brunner, M., Loarer, E., & Houssemand, C. (2010). Incomplete psychometric equivalence of scores obtained on the manual and the computer version of the Wisconsin Card Sorting Test? *Psychological Assessment*, *22*(1), 199–202. <https://doi.org/10.1037/a0017661>

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary*. Oxford University Press.

Strauss, E., & Spreen, O. (1990). A comparison of the Rey and Taylor figures. *Archives of Clinical Neuropsychology*, *5*(4), 417–420. [https://doi.org/10.1016/0887-6177\(90\)90020-P](https://doi.org/10.1016/0887-6177(90)90020-P)

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. <https://doi.org/10.1037/h0054651>

Stuss D. T. (2011). Functions of the frontal lobes: relation to executive functions. *Journal of the International Neuropsychological Society*, *17*(5), 759–765. <https://doi.org/10.1017/S1355617711000695>

Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., Borowiec, A., & Katz, D. I. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia*, *43*(3), 396–417. <https://doi.org/10.1016/j.neuropsychologia.2004.06.010>

- Stuss, D. T., & Benson, D. F. (1986). *The Frontal Lobes*. Raven Press.
- Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A multidisciplinary approach to anterior attentional functions. *Structure and Functions of the Human Prefrontal Cortex*, 769(1), 191–212.
<https://doi.org/10.1111/j.1749-6632.1995.tb38140.x>
- Sürücü, L. & Maslakçı, A. (2020). Validity and reliability in quantitative research. *Business & Management Studies: an International Journal*, 8(3): 2694–2726.
<https://doi.org/10.15295/bmij.v8i3.1540>
- Sylvester, C. C., Wager, T. D., Lacey, S. C., Hernandez, L., Nichols, T. E., Smith, E. E., & Jonides, J. (2003). Switching attention and resolving interference: fMRI measures of executive functions. *Neuropsychologia*, 41(3), 357–370.
[https://doi.org/10.1016/S0028-3932\(02\)00167-7](https://doi.org/10.1016/S0028-3932(02)00167-7)
- Šnoblová, V., & Krejčová, L. (2017). *Barevný test cesty pro děti*. Propsyco.
- Štorková, P, Preiss, M., & Kopeček, M. (2004). Efekt nácviku testu verbální fluence a testování alternativní verze. Pilotní studie. *Psychiatrie*, 8(3), 187–190.
<https://doi.org/10.14735/amcsnn2015292>
- Taylor, L. B. (1969) Localisation of cerebral lesions by psychological testing: Chapter XIV. *Neurosurgery* 16, 269–287. https://10.1093/neurosurgery/16.CN_suppl_1.269
- Telecká, S. (2013). WCST–Wisconsinský test třídění karet. 1. české vydání. Hogrefe–Testcentrum.
- Templin, J. (2016). Item response theory. *The Encyclopedia of Adulthood and Aging*.
<https://doi.org/10.1002/9781118521373.wbeaa320>
- Thompson, A., & Steinbeis, N. (2020). Sensitive periods in executive function development. *Current Opinion in Behavioral Sciences*, 36, 98–105.
<https://doi.org/10.1016/j.cobeha.2020.08.001>
- Thurstone, L. L., & Thurstone, T. G. (1962). *Primary Mental Abilities*. Science Research Associates.

Tirapu-Ustárroz, J., Bausela-Herrerias, E., & Cordero-Andrés, P. (2018). Model of executive functions based on factorial analyses in child and school populations: A meta-analysis. *Revista de Neurología*, 67(6), 215–225. <https://doi.org/10.33588/rn.6706.2017450>

Tirapu-Ustárroz, J., Cordero-Andrés, P., Luna-Lario, P., & Hernández-Goñi, P. (2017). Proposed model of executive functions based on factorial analyses. *Revista de Neurología*, 64(2), 75–84. <https://doi.org/10.33588/rn.6402.2016227>

Tombaugh, T. N. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology*, 19, 203-214.

Tombaugh, T. N., Faulkner, P., & Hubley, A. M. (1992). Effects of age on the Rey-Osterrieth and Taylor complex figures: Test-retest data using an intentional learning paradigm. *Journal of Clinical and Experimental Neuropsychology*, 14(5), 647–661. <https://doi.org/10.1080/01688639208402853>

Tremblay, M. P., Potvin, O., Callahan, B. L., Belleville, S., Gagnon, J. F., Caza, N., Ferland, G., Hudon, C., & Macoir, J. (2015). Normative data for the Rey–Osterrieth and the Taylor Complex Figure tests in Quebec-French people. *Archives of Clinical Neuropsychology*, 30(1), 78–87, <https://doi.org/10.1093/arclin/acu069>

Tsuchida, A., & Fellows, L. K. (2009). Lesion evidence that two distinct regions within prefrontal cortex are critical for n-back performance in humans. *Journal of Cognitive Neuroscience*, 21(12), 2263–2275. <https://doi.org/10.1162/jocn.2008.21172>

Tsuchida, A., & Fellows, L. K. (2013). Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 49(7), 1790–1800. <https://doi.org/10.1016/j.cortex.2012.10.014>

Urbánek, T. (2000): *Strukturální modelování v psychologii*. Psychologický ústav AV ČR a Nakladatelství Pavel Křepela.

Urbánek, T., Denglerová, D., & Širůček, J. (2011). *Psychometrika: Měření v psychologii*. Portál.

- Ústav zdravotnických informací a statistiky ČR. (2021). *MKN-10: Mezinárodní klasifikace nemocí a přidružených zdravotních problémů: desátá revize. Aktualizované vydání k 1. 1. 2022*. <https://mkn10.uzis.cz>
- Villalobos, D., Povedano-Montero, J., Fernández, S., López-Muñoz, F., Pacios, J., & del Río, D. (2022). Scientific research on verbal fluency tests: A bibliometric analysis. *Journal of Neurolinguistics*, 63. <https://doi.org/10.1016/j.jneuroling.2022.101082>
- Wagner, S., Helmreich, I., Dahmen, N., Lieb, K., & Tadic, A. (2011). Reliability of three alternate forms of the trail making tests A and B. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 26(4), 314–321. <https://doi.org/10.1093/arclin/acr024>
- Wager, T. D., Sylvester, C. C., Lacey, S. C., Nee, D. E., Franklin, M., & Jonides, J. (2005). Common and unique components of response inhibition revealed by fMRI. *NeuroImage*, 27(2), 323–340. <https://doi.org/10.1016/j.neuroimage.2005.01.054>
- Wechsler D. (2011) *Wechsler Abbreviated Scale of Intelligence (2nd ed.)*. Pearson.
- Wechsler D. & Psychological Corporation. (1997). *WAIS III: Administration and Scoring Manual: Wechsler Adult Intelligence Scale (3rd ed.)*. Psychological Corporation.
- Whiston, S. C. (2012). *Principles and Applications of Assessment in Counseling*. Cengage Learning.
- Wild, K. V., & Musser, E. D. (2013). The Cambridge Neuropsychological Test Automated Battery in the assessment of executive functioning. In *Handbook of Executive Functioning*, 171-190. Springer.
- Yamashita H. (2009). One-year delayed recall performance of the Rey-Osterrieth Complex Figure in a healthy young adult sample. *Applied Neuropsychology*, 16(2), 141–143. <https://doi.org/10.1080/09084280802623064>
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1(1), 297-301. <https://doi.org/10.1038/nprot.2006.46>

Seznam obrázků

- Obrázek 1 Unity and Diversity model exekutivních funkcí
- Obrázek 2 Tvorba alternativní verze TMT
- Obrázek 3 Kongruentní, inkongruentní a neutrální vztah

Seznam tabulek

Tabulka 1	Četnost povolání rodičů
Tabulka 2	Třídění – deskriptivní statistika a Shapirův-Wilkův test
Tabulka 3	Pojmenování – deskriptivní statistika a Shapirův-Wilkův test
Tabulka 4	Inhibice – deskriptivní statistika a Shapirův-Wilkův test
Tabulka 5	Přesun pozornosti – deskriptivní statistika a Shapirův-Wilkův test
Tabulka 6	Pracovní paměť – deskriptivní statistika a Shapirův-Wilkův test
Tabulka 7	Fonemická fluence – deskriptivní statistika a Shapirův-Wilkův test
Tabulka 8	Kategorická fluence – deskriptivní statistika a Shapirův-Wilkův test
Tabulka 9	Prospektivní paměť – deskriptivní statistika a Shapirův-Wilkův test
Tabulka 10	Třídění – Pearsonův korelační koeficient
Tabulka 11	Třídění – <i>t</i> -test
Tabulka 12	Inhibice a přesun pozornosti – Spearmanův korelační koeficient
Tabulka 13	Inhibice a přesun pozornosti – Wilcoxonův test pro párové vzorky
Tabulka 14	Verbální fluence – Pearsonův korelační koeficient
Tabulka 15	Verbální fluence – <i>t</i> -test
Tabulka 16	Prospektivní paměť – Spearmanův korelační koeficient
Tabulka 17	Prospektivní paměť – Wilcoxonův test pro párové vzorky