

Test Kvantitativního usuzování Triton a hladový oceán – technický manuál – verze 1.0 (9.2.2022)

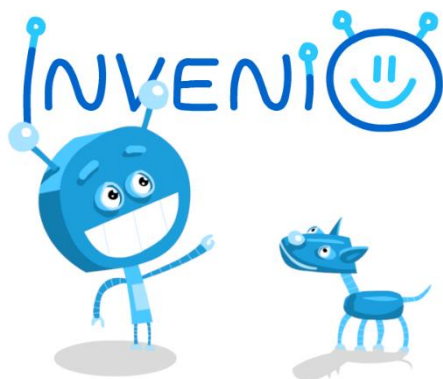
Aktuální verzi citujte jako: Ťápal A., Jabůrek, M., Palíšek, P., & Portešová, Š. (2022). *Test kvantitativního usuzování Triton a hladový oceán - technický manuál, verze 1.0*. Invenio. Dostupné z www.invenio.muni.cz.

Psychodiagnostický systém Invenio

Úvodní představení

Invenio je online psychodiagnostický systém určený ke screeningu profilu kognitivních schopností žáků 1.–6. tříd. Vychází z empiricky podložené psychologické teorie C-H-C (Cattell–Horn–Carroll, např. Schneider & McGrew, 2018), dle níž se inteligence skládá z řady dílčích schopností jako je např. *Fluidní inteligence*, *Krátkodobá pracovní paměť*, *Vizuální zpracování* a dalších. Každý test systému je přitom zaměřen na jednu z těchto schopností. Při testování je současně využíván moderní přístup tzv. game-based assessmentu, kdy jsou jednotlivé testy pojaty jako jednoduché počítačové hry, do kterých jsou integrovány herní prvky jako např. příběh, komiksová grafika, zvuky, sbírání bodů apod. Díky tomu obvykle dochází ke zvyšování motivace a hravosti a redukci případných obav u testovaných dětí.

Níže popsány test s názvem Triton a hladový oceán (zkráceně *Triton*) je jednou z řady diagnostických her, které jsou součástí systému Invenio. V aktuální podobě je určen žákům 3.–6. tříd základních škol.



Obrázek 1. Logo a hlavní postavy systému Invenio.

Autoři

Systém Invenio vyvinuli odborníci z Katedry psychologie Fakulty sociálních studií Masarykovy univerzity v Brně. Více informací o systému, jednotlivých členech týmu a dalších testech naleznete na www.invenio.muni.cz.

Odborný tým můžete kontaktovat na info@invenio.muni.cz.

Účel a možnosti použití

Systém Invenio je screeningovým nástrojem, který je zaměřen na jednotlivé úzké schopnosti v rámci inteligenční teorie C-H-C. Může proto být využit všude, kde je účelné získat orientační představu o jednotlivých kognitivních schopnostech žáka.

Získané výsledky mohou pomoci dětem, pracovníkům školy (pedagogům, školním psychologům či speciálním pedagogům) či rodičům při individualizaci vzdělávacího přístupu, k zamyšlení nad dalším rozvojem měřených schopností, při zvažování dalšího profesního směřování (volba školy či vzdělávacího kurzu) apod. Hlavním účelem systému je přitom vyhledávání nadaných žáků základních škol. Systém je vyvinut pro plošné testování ve školách a je proto koncipován tak, aby se ho mohli účastnit i žáci, u nichž lze očekávat spíše průměrný nebo podprůměrný výkon. Testy proto obsahují položky různé obtížnosti, jejich normy jsou vytvořeny pro běžnou populaci žáků, a celý systém využívá herní principy, které obvykle snižují testovou úzkost.

Samotné zadávání testu je velmi snadné, protože žáci postupují dle instrukcí, které jsou součástí hry. Úkolem zadávajícího je především zajistit standardní podmínky. Testy tak může zadávat jakákoliv zaškolená osoba (např. pracovník školy). V aktuální podobě není možné testy zadávat v domácím prostředí žáka, aby nebyla ohrožena standardnost a validita testování.

Testování nemohou být žáci se závažnými poruchami zraku, sluchu a jemné motoriky, děti se závažným mentálním znevýhodněním či žáci, kteří nerozumí češtině, protože by mohla být ohrožena validita získaných výsledků. Testy je možné administrovat i žákům méně zběhlým v užívání počítače. Je pouze nutné se ujistit, že jsou děti (především z prvního ročníku) schopny samostatně ovládat počítačovou myš.

Obecná specifikace testů

Jednotlivé testy jsou administrovány prostřednictvím počítače. Jsou spouštěny online v rámci webového prohlížeče. Je možné je zadávat skupinově, nicméně každý žák pracuje individuálně na svém počítači. Hry ovládá pomocí myši a klávesnice, zvukové podněty přijímá prostřednictvím sluchátek.

Celý diagnostický systém je z pohledu hráče situován do vesmírného prostředí. Obecným úkolem je plnit různorodé úkoly na jednotlivých vesmírných planetách. Každá planeta přitom představuje jeden psychodiagnostický test zaměřený na jednu z úzkých schopností teorie C-H-C. V základním hráčském menu jsou zobrazeny všechny planety (testy), které má hráč v rámci daného testového sezení prozkoumat. Již ukončené testy jsou graficky odlišeny od těch, které bude hráč teprve řešit. Tím je dosahováno větší názornosti jak pro hráče, tak pro administrátora.

V rámci herního příběhu vystupují dvě postavy – robot P.U.K (První Univerzální Kosmonaut) a jeho psi robotický kamarád Vincent van Dog. Ti provádí hráče testem, předávají mu instrukce k řešení jednotlivých testových položek a upozorňují jej na případné chyby. Větší srozumitelnosti je dosahováno jednoduchými animacemi (ukazování na některý prvek, přesun prvku po obrazovce apod.).

Kromě řešení jednotlivých úloh může hráč využít dvě tlačítka (viz Obrázek 2) – pauzu, která slouží k dočasnému zastavení hry, a vypnutí zvuků, kterým se dají eliminovat doprovodné herní zvuky jako je hudba, klikání apod. Zvuky, které jsou zásadní pro ovládání hry a řešení úloh, jsou slyšet stále a není možné je vypnout.



Obrázek 2. Ovládací prvky totožné pro všechny testy systému Invenio – pauza a vypnutí doprovodných zvuků.

Podoba výstupních zpráv

Po testování systém Invenio vygeneruje dva dokumenty s výsledky testovaného dítěte – *Zprávu pro rodiče* a *Odbornou zprávu*.

Cílem *Zprávy pro rodiče* je poskytnout zákonným zástupcům dítěte základní informace o měřených schopnostech, výsledcích, jejich interpretaci a na ně navázaných doporučeních. Výsledky jsou uváděny ve formě zařazení do jedné ze čtyř verbálních kategorií, které jsou odvozené od bodové hodnoty standardního skóre (s průměrem 100 a směrodatnou odchylkou 15). Tyto verbální kategorie slouží uživatelům pro snadnější interpretaci výsledků. Jednotlivé verbální kategorie s příslušným rozmezím standardních skóre jsou k dispozici v následující Tabulce 1.

Tabulka 1. Verbální kategorie a jejich vztah ke standardnímu skóru.

Verbální kategorie	Rozmezí standardního skóru
Snížená schopnost	85 a méně*
Průměrná schopnost	85-114
Rozvinutá schopnost	115-129
Mimořádně rozvinutá schopnost	130 a více

Pozn.: *Kategorie snížená schopnost je uvedena ve chvíli, kdy bodový odhad standardního skóre činí 85 a méně bodů a zároveň horní mez intervalu spolehlivosti činí 90 a méně bodů.

Dokument označený jako *Odborná zpráva* je odborným výstupem z psychologické diagnostiky, který je uživatelům poskytován pro účely případných navazujících vyšetření a odborných konzultací (např. v pedagogicko-psychologické poradně). Obsahuje technické informace o použitých diagnostických hrách, měřených schopnostech a výsledky dítěte ve formě standardních skóre. Ty jsou uváděny v podobě 95% intervalu spolehlivosti, který je odvozen od chyby měření testu.

Triton a hladový oceán



Obrázek 3. Medailonek testu Triton a hladový oceán.

Východiska a zaměření

Teoretická východiska

Jak již bylo uvedeno výše, systém Invenio a jeho jednotlivé testy vychází primárně z teorie inteligence C-H-C.

C-H-C teorie inteligence

C-H-C (Cattell–Horn–Carroll) teorie (např. Schneider & McGrew, 2018) je integrací dvou předních psychometrických přístupů k inteligenci – Cattell-Hornovy Gf-Gc teorie (např. Horn & Noll, 1997) a Carrollovy teorie tří vrstev (v angl. The Three-Stratum Theory of Cognitive Abilities; Carroll, 1993). Představuje ucelenou taxonomii kognitivních schopností, které rozděluje do tří hierarchických vrstev na

základě jejich obecnosti. První vrstvu tvoří velké množství vysoce diferencovaných schopností (tzv. úzké schopnosti neboli narrow abilities) jako je např. *Slovní plynulost*, *Asociační paměť*, *Rychlost usuzování* apod. Tyto dílčí schopnosti jsou ve druhé vrstvě sdružovány do schopností obecnějších, jako je *Krátkodobá pracovní paměť* či *Fluidní inteligence* (tzv. široké schopnosti neboli broad abilities). Třetí, nejvyšší vrstvu, tvoří jediná obecná schopnost obvykle označována jako „g“ faktor, případně jako *Obecná inteligence*.

Označení jednotlivých schopností v rámci teorie C-H-C se může napříč odbornými zdroji a verzemi teorie lišit (např. schopnost *Gc* bývá označována jako *Krystalická inteligence*, *Porozumění-znalosti* apod.). Současně jsou názvy schopností označovány zkratkami, které jsou naopak konzistentní a mohou proto sloužit jako jednoznačný identifikátor. Z těchto důvodů uvádíme v dalším textu vždy také zkratky úzkých i širokých schopností.

V rámci C-H-C teorie je test *Triton* zaměřen na širokou schopnost s názvem *Fluidní inteligence* (Fluid reasoning; *Gf*) a v rámci ní na úzkou schopnost s názvem *Kvantitativní usuzování* (Quantitative reasoning; *RQ*).

Široká schopnost – Fluidní inteligence (Gf)

Fluidní inteligence představuje jádro toho, co bylo a stále je tradičně pokládáno za inteligenci. Ačkoliv je *Gf* v dnešní době v rámci teorie C-H-C považována pouze za jednu z řady širokých schopností, je možné ji označit jako základní kámen této teorie. Výsadní postavení této schopnosti se projevuje také na úrovni praktické diagnostiky. Podle Hunta (2011) patří k nejčastěji hodnoceným schopnostem v testech inteligence.

Fluidní inteligence označuje způsob myšlení, který využíváme při setkání s novými úkoly, jejichž řešení nemáme zautomatizované a u kterých není možné se spolehnout na dříve naučené postupy či předchozí znalosti. Tento typ myšlení zahrnuje celou řadu rozličných procesů jako je chápání podstaty konceptů a podobností mezi nimi (např. *co mají společného pes, kočka a kráva*), rozpoznávání vztahů mezi podněty (např. při analogiích typu - *matka k babičce se má stejně jako otec k dědečkovi*), vyvozování závěrů, reorganizaci, transformaci a propojování informací (např. vyvození pravidla, podle kterého je uspořádána číselná či obrázková řada a doplnění chybějícího prvku), tvorbu hypotéz a jejich potvrzování či vyvracení apod.

Obecně lze tuto schopnost chápat jako logické usuzování zaměřené na řešení nových problémů, ve kterém se primárně uplatňují dva procesy - indukce a dedukce. Pokud se tyto procesy odehrávají na pozadí matematických vztahů a čísel, mluvíme o tzv. *Kvantitativní usuzování (RQ)*.

Úzká schopnost – Kvantitativní usuzování (RG)

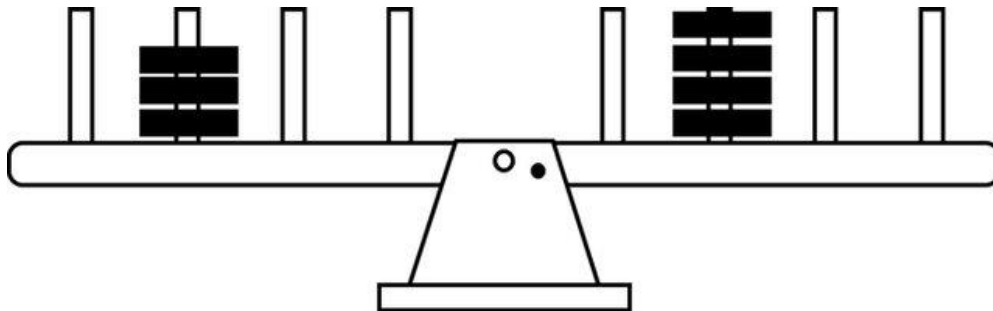
Kvantitativní usuzování je možné vymezit jako “schopnost provádět logické operace s matematickými veličinami, vztahy a operátory” (Schneider & McGrew, 2018, str. 94). Z tohoto popisu je zřejmé, že tuto schopnost definujeme s ohledem na obsah, tedy povahu prvků, kterých se usuzování týká. V tomto ohledu je *RQ* odlišné od dalších dvou hlavních úzkých schopností *Fluidní inteligence* – indukce a dedukce, které jsou vymezeny procesuálně. Při řešení úloh, u kterých je uplatňována schopnost *RQ*, je tedy nutné také zapojení alespoň jednoho z těchto procesuálních faktorů. V případě hry *Triton* se jedná jak o indukci, tak dedukci. S ohledem na povahu testu však považujeme za primární úzkou schopnost, která se při řešení úloh uplatňuje, *Kvantitativní usuzování*.

Důležité je přitom zdůraznit, že při řešení úloh v testu *Triton* (a obecně při řešení testů zaměřených na *RQ*) hráč využívá jednoduché matematické operace jako je sčítání, odečítání a násobení. Tyto úlohy však nevyžadují hlubší matematické znalosti a složité aritmetické operace. Jde spíše o aplikaci jednoduchých matematických principů, které stojí v pozadí logického úsudku. Náročnost úloh je primárně zvyšována prostřednictvím nároků na komplexitu usuzování.

Existující testové principy a podobně zaměřené testy

Úkol s kladinou (*Balance beam task*)

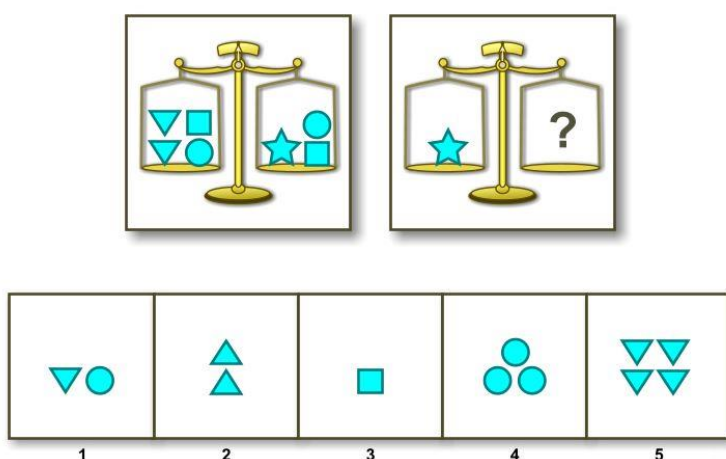
Hra *Triton* je inspirována principem *Balance beam task* od Inhelderové a Piageta (1958), který byl využíván k zachycení vývojových stádií usuzování u dětí. Podstatou úkolu, který byl dále rozpracován a využíván v celé řadě dalších (i aktuálních) výzkumů (např. Hofman et al., 2015; Schapiro & McClelland, 2009; Siegler, 1976), byla dvouramenná kladina se 4 kolíčky na každé straně v pravidelné vzdálenosti od středu (viz Obrázek 4). Na každý kolíček bylo možné umístit několik kovových závaží. Úkolem dítěte bylo s ohledem na počet závaží a jeho vzdálenosti od středu rozhodnout, zda jsou ramena kladiny v rovnováze či některé z nich klesne dolů. V případě nerovnováhy mělo dítě určit, které z ramen je zatíženo více.



Obrázek 4. Ukázka principu “Balance beam task” (Jansen et al., 2007).

Testový princip *Figure Weights*

Na podobném principu staví Williams a Raifordová, kteří v roce 2005 položili základ testu *Figure Weights* (Lichtenberger & Kaufman, 2009), který je aktuálně využíván v inteligenčních bateriích *Wechsler Adult Intelligence Scale 4 (WAIS-IV; Wechsler, 2008a)* a *Wechsler Intelligence Scale for Children 5 (WISC-V; Wechsler, 2014)*. V této podobě úkol již nepracuje se vzdáleností závaží od středu kladiny a přidává nový prvek – na vahách se objevují různé typy závaží, které se liší hmotností. Váha závaží je přitom vyjádřena pomocí tzv. podmínek. Ty jsou znázorněny prostřednictvím vah, jejichž ramena jsou již vyrovnaná, a slouží k zachycení relací mezi různým typem závaží. Úkolem testovaného je vyjít z těchto podmínek (v každé položce je buď 1, nebo 2 podmínky) a vybrat z 5 možností ten soubor závaží, který vyrovná prázdné rameno zbývající váhy (viz Obrázek 5).



Obrázek 5. Ukázka položky testu *Figure Weights*¹.

¹ Jedná se o volně dostupnou položku zveřejněnou na stránce <https://www.testingmom.com/tests/wisc-test/sample-wisc-practice-questions/>, nejde tedy o konkrétní úlohu z baterie *WISC V* či *WAIS IV*.

Specifika testu Triton a hladový oceán

Test *Triton* navazuje na obě výše popsaná pojetí, která dále obohacuje o nové herní i testové prvky (viz sekce *Struktura a jednotlivé prvky*). Velkým rozdílem oproti testu *Figure Weights* je absence striktního časového limitu na položku. V testu *Triton* je řešení položky časově neomezené a to hned z několika důvodů.

- 1) Časový limit v testu *Figure Weights* je kritizován odborníky z praxe s ohledem na to, že upozorňování na blížící se konec limitu ruší probandy při řešení úloh (Lichtenberger & Kaufman, 2009).
- 2) Z hlediska obsahové validity je účelem testu *Triton* měřit *Kvantitativní usuzování*. V rámci C-H-C teorie je přitom *Rychlost usuzování (RE)* samostatnou úzkou schopností. Naším cílem je měřit jednotlivé schopnosti tak, aby byl vliv ostatních schopností na výsledek v testu co nejvíce omezen.
- 3) Primárním účelem systému Invenio je vyhledávání nadaných žáků. Souhlasíme přitom s názorem odborníků (Gridley et al., 2003), že je vhodnější vyhnout se časovým limitům a bonusům za rychlé řešení v testech, u kterých je to možné a rychlost reakce není součástí měřeného konstruktů, a to např. proto, že někteří perfekcionistační nadaní žáci preferují správnost před rychlostí. Současně se ukazuje, že se kognitivní nadání nemusí vždy projevovat v rychlosti zpracování informací (Reams et al., 1990).
- 4) Kromě toho předpokládáme, že časový limit může zvyšovat úzkost z testování. Ukazuje se přitom, že časový limit může v testových situacích znevýhodňovat úzkostné jedince (Onwuegbuzie & Seaman, 2010).

Popis a struktura

V následující Tabulce 1 je zachycen proces vývoje testu a navazujících studií. V této sekci je představena finální podoba metody. Postupný vývoj testu je popsán níže.

Tabulka 1. Časová osa vývoje testu a navazujících studií.

Vývojová fáze	Časové období
Pilotní verze 1 – papírová verze s provizorní grafikou	Leden, 2016 – Červen, 2016
Pilotní verze 2 – počítačová verze s finální grafikou	Červen, 2016 – Duben, 2017
Pilotní verze 3 – doplnění instrukcí přímo do testu	Duben, 2017 – Květen, 2017
Pilotní verze 4 – tvorba a ověřování položkové banky	Květen, 2017 – Listopad, 2017
Pilotní verze 5 – úprava instrukcí a tvorba finálního setu	Listopad, 2017 – Leden, 2018
Pilotní verze 6 a studie Buráňové (2018) – sběr dat	Březen, 2018 – Duben, 2018
Standardizace verze 7 – sběr dat	Červen, 2018 – Prosinec, 2018
Studie Kubrického (2019) – sběr dat	Říjen, 2018 – Listopad, 2018
Studie Portešové, Jabůrka, Ťápala a Straky (2020) – dosběr dat	Prosinec, 2018 – Leden, 2019
Studie na ZŠ Trávník, Přerov a ZŠ Novoměstská, Brno – sběr dat	Duben, 2019 – Červen, 2019
Studie Portešové, Jabůrka, Rečky a Straky (2020) – sběr dat	Květen, 2019 – Červen, 2019
Studie Michalka (2020) – sběr dat	Listopad, 2019 – Prosinec, 2019

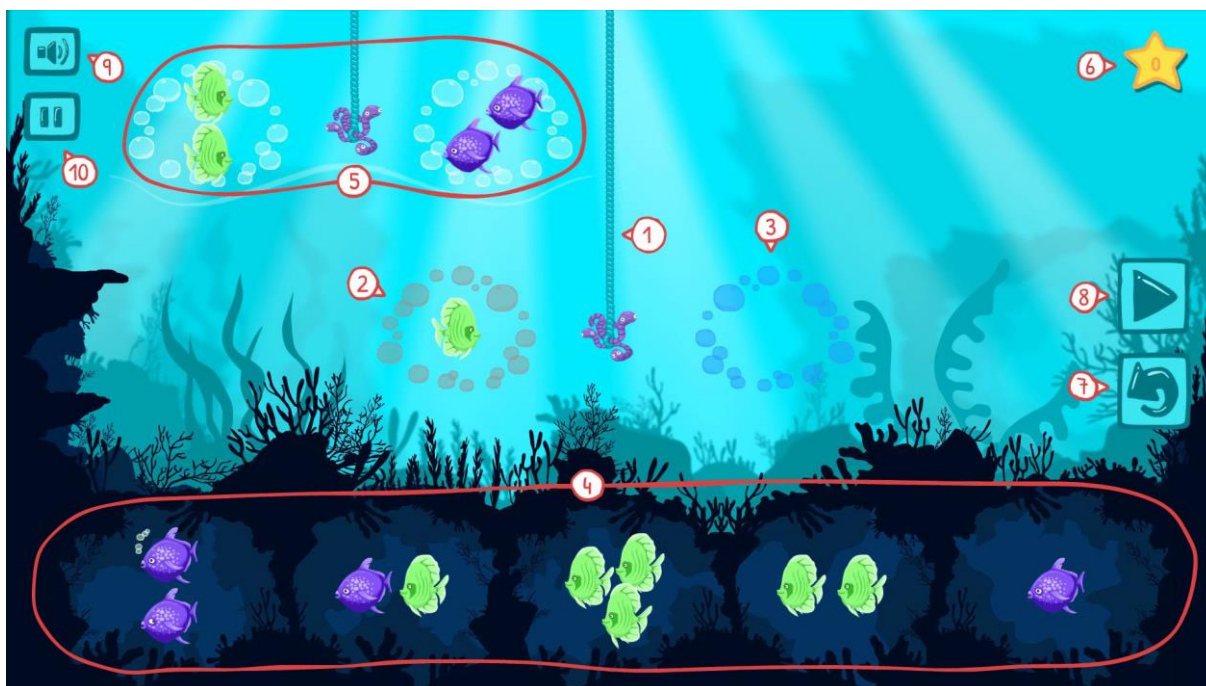
Herní příběh

Planeta, kam hráč přilétá, je celá pokrytá vodou. Potkává zde Tritona, mořského biologa, který zkoumá místní faunu. Během svých výzkumů zjistil, že v moři začíná ubývat živočichům potrava a snaží se tuto záhadu rozluštit. Aby měl Triton dost času na svůj výzkum, musí mu hráč pomoci hladové mořské živočichy krmit pomocí červíků.

Struktura a jednotlivé prvky

V této sekci je popsána aktuálně finální 8. verze testu *Triton*. Tato verze je složena z 29 položek, z nichž 26 je vyhodnocovaných pro výpočet standardního skóru. První 3 velmi snadné položky slouží jako zácvik.

Ukázka položky se znázorněním jednotlivých prvků testu je uvedena na Obrázku 6.



Obrázek 6. Příklad položky testu *Triton* se znázorněním jednotlivých prvků.

Základní prvky testu jsou označeny čísly 1-5. V hlavní části obrazovky se nachází dlouhý háček s červíky (1). Na obou stranách háčku je kruh tvořený z bublin, tzv. slot. Na levé straně je v kruhu z hnědých bublin zobrazen určitý počet mořských živočichů (2). Na druhé straně je prázdný slot vytvořený z tmavě modrých bublin (3). Úkolem hráče je zaplnit tento prázdný slot přesunutím jedné z pěti skupin živočichů z jeskyní, které se nachází v dolní části obrazovky (4), a to tak, aby vyrovnal sílu živočichů na levé straně hlavního háčku (2). Základním pravidlem je to, že živočichové stejné barvy a tvaru mají stejnou sílu a síla živočichů na jedné straně háčku se sčítá. V obtížnějších položkách je síla jednotlivých živočichů vyjádřena tzv. podmínkami, které jsou zobrazeny ve formě menších háčků se sloty tvořenými světle modrými bublinami (5). V těchto podmínkách jsou obě strany háčku již vyrovnané, čímž je vyjádřen vztah mezi jednotlivými mořskými živočichy. Nejobtížnější položky obsahují až dvě takové podmínky.

V pravé horní části obrazovky je hvězdička signalizující počet získaných bodů (6).

Kromě přesouvání skupin živočichů z jeskyní do prázdných slotů a zpět může hráč test ovládat pomocí tlačítka reset (7), kterým navrátí všechny živočichy do původní pozice. Přejít k další položce umožňuje tlačítko play (8), které se objeví až po zaplnění prázdného slotu (3). Obrázek 6 je pro tyto účely upravený, aby obsahoval zároveň prázdný slot (3) a tlačítko play (8).

Vlevo nahoře jsou dále umístěny ovládací prvky společné pro všechny hry systému Invenio, které však nejsou zásadní pro řešení úloh – možnost ztlumit nedůležité zvuky (9) a pauza k dočasnému zastavení hry (10).

Položky jsou v rámci testu *Triton* seřazeny tak, aby se zvyšovala jejich obtížnost, tedy náročnost logických operací nutných k vyřešení položky. Současně se od sebe položky liší svou komplexitou. Komplexita položek je přitom vyjádřena množstvím použitých prvků (živočichů a slotů) a testových principů. Každý nový testový princip je vysvětlen pomocí instrukcí, které jsou kombinací mluveného slova a jednoduchých animací. Součástí představování nového principu je také ukázka řešení konkrétní položky, ve které je tento princip použit. V Tabulce 2 jsou uvedeny jednotlivé charakteristiky položek a testových principů.

Tabulka 2. Charakteristika položek a principů testu *Triton*.

Číslo položky	Charakteristika položek a použité testové principy
1–3	Zácvičné položky, jejichž podstata spočívá v pouhém kopírování levé strany háčku. Je vyžadováno pochopení základního pravidla testu – živočichové se stejnou barvou, tvarem a v totožném množství jsou stejně silní.
4–5	Nový princip, který vychází z pojetí matematických množin – větší počtu slotů na každé straně háčku. Každá strana může být členěna až na 4 podmnožiny (sloty) na každé straně. Je vyžadováno sčítání síly oddělených skupin živočichů na každé straně. Současně hráč nevybírá jen jednu odpověď z pěti možností, ale zaplňuje všechny prázdné sloty na pravé straně háčku. Tlačítko play se objevuje až po zaplnění všech prázdných slotů. Některé sloty na pravé straně háčku mohou být zaplněné již při započetí úlohy.
6–7	Nový princip – lektvar s mrtvou vodou. Přetažením na kterýkoli slot na levé straně hlavního háčku dojde k odstranění všech živočichů v dané skupině. Tlačítkem reset může hráč navrátit položku do původního stavu.
8–9	Nový princip – lektvar s živou vodou. Přetažením na kterýkoli slot na pravé straně hlavního háčku dojde ke zdvojnásobení všech živočichů v dané skupině. Tlačítkem reset může hráč navrátit položku do původního stavu.
10–18	Nový princip – 1. podmínka. Nový háček v levé horní části obrazovky s již vyrovnanými stranami, který určuje relaci mezi jednotlivými živočichy. V úlohách postupně roste obtížnost a střídá se povaha operací, které jsou k řešení nezbytné. Jedná se o jednoduché přičítání, odečítání, násobení a dělení a nahrazování prvků za jiné dle prezentované podmínky. V úlohách jsou v některých případech využívány již představené testové principy (sloty a oba typy lektvarů), případně jejich kombinace. V žádné položce však nejsou přítomny oba typy lektvarů najednou.
19–29	Nový princip – 2. podmínka. Nový háček v pravé horní části obrazovky s již vyrovnanými stranami. Je vyžadováno využívat informace z obou podmínek najednou a aplikovat složitější matematické a logické operace. V položkách jsou v některých případech využívány již představené testové principy.

Každá položka má právě jedno správné zaplnění slotů a na řešení není stanoven časový limit. V úvodu každé položky, která obsahuje jeden z lektvarů, je hráč upozorněn na přítomnost tohoto testového prvku psem Vincentem. Cílem je minimalizovat možnost chybné odpovědi vlivem přehlédnutí a nepozornosti. V případě, že hráč provede nějaký nepovolený úkon (např. snaží se použít lektvar na nevhodný slot, přesunout skupinu živočichů do již obsazeného slotu apod.), postava Vincenta mu s pomocí obrázkového piktogramu vysvětlí, jak je třeba postupovat (např. „Lektvar s živou vodou můžeš použít jen na rybky v tmavě modrých bublinách.“).

Po stisknutí tlačítka play dostává hráč zpětnou vazbu, zda byla jeho odpověď správná či chybná. Jedním z herních prvků, který se objevuje ve všech testech systému *Invenio*, je sbírání bodů. V případě metody *Triton* získává hráč za každou správnou odpověď 3 body, v případě chyby 1 bod. Cílem tohoto prvku je zvýšit motivaci a zájem o test. Nepředpokládáme přitom, že má tento herní prvek vliv na samotné testování.

Pilotní verze

Verze 1

V první polovině roku 2016 vznikala základní koncepce hry na základě studia relevantní odborné literatury a výše zmíněných testových principů – konkrétně *Balance beam task* (Inhelder & Piaget, 1958) a *Figure Weights* (např. Wechsler, 2014). V červnu 2016 bylo 20 žákům prvních a druhých tříd ZŠ a

MŠ Brno, Vedlejší v papírové podobě administrována prvotní verze testu v několika mírně odlišných variantách s cílem ověřit, zda rozumí úlohám v novém prostředí (podmořský svět místo jednoduchých vah), odlišně vyjádřeným vztahům mezi prvky (síla ryb místo hmotnosti závaží) a novým principům (lektvar života a smrti a princip slotů). V této fázi vývoje byl součástí testu také princip nerovnováhy, který spočíval v tom, že živočichové na některé straně lana byli silnější. Současně bylo dětem prezentováno několik grafických variant prostředí s cílem zjistit jejich estetickou preferenci. Na základě zpětné vazby dětí byla zvolena grafická podoba testu a jeho finální pojetí, které je totožné se standardizační verzí. Vzhledem k tomu, že ani u nejmladších dětí se neobjevil problém v pochopení instrukcí, jednotlivých testových principů či řešení úloh v rámci inovativního pojetí s mořskými živočichy, byl zahájen vývoj testu v počítačové podobě.

Verze 2

Na konci dubna 2017 již byla první počítačová verze testu ověřována na 12 žácích 1.-5. ročníků ZŠ a MŠ Brno, Vedlejší při individuální administraci s primárním cílem otestovat základní ovládnutí hry a instrukce, které byly v této podobě administrovány ústně přítomným odborníkem. Na základě tohoto testování došlo k četným úpravám instrukcí a samotné metody.

Verze 3

Následně byla vytvořena instruktážní videa s provizorním dabingem, která byla zakomponována přímo do testu. Nová verze metody s těmito videi byla administrována 5 žákům první třídy ZŠ a MŠ Brno, Vedlejší na konci května 2017. S ohledem na to, že žádný z žáků neměl problém s chápáním počítačově administrovaných instrukcí, ani s ovládnutím metody, byl test v červnu 2017 administrován skupinově 30 žákům (15 čtvrtákům a 15 pátákům) ZŠ Horní Beřkovic. Po analýze dat z této administrace došlo k několika drobným úpravám instrukcí a dílčích prvků testu.

Verze 4

V následujícím období bylo vytvořeno více než 150 položek testu. Při jejich vývoji jsme vycházeli z několika východisek:

- 1) Identifikovat dílčí kognitivně-matematické operace, které jsou využívány při řešení úloh, a designovat položky tak, aby žádná z těchto operací nebyla opomenuta.
- 2) Pokrýt všechny obtížnostní úrovně od velmi jednoduchých po extrémně obtížné položky.
- 3) Zajistit pokrytí škály položek v závislosti na jejich vizuální komplexitě – od položek s několika málo prvky po vizuálně náročné položky s řadou prvků a testových principů.
- 4) Zajistit, aby každá položka měla právě jedno správné řešení.
- 5) Identifikovat možné typy distraktorů (možností k výběru, které nejsou správnou odpovědí) a zajistit přítomnost všech různých typů napříč položkami.
- 6) Volit barvu a tvar jednotlivých mořských živočichů, aby položky nebyly vizuálně nepříjemné a nepřehledné.

Ze vzniklé položkové banky bylo vybráno nejlepších 112 položek, které byly v říjnu a listopadu 2017 pilotovány v rámci 3 setů. Set A obsahoval 36 položek, zbývající dva sety byly tvořeny v obou případech 38 položkami. Instruktážní videa byla ve všech třech případech totožná. Setem A bylo otestováno 7 tříd ($n = 170$ dětí), setem B 5 tříd ($n = 130$ dětí) a setem C také 5 tříd ($n = 115$ dětí). V každém setu byli přibližně rovnoměrně zastoupeni žáci 1.-5. ročníků ZŠ. Testování probíhalo celkově na 5 školách v Praze (ZŠ Květnového vítězství) a Jihomoravském kraji (ZŠ a MŠ Náměstí Svornosti, Brno; ZŠ Tišnov, Smíškova; ZŠ a MŠ Doubravník; ZŠ a MŠ Ostrovačice).

Po sérii statistických analýz, která zahrnovala položkovou analýzu, screening chybějících dat, analýzu časů řešení jednotlivých položek, Raschovy modely, exploratorní IRT a analýzu reliability, došlo k vyřazení principu nerovnováhy, redukci instruktážních videí z 10 na 6 a jejich zestručnění a vytvoření nové verze testu.

Verze 5

Tímto způsobem vznikl hlavní set s 30 položkami a doplňkový set s 20 položkami s nejlepšími parametry. Oba sety byly v lednu 2018 pilotovány na 62 žácích 3.-5. třídy ZŠ a MŠ Brno, Husova. Na základě získaných dat bylo ověřeno, že hlavní set funguje bezproblémově. Jedinou změnou bylo rozšíření hlavního setu o 2 obtížné položky z doplňkového setu, aby lépe diskriminoval v nadprůměru.

Verze 6

Tento hlavní set s 32 položkami byl dále pilotován, částečně v rámci studie Buráňové (2018), která ověřovala souvislost testu *Triton* s metodami CFT 20-R a TIM³⁻⁵ (viz sekci *Souběžná validita*) a vliv typu zpětné vazby na testový výkon. Test byl v této fázi vývoje ověřován v 5 různých variantách, které se lišily podobou zpětné vazby na odpověď testovaného žáka (detailněji viz sekci *Doplňkový výzkum*). Části žákům byl kromě hlavního setu administrován také doplňkový set verze 5 se zbývajících 18 položkami.

Sběr dat probíhal v březnu a dubnu 2018 celkem na 4 školách – ZŠ Polešovice; ZŠ Bílovice nad Svitavou; ZŠ Úvoz, Brno; ZŠ a MŠ Pastviny, Brno. Hlavním setem bylo otestováno celkem 363 žáků 1.-5. tříd ZŠ, z toho 134 žáků řešilo také doplňkový set.

Na základě analýzy získaných dat byly 2 položky z hlavního setu vyřazeny a nahrazeny 2 položkami z vedlejšího setu. Položky byly dále seřazeny dle obtížnosti. Došlo k estetickému vylepšení instruktážních videí (jak po vizuální, tak zvukové stránce) a implementaci nové funkce – nově je hráč při chybných úkonech (např. při pokusu použít lektvar na nevhodný slot) upozorněn, že tímto způsobem daný prvek hry používat nelze. Tímto způsobem byla dokončena verze 7, která byla použita během standardizace..

Standardizace

Tato sekce popisuje standardizační studii testu *Triton* ve verzi 7. Analýzy byly prováděny v paradigmatech klasické testové teorie (CTT) a teorie odpovědi na položku (IRT). Reliabilita testu, normy a validita jsou diskutovány v následujících samostatných sekcích.

Sběr dat a vzorek

Standardizace 7. verze testu *Triton* byla provedena na vzorku 1087 žáků (1.-6. ročníku) z osmnácti českých základních škol. Popis vzorku je k dispozici v Tabulce 3; 63 žáků nevedlo ročník.

Tabulka 3. Popisné statistiky standardizačního vzorku.

Ročník	N	Typický věk	% dívek
1	158	6	46
2	129	7	50
3	137	8	38
4	237	9	50
5	247	10	45
6	116	11	36

Sběr dat probíhal mezi červnem a prosincem 2018. Účast byla dobrovolná a vyžadovala informovaný souhlas podepsaný zákonným zástupcem. Test byl administrován v počítačových učebnách vyškolenými učiteli nebo školními psychology.

Školy podílející se na standardizaci jsou uvedeny v Tabulce 4. Vybrány byly příležitostně na základě ochoty podílet se na studii. Potenciálně participující instituce však byly voleny tak, aby pokrývaly velikosti sídel typické pro Českou republiku. Rozdělení participantů dle velikosti sídla je k dispozici v Tabulce 5.

Tabulka 4. Školy podílející se na standardizaci.

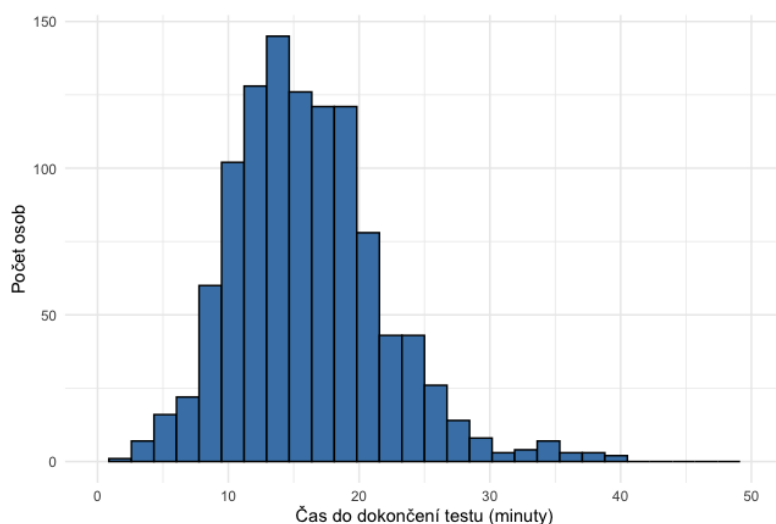
Název školy	Kraj
MŠ a ZŠ Dolní Bukovsko	Jihočeský
ZŠ a MŠ Jihomoravské náměstí, Brno	Jihomoravský
ZŠ Nad Rybníkem, Jedovnice	Jihomoravský
ZŠ T.G.M. Rajhrad	Jihomoravský
ZŠ Masarova, Brno	Jihomoravský
ZŠ Vančurova, Hodonín	Jihomoravský
ZŠ Očovská, Hodonín	Jihomoravský
ZŠ Školní, Horní Slavkov	Karlovarský
ZŠ a MŠ Tyršova, Frenštát pod Radhoštěm	Moravskoslezský
Fakultní ZŠ dr. Milady Horákové a MŠ, Rožňavská, Olomouc	Olomouc
ZŠ Zámecká, Litomyšl	Pardubický
ZŠ nám. Curieových, Praha 1	Praha
ZŠ sv. Vršily, Ostrovní, Praha 1	Praha
ZŠ a MŠ Norská, Kladno	Středočeský
ZŠ Želiv	Vysočina
ZŠ Otokara Březiny, Jihlava	Vysočina
ZŠ a MŠ Herálec	Vysočina
ZŠ Hálkova, Humpolec	Vysočina

Tabulka 5. Rozdělení participantů dle velikosti sídla, kde se nachází jejich školy.

Velikost sídla	Méně než 1 999	2 000 až 9 999	10 000 až 49 999	50 000 a více	Praha
Participantů	162	130	351	317	73

Odpovědi 6 participantů byly odstraněny pro podezření z nepochopení testových instrukcí. Jinak byla využita data všech účastníků, a to i 106 žáků, kteří test nedokončili.

Mediánový čas řešení celého testu činil 15,5 minut. Rozložení je prezentováno níže na Obrázku 7.

**Obrázek 7.** Rozložení času řešení celého testu.

Položková analýza

Položková analýza popisuje statistické vlastnosti testových položek. Tyto parametry (podíly správných odpovědí, medián délky řešení v sekundách a korigovaná korelace s celkovým skórem) jsou prezentovány níže v Tabulce 6.

Tabulka 6. Položková analýza.

Pořadí položky	popularita							čas řešení (Me)	r_cor
	celkem	1. ročník (n = 156)	2. ročník (n = 129)	3. ročník (n = 137)	4. ročník (n = 196)	5. ročník (n = 221)	6. ročník (n = 116)		
1*	0,97	0,96	0,98	0,98	0,94	0,98	0,99	9,25	0,48
2*	0,97	0,97	0,95	0,96	0,97	0,98	0,98	7	0,35
3*	0,93	0,95	0,88	0,88	0,91	0,99	0,86	9	0,61
4	0,67	0,32	0,55	0,69	0,71	0,81	0,74	16	0,84
5	0,71	0,35	0,5	0,77	0,77	0,87	0,78	15,5	0,88
6	0,82	0,45	0,72	0,86	0,89	0,93	0,91	26	0,81
7	0,84	0,62	0,74	0,88	0,89	0,92	0,86	18	0,71
8	0,64	0,27	0,48	0,64	0,71	0,78	0,78	36	0,89
9	0,75	0,4	0,59	0,76	0,81	0,9	0,85	17	0,9
10	0,81	0,69	0,78	0,77	0,85	0,86	0,86	11,5	0,56
11	0,81	0,67	0,7	0,84	0,84	0,91	0,81	20	0,7
12+	0,54	0,3	0,35	0,53	0,59	0,64	0,68	11,25	0,54
13	0,66	0,31	0,56	0,7	0,69	0,82	0,71	30	0,75
14	0,64	0,46	0,56	0,67	0,67	0,67	0,79	17,5	0,66
15	0,72	0,41	0,57	0,68	0,8	0,83	0,86	9	0,71
16	0,23	0,05	0,06	0,21	0,23	0,3	0,52	12	0,69
17	0,47	0,17	0,27	0,5	0,54	0,6	0,67	44	0,73
18	0,45	0,26	0,22	0,48	0,46	0,54	0,67	20,5	0,57
19	0,41	0,08	0,26	0,45	0,46	0,48	0,7	41,5	0,79
20+	0,46	0,51	0,38	0,42	0,41	0,47	0,55	25,5	0,38
21	0,38	0,15	0,21	0,41	0,38	0,46	0,63	19,25	0,68
22	0,43	0,2	0,26	0,47	0,47	0,5	0,62	33,5	0,69
23	0,34	0,14	0,22	0,33	0,34	0,41	0,52	31	0,58
24	0,31	0,18	0,19	0,31	0,33	0,33	0,5	21	0,57
25+	0,19	0,15	0,17	0,29	0,25	0,1	0,24	39,5	0,33
26	0,35	0,13	0,2	0,37	0,39	0,44	0,51	18,5	0,7
27	0,17	0,1	0,1	0,17	0,19	0,18	0,26	30,5	0,42
28	0,15	0,05	0,05	0,15	0,18	0,16	0,29	40	0,6
29	0,13	0,03	0,07	0,18	0,15	0,12	0,2	23	0,47
30	0,26	0,14	0,23	0,29	0,27	0,28	0,35	19,5	0,37
31	0,07	0,04	0,08	0,07	0,07	0,06	0,1	29,5	0,3
32	0,04	0,06	0,04	0,04	0,03	0,04	0,03	32,75	0,03

* = položka není skórována

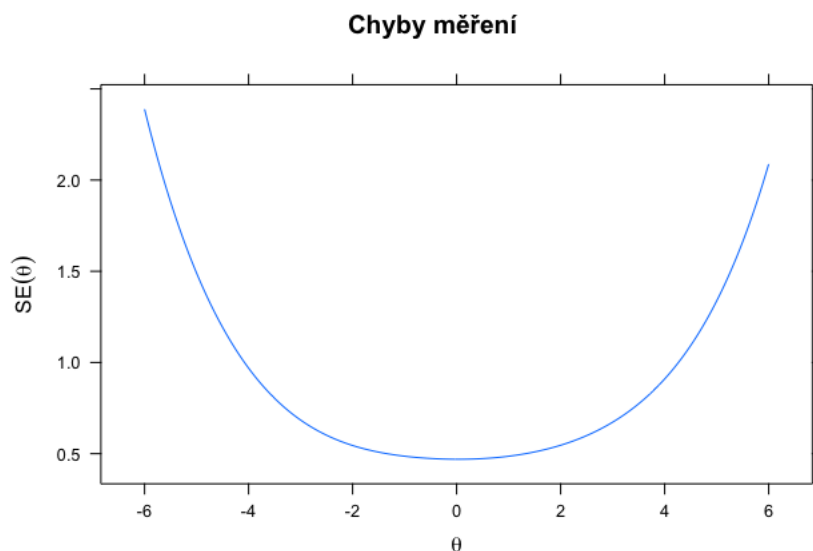
+ = položka byla po standardizaci vyřazena

IRT modely

Standardizační data byla nejprve analyzována pomocí Raschova (Bond, Yan, & Heene, 2020) modelu pro všechny účastníky společně. Do modelu vstupovaly všechny položky až na první tři (1 – 3), které plní pouze zácvičnou funkci a nejsou skórovány. Model byl odhadnut pomocí balíčku *mirt* (1.32.1) v prostředí R (R Core Team, 2020). Raschův model předpokládá jednodimenzionální charakter položek. Do modelu nevstoupilo 8 % osob, jejichž odpověďové vzorce vykazovaly příliš vysoký misfit.

Shoda modelu s daty byla adekvátní (RMSEA = 0,070 ₉₀CI [0,066 ; 0,072], TLI = 0,907, SRMSR = 0,098).

Křivka chyby měření Raschova modelu (znázorněna níže na Obrázku 8) reprezentuje chybu měření v závislosti na schopnosti účastníka. Schopnost účastníka je latentním skórem, pro nějž je v rámci Raschova modelu předpokládáno normální rozložení. Toto rozložení má v případě jednoskupinového modelu průměr 0 a modelem odhadovanou směrodatnou odchylku, v tomto případě $s = 1,37$. Je zřejmé, že měření je pro většinu osob relativně přesné. V pásmu od -2,8 do +2,8 na ose x, což přibližně odpovídá ± 2 SD latentního skóru, se chyba měření pohybuje přibližně od 0,5 do 0,6. Nad 3 SD poté začíná strmě narůstat (více viz sekce *Reliabilita*).



Obrázek 8. Křivka chyby měření (inverzní k lokální reliabilitě z Obrázku 8).

Indexy položkového fitu (*infit* a *outfit*), popisující úroveň, s jakou se každá z položek shoduje s modelem, jsou uvedeny v Tabulce 7. Za nepřijatelné jsou obecně považovány položky s hodnotou *outfit* či *infit* větší než 2 (Linacre, 2002).

Tabulka 7. Infit a outfit položek (Raschův model).

Pořadí položky	outfit	infit
1*	0,94	1,01
2*	1,34	1,07
3*	0,79	1,00
4	0,75	0,82
5	0,66	0,77
6	0,63	0,83
7	0,93	0,91
8	0,68	0,76
9	0,59	0,76
10	1,33	1,05
11	0,84	0,93
12+	1,18	1,11
13	0,97	0,91
14	0,98	1,02
15	0,82	0,95
16	0,89	0,96
17	0,92	0,93
18	1,20	1,10
19	0,82	0,86
20+	1,49	1,25
21	1,09	0,96
22	1,04	0,98
23	1,20	1,03
24	1,32	1,03
25+	2,00	1,16
26	0,98	0,96
27	1,85	1,10
28	1,21	1,00
29	1,47	1,08
30	1,50	1,22
31	2,25	1,11
32	3,52	1,13

* = položka není skórována

+ = položka byla po standardizaci vyřazena

Na základě výsledků analýz uvedených v Tabulkách 6 a 7 byly vyřazeny položky 12, 20 a 25. Toto rozhodnutí nebylo založeno na jediném kritériu, ale vždy šlo o kombinaci faktorů. V případě položky 12, která byla navržena jako velmi snadná úloha s jednou podmínkou, se jednalo o překvapivě vysokou obtížnost, relativně horší hodnoty *infit* a *outfit* a nižší korigovanou korelaci s celkovým skórem. V případě položky 20 se ukázalo, že v důsledku nevhodně zvolených distraktorů je hned po žácích 6. třídy nejněsnější pro žáky 1. ročníku, což se projevilo zhoršenou hodnotou *outfit* a nižší korigovanou korelací této položky. Poslední vyřazená položka 25 byla pro testované žáky nečekaně obtížná. Tato obtížnost byla navíc nekonzistentní napříč jednotlivými ročníky, kdy žáci 5. ročníku oproti ostatním (i mladším žákům) při řešení této úlohy výrazně selhávali. To se projevilo i zhoršenou hodnotou *outfit* u této položky. Problematické jsou i poslední dvě položky testu (31 a 32), které mají nižší korigovanou

korelaci a příliš vysokou hodnotu *oufit*. To je však dáno především jejich extrémní obtížností i v případě nejstarších žáků. Vzhledem k tomu, že primárním cílem testu *Triton* je vyhledávání nadaných žáků, rozhodli jsme se tyto položky v testu ponechat s cílem zvýšit reliabilitu měření v pásmu nadprůměru. Tímto způsobem vznikla 8. verze testu s 29 položkami.

Reliabilita

Reliabilita byla odhadována pomocí dvou metod – split-half (klasická testová teorie) a IRT.

Split-half reliabilita

Split-half reliabilita je metodou výpočtu reliability zohledňující narůstající obtížnost položek. Počítá se jako korelace počtu správných odpovědí na liché a sudé položky. Split-half reliabilita hry *Triton* odpovídá $r = 0,82$, což ukazuje na spolehlivost odpovídající účelu testu. Tabulka 8 shrnuje hodnoty split-half reliability pro jednotlivé ročníky.

Tabulka 8. Odhad split-half reliability pro jednotlivé ročníky.

Ročník	Split-half reliabilita
1	0,67
2	0,69
3	0,86
4	0,76
5	0,78
6	0,86

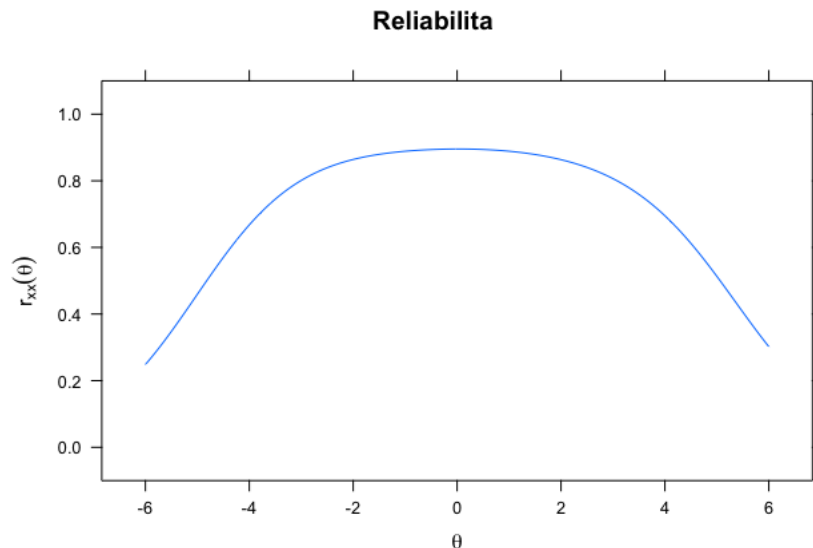
Z Tabulky 8 plyne, že relativně nízké reliability ($r < 0,75$) test dosahuje v první a druhé třídě. V ostatních ročnících lze měření z hlediska split-half reliability považovat za spolehlivé adekvátně skříninkovému účelu testu.

IRT reliabilita

Odhad IRT reliability je založen na Raschově modelu aplikovaném na sebraná data. Takto odhadnutá marginální reliabilita je podobná odhadu pomocí split-half (0,81; empirická reliabilita: 0,87²).

Na Obrázku 9 je uveden odhad lokální reliability testu. IRT přístup umožňuje odhadovat reliabilitu testu v závislosti na schopnosti participanta (tj. brát v úvahu chyby měření lišící se pro různé úrovně latentní schopnosti). Průměrné schopnosti a jejich odchylky pro jednotlivé ročníky jsou k dispozici v kapitole *Normy*.

² Pro popis rozdílu mezi marginální a empirickou reliabilitou z balíčku *mirt* pro jazyk R viz např. <https://stats.stackexchange.com/questions/427631/difference-between-empirical-and-marginal-reliability-of-an-irt-model>



Obrázek 9. Lokální reliabilita testu získaná z Raschova modelu. Řecké písmeno theta symbolizuje latentní schopnost.

Obrázek 9 demonstruje přijatelnou reliabilitu testu v rozpětí schopností většiny participantů (tj. od -2,8 do +2,8 na ose x, což zhruba odpovídá ± 2 SD latentního skóru). Extrémně schopní participanté jsou měřeni se sníženou spolehlivostí. Test si nicméně zachovává solidní standard pro většinu nadaných žáků ($> 0,7$), byť reliabilita u mimořádně nadprůměrných žáků (3 SD nad průměrem) klesá do pásma [0,7 – 0,4].

Skupinový model

Pro odhady průměrů distribucí latentních schopností a jejich směrodatných odchylek pro jednotlivé ročníky jsme vycházeli ze skupinového Raschova modelu s ročníky jako skupinami. Obtížnosti položek byly napříč skupinami omezeny na stejnou hodnotu, průměr latentní distribuce pro první ročník byl omezen na hodnotu 0. Průměry latentních distribucí ostatních ročníků a směrodatné odchylky všech latentních distribucí byly pak volně odhadovány. V modelu nebyly použity zácvičné ani vyřazené položky (viz Tabulka 7).

Skupinový model se všemi ročníky (1 – 6) jako skupinami vykazoval dostatečnou shodu s daty (RMSEA = 0,027 $_{90}CI$ [0,025 ; 0,028], SRMSR (1. ročník) = 0,193, SRMSR (2. ročník) = 0,127, SRMSR (3. ročník) = 0,105, SRMSR (4. ročník) = 0,090, SRMSR (5. ročník) = 0,099, SRMSR (6. ročník) = 0,126, TLI = 0,87). Zejména u 1. ročníku je ovšem velikost průměrných reziduí relativně vyšší než v případě ostatních, index TLI je pak adekvátní. Jestliže do modelu zahrneme pouze 3. - 6. ročníky, shoda modelu s daty se očekávaně zlepší (RMSEA = 0,025 $_{90}CI$ [0,022 ; 0,028], SRMSR (3. ročník) = 0,135, SRMSR (4. ročník) = 0,090, SRMSR (5. ročník) = 0,010, SRMSR (6. ročník) = 0,125, TLI = 0,92).

Výše uvedená zjištění ve spojení s nízkou reliabilitou u prvních a druhých ročníků nás vedla k závěru, že test je vhodný pouze pro žáky třetích až šestých ročníků.

Odhadnuté hodnoty průměrů a směrodatných odchylek latentních skóru na základě skupinového Raschova modelu pro 3. – 6. ročníky jsou uvedeny v Tabulce 9. Pro odhad normativního modelu byly ze vzorku odstraněny děti s diagnostikovaným mimořádným nadáním ($n = 19$), aby vzhledem ke svému nerovnoměrnému zastoupení v ročnících nezkrášlovaly ročníkové výsledky.

Tabulka 9. Popisné statistiky latentních skóru dle ročníku.

Ročník	Průměrný skór	SD skóru
3	0	1,3
4	0,1	1,2
5	0,4	1,1
6	0,9	1,3

Pozn.: Průměrné skóry jsou fixovány na hodnoty větší nebo rovno nule.

Validita

V této sekci předkládáme důkazy o tom, že je test *Triton* validním nástrojem pro měření *Fluidní inteligence* a *Kvantitativního usuzování*. Níže uvedené korelace odpovídají Pearsonovu korelačnímu koeficientu, není-li uvedeno jinak.

Obsahová validita

Jak je uvedeno v sekci *Východiska a zaměření*, test *Triton* byl vyvíjen primárně na základě C-H-C teorie inteligence (např. Schneider & McGrew, 2018) a v souladu s již ověřenými testovými principy, konkrétně Balance beam task od Inhelderové a Piageta (1958) a testu Figure Weights z baterií WAIS-IV (Wechsler, 2008a) a WISC-V (Wechsler, 2014).

Zaměření testu Figure Weights na *Fluidní inteligenci* bylo potvrzeno empiricky např. v rámci konfirmační faktorové analýzy na datech baterie WAIS-IV, kdy tento test vykazoval silný náboj na faktor *Gf* (Wechsler, 2008b). Zapojení *Kvantitativního usuzování* (konkrétně využívání numericko-symbolických asociací) při řešení úloh tohoto testu bylo ověřováno prostřednictvím pilotní klinické studie (McCrea & Robinson, 2011) s pěti dospělými s fokálními poškozeními mozkové kůry způsobenými cévní mozkovou příhodou či chirurgickým zákrokem, kterým byla administrována baterie WAIS-IV. Výsledky této pilotní studie poukazují na možné zapojení pravostranného temporoparietálního pomezí při řešení testu *Figure Weights*. Toto pomezí se přitom zapojuje do odhadu velikosti čísel.

Kromě *Fluidní inteligence* je při řešení úloh testu *Figure Weights* využívána také *Krátkodobá pracovní paměť* (Short-term working memory; *Gwm*). V manuálech testů WISC-V (Wechsler, 2014) a WAIS-IV (Wechsler, 2008a; 2014) se uvádí, že je při řešení úloh testu její zapojení minimalizováno na úroveň, která je nutná pro zvládnutí typických numerických úloh (např. počítání z paměti). Toho je dle autorů dosaženo díky tomu, že jsou úlohy předkládány vizuálně a dítě má tak možnost při jejich řešení jednotlivé podněty permanentně sledovat a průběžně obnovovat jejich reprezentace v pracovní paměti. Jak naznačují výsledky konfirmační faktorové analýzy testu WAIS-IV, její zapojení je poměrně značné (Wechsler, 2008b). Na možnou účast pracovní paměti také poukazují výsledky již zmíněné pilotní klinické studie (McCrea & Robinson, 2011). Při řešení testu *Figure Weights* se dle autorů zapojují převážně frontoparietální síť pravé hemisféry, jež se podílejí na vizuální pracovní paměti.

Souběžná validita

Souběžná validita s TIM³⁻⁵ – Studie Buráňové (2018)

Buráňová (2018) ověřovala souběžnou validitu 6. verze metody *Triton*³ a *Testu pro identifikaci nadaných žáků v matematice pro 3.–5. třídu (TIM³⁻⁵)* (Cígler et al., 2017).

V případě testu *Triton* používala 4 různé verze metody, které se lišily podobou zpětné vazby po každé položce (detailněji viz sekci *Vliv zpětné vazby*). Vzhledem k tomu, že se neprokázalo, že by podoba

³ Ve studii Buráňové (2018) je 6. verze tetu *Triton* označována pracovním názvem *Feed the Fish*.

zpětné vazby ovlivňovala výkon v testu, byla data ze všech čtyř verzí testu pro účely ověřování souběžné validity sloučena.

TIM³⁻⁵ je v rámci C-H-C teorie zaměřen na *Fluidní inteligenci* a v rámci ní na úzkou schopnost zvanou *Kvantitativní usuzování*. Test má dvě formy a je administrovaný skupinově. Každá z forem je tvořena 25 položkami, které jsou pojaty jako slovní úlohy, jejichž řešení vyžaduje logický úsudek v matematickém kontextu. V tomto ohledu je *TIM³⁻⁵* podobný testům typu *Applied Problems*, např. z baterie *Woodcock-Johnson IV Tests of Achievement* (Schrang et al., 2014). Reliabilita testu dosahuje hodnot Cronbachovy alfy od 0,67 do 0,91 (Cígler et al., 2017). Rozdíly ve skórech dětí, které byly předtím pedagogicko-psychologickou poradnou identifikovány jako mimořádně nadané, a těmi, které nikoliv, sahají od Cohenova $d = 1,2$ do 1,8, což ukazuje na velmi dobrou kritériální validitu testu.

Vzorek Buráňové (2018) sestával z $N = 202$ dětí třetích až pátých tříd (průměrný věk = 10,1, $SD = 0,86$) dvou základních škol z Jihomoravského kraje (ZŠ Bílovice nad Svitavou a ZŠ Úvoz, Brno). Autorka administrovala test *TIM³⁻⁵* v běžných třídách v průběhu ledna a února 2018, metodu *Triton* v počítačových učebnách v dubnu 2018. Oba testy zadávala skupinově.

Korelace příslušných hrubých skóreů metody *Triton* a testu *TIM³⁻⁵* pro třetí až pátou třídu činila postupně $r = 0,56$; $0,68$ a $0,82$, což ukazuje, že spolu oba nástroje u mladších dětí souvisí a u žáků pátých ročníků měří z velké části souběžné konstrukty, tedy *Kvantitativní usuzování*.

Souběžná validita s CFT 20-R

Studie Buráňové (2018)

Buráňová (2018) ve stejné studii rovněž posuzovala vztah 6. verze metody *Triton* a *Cattellova testu fluidní inteligence (CFT 20-R)*; Fajmonová et al., 2015).

CFT 20-R je neverbální test *Fluidní inteligence* zaměřený primárně na *Induktivní usuzování*. Je administrovaný skupinově dětem ve věku od 7,5 do 15 let. Test sestává ze dvou setů, z nichž každý má čtyři subtesty. Administrovat je možné pouze první set či oba. První set je také možné použít se standardním, či s prodlouženým časovým limitem. V případě studie Buráňové (2018) byly administrovány oba sady, u prvního byl použit prodloužený časový limit. Vnitřní konzistence celého testu je odhadována na úrovni Cronbachovy alfy = 0,88.

Vzorek Buráňové (2018) odpovídal vzorku popsanému v předchozí sekci. Autorka administrovala *CFT 20-R* v běžných třídách a metodu *Triton* v počítačových učebnách. Oba testy zadávala skupinově – v lednu a únoru 2018 test *TIM³⁻⁵*, v březnu a dubnu 2018 metodu *Triton*.

Korelace testu *Triton* s metodou *CFT 20-R* pro třetí až pátou třídu činila postupně $r = 0,43$; $0,57$ a $0,68$.

Studie na ZŠ Trávník, Přerov a ZŠ Novoměstská, Brno

Souběžná validita s testem *CFT 20-R* byla posuzována také pro 9. verzi testu, a to na vzorku 134 žáků (průměrný věk = 10,9, $SD = 0,62$) 4. a 5. ročníků ZŠ Trávník, Přerov, a ZŠ Novoměstská, Brno. Tato verze testu *Triton* se liší od 8. verze testu (popsané v sekci *Popis a struktura*) tím, že byla rozšířena o 4 nové položky a hráči v ní nezískávali za své odpovědi body. V případě testu *CFT 20-R* byl administrován pouze první set a to s prodlouženým časovým limitem. Sběr dat probíhal v průběhu dubna až června 2019. Časový interval mezi oběma testy nebyl delší než 23 dní, ve většině případů byl menší než 10 dní. Korelace mezi IQ skóre z *CFT 20-R* a standardními skóre z hry *Triton* byla v této studii odhadnuta na úrovni $r = 0,45$.

Závěr

Výše uvedené korelace z obou studií lze považovat za jeden z důkazů validity testu *Triton*. Je zřejmé, že tato metoda měří velmi podobný konstrukt jako etablovaný test *Fluidní inteligence CFT 20-R*. Domníváme se navíc, že zjištěné korelace mezi oběma testy mohou být podhodnocené s ohledem na odlišnosti obou metod. První skupina rozdílů je dána formou nástrojů – *Triton* obsahuje oproti testu *CFT*

20-R gamifikační prvky, je administrován na počítači a na jednotlivé položky není stanoven časový limit. Testy se však liší také svým zaměřením na úrovni úzkých schopností – *Triton* je zaměřen na *Kvantitativní usuzování*, *CFT 20-R* na *Induktivní usuzování*. Zjištěné souvislosti jsou o něco vyšší i ve srovnání s korelací mezi podobně zaměřenými testy v rámci baterie *Woodcock-Johnson IV (WJ IV; McGrew et al., 2014) – Číselné řady* (svým zaměřením podobné testu *Triton*) a *Formování konceptů* (podobné testu *CFT 20-R*) korelují ve věkové skupině 9-13 let $r = 0,42$.

Souběžná validita s testy Krátkodobé pracovní paměti

Studie na ZŠ Trávník, Přerov a ZŠ Novoměstská, Brno

Ve shora zmíněné studii na ZŠ Trávník, Přerov, a ZŠ Novoměstská, Brno, byl u stejných 134 žáků 4. a 5. ročníku individuálně administrován také test *Obrácené číselné řady* z baterie *Woodcock-Johnson: International edition II – Tests of Cognitive Abilities (WJ IE II COG; Ruef et al., 2010)*. Participanti zde mají za úkol opakovat administrátorem předčítané číselné řady v obráceném pořadí. Tento test je tedy měřítkem *Krátkodobé pracovní paměti*, konkrétně úzkých schopností *Kapacita pracovní paměti (WM)* a *Kontrola pozornosti (AC)*.

Časový interval mezi zadáním metody *Triton* a testu *Obrácené číselné řady* typicky činil 3-14 dní. korelační koeficient mezi standardními skóry obou nástrojů nabýval hodnoty $r = 0,3$ ($p < 0,001$), což ukazuje na slabý pozitivní vztah mezi oběma testy.

Studie Kubrického (2019)

Kubrický (2019) také posuzoval vztah mezi *Krátkodobou pracovní pamětí* a skóry v 7. verzi testu *Triton*⁴. V rámci strukturního modelování použil jako prediktory dvě metody – upravený test *Verbální pozornost* z baterie *WJ IV* a *Corsiho kostky* (např. Brunetti et al., 2014). Data z jeho studie byla rovněž použita ke standardizaci 7. verze testu *Triton*.

Autor modifikoval test *Verbální pozornost* z české verze baterie *WJ IV* (Schrank et al., 2019) do skupinově administrovatelné podoby. V rámci *Krátkodobé pracovní paměti* měří test dvě úzké schopnosti – *Kapacitu pracovní paměti (WM)* a *Kontrolu pozornosti (AC)*. Vnitřní konzistence testu byla v této studii odhadnuta na úrovni Cronbachova alfa = 0,88.

Corsiho kostky jsou počítačovou verzí testu *Corsi block-tapping task (CBTT)* – neuropsychologického nástroje, jenž měří krátkodobou prostorovou paměť a slouží k diagnostice vizuální krátkodobé paměti. Počítačová verze má podobné psychometrické vlastnosti jako verze původní (Brunetti et al., 2014). Vnitřní konzistence metody v této studii nicméně odpovídala hodnotě Cronbachova alfa = 0,62, což ukazuje na hraniční reliabilitu, omezující úsudky vytvořené na základě popisovaného výzkumu.

Corsiho kostky vyplnilo $N = 127$ žáků a modifikovaný test *Verbální pozornosti* $N = 120$ žáků 4. a 5. tříd ZŠ (typický věk 9 a 10 let). Testování probíhalo v říjnu a listopadu 2018 na dvou školách z Jihomoravského kraje (ZŠ Vančurova, Hodonín a ZŠ Očovská, Hodonín) společně se standardizací 7. verze testu *Triton*. Nejdříve byly bezprostředně za sebou administrovány *Corsiho kostky* a test *Triton*, modifikovaný test *Verbální pozornosti* pak následující den.

Korelace modifikovaného testu *Verbální pozornosti* z *WJ IV* a skóru z testu *Triton* byla odhadnuta na úrovni $r = 0,41$. *Corsiho kostky* a skóry z testu *Triton* byly asociovány na podobné úrovni, $r = 0,4$. Vzhledem k omezené reliabilitě použité verze *Corsiho kostek* je však nutné tento výsledek interpretovat s obezřetností.

Závěr

Výše uvedené korelace z obou studií lze považovat za jeden z důkazů validity testu *Triton*. Zjištěné hodnoty jsou v souladu s korelacemi reportovanými mezi testem *Figure Weights* (ze kterého metoda *Triton* vychází) a dalšími testy *Krátkodobé pracovní paměti* (konkrétně *Digit Span*, *Picture Span* a

⁴ Ve studii Kubrického (2019) je 7. verze testu *Triton* označována pracovním názvem *Feed the Fish*.

Letter-Number Sequencing) baterie *WISC-V* (Wechsler, 2014) – pro věk 8-10 let jsou uváděny hodnoty v rozmezí $r = 0,3-0,45$ (Wechsler et al., 2014). Dle očekávání se tedy na řešení úloh metody *Triton* podílí kromě *Fluidní inteligence* také *Krátkodobá pracovní paměť*. V kontextu těchto výsledků lze *Fluidní inteligenci* v rámci testu chápat jako schopnost specifické mentální manipulace s prvky z úloh, které jsou převedeny do podoby dílčích mentálních modelů, *Krátkodobou pracovní paměť* pak jako velikost mentální plochy, na které se tato manipulace odehrává, respektive počet mentálních modelů a jejich částí, kterými je v dané chvíli možné manipulovat.

Specifické populace

Žáci se specifickými poruchami učení - Studie na ZŠ Trávník, Přerov a ZŠ Novoměstská, Brno

V rámci studie na ZŠ Trávník, Přerov a ZŠ Novoměstská, Brno jsme na stejných žácích 4. a 5. ročníku (kromě vztahu s testy *CFT 20-R* a *Obrácené číselné řady*) ověřovali také souvislost 9. verze metody *Triton* s testy standardně používanými při diagnostice specifických poruch učení. Konkrétně se jednalo o dva testy *Baterie diagnostických testů gramotnostních dovedností pro žáky 2. až 5. ročníků ZŠ* (Caravolas & Volín, 2005) – *Elize hlásek* a *Test čtení s porozuměním*.

Elize hlásek je testem pro měření fonemického porozumění, které se odráží ve fonemické analýze pseudoslov. Posuzuje se přesnost a rychlost odpovědi respondenta. Test má dvě části – v první z nich má testovaný vynechávat druhou hlásku z (administrátorem verbálně prezentovaného) pseudoslova, ve druhé pak poslední hlásku. Vnitřní konzistence v případě žáků čtvrtých tříd je odhadována jako Cronbachova alfa = 0,68, u žáků pátých tříd pak 0,59.

Test čtení s porozuměním poskytuje odhad obecné čtenářské úrovně. Úkolem testovaného je v limitu 7 minut přečíst co nejvíce pasáží (maximálně 20) o délce jedné až tří vět a výběrem z možností doplnit do každé pasáže dvě chybějící slova. Autoři v manuálu metody neuvádí žádné informace o reliabilitě testu.

Časový interval mezi zadáním metody *Triton* a individuální administrací testu *Elize hlásek* typicky činil 3-14 dní. Standardní skóry přesnosti v *Elizi hlásek* korelovaly se standardními skóry testu *Triton* $r = 0,23$ ($N = 130$, $p = 0,01$). Standardní skóry rychlosti odpovědi v *Elizi hlásek* korelovaly se standardními skóry testu *Triton* $r = 0,04$ ($N = 130$, $p = 0,66$).

Časový interval mezi zadáním metody *Triton* a skupinovou administrací *Testu čtení s porozuměním* typicky činil 0-10 dní. Standardní skóry počtu správných odpovědi v *Testu čtení s porozuměním* korelovaly se standardními skóry metody *Triton* $r = 0,06$ ($N = 141$, $p = 0,47$). Standardní skóry podílu správných odpovědi k řešeným pak $r = 0,1$ ($N = 139$, $p = 0,23$).

U většiny sledovaných proměnných testů, které bývají standardně používány při diagnostice specifických poruch učení, nebyla nalezena korelace se skóry metody *Triton*. Tyto výsledky naznačují, že test *Triton* zřejmě u dětí se specifickými poruchami učení nefunguje rozdílně než u dětí bez těchto potíží. Mezi přesností v *Elizi hlásek* a skóry v testu *Triton* byla ovšem pozorována slabá pozitivní asociace, kterou lze dle našeho názoru přičíst vlivu obecné inteligence.

Další výzkumy

Vliv zpětné vazby

Ve zmíněné studii Buráňové (2018; viz sekci *Souběžná validita*) autorka zkoumala rovněž efekt zpětné vazby na skóry v 6. verzi testu *Triton*. Konkrétně byly použity čtyři podmínky (typy zpětné vazby), které byly prezentovány participantům po odpovědi na každou z položek. Jednalo se o:

- A. Žádnou zpětnou vazbu – participant nezískává žádnou informaci o správnosti či nesprávnosti své odpovědi.
- B. Jednoduchou/utvrzující zpětnou vazbu – participant je informován, zda byla jeho odpověď správná, či chybná.

- C. Podrobnou/animovanou zpětnou vazbu – participantovi je prezentován celý správný postup řešení úlohy prostřednictvím animace a to nezávisle na tom, zda byla jeho odpověď správná, či chybná.
- D. Participantem ovládanou zpětnou vazbu – participantovi je umožněno, aby si vybral kteroukoliv z předešlých alternativ (A, B, C), a to v jakémkoliv pořadí a počtu opakování.

Vzorek Buráňové (2018) byl následně v průběhu prosince 2018 a ledna 2019 rozšířen Portešovou, Jabůrkem, Ťápalem a Strakou (2020). Poté sestával z $N = 447$ dětí od 3. do 6. třídy ZŠ.

Portešová, Jabůrek, Ťápal a Straka (2020) uzavřeli svou studii tím, že ze získaných dat nevyplývá, že by některý z typů zpětné vazby napomáhal testovému výkonu. Ve skupině, kde si participant sami vybírali typ zpětné vazby (varianta D), byla nejčastěji volena podrobná zpětná vazba (alternativa C).

Perfekcionismus

Portešová, Jabůrek, Rečka a Straka (2020) ve své studii na vzorku 155 studentů prvního ročníku střední školy (47 % dívek, M věk = 12,4, $SD = 0,39$, průměrné IQ dle *CFT 20-R* = 120) prověřovali chování nadaných žáků při postupu testem *Triton* ve vztahu k jejich profilu perfekcionismu dle Parkera (1997; absence perfekcionismu, funkční perfekcionismus, nebo dysfunkční perfekcionismus), pohlaví a experimentální podmínce vyvolávající kompetitivnost. Výhradně pro účely této studie byla vytvořena nová verze testu *Triton* s odlišnými položkami, ve kterých nebyl použit princip více slotů, lektvar života a lektvar smrti (vysvětlení principů viz Tabulku 2). Jako závislé proměnné autoři sledovali četnost sledování bodů (body se zobrazily kliknutím na hvězdičku) a dobu vyplňování testu.

Obecně se ukázalo, že experimentální navození soutěže vede u všech respondentů k větší frekvenci monitorování získaných bodů a k nárůstu délky řešení hry. Pomocí multilevel lineární regrese autoři našli důkazy pro vyšší četnost sledování bodů u osob spadající do funkčně perfekcionistického profilu ve srovnání s dysfunkčním ($\beta = -1,36$, $p < 0,05$). Kompetitivní experimentální podmínka pak četnost sledování bodů zvyšovala méně u osob spadajících do profilu dysfunkčního perfekcionismu než u funkčního ($\beta = -0,64$, $p < 0,05$). U chlapců experimentální podmínka sledování skóre zvyšovala více než u dívek ($\beta = -0,55$, $p < 0,05$). Ve studii jsou dále uvedeny a komentovány další závěry.

Vliv úzkosti

Michalko (2020) ve své studii posuzoval vztah mezi úzkostí a výkonem v testu *Triton*⁵. Pro tyto účely vytvořil novou verzi metody se 17 položkami (z toho 12 novými), ve kterých nebyl použit princip více slotů, lektvar života a lektvar smrti (vysvětlení principů viz Tabulku 2).

Úzkost byla posuzována pomocí *Škály na měření úzkosti a úzkostnosti dětí – revize (ŠAD-R)*. Metodu vytvořil Toman (2013) jako revizi původní škály na měření jmenovaných konstruktů (Müllner et al., 1983), která vznikla na základě Spielbergova (1972) pojetí úzkosti. Metoda má dvě části (měřící stavovou úzkost, a úzkostnost jako osobnostní rys) o 17 položkách. Reliabilitu ve smyslu vnitřní konzistence vykazuje škála úzkostnosti Cronbachova alfa = 0,86-89. Škála stavové úzkosti pak Cronbachova alfa = 0,89-0,92.

Sběr dat probíhal v listopadu a prosinci 2019. Vzorek sestával z $N = 64$ žáků 6. tříd ve věku od 10 do 13 let ($M = 11,47$, $SD = 0,56$). Skupině A ($N = 32$) byla administrována verze testu *Triton* s korektivní zpětnou vazbou (tedy potvrzením, zda byla odpověď správná či chybná), zatímco skupina B takovou zpětnou vazbu neobdržela. Dotazník úzkosti byl administrován před zadáním testu *Triton* a znovu po jeho dokončení.

Regresní model s pre-testovou úzkostí, podmínkou (zpětnou vazbou) a vzájemnou interakcí jako prediktory nabídl důkazy směrem ke znatelnému efektu zpětné vazby na výkon. Úzkost ovšem

⁵ Ve studii Michalka (2020) je metoda *Triton* označována pracovním názvem *Feed the Fish*.

s testovými skóry nesouvisela [podmínka: $\beta = 3,43$, $t(60) = 2,22$, $p = 0,03$; pre-testová úzkost: $\beta = -0,07$, $t(60) = -0,84$, $p = 0,41$; interakce: $\beta = -0,11$, $t(60) = -0,95$, $p = 0,35$]. Závěry Michalka o vlivu zpětné vazby na výkon v testu *Triton* nepotvrzují výsledky studie Portešové, Jabůrka, Tápala a Straky (2020), která je blíže popsána v sekci *Vliv zpětné vazby* a která žádný efekt neobjevila. Bylo by proto vhodné se této problematice dále výzkumně věnovat.

Poděkování

Rádi bychom srdečně poděkovali Daně Seryjové Juhové, Svatavě Buráňové, Richardu Kubrickému a Peterovi Michalkovi, kteří se v rámci svých závěrečných prací podíleli na dalším výzkumném ověřování a zdokonalování testu.

Děkujeme rovněž Jihomoravskému kraji za finanční podporu při vývoji metody v rámci projektu „JMK 181885/2016 - Systematické vyhledávání a rozvoj nadaných dětí v JMK“.

Literatura

- Bond, T., Yan, Z., & Heene, M. (2020). *Applying the Rasch model: Fundamental measurement in the human sciences*. Routledge. <https://doi.org/10.4324/9780429030499>
- Brunetti, R., Del Gatto, C., & Delogu, F. (2014). eCorsi: implementation and testing of the Corsi block-tapping task for digital tablets. *Frontiers in psychology*, 5, 939. <https://doi.org/10.3389/fpsyg.2014.00939>
- Buráňová, S. (2018). Ověření validity nově vyvíjené počítačové hry k měření fluidní inteligence a souvisejících metakognitivních dovedností. Unpublished master's thesis. Masarykova univerzita.
- Caravolas, M., & Volín, J. (2005): *Baterie diagnostických testů gramotnostních dovedností pro žáky 2. až 5. ročníků ZŠ*. Institut pedagogicko-psychologického poradenství ČR.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511571312>
- Cígler, H., Jabůrek, M., Straka, O., & Portešová, Š. (2017). *Psychometrická analýza TIM3–5 – Testu pro identifikaci nadaných žáků v matematice pro 3.–5. třídu*. Masarykova univerzita. <https://munispace.muni.cz/library/catalog/book/968>
- Gridley, B. E., Norman, K. A., Rizza, M. G., & Decker, S. L. (2003). Assessment of gifted children with the Woodcock Johnson III. In F. A. Schrank & D. P. Flanagan (Eds.), *WJ III clinical use and interpretation: Scientist–practitioner perspectives* (pp. 285–317). Academic Press. <https://doi.org/10.1016/B978-012628982-4/50010-6>
- Hofman, A. D., Visser, I., Jansen, B. R., & van der Maas, H. L. (2015). The balance-scale task revisited: a comparison of statistical models for rule-based and information-integration theories of proportional reasoning. *PloS one*, 10(10), e0136449. <https://doi.org/10.1371/journal.pone.0136449>
- Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 53–91). Guilford Press.
- Hunt, E. (2011). *Human intelligence*. New York: Cambridge University Press
- Jansen, B. R., Raijmakers, M. E., & Visser, I. (2007). Rule transition on the balance scale task: a case study in belief change. *Synthese*, 155(2), 211-236. <https://doi.org/10.1007/s11229-006-9142-9>
- Kubrický, R. (2019). *Význam krátkodobé paměti při hraní počítačové hry zaměřené na fluidní inteligenci*. Unpublished bachelor's thesis. Masarykova univerzita.
- Lichtenberger, E. O., & Kaufman, A. S. (2009). *Essentials of WAIS-IV assessment*. John Wiley.
- Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean. *Rasch Measurement Transactions*, 16(2), 878.
- McCrea, S., & Robinson, T. P. (2011). Visual puzzles, figure weights, and cancellation: Some preliminary hypotheses on the functional and neural substrates of these three new WAIS-IV subtests. *Neurology*, 1–19. <https://doi.org/10.5402/2011/123173>
- McGrew, K. S., LaForte, E. M., & Schrank, F. A. (2014). *Woodcock Johnson IV: Technical manual*. Riverside.

- Michalko, P. (2020). *Korektivní zpětná vazba a časový management v prostředí diagnostické hry Feed the Fish*. Unpublished bachelor's thesis. Masarykova univerzita.
- Müllner, J., Ruisel, I., & Farkaš, G. (1983). *Škály na meranie úzkosti a úzkostlivosti u detí*. Psychodiagnostické a didaktické testy.
- Onwuegbuzie, A. J., & Seaman, M. A. (1995). The effect of time constraints and statistics test anxiety on test performance in a statistics course. *The Journal of Experimental Education*, 63(2), 115-124. <https://doi.org/10.1080/00220973.1995.9943816>
- Parker, W. D. (1997). An empirical typology of perfectionism in academically talented children. *American Educational Research Journal*, 34(3), 545-562. <https://doi.org/10.3102/00028312034003545>
- Portešová, Š., Jabůrek, M., Rečka, K., & Straka, O. (2020). *To click, or not to click? How perfectionism affects game points monitoring during a logic videogame in high-able students* [Manuscript submitted for publication]. Masarykova univerzita.
- Portešová, Š., Jabůrek, M., Tápal, A., & Straka, O. (2020). *The effect of computer-based feedback in game-like fluid reasoning tasks* [Manuscript submitted for publication]. Masarykova univerzita.
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Reams, R., Chamrad, D., & Robinson, N. M. (1990). The race is not necessarily to the swift: Validity of WISC-R bonus points for speed. *Gifted Child Quarterly*, 34(3), 108-110. <https://doi.org/10.1177/001698629003400304>
- Ruef, M., Furman, A., & Muñoz-Sandoval, A. (2010). *Woodcock-Johnson Czech International Edition II*, Czech editors: J. Mareš, E. Vondráková, J. Zapletalová, D. Heider & J. Burešová. The Woodcock-Muñoz Foundation.
- Schapiro, A. C., & McClelland, J. L. (2009). A connectionist model of a continuous developmental transition in the balance scale task. *Cognition*, 110(3), 395-411. <https://doi.org/10.1016/j.cognition.2008.11.017>
- Schneider, W. J., & McGrew, K. S. (2018). The Cattell-Horn-Carroll Theory of Cognitive Abilities. In D. P. Flanagan, & E. M. McDonough (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (4th edition) (pp. 73-163). The Guilford Press.
- Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *The Woodcock-Johnson IV: Tests of Achievement*. Riverside.
- Schrank, F. A., McGrew, K. S., & Mather, N. (2019). *Woodcock-Johnson® IV Testy kognitivních schopností*, Czech editors: A. Furman, L. Krejčová, J. Laciga, T. Urbánek & M. Víchová. Propsyco.
- Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive psychology*, 8(4), 481-520. [https://doi.org/10.1016/0010-0285\(76\)90016-5](https://doi.org/10.1016/0010-0285(76)90016-5)
- Spielberger, C. D. (1972). Conceptual and methodological issues in anxiety research. In C. D. Spielberger (Ed.), *Anxiety: Current trends in theory and research*, Vol. 2 (pp. 481-493). Academic Press. <https://doi.org/10.1016/B978-0-12-657402-9.50013-2>
- Toman, J. (2013). *Škály na měření úzkosti a úzkostlivosti u dětí – revize*. Institut pedagogicko-psychologického poradenství.

Wechsler, D. (2008a). *Wechsler Adult Intelligence Scale* (4th ed.). Pearson.
<https://doi.org/10.1037/t15169-000>

Wechsler, D. (2008b). *WAIS-IV Technical and Interpretive Manual*. Pearson.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children* (5th ed.). Pearson.

Wechsler, D., Raiford, S. E., & Holdnack, J. A. (2014). *WISC-V technical and interpretive manual supplement: Special group validity studies with other measure and additional tables*.
<http://downloads.pearsonclinical.com/images/Assets/WISC-V/WISC-V-Tech-Manual-Supplement.pdf>