MASARYKOVA UNIVERZITA

FAKULTA SOCIÁLNÍCH STUDIÍ

# Subtest Líná Džina: Ověření metody pro měření schopnosti rozpoznat emoce z intonace u dětí mladšího školního věku

Bakalářská práce

KAROLÍNA UMOVÁ

Vedoucí práce: Mgr. Veronika Šromová

Katedra psychologie
obor Psychologie

Brno 2023

MUNI
FSS

# Bibliografický záznam


| | |
|---|---|
| **Autor:** | Karolína Umová<br>Fakulta sociálních studií<br>Masarykova univerzita<br>Katedra psychologie |
| **Název práce:** | Subtest Líná Džina: Ověření metody pro měření schopnosti rozpoznat emoce z intonace u dětí mladšího školního věku |
| **Studijní program:** | Psychologie |
| **Studijní obor:** | Psychologie |
| **Vedoucí práce:** | Mgr. Veronika Šromová |
| **Konzultant práce:** | Mgr. et Mgr. Petr Palíšek |
| **Rok:** | 2023 |
| **Počet stran:** | 74 |
| **Klíčová slova:** | rozpoznání emoce z intonace, emoční inteligence, Invenio, mladší školní věk, intonace, výkonový test |

# Bibliographic record

**Author:** Karolína Umová
Faculty of Social Studies
Masaryk University
Department of Psychology

**Title of Thesis:** Subtest Líná Džina: Validation of a method measuring emotion recognition from intonation in school-age children

**Degree Programme:** Psychology

**Field of Study:** Psychology

**Supervisor:** Mgr. Veronika Šromová

**Consultant:** Mgr. et Mgr. Petr Palíšek

**Year:** 2023

**Number of Pages:** 74

**Keywords:** emotion recognition from voice, emotional intelligence, Invenio, school-age children, intonation, performance test

# Anotace

Tato bakalářská práce se zabývá ověřením nového subtestu s názvem *Líná Džina* určeného k měření schopnosti rozpoznávat emoci z intonace u dětí mladšího školního věku. Východiskem pro tvorbu subtestu je Salovey-Mayerova teorie emoční inteligence. Subtest je součástí testu *Frída a podivuhodná galerie,* který by měl měřit dílčí schopnosti v rámci emoční inteligence, jedné z možných širokých schopností v CHC teorii inteligence, ze které vychází také zbytek testů Invenio. Spolu se subtestem byl administrován test měřící verbální schopnost pro ověření diferenciální validity. Pilotáže se účastnilo 462 žáků 2.-5. ročníku základních škol. Dalších 6 participantů se účastnilo kognitivní pilotáže. Zjištěná vnitřní konzistence testu byla $\alpha = 0{,}68$, 95%CI [0,43; 0,85] a $\omega = 0{,}66$. Výsledky IRT analýzy poukazují na to, že je subtest relativně spolehlivým měřítkem pouze pro participanty s podprůměrnou úrovní schopnosti, nikoliv však průměrnou. Na základě výsledků byla formulovaná konkrétní doporučení na úpravu položek a diskutovány možné směry vývoje testu.

# Abstract

This bachelor thesis deals with the validation of a new method measuring emotion recognition from intonation in school-aged children called *Líná Džina.* It builds upon the theory of emotional intelligence proposed by Salovey and Mayer. The method is part of a test called *Frída a podivuhodná galerie* that aims to measure different aspects of emotional intelliegnce, which is one of the possible broad factors in the CHC theory of intelligence. The method was administered alongside a method measuring verbal ability in order to assess diferential validity. The sample consisted of 462 elementary school pupils (2nd through 5th grade). Six more children participated in cognitive interviews. The internal consistency estimates of the test were Cronbach's $\alpha$ = 0,68, 95%CI [0,43; 0,85] and McDonald's $\omega$ = 0,66. The results of IRT analysis indicate that the test is not a reliable measure for participants with average ability and is only somewhat reliable for below average particpants. Specific recommendations for improvement of the method are discussed, as well as future possible directions.

# Čestné prohlášení

Prohlašuji, že jsem bakalářskou práci na téma **Subtest Líná Džina: Ověření metody pro měření schopnosti rozpoznat emoce z intonace u dětí mladšího školního věku** zpracovala sama. Veškeré prameny a zdroje informací, které jsem použila k sepsání této práce, byly citovány v textu a jsou uvedeny v seznamu použitých pramenů a literatury.

V Brně 10. května 2023 .....

Karolína Umová

# Poděkování

V první řadě děkuji vedoucí této práce, Veronice Šromové, za její báječný přístup a vedení, díky kterému jsem neměla nikdy pocit, že jsem na cokoliv sama. Děkuji také konzultantovi Petrovi Palíškovi za všechny moc cenné připomínky ohledně designu a analýzy dat a společné fandění IRT a nadávání na *mirt*. Díky mu patří také za jeho přátelství, kterého si moc vážím. Dále za super připomínky k práci a skvělý přístup děkuji Michalovi Jabůrkovi a celému Inveniu. Jsem také vděčná Janu Širůčkovi za konzultaci, která mi pomohla určit směřování práce a zároveň mě hodně povzbudila.

Mockrát děkuji všem neskutečně ochotným ředitelům a zástupcům škol, kteří se o administraci postarali, a také všem dětem, které se výzkumu zúčastnily.

Veliké díky patří mojí rodině, mým spolužákům, přátelům, spolubydlícím, kolegům a známým, kteří můj život i mě samotnou dělají mnohem lepší. Zvlášť pak díky těm, co mi půjčili své děti na kognitivní pilotáže – Honzovi s Lenkou, Evči a Davidovi. Hrozně ráda bych poděkovala všem jmenovitě, ale bohužel by se to na jednu stránku nevešlo. Zcela nevyčerpávajícím způsobem tak děkuji Anežce, Kačce, Lence, Aničce, Barče, Báře, Nince, Domče a Adélce, že u nich vždy nacházím podporu.

Nejvíc ze všeho děkuji svému příteli Adíčkovi za všechnu lásku a mimořádně podpůrné prostředí, za rady i slova útěchy a neustálé ujišťování, že *všechno dává smysl*.

*At its most advanced level, intrapersonal knowledge allows one to detect and to symbolize complex and highly differentiated sets of feelings, to attain a deep knowledge of (...) feeling life.*

*Gardner (1983), s. 239*

# Obsah

**Seznam zkratek** 13

**1 Úvod** 15

1.1 Emoční inteligence 16
1.2 Cattel-Horn-Carrollova teorie inteligence 19
1.3 Schopnost rozpoznání emoce z intonace 21
1.4 Existující měřící nástroje 23
1.5 Vztah s verbální inteligencí 26
1.6 Specifika testování dětí 28
1.7 Cíle studie 28

**2 Metoda** 30

2.1 Líná Džina 30
2.2 Věřte, nevěřte! 30
2.3 Procedura 31
2.4 Tvorba vzorku 32
2.5 Popis výzkumného souboru 33
2.6 Analytický plán 34

**3 Výsledky** 35

3.1 Čištění a příprava dat 35
3.2 Doba vyplňování 35
3.3 Deskriptivní statistiky subtestu Líná Džina 37
3.4 Kognitivní pilotáže 39
3.5 CTT analýza subtestu Líná Džina 40
3.6 IRT analýza subtestu Líná Džina 43
3.7 Analýza testu Věřte, nevěřte! 48
3.8 Testy hypotéz 50

**4 Diskuze** 52

4.1 Shrnutí výzkumných zjištění ..... 52
4.2 Celkové zhodnocení testu ..... 53
4.3 Možné způsoby dalšího vývoje testu ..... 54
4.4 Limity ..... 55

**Použité zdroje** **58**

**Příloha A Náhled v platformě Qualtrics** **69**

**Příloha B Tetrachorické korelace subtestu Líná Džina** **70**

**Příloha C Tetrachorické korelace smysluplných položek** **71**

**Příloha D Tetrachorické korelace nesmyslných položek** **72**

# Seznam zkratek

| | | |
|---|---|---|
| 1PL | – | jednoparametrový logistický model |
| 2PL | – | dvouparamterový logistický model |
| BBC | – | British Broadcasting Company |
| Cam-C | – | Cambridge Face-Voice Battery for Children |
| CHC teorie | – | Cattel-Horn-Carrollova teorie inteligence |
| CTT | – | klasická testová teorie |
| DANVA | – | Diagnostic Analysis of Nonverbal Accuracy |
| EI | – | emoční inteligence |
| ERI | – | Emotion Recognition Index |
| FSS | – | Fakulta sociálních studií |
| Gc | – | široká schopnost Porozumění-znalosti v CHC |
| Gei | – | široká schopnost Emoční inteligence v CHC |
| GERT | – | Geneva Emotion Recognition Test |
| MERT | – | Multimodal Emotion Recognition Test |
| MSCEIT (-YV) | – | Mayer-Salovey-Caruso Emotional Intelligence Test (- Youth Version) |
| MU | – | Masarykova univerzita |
| RE | – | rozpoznání emoce |
| REI | – | rozpoznání emoce z intonace |
| RMSEA | – | Root Mean Square Error of Approximation |
| RMV | – | Reading the Mind in the Voice Test |
| SRMR | – | Standardized Root Mean Square Residual |
| TLI | – | Tucker-Lewis Index |

# 1 Úvod

Tým Invenio, který pracuje na Katedře psychologie Fakulty sociálních studií Masarykovy univerzity, vyvíjí množství testů, které slouží k identifikaci nadaných dětí prostřednictvím měření úrovně jejich schopností v různých doménách. Vychází přitom z Cattel-Horn-Carrollovy teorie inteligence (McGrew, 2009). Test *Frída a podivuhodná galerie* (Portešová et al., 2022) by měl zjišťovat schopnosti dětí na prvním stupni základních škol v socio-emočních dovednostech. Východiskem je mu teorie emoční inteligence konceptualizovaná Mayerem et al. (1999). Tato bakalářská práce se zabývá ověřením subtestu s názvem *Líná Džina*, který by měl měřit schopnost rozpoznat emoci z intonace v hlase. Spolu s ním byl administrován také druhý test s pracovním názvem *Věřte, nevěřte!*, který je zaměřen na verbální schopnost dětí.

Umět identifikovat děti s různými odchylkami ve schopnostech řazených pod pojem emoční inteligence je důležité z toho důvodu, že se jedná o korelát schopností, které mohou sloužit dítěti k úspěšné adaptaci na školní prostředí i zvládání školní docházky. Podle různých autorů například souvisí s chováním vůči vrstevníkům (Qualter et al., 2018) či akademickým úspěchem (MacCan et al., 2020). Obecně lze říct, že děti, které lépe rozumí emočnímu prožívání druhých, by měly být schopny lépe používat a regulovat své vlastní emoce (Saarni et al., 2007).

## 1.1 Emoční inteligence

### 1.1.1 Historie EI a příbuzné koncepty

Konstrukt emoční inteligence byl poprvé konceptualizován počátkem devadesátých let 20. století (O'Connor et al., 2019). Za konceptem stáli dva výzkumníci z Yaleovy univerzity, Peter Salovey a John D. Mayer, kteří původně definovali emoční inteligenci jako *„schopnost monitorovat emoce a pocity své i druhých lidí, rozlišovat mezi nimi a používat je k řízení vlastního myšlení a jednání"* (Salovey & Mayer, 1990, s. 189).

Zkoumání schopností a dovedností, které bychom v současnosti zařadili pod pojem emoční inteligence, ovšem trvá mnohem déle. Mezi její teoretické předchůdce můžeme zařadit sociální inteligenci, kterou E. L. Thorndike vymezil jako *„schopnost porozumět mužům a ženám, chlapcům i dívkám – a v mezilidských vztazích se chovat moudře."* (Thorndike, 1920, s. 228). Pokusy o její měření však typicky selhávaly na nemožnosti schopnost izolovat od verbálního porozumění a na velmi vysoké korelaci měřítek s testy obecné inteligence (Conzelmann et al., 2013).

Emoční inteligence dále bývá spojována s kontextuální (praktickou) inteligencí z triarchické teorie inteligence Roberta Sternberga. Jedná se o druh inteligence, pomocí nějž bychom měli být schopni se orientovat ve světě, tedy chovat se co možná nejefektivnějším způsobem (Sternberg, 1984). Analogii k emoční inteligenci nacházíme právě ve schopnosti dokázat uchopit kontext, který se okolo nás odehrává. Dalším podobným konstruktem je pak personální inteligence, tedy schopnost uvažovat o osobnosti své a svého okolí (Mayer et al., 2016).

Mayer et al. (2016) zastřešují všechny zmíněné konstrukty termínem *horké* inteligence. Mělo by se jednat o druhy inteligence, které pracují s informacemi, jež jsou relevantní osobně (týkají se identity, sociálního statutu atd.). Protikladem jsou *chladné* druhy inteligence, jak označujeme takové, které pracují s neosobními znalostmi a dovednostmi (např. matematické či vizuálně-prostorové schopnosti). Vzájemná pozice jednotlivých horkých inteligencí není dodnes jasná, autoři argumentují pro to, aby byly jednotlivé druhy prozatím posuzovány zvlášť, dokud se nenajdou důkazy pro jejich hierarchické uspořádání.

Pojem emoční inteligence se rychle rozšířil mezi laickou veřejnost, k čemuž bezpochyby přispěla kniha napsaná Danielem Golemanem v roce 1995, *Emoční inteligence (Proč může být emoční inteligence důležitější než IQ)*, která se stala jednou z nejznámějších pop-psychologických

knih všech dob. Emoční inteligence se sice těší neutuchajícímu zájmu výzkumníků, ale také je velmi často podrobována kritice. Waterhouse (2006) navrhuje, že zároveň výhodou i nevýhodou EI je fakt, že se jedná o velmi jednoduché, ale rovněž přesvědčivé a elegantní vysvětlení toho, proč se lidé liší ve schopnosti nakládat s emocemi. Zdá se však, že jednotně definovat tento konstrukt je složitější, než se uvažovalo v devadesátých letech, a do dnešního dne neexistuje univerzálně přijímaná konceptualizace.

### 1.1.2 Odlišení výkonnostního modelu EI od souvisejících konstruktů

V následujících kapitolách operujeme s konstrukty s odlišnými názvy, které ale často odkazují k výzkumům s podobným designem a měřícím nástrojům s velmi podobným testovým principem. Podle Riggio et al. (2010) byl vůbec prvním souhrnným názvem pro tyto schopnosti pojem *neverbální komunikace* či *neverbální schopnost*. Zahrnoval schopnosti kódování či dekódování emoce do či z výrazu v obličeji či hlasu, ale i komplexnější schopnosti jako detekce klamání. O něco později začal být tentýž konstrukt nazýván socio-emoční dovednosti. V 90. letech pak s nástupem emoční inteligence byla sociální složka vypuštěna (vzhledem k již zmiňované problematice jejího měření). Do současnosti tedy existují dvě linie výzkumu zkoumající podobné konstrukty, neverbální schopnost a emoční inteligenci, málokdy však své výsledky berou navzájem v potaz (Riggio et al., 2010).

Podle Mayera et al. (1999) existují dva druhy modelů EI. Prvním je tzv. výkonnostní model. Emoční inteligenci považují za výkonnostní charakteristiku, jejíž jednotlivé užší schopnosti se rozdělují do celkem čtyř větví. Druhým případem jsou smíšené modely, které mimo schopnosti EI zahrnují také nevýkonnostní složky osobnosti, jako jsou rysy či motivace. Jedná se například o Bar-On model inteligence (Bar-On, 2006), který k EI přistupuje jako k nekognitivní osobnostní charakteristice (Qualter et al., 2008) či již zmiňovaný Golemanův přístup (Goleman, 1995).

Právě vymezením EI jako výkonnostní charakteristiky se dlouhodobě Salovey a Mayer snaží docílit toho, aby EI byla etablována jako druh inteligence na úrovni např. fluidní inteligence (Mayer et al., 1999).

### 1.1.3 Salovey-Mayerův model emoční inteligence

Pravděpodobně nejznámější konceptualizací emoční inteligence, z níž vychází také ověřovaný subtest *Líná Džina* i celá testová baterie *Frída a podivuhodná galerie*, je již zmiňovaný Salovey-Mayerův model emoční inteligence (Mayer et al., 1999; 2016). Po různých úpravách model současně rozděluje EI na čtyři větve, a to jsou:

1) vnímání emocí
2) facilitace myšlení pomocí emocí
3) porozumění emocím
4) řízení emocí sebe a druhých

Každá z větví pak obsahuje dílčí schopnosti, které jsou seřazeny vzestupně (odspodu) v tabulce 1. Tabulka je adaptována z článku Mayera et al. (2016) a obsahuje pouze schopnosti 1. a 3. větve.

**Tabulka 1.** Schopnosti 1. a 3. větve Salovey-Mayerova modelu

| **1. větev**<br>**Vnímání emocí** | **3. větev**<br>**Porozumění emocím** |
|---|---|
| Identifikovat emoce ve vlastních fyzických stavech, myšlenkách a pocitech | Pojmenovat emoce a rozpoznat vztahy mezi nimi |
| Vnímat emoce u druhých za pomoci vokálních nápovědí, výrazu v obličeji, jazyka a chování | Určit antecedenty, význam a následky emocí |
| Vnímat emocionální obsah v prostředí, vizuálním umění a hudbě | Zhodnotit situace, které mají potenciál vyvolat emoci |
| Pokud je to potřeba, adekvátně vyjadřovat emoce | Rozlišovat mezi náladami a emocemi |
| Porozumět, jak jsou emoce vyjadřovány v závislosti na kontextu a kultuře | Porozumět komplexním a smíšeným emocím |
| Odlišovat adekvátních a neadekvátních vyjádření emocí | Rozpoznat pravděpodobné přechody mezi emocemi, jako například ze vzteku ke spokojenosti |
| Identifikovat klamavá či neupřímná vyjádření emocí | Pochopit, jak se osoba může cítit v budoucnu nebo při určitých podmínkách (*affective forecasting*) |
| | Rozpoznávat kulturní rozdíly v hodnocení emocí |

Způsob, jakým je subtest *Líná Džina* konstruován, po participantovi vyžaduje zapojení schopností z těchto dvou větví Salovey-Mayerova modelu inteligence. Pochopitelně jde o schopnost vnímat emoce u druhých

za pomoci vokálních nápověd z první větve. Za druhé, protože jsou scénky kreslené a na konci má postava nějaký výraz v tváři, uplatňuje se také schopnost rozpoznání emoce z výraz u (byť kresleného) v obličeji. Je pravděpodobné, že do odpovídání vstupuje také kulturní kontext, kdy participanti spíše volí kulturně normativní odpověď, což souvisí také se schopností odlišit adekvátní a neadekvátní vyjádření emocí.

Dalo by se taktéž argumentovat, že subtest vychází z některých dílčích schopností třetí větve *porozumění emocím*. Dítě totiž musí správně vybrat, s jakou intonací byla věta pronesena, nikoliv pouze označit emoci, kterou intonace vyjadřovala. Uplatňuje se zde schopnost zhodnotit, zda má situace potenciál vyvolat emoci a také pochopit, co emoci předcházelo, stejně jako porozumět emoci samotné. Scénky jsou navrženy tak, aby bylo vyznění jednoznačné, a tedy založené pouze na rozpoznání emoce, nikoliv na nutnosti porozumět nuancím a vztahům mezi nimi, je ale otázkou, do jaké míry je možné tyto jevy oddělit. Naopak (na rozdíl od jiných testů) neobsahují nutnost emoci pojmenovávat.

## 1.2 Cattel-Horn-Carrollova teorie inteligence

Uvažování o emoční inteligenci jako o schopnosti vedlo také k otázce jejího zařazení do jedné z nejuznávanějších teorií inteligence, a to Cattell-Horn-Carrollovy (CHC) teorie (Schneider & McGrew, 2018). CHC je taxonomickou teorií, která spojuje poznatky více teoretiků. Používaná je zejména v oblastech výzkumu, ve kterých jsou vytvářeny individuální baterie kognitivních schopností (McGrew, 2009), což je také případ týmu Invenio.

Snahy o kvantifikaci lidského intelektu sahají do počátku 20. století, mezi předchůdce CHC teorie můžeme řadit Thurstonův výzkum primárních mentálních schopností a Spearmanovu snahu o nalezení faktoru g, tedy faktoru obecné inteligence (Buchanan & Finch, 2005). Za autora CHC teorie pak lze považovat Kevina McGrewa, který dohromady spojil myšlenky dvou velkých teorií inteligence, po jejichž autorech je CHC teorie pojmenovaná. Konkrétně se jedná o Carrollovu teorii tří vrstev a Cattel-Hornovu *Gf-Gc* teorii.

Horn vytvořil na základě velkého množství faktorových analýz seznam devíti faktorů, jako jsou fluidní (*Gf*) a krystalická (*Gc*) inteligence, ale například také rychlost kognitivního zpracování. V této práci navázal na snahy svého učitele, Raymonda Cattella. John Carroll vytvořil v zásadě

podobnou teorii, ale s rozdílem, že v jeho modelu jsou schopnosti organizovány do tří vrstev (nazvaných *strata*) a zastřešuje je jeden faktor, nazývaný obecná inteligence (g), s jehož existencí Horn nesouhlasil (Schneider & McGrew, 2018).

Spojením těchto dvou teorií vznikla CHC teorie. Pod zastřešujícím obecným faktorem inteligence se nachází tzv. široké schopnosti, což je například právě zmiňovaná fluidní inteligence. Poslední vrstvu pak tvoří tzv. úzké schopnosti, z nichž se široké schopnosti skládají. V případě *Gf* by se jednalo například o verbální usuzování. Co se týče subtestu *Líná Džina*, v kontextu emoční inteligence by se jednalo o širokou schopnost *Gei* a úzkou schopnost *vnímání emocí.*

Pochopitelně existují také kritické hlasy ve směru CHC teorie. Mezi zmiňované problémy patří například to, že se jedná spíše o taxonomii než teorii (tj. neumožňuje dělat nové predikce), otázka její replikovatelnosti a také velmi snadná modifikovatelnost (Wassermann, 2019), kdy počet širokých schopností vzrostl od roku 1997 o celých sedm (Schneider & McGrew, 2018). Zmíněná modifikovatelnost je problematická v tom, že může vést k argumentaci kruhem, při které může dojít k tomu, že se teorie bude validizovat sama sebou tím, že do ní mohou být přidávány stále nové schopnosti (McGill & Dombrowski, 2019).

### 1.2.1 EI jako kandidátská schopnost v CHC

Zmíněná snadná modifikovatelnost CHC je důležitá v souvislosti s emoční inteligencí, neboť je některými EI považována za kandidátskou širokou schopnost. Čtyřvětvý model EI splňuje některé podmínky pro zařazení do CHC (Schneider & McGrew, 2018).

Důvodů, proč je zařazení EI problematické, je však několik. Jedná se zejména o nedostatek důkazů pro ověření struktury modelu EI. Samotní autoři v novějších studiích uznávají, že se strukturu nedaří empiricky ověřit a původní faktory nazývají *oblastmi zájmu* (Mayer et al., 2016). Emoční inteligence je ve všech jejich studiích ověřována testovou baterií *Mayer-Salovey-Caruso Emotional Intelligence Test* (dále MSCEIT; Mayer et al., 2003). Bylo by vhodné teorii ověřovat také pomocí jiných testů, například těch vyvíjených týmem Invenio. Samotný MSCEIT obecně není zcela optimálním měřítkem z psychometrických důvodů, často jsou například v subtestech faktory syceny pouze dvěma položkami (Schneider & McGrew, 2018).

## 1.3 Schopnost rozpoznání emoce z intonace

Na schopnost rozpoznat emoci z intonace (dále REI) lze nahlížet pomocí Brunswikova modelu dyadického sociálního vnímání (Brunswik, 1956). Podle Morningstar et al. (2018) je možné oddělit dvě fáze, a to kódování (pocházející od člověka, který produkuje emocionálně zabarvené vokalizace) a dekódování, které se odehrává na straně příjemce. Na straně producenta do výsledné podoby vokalizace vstupují jak nevědomé fyziologické reakce organismu, tak sociální a naučené faktory. Intonaci lze hodnotit podle tří základních měřítek, a to výšky, intenzity a časové souslednosti. Co se týče strany příjemce, zdá se, že vokální vyjádření některých základních emocí je univerzální, u dalších už záleží na kontextu.

Podle Ekmana a Davidsona (1994) existují biologicky dané, základní emoce. Takové emoce jsou vrozené, univerzální, nastávající za stejných podmínek pro všechny a vyvolávající jasně oddělitelnou fyziologickou reakci. Jedná se o radost, smutek, strach, vztek a znechucení, někteří autoři k nim dále řadí překvapení, pohrdání a stud. Oproti tomu existují také sekundární emoce, které z primárních vycházejí. Předpokládá se, že základní emoce jsou velkými „rodinami“, do kterých spadají jednotlivé sekundární emoce, z toho důvodu jejich rozpoznání vývojově předchází rozpoznání sekundárních emocí (Reeve, 2018).

Je potřeba odlišit moment, kdy dítě emoce dokáže rozpoznat, a moment, kdy je začne pojmenovávat. Typicky u dětí pojmenovávání začíná ve věku dvou let (Izard & Harris, 1995). Od té doby se velikost emočního slovníku násobně zvětšuje a roste až do období pozdní adolescence (Nook et al., 2020).

Subtest *Líná Džina* pracuje právě s Ekmanovými primárními emocemi (Ekman & Davidson, 1994), se kterými rovněž operovali Salovey a Mayer (Maul, 2012). K těmto emocím jsme přidali ještě hanbu (která je velmi podobná studu) a hrdost. Jedná se o tzv. sebevztažné emoce (*self-conscious*), které by měly souviset s rozvojem sebeuvědomění (Saarni et al., 2007). Podle Lewise (1993) se sebevztažné emoce začínají objevovat už mezi 15 a 18 měsíci a jejich rozvoj dále pokračuje spolu s tím, jak se dítě dostává do sociálních situací. Pro ověřovaný subtest jsou relevantní zejména z toho hlediska, že sociální situace, v nichž jsou tyto emoce běžné, jsou v subtestu užity jako podnětový materiál.

### 1.3.1 Vývoj schopnosti rozpoznání emoce z intonace

Schopnost rozpoznání emocí z nápovědí v hlase se zlepšuje v průběhu ontogeneze. Podle Grossman et al. (2010) mají lidé specializované oblasti v superiorním temporálním sulcu, jejichž funkcí je zpracování hlasových vjemů a také detekce emoční prozodie. Schmirer a Adolphs (2017) tvrdí, že sice nejde o oblasti výhradně určené k detekci emocí v hlase (aktivují se také například při poslechu hudby), ale na prozodii hlasu reagují silněji než na jiné zvukové obsahy. Okolo sedmého měsíce už je mozek dítěte vyvinutý natolik, že na detekci vokálních stimulů reaguje stejně jako mozek dospělého. Není zcela jasné, zda jsou procesy následující po detekci totožné u dětí a dospělých, vzhledem k překvapivé absenci studií na toto téma (Morningstar et al., 2018). Emoční prozodie je každopádně důležitá u kojenců, protože má primát před verbálním obsahem, kterému se kojenci teprve učí rozumět (Aguert et al., 2013).

V průběhu předškolního věku však naopak dochází k tzv. lexikálnímu zkreslení, kdy děti preferují lexikální nebo situační nápovědi na úkor prozodických. Tímto jevem se zabýval experiment Morton a Trehub (2001), ve kterém byly participantům v mladším školním věku a mladým dospělým pouštěny nahrávky, které byly intonací v konfliktu s obsahem řečeného. Děti mladší 8 let daly spíše na verbální obsah, děti ve věku 9 a 10 let neodpovídaly jednoznačně a u dospělých převažovalo spolehnutí se na prozodický obsah. Covic et al. (2021) argumentují, že se jedná o důsledek omezené kapacity vnímání, kdy předškolní děti a děti v mladším školním věku tíhnou spíše ke zpracování verbálních informací na úkor inference z prozodie, která je relativně komplexnější. Rozpoznání emoce z vokálních stimulů obecně není zcela jednoduché, na což poukazuje také fakt, že se u dětí schopnost rozpoznat emoci z intonace obecně vyvíjí pomaleji než schopnost rozpoznat emoce z výrazu obličeje (Chronaki et al., 2014). Morningstar et al. (2018) zároveň tvrdí, že vyšší obtížnost souvisí s tím, že se hlasem dají efektivněji komunikovat komplexnější emoce, které výrazem v obličeji nelze tak snadno zprostředkovat. V mezikulturním srovnání se pak ukazuje, že lexikální zkreslení je silnější a déle trvající u dětí ze západních kultur oproti dětem japonským, což lze připsat sociálním faktorům, kdy obecně ve východních kulturách je na paralingvistické nápovědy kladen větší důraz a děti se s nimi tedy učí pracovat dříve (Ikeda et al., 2021).

Sauter et al. (2013) poukázali na to, že se schopnost dekódování prozodických nápovědí liší podle typu stimulu, který je vokálně prezentován. Ve svém experimentu prezentovali dětem buď stimuly v podobě

citoslovcí, nebo v podobě trojmístného čísla řečeného s emoční valencí. Ukázalo se, že děti kategorizovaly emoce z citoslovcí bez zásadních chyb už v pěti letech, zatímco u čísel chybovaly více. U starších dětí pak tento rozdíl nebyl tolik patrný.

## 1.4 Existující měřící nástroje

Podobně jako existuje dělení EI na výkonnostní a smíšené modely, existují také dva způsoby, jak EI měřit, a to sice pomocí výkonnostního či rysového přístupu. Výkonnostní měřítka používají úkoly, které jsou podobné tradičním inteligenčním testům, ve kterých je snahou dostat informaci o maximálním výkonu participanta (Fiori et al., 2014). Dotazníky měřící EI jako rys se zaměřují spíše na typický výkon a mívají charakter self-reportového dotazníku. Tato distinkce je nejčastější (O'Connor et al., 2019). Všechny testy ze systému Invenio jsou založeny na výkonnostním principu, totožný přístup tedy zaujímáme k měření schopnosti rozpoznání emocí. Mezi výhody výkonnostního přístupu patří nižší interference sociálně žádoucího odpovídání, stejně jako fakt, že participant nemusí zhodnocovat své vlastní schopnosti, které často máme tendenci hodnotit neadekvátně (Rivers et al., 2012). Na druhou stranu se nabízí otázka ekologické validity výkonnostních testů, protože často obsahují úkoly, které v běžných sociálních interakcích nemusíme řešit (Odukoya et al., 2020). Zároveň je v nich vyjádření emocí často velmi expresivní, což během běžných reálných interakcí být nemusí (Morningstar et al., 2018).

V této sekci se zaměřím na výkonnostní testy, které by spadaly či přímo spadají do první větve Salovey-Mayerova modelu, tedy rozpoznání emocí s přihlédnutím k rozpoznání emocí z intonace, zejména u dětí. Počet testů, který byl dosud vyvinut, je vzhledem k popularitě konstruktu překvapivě malý (Schlegel & Scherer, 2015), proto je seznam téměř vyčerpávající.

### 1.4.1 Přehled nástrojů používaných k měření EI

Již zmíněný MSCEIT (Mayer et al., 2003) je patrně nejrozšířenějším testem měřícím EI. Obsahuje subtesty zaměřené na každou ze čtyř větví původního modelu (viz podkapitola Salovey-Mayerův model EI), nad rámec

původního modelu je dělí na *zkušenostní* (obsahující první dvě větve Vnímání emocí a Používání emocí k řízení myšlení) a *strategickou* (Porozumění emocí a Řízení emocí). Co se týká první větve vnímání emocí, obsahuje pouze dva subtesty, a to identifikaci emoce z výrazu obličeje a z obrázků. Subtest s obličeji má takovou podobu, že participant vidí fotografii a má popsat rozsah a hloubku v obrázku pro každou emoci zvlášť. Test pracuje s pěti emocemi, a to štěstím, strachem, překvapením, znechucením a nadšením. Subškála s obrázky pak ukazuje určitou krajinu nebo vzor a participant posuzuje, jakou míru emoce vyjadřuje každý obrázek pomocí výrazu kreslené tváře (Odukoya et al., 2020).

Jedním z problémů, kterým měření EI podle MSCEIT čelí, je otázka vyhodnocování. V zásadě existují dva způsoby, jak určit, která odpověď je správná. První je pomocí konsenzu, kdy je participantovi uděleno skóre podle toho, kolik procent dalších lidí zvolilo stejnou odpověď. Potom ale paradoxně nejvíce zvyšují celkový skór jednoduché položky, na které odpovídají všichni stejně. Druhou možností je pak posouzení správné odpovědi skupinou expertů. Výsledky získané oběma způsoby však dávají relativně podobné výsledky (Fiori et al., 2014).

Specifikem testu MSCEIT je to, že na rozdíl od většiny ostatních testů měřících rozpoznání emoce (dále RE) byla vytvořena také dětská verze, tzv. MSCEIT-YV (Mayer et al., 2012) pro děti od 10 do 17 let. Rozdíl oproti testům pro dospělé však tkví pouze v tom, že obsahuje test s fotografiemi dětských obličejů místo dospělých (Rivers et al., 2012). MSCEIT je tak sice nejhojněji používaným nástrojem měřícím EI, ale u schopnosti vnímání emocí nabízí pouze dva subtesty a mimo vizuálních informací se nezaměřuje na další kanály, pomocí nichž lidé komunikují emočně nabité informace. Existují však jiné testy, které měří také schopnost rozpoznání emocí komunikovaných intonací, postojem či pohybem.

Nejstarší z těchto nástrojů je PONS (Rosenthal et al., 1979), který vychází z konstruktu interpersonální senzitivity (Bernieri, 2001). Test vznikl ještě předtím, než byl vůbec pojem EI používán. Jedná se 47 minut dlouhé video obsahující 220 položek, kde testovaný vybírá ze dvou popisů situace ten, který je blíž realitě toho, co vidí nebo slyší. Položky jsou buď pouze audio, obrazové či video, nebo se vyskytují v různých kombinacích. Audionahrávky byly upraveny dvěma způsoby, buď je prezentovaná náhodně sestříhaná a seskládaná řeč (čímž se však ztrácí informace o tom, jak hlas v čase stoupá a klesá), nebo je hlas elektronicky upraven tak, aby nebyl srozumitelný. Úkolem participanta je určit, zda se jedná o pozitivní nebo negativní afekt a dominantní versus submisivní pozici.

K dříve vyvinutým testům řadíme též test DANVA (Baum & Nowicki, 1998). Receptivní část testu DANVA obsahuje měření RE z výrazů obličeje, hlasu a držení těla (pro výrazy obličeje a intonace existuje také v dětské verzi). Test DANVA2-CP (Rothman & Nowicki, 2004) obsahuje nahrávky dětských herců, kteří říkají obsahově neutrální větu v různých intonacích. Testovaný vybírá emoci z možností radost, smutek, naštvání a strach.

Dalším testem je MERT (Bäzinger et al., 2009). Jedná se o multimodální test. Úkolem participanta není určit, jakou emoci osoba vyjadřuje, jako spíš to, jaká emoce je reprezentována (to je podle autorů to jediné, co lze zjistit z dekontextualizovaných projevů). Celkem je k rozpoznání deset emocí (včetně některých sekundárních) s tím, že každá emoce je zprostředkovaná ve čtyřech módech (jen video, jen audio, audio + video, fotografie). Používány jsou věty, které sice znějí jako němčina, nenesou však sémantický význam.

Velmi podobným testem je GERT (Schlegel & Scherer, 2015), který je založený pouze na multimodálním vyjádření emocí. Obsahuje 83 krátkých videoklipů, ve kterých deset herců předvádí 14 emocí, z nichž je šest pozitivních. Opět vyslovují slova, která nedávají věcný smysl. Specifikem je také to, že byl vytvořen na základě Raschova modelu (Schlegel & Scherer, 2015).

Test ERI vytvořili Scherer a Scherer (2011) jako měřítko konstruktu, který nazývají emoční kompetence, tedy *„jak dobře fungují emoce vzhledem k jejich evoluční funkci a strategickým cílům v kulturně definovaných sociálních situacích*“ (p. 306). Příslušná součást testu se nazývá Vocal-I a zachycuje expresi nějaké věty bez významu (podobně jako v případě MERT) v čtyřech emocích (smutku, štěstí, strachu, vzteku). Úkolem participanta pak je vybrat jednu z těchto emocí.

Všechny zmíněné testy (ERI, PONS, MERT, GERT a DANVA) výzkumníci často administrují souběžně pro ověření konstruktové validity a testy mezi sebou vykazují korelace od $r = 0{,}3$ (Scherer & Scherer, 2011) do $r = 0{,}5$ (Schlegel & Scherer, 2015).

Další kategorií testů, které mají podobný charakter, avšak nejsou určeny k testování na běžné populaci, jsou testy určené k zachycení deficitu ve vnímání emocí ve spojení s nějakou diagnózou. Příkladem je *Reading the Mind in the Voice* test (RMV; Golan et al., 2006), který se používá k identifikaci sociálně-emočních deficitů u participantů s vysokofunkčním autismem. Podle autorů je absence rozpoznání emocí u druhých důsledkem deficientní teorie mysli u lidí s autismem (Baron-Cohen et al.,

2001). Tento test je méně známou alternativou testu *Reading the Mind in the Eyes* (Baron-Cohen et al., 2001). Participantům je přehrána nahrávka obsahově neutrální, emočně zabarvené věty, vystřižené z dramatu BBC, a za úkol má vybrat ze čtyř název emoce, kterou nahrávka nejvíce evokuje. V podstatně podobným testem určeným dětské populaci je baterie The Cambridge Face-Voice Battery for Children (Cam-C; Golan et al, 2015). V něm dětští účastníci (pilotáž proběhla na vzorku 8 až 11 let starých) mají za úkol zvolit název nejvíce odpovídající emoce na obsahově neutrální větu.

Současný stav měření emoční inteligence, přesněji pak vnímání emocí, nabízí relativně homogenní výčet testů, z nichž jsou některé adaptovány na dětskou populaci, valná většina však ne. Některé metody afekty, které mají být rozpoznány, dichotomizují (např. metoda Aguert et al. 2013), případně rozdělují na tři kategorie, např. ve studii Covic et al. (2021) měly děti rozeznávat mezi třemi možnostmi (radost, smutek a naštvání). Takové testy jsou sice jednodušší pro participanta, ale nevypovídají mnoho o schopnosti rozpoznání konkrétní emoce jako spíš o rozpoznání emoční valence. Naopak například Bäzinger et al. (2009) v MERT pracují s celkem osmi emocemi (rozdělují např. horký a chladný hněv) na rozpoznání se zdůvodněním, že pouze dělením na mnoho emocí lze identifikovat i participanty s vysokou mírou schopnosti. Jednotlivé metody se dále liší v tom, zda pracují s verbálními stimuly, či nějakým způsobem řeč převádí do nesrozumitelné podoby. V případě všech testů však platí předpoklad, že participant zná název konkrétní emoce a dostatečně rozumí významu dané emoce. Existují dva důvody, proč je tento předpoklad problematický. Prvním je to, že test neměří pouze rozpoznání emoce, ale také schopnost ji pojmenovat. Druhou otázkou je pak vývojové hledisko, protože slovní zásoba v mladším školním věku je menší než u dospělého člověka (Grosse et al., 2021). Dítě může rozpoznat to, co druhý cítí, skrze implicitní mechanismy, které nutně nevyžadují znalost daného slova.

## 1.5 Vztah s verbální inteligencí

Verbální schopnost patří v rámci CHC teorie do široké schopnosti porozumění-znalosti *(comprehension-knowledge;* Gc*).* Jedná se tedy o součást krystalizované inteligence. Řadíme sem jak deklarativní (znalost kon-

ceptů), tak procedurální (vědět, jak něco provést) znalosti, které získáváme učením. Jde tedy o schopnost porozumění a komunikace kulturně ceněných znalostí. Lidé s lepšími jazykovými schopnostmi pak obvykle mají lepší faktické znalosti (Schneider & McGrew, 2018). Někteří autoři se dokonce přiklánějí k takové konceptualizaci, že je *Gc* a verbální schopnost prakticky totožný konstrukt (Kan et al., 2011).

Verbální inteligence či schopnost je často skloňovaným pojmem ve spojitosti s emoční inteligencí. Všechna měřítka REI jsou do určité míry verbální (Rivers et al., 2012), což často bývá součástí jejich kritiky (Maul, 2012). Zeidner et al. (2001) argumentují, že značná část emočních a sociálních znalostí může být implicitní a obtížná k verbalizaci. Tím spíše v případě dětí, které se verbální pojmenování pro některé komplexnější emoce (jako je například stud) obvykle učí až okolo sedmi let (Baron-Cohen et al., 2010).

Mayer et al. (1999) tvrdí, že jsou nalézané korelace mezi testem MSCEIT a skóry na škálách verbální schopnosti odůvodnitelné tím, že podle Spearmanova *law of positive manifold* by spolu měly všechny druhy inteligence korelovat. Nejde o zcela uspokojivé vysvětlení, neboť například ve studii Golan et al. (2015) spolu středně silně souvisela pouze schopnost rozpoznat emoci z hlasu a verbální škály ($r$ = 0,42; $p$ <0,01), nikoliv schopnost rozpoznat emoci z výrazu obličeje, takže je nutné se zamýšlet také nad tím, že se souvislost liší podle toho, na kterou součást emoční inteligence se zaměřujeme.

Podle Baron-Cohen et al. (2010) je jedním z důvodů, proč neexistuje mnoho studií měřících rozpoznání komplexních emocí z hlasu, právě neoddělitelnost těchto stimulů od jazyka. Obecně nalézané korelace bývají různé, podle Schneidera a McGrewa (2018) se pohybují od $r$ = 0,2 do $r$ = 0,6, přičemž nejsilnější vztah typicky nacházíme mezi testy měřícími porozumění emocím a měřítky krystalizované inteligence.

Subtest *Líná Džina* je konstruovaný tak, aby se nutnosti verbálně pojmenovat emoci vyhnul. Zároveň částečně vyžaduje dovednosti z větve porozumění emocí Salovey-Mayerova modelu. Díky tomu by mohl přinášet nový pohled na problematiku testů rozpoznání emoce z intonace, vzhledem k absenci nutnosti pojmenování by totiž mohl s verbální schopností souviset méně.

## 1.6 Specifika testování dětí

Specifikem našeho měření je věková skupina, pro níž je subtest *Líná Džina* určen, tedy děti prvního stupně základních škol. U dětí se často setkáváme s testovou úzkostí, což je stav, který je spojen se sníženou výkonností dětí během testování (Mavilidi et al., 2013). Podobně pak musíme brát v úvahu menší velikost dětského slovníku, která je také v případě emocí menší než slovník dospělého člověka (Grosse et al., 2021).

### 1.6.1 Specifika počítačového testování dětí

Důležitým faktorem ke zvážení je specifikum testování dětí pomocí počítače. Podle Barnes (2015) je úspěch dítěte v kognitivním testu do určité míry závislý na tom, zda má dítě přístup k počítači v domácím prostředí. Naopak je třeba zvážit i to, že pokud se dítě doma k počítači dostane, je to obvykle v kontextu volnočasové aktivity, proto může pro dítě být nezvyklé na počítači vyplňovat test, na který je zapotřebí se soustředit a dokončit ho.

Zároveň je také otázkou míra pozornosti, jíž je dítě schopno úkolu věnovat. Mladší děti mají obecně kratší časové úseky, v nichž jsou schopny udržovat pozornost. Podle experimentu Carpenter et al. (2018) pak mají děti ve věku 4 až 11 let při kontrole jiných proměnných obecně horší výkon v testech měřících pracovní paměť (která je na pozornosti závislá), pokud jsou administrovány na počítači oproti papírové formě. S pozorností se pojí také již zmiňovaná testová úzkost, která může vést k omezení kapacity pracovní paměti (Mavilidi et al., 2021).

Počítačové testování však skýtá mnoho výhod, jako je možnost administrace na dálku či využití metod, které by při hromadné papírové administraci byly náročné či téměř nemožné k provedení (Širůček, 2010). Tuto výhodu uplatňujeme také v případě subtestu *Líná Džina*, ve kterém jsou odpovědi ve formátu audionahrávek emočně zabarvených vyjádření.

## 1.7 Cíle studie

Cílem studie je ověření explorační ověření fungování nové metody měřící schopnost rozpoznání emoce z intonace. Metoda s názvem *Líná Džina* vznikla v rámci screeningového systému Invenio v kontextu CHC teorie

inteligence. Administrujeme ji dětem navštěvujícím 2. až 5. ročník základních škol v České republice, přičemž očekáváme, že úroveň schopnosti roste s věkem (a tedy navštěvovaným ročníkem) participantů. Předpokládáme zároveň, že má subtest jednodimenzionální strukturu.

V rámci pilotáže bude administrována spolu s metodou *Věřte, nevěřte!,* která by měla měřit verbální schopnost. Účelem souběžné administrace je ověření diferenciální validity, přičemž očekáváme výslednou korelaci na spodní hranici ve srovnání s jinými metodami, přesněji korelaci okolo $r = 0{,}2$.

# 2 Metoda

## 2.1 Líná Džina

Subtest *Líná Džina* je součástí testové baterie *Frída a podivuhodná galerie* (Portešová et al., 2022). Jedná se o test podaný formou počítačové hry. Participant dostává postupně 20 kreslených scének v podobě krátkých videí, které vždy vyústí v nějakou sociální situaci vyžadující reakci, jíž vypravěč uvede *"(protagonista) říká:..",* načež je úkolem participanta si přehrát tři možnosti, s jakou intonací mohla být obsahově neutrální věta pronesena. K dispozici má možnosti, které se navzájem liší jenom emočním zabarvením. Ukázková položka subtestu je dostupná pod tímto odkazem[1]. Autorka práce se na konceptualizaci a tvorbě jednotlivých scének nepodílela, do procesu vývoje se zapojila až při finální volbě distraktorů pro některé položky.

## 2.2 Věřte, nevěřte!

Metoda s pracovním názvem *Věřte, nevěřte!* vznikla na Katedře psychologie FSS MUNI v rámci diplomového projektu Mgr. Kláry Hečlové (Hečlová, 2020). K použití metody nám autorka dala svolení. Jedná se o metodu založenou na principu verbálních absurdit, ve které má participant za úkol posuzovat smysluplnost předloženého výroku (má na výběr pouze dvě možnosti). Příkladem použitého smysluplného výroku může být věta *„Tatínek si v zimě na horách obléká pod kalhoty ještě podvlíkačky"*, příkladem absurdity věta *„Margarín je blyštivý modrý kámen".*

Položky do této verze testu jsme vybírali z celkového počtu 54 položek, které prošly pilotáží v roce 2020. Vybrali jsme celkově 15 položek zastupující oba principy (smysluplné věty a nesmysly). Stejný počet jsme zvolili z toho důvodu, aby nebyl nepřiměřeně zastoupen jeden princip. Celkový počet 30 jsme zvolili, aby byl test dostatečně dlouhý vzhledem k jeho struktuře, protože dichotomický formát odpovědi znamená 50% šanci, že si participant správnou odpověď tipne. Položky jsme vybírali na základě obtížností a korelací, obdržených z položkové analýzy dat z března 2020.



[1]https://masaryk.eu.qualtrics.com/jfe/form/SV_9oepNDF9ApBLGXc

V původní verzi byl v roce 2020 test administrován s dichotomickou odpověďovou možností *pravda* či *lež*, ale my jsme po konzultaci s vedoucím původní práce, Mgr. Michalem Jabůrkem, Ph.D., změnili formát odpovědí na výroky *ano, to dává smysl* a *ne, to je nesmysl.* Změnu jsme provedli z toho důvodu, že pravda a lež jsou abstraktní slova, jejichž význam některé děti na prvním stupni nemusí chápat. Vycházeli jsme také z pilotáží provedených Mgr. Klárou Hečlovou, z nichž vyplývalo, že je pro děti v případě možností pravda či lež obtížné zhodnotit, zda mají posuzovat svou osobní zkušenost, či formulaci na základě obecné platnosti (Hečlová, 2020). Mimo psané verze výroků participanti měli k dispozici nahrávky jednotlivých vět. Audionahrávky mohly pomoci zejména mladším participantům, pro které by množství vět ke čtení mohlo být překážkou.

## 2.3 Procedura

Před samotným testováním byly provedeny dvě pilotáže, pokaždé se dvěma dětmi pocházejícími z cílové populace. Smyslem první pilotáže bylo zhodnotit časovou náročnost testu *Lína Džina* a také zkontrolovat, zda děti rozumí zadání testu. Cílem druhé pilotáže pak bylo zhodnotit časovou náročnost při administraci s metodou *Věřte, nevěřte!* a srozumitelnost obou testů.

Kromě těchto pilotáží jsem (autorka práce) provedla s celkem 6 dětmi ve věku 7 let (3 děti), 8, 9 a 11 let kognitivní pilotáže. V nich jsem s každým procházela individuálně celý subtest *Líná Džina*, po každé přehrané scénce jsem se ptala, jakou emoci pravděpodobně situace v dotčené postavě vyvolala *(„Jak se teď cítí? Jak by asi mohl/a reagovat?“*) za účelem zjištění, jak jasné je vyznění položky. Následně jsem jim přehrávala po jedné možnosti a ptala se, jaký pocit či tón hlasu v dané nahrávce slyší *(„Jak ti zní tohle? Jaký má ten člověk asi pocit?“).* Snažila jsem se přitom vyhnout nutnosti emoci nějak pojmenovat, děti mohly situaci či emoci popsat svými slovy. Děti jsme za účast v pilotáži odměňovala čokoládou.

V kvantitativní části jsme položky zadávali pomocí dotazníku na platformě *Qualtrics*. Jako první dostali participanti demografické otázky (osobní kód, kód školy, třídu a věk). Následně si všichni nasadili sluchátka a dostali k poslechu zvukovou stopu s instrukcemi k části. Všichni nejprve vyplňovali v náhodném pořadí položky subtestu *Líná Džina*, a to

z toho důvodu, že byl primárním zájmem tohoto výzkumu a v rámci toho jsme chtěli eliminovat možné efekty únavy, které by se mohly dostavit, pokud by ho vyplňovali až jako druhý v pořadí. Náhodné pořadí položek jsme zvolili z toho důvodu, aby se děti navzájem při testování co nejméně vyrušovaly společnou snahou najít správnou odpověď. Po první části následovaly nahrané instrukce ke druhému bloku a otázky z testu *Věřte, nevěřte!,* opět v náhodném pořadí. Na konci dotazníku bylo vyhrazené místo pro poznámky. Zadávající učitelé v instrukcích dostali pokyn, aby sem uváděli skutečnost, že se jedná o dítě cizinců (tj. čeština není jeho mateřský jazyk) či má dané dítě nějaké poruchy učení či chování.

V testu bylo možné se k otázkám vracet zpátky a zároveň byl dotazník nastaven tak, že nebylo možné položku nechat bez odpovědi. K tomuto kroku jsme přistoupili z toho důvodu, že zejména mladší děti nemusí být v ovládání počítače zcela zdatné a mohly by mezi položkami přeskakovat, aniž by si všimly, že nezvolily žádnou odpověď.

## 2.4 Tvorba vzorku

Zvolenou populací jsou děti navštěvující druhý až pátý ročník základních škol v České republice. Důvodů k volbě takového rozmezí jsme měli více. Testy systému Invenio jsou navrženy pro děti od 1. do 6. třídy základních škol. Účast prvních tříd jsme vyloučili z toho důvodu, že test *Věřte, nevěřte!* obsahuje větší množství textu na čtení, který by pro žáky prvních tříd mohl být zahlcující. Zároveň jsme předpokládali, že pro děti bude subtest *Líná Džina* spíše jednodušší, z toho důvodu jsme se rozhodli vzorek omezit na první stupeň ZŠ.

Co se týče plánované velikosti vzorku, kritérií pro její stanovení jsme měli více. Zaprvé se opírá o výsledky power analýzy, která byla provedena v programu G*Power (Faul et al., 2007). Pro odhalení plánované slabší korelace ($r$ = 0,20) při síle testu β = 0,80 program doporučil velikost vzorku aspoň 314 respondentů. Druhým faktorem hrajícím roli je zamýšlené použití IRT analýzy. DeMars (2010) uvádí, že pro test s 20 položkami při použití 2PL modelu (viz kapitola *Analytický plán*) je potřebný vzorek 250-500 respondentů, plánovanou velikost vzorku jsme tak stanovili na 300-500 respondentů, podle finálního zájmu škol.

Celkově jsme s nabídkou testování oslovili 49 ředitelů a školních psychologů ze základních škol po celé České republice. Při výběru škol hrála roli úvaha nad tím, aby byly zastoupeny děti z větších i menších

obcí. Důležité se nám to zdálo s ohledem na to, že některá slova použitá v subtestu *Věřte, nevěřte!* (např. *rýč*) by mohla být známější určité skupině dětí a výsledek systematicky zkreslovat.

Do testování se zapojilo 6 škol, které o něj projevily zájem: ZŠ Chaplinovo náměstí v Praze 5, ZŠ Moravský Písek, ZŠ Dobřichovice, ZŠ Křídlovická Brno, Krakonošova ZŠ a MŠ Loukov a ZŠ Všenory. Těmto školám jsme zasílali odkaz na test na platformě *Qualtrics*, pokyny k administraci a informovaný souhlas k vyplnění zákonnými zástupci. Sběr dat probíhal od 12. 12. do 21. 12. 2022 a po vánočních prázdninách 11. a 12. 1. 2023, autorka práce byla osobně přítomná u jednoho sběru (na ZŠ Chaplinovo náměstí). Školám jsme za účast slíbili zpětnou vazbu v podobě průměrného výsledku třídy a kódu třech řešitelů s nejvíce body, kterou jsme dodali v průběhu února 2023.

## 2.5 Popis výzkumného souboru

Finální vzorek čítal 462 respondentů ze 2. až 5. tříd základních škol (viz kapitola *Čištění dat*).

**Tabulka 2.** Počty respondentů podle škol a ročníků

| **Škola/ročník** | **2** | **3** | **4** | **5** | **Σ** |
|---|---|---|---|---|---|
| ZŠ Chaplinovo náměstí | 34 | 34 | 28 | 30 | 126 |
| ZŠ Moravský Písek | 0 | 0 | 15 | 15 | 30 |
| ZŠ Dobřichovice | 0 | 0 | 48 | 78 | 126 |
| ZŠ Křídlovická | 0 | 0 | 56 | 62 | 118 |
| ZŠ Všenory | 0 | 0 | 16 | 21 | 37 |
| ZŠ Loukov | 6 | 5 | 6 | 8 | 25 |
| Σ | 40 | 39 | 169 | 214 | 462 |

Nedopatřením z finální verze dotazníku vypadla položka týkající se pohlaví, proto nám tento údaj k popisu vzorku chybí. Na rovnoměrné genderové složení vzorku můžeme pouze usuzovat z toho, že běžné třídy na základní škole jsou složeny z přibližně rovného dílu dívek a chlapců a testování se účastnily vždy celé třídy.

Ve vzorku jsou více zastoupeny 4. a 5. ročníky, což souvisí s ochotou škol vyhradit na testování hodiny informatiky, kterou mají žáci v rozvrhu až ve vyšších ročnících.

## 2.6 Analytický plán

Získaná data budu[2] analyzovat v paradigmatech klasické testové teorie (dále CTT) a také teorie odpovědi na položku (dále IRT).

Nejprve zjistím popisné statistiky položek subtestu *Líná Džina* a provedu položkovou analýzu, kde se zaměřím na to, jaké možnosti byly s jakou frekvencí voleny. Reliabilitu subtestu budu odhadovat pomocí koeficientu Cronbachova alfa a McDonaldova omega.

Následně provedu analýzu obou testů vycházející z IRT přístupu. Vyzkouším dvě varianty IRT modelů, konkrétně 1PL a 2PL model. Budu pracovat s tím modelem, který bude data lépe popisovat. V rámci IRT analýzy budu zjišťovat vlastnosti jednotlivých položek, jejich fungování jako celku a IRT odhad reliability.

Výsledky finálních verzí testů *Líná Džina* a *Věřte, nevěřte!* budu korelovat. K analýze využiji software *Microsoft Excel* verze *16.69.1., SPSS* verze *28.0*, *JASP* verze *0.16.1* a *R* verze *4. 2. 1.*, konkrétně balíčky *tidyverse* (Wickham et al., 2019), *psych (*Revelle, 2022*), mirt (*Chalmers, 2012), *lavaan* (Rossel, 2012), *subscore* (Dai et al., 2022) a *semTools* (Jorgensen et al., 2022).



[2] Výsledková část práce je na rozdíl od zbytku práce psaná v první osobě, protože zatímco na vývoji testu se podílel celý tým, analýzu dat prováděla autorka práce sama.

# 3 Výsledky

## 3.1 Čištění a příprava dat

Na platformě *Qualtrics* jsem zaznamenala celkově 511 odpovědí. Po odstranění odpovědí pocházejících z pilotáží či od učitelů jsem pracovala se 480 odpověďmi.

Předem jsem stanovila několik kritérií, podle kterých budu odpovědi respondentů vyřazovat: 1) vyplnění méně než 30 % položek v testu, 2) celkový čas strávený na testu kratší než 15 minut (900 sekund), nebo naopak 3) čas delší než 120 minut (7200 sekund). K volbě těchto kritérií jsem přistoupila na základě zkušenosti s administrací. Některým dětem trvalo dokončení testu dvojnásobek průměrného času, ačkoliv jej vyplňovaly standardním způsobem.

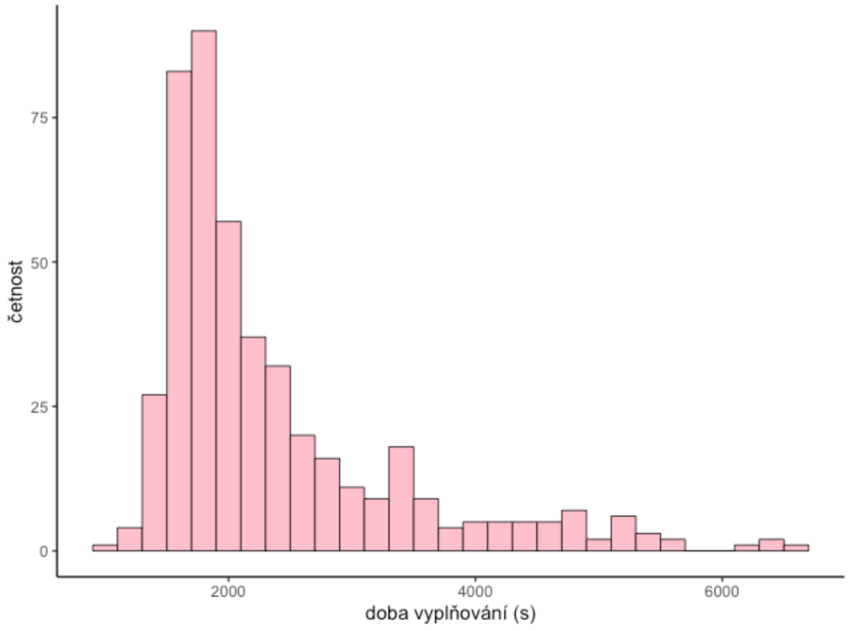
Takto jsem vyřadila data od 15 respondentů. Tři záznamy jsem vyřadila z důvodu duplicity (ponechala jsem vždy pouze jednu z totožných variant). Celkově jsem pracovala s daty od 462 respondentů. Osm respondentů nestihlo vyplnit druhou část testu, pro test *Věřte, nevěřte!* tedy pracuji s 454 participanty.

Výsledky subtestu *Lína Džina* jsem dichotomizovala na *1 = správná odpověď* a *0 = špatná odpověď*, výsledky pro *Věřte, nevěřte!* jsou dichotomické z podstaty testu, správnost odpovědí jsem rekódovala analogicky.

## 3.2 Doba vyplňování

Průměrná doba vyplňování testu byla 39,3 minuty (2326 sekund). Vzhledem k tomu, že je zešikmení rozložení času vyplňování relativně silné, informativnější v tomto kontextu může být medián, který je roven přibližně 33 minutám (1995 sekund). V potaz beru pouze kompletní odpovědi, tj. od participantů, kteří stihli vyplnit obě části testu.

**Obrázek 1.** Rozložení celkové doby trvání vyplňování testu.




**Tabulka 3.** Průměrná doba vyplňování v minutách v závislosti na věku respondenta.

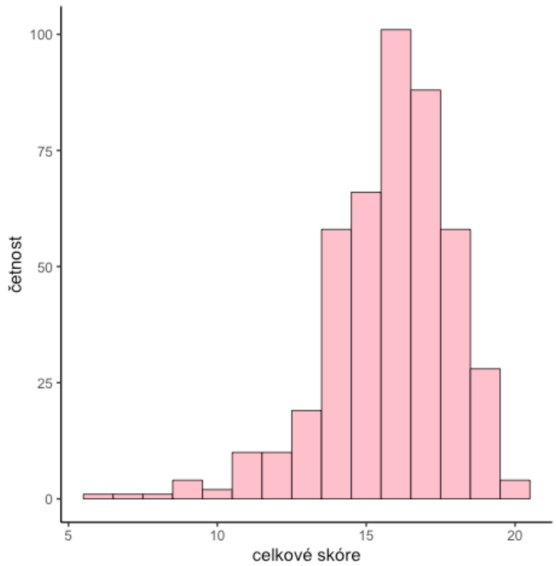
| **věk** | **7** | **8** | **9** | **10** | **11** | **12** |
|---|---|---|---|---|---|---|
| ***M*** | 51,25 | 63,94 | 42,70 | 35,60 | 32,38 | 24,81 |
| ***SD*** | 11,68 | 18,97 | 19,65 | 12,47 | 10,77 | 1,21 |

Celkově je patrná delší doba vyplňování u mladších participantů, která se s postupem věku snižuje. Důležitým faktorem zde může být to, jakou metodu participanti zvolili k vyplňování testu *Věřte, nevěřte!,* tedy zda spíše poslouchali nahrávky věty, či si je četli sami. Je také potřeba brát v úvahu odlišnou rychlost čtení u starších a mladších dětí. Z dat bohužel není možné zjistit, kdo si pouštěl audiostopy a kdo věty četl. Dalším faktorem, který vstupuje do doby vyplňování, by mohla být celková úroveň práce s počítačem, kterou u starších rovněž předpokládáme vyšší.

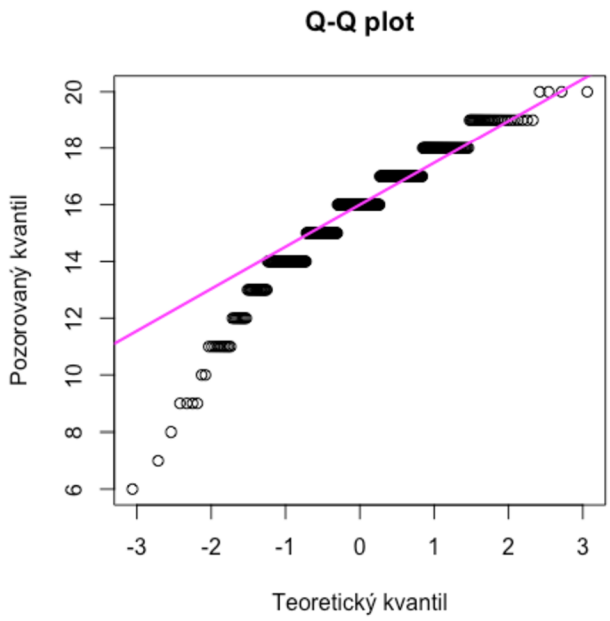
## 3.3 Deskriptivní statistiky subtestu Líná Džina

Nejprve se zaměřím na celkové skóre dosažené v testu. Mimo analýzu četností volených distraktorů výsledky prezentuji pro dichotomizovanou podobu dat. Ze samotného rozložení hodnot je patrné, že je subtest podle očekávání pro děti do značné míry jednoduchý. Celkově 4 respondenti dosáhli plného počtu bodů, nejnižších skóre 6, 7 a 8 bodů pak dosáhlo vždy po jednom respondentovi. Průměrné skóre je pak *M* = 15,8; *SD* = 2 body. Variabilita skórů je tak relativně malá. Rozložení se nepodobá normálnímu, hodnoty *šikmost* = -1 *špičatost* = 2 poukazují na leptokurtické, zleva zešikmené rozložení, což dokládá i Q-Q plot na obrázku 3.

**Obrázek 2.** Rozložení celkových skórů subtestu Líná Džina.



**Obrázek 3.** Q-Q plot celkových skórů subtestu Líná Džina.




Dále se zaměřím také na to, jak se lišilo průměrné skóre podle toho, jakou třídu participant navštěvuje, což souvisí s předpokladem, že schopnost rozpoznat emoci z intonace roste s věkem. Nejprve jsem průměry zkusila dělit podle ročníku, které respondenti navštěvují.

**Tabulka 4.** Průměry v jednotlivých ročnících.

| **ročník** | ***N*** | ***M*** | ***SD*** |
|---|---|---|---|
| 2. | 36 | 14,36 | 2,49 |
| 3. | 37 | 15,35 | 2,44 |
| 4. | 164 | 15,79 | 2,10 |
| 5. | 214 | 16,12 | 1,89 |

Při hodnocení je však třeba brát v potaz nerovnoměrné zastoupení participantů v jednotlivých ročnících, více informativní v tomto kontextu může být rozdělení podle věku participantů.

**Tabulka 5.** Průměry podle věku respondentů.

| věk | *N* | *M* | *SD* |
|---|---|---|---|
| 7 | 23 | 14,52 | 2,54 |
| 8 | 32 | 14,94 | 2,53 |
| 9 | 104 | 15,81 | 1,92 |
| 10 | 182 | 15,79 | 2,17 |
| 11 | 108 | 16,30 | 1,79 |
| 12 | 2 | 17,00 | 1,41 |

Stejně jako s ohledem na třídu, zdá se, že velmi mírnou rostoucí tendenci pozorujeme i v tomto případě, na druhou stranu opět platí problém s nevyváženostní vzorku. Průměry jsou především důkazem pro to, že i sedmileté děti průměrně dosahovaly průměrně vysokého skóre.

Zadávající učitelé dostali v instrukcích pokyn, aby do sekce určené pro poznámky pomohly dětem napsat, že mají nějakou poruchu učení či skutečnost, že jsou jeho rodiče cizinci. Bohužel se zdá, že ne všichni učitelé instrukce četli, a z toho důvodu tyto informace nemáme od všech tříd, na druhou stranu výsledky těch, u kterých tak uvedeno bylo, se nelišily od výsledků ostatních dětí.

## 3.4 Kognitivní pilotáže

Kognitivních pilotáží se účastnilo celkem 6 dětí ve věku 7 let (3 děti), 8 let, 9 let a 11 let. Jejich hodnocení jednotlivých položek bylo relativně konsenzuální, ve většině případů jim potíže působily stejné části.

Z pilotáží vyplývá, že se u položek vyskytují tři druhy problémů, které do odpovědi nějakým způsobem vstupují. První potíží je nejasné celkové vyznění scénky (není z ní zcela patrné, jakou emoci protagonista cítí), tento problém se vyskytl u položek QD8 a QD19, v obou případech by reakcí mohl opodstatněně být *strach* či *vztek*. Druhým úskalím pak je nejednoznačnost mezi tím, jak se postava pravděpodobně cítí a jak se tváří či s kým je zrovna ve scénce vyobrazená. Tento problém se vyskytl u položek QD4 a QD5. Poslední nejasností pak byla nejasnost distraktorů, které u dětí nedostatečně jasně evokovaly emoci, kterou měly. Problém se vyskytl u položek QD18, QD19 a QD20.

## 3.5 CTT analýza subtestu Líná Džina

### 3.5.1 Položková analýza

V tabulce 6 jsou uvedeny popisné statistiky jednotlivých položek subtestu *Líná Džina*.

**Tabulka 6.** Popisné statistiky jednotlivých položek subtestu Lína Džina.

| položka | *M* | *SD* | *korig. r* | položka | *M* | *SD* | *korig. r* |
|---|---|---|---|---|---|---|---|
| **QD1** | 0,80 | 0,40 | 0,33 | **QD11** | 0,55 | 0,50 | 0,23 |
| **QD2** | 0,97 | 0,16 | 0,49 | **QD12** | 0,90 | 0,31 | 0,61 |
| **QD3** | 0,80 | 0,40 | 0,35 | **QD13** | 0,90 | 0,31 | 0,20 |
| **QD4** | 0,91 | 0,28 | 0,39 | **QD14** | 0,86 | 0,35 | 0,26 |
| **QD5** | 0,94 | 0,23 | 0,44 | **QD15** | 0,92 | 0,27 | 0,33 |
| **QD6** | 0,60 | 0,49 | 0,43 | **QD16** | 0,65 | 0,48 | 0,35 |
| **QD7** | 0,71 | 0,45 | 0,23 | **QD17** | 0,89 | 0,31 | 0,17 |
| **QD8** | 0,74 | 0,44 | 0,11 | **QD18** | 0,71 | 0,45 | 0,17 |
| **QD9** | 0,89 | 0,31 | 0,56 | **QD19** | 0,53 | 0,50 | -0,10 |
| **QD10** | 0,63 | 0,48 | 0,25 | **QD20** | 0,76 | 0,43 | 0,14 |

Na první pohled je patrné, že většina položek má vysoký podíl správných odpovědí. Nižší průměry (a tedy větší obtížnost v CTT smyslu) nacházíme u položek QD6, QD10, QD11 a QD19. Naopak u položek QD2 a QD5 téměř všichni respondenti volili stejně, což vede k nízké variabilitě odpovědí. Minimální očekávaná obtížnost je přitom 0,33, protože taková je pravděpodobnost uhádnutí správné odpovědi při třech možnostech volby v testu.

Zijlmans et al. (2019) doporučují jako spodní přijatelnou hranici korelace položky se zbytkem testu $r$ = 0,30. Pod touto hranicí se nachází vícero položek v subtestu, celkově více než polovina. Položka QD19 dokonce se škálou koreluje negativně, což poukazuje na nutnost jejího přezkoumání. Druhou nejnižší korelaci (byť aspoň pozitivní) nacházíme u položky QD8.

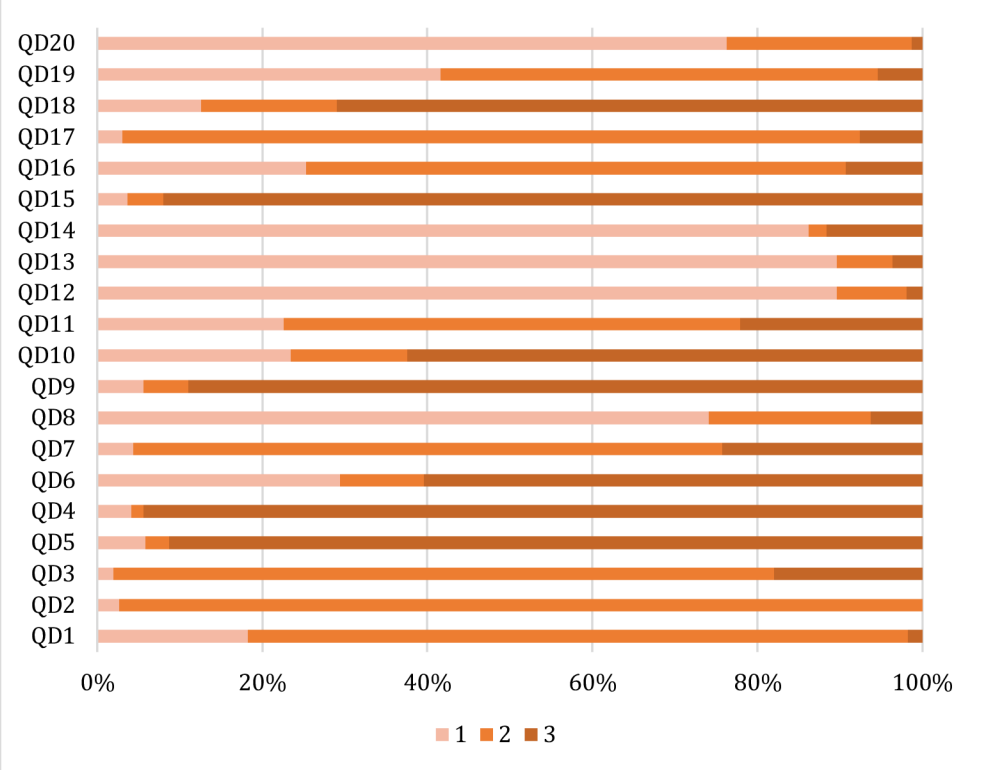
Vzhledem k dichotomické povaze výsledných dat budeme odhad reliability zakládat na tetrachorických korelacích, celou korelační matici lze najít v příloze B. Odhad Cronbachovy alfy pro subtest je α = 0,68, 95 %CI [0,43; 0,85], což je spíše nižší hodnota, kterou bychom ale pro nediagnostické účely mohli považovat za relativně přijatelnou, je však

nutné brát v potaz široký interval spolehlivosti. V testu se nachází jedna položka, jejíž vyřazení by reliabilitu škály zvýšilo, a to QD19, korelující se zbytkem škály negativně. Bez položky Q19 odhad vnitřní konzistence lehce stoupne na α = 0,70, 95 %CI [0,46; 0,86]. Odhad McDonaldovy omegy pro subtest *Lína Džina* je ω = 0,66.

### 3.5.2 Analýza distraktorů

Jednou z příčin jednoduchosti testu může být volba distraktorů a dále také jejich počet (participanti volí pouze ze tří možností), díky kterému je relativně velká šance správné odpovědi při hádání. Test byl navrhován tak, aby pro dospělého člověka byla správná odpověď evidentní, podle čehož byly také voleny distraktory.

**Obrázek 4.** Rozložení četnosti odpovědí na položky.




Z grafu je patrné, že pro některé položky byla s velkou převahou volena jedna odpověď a podíl ostatních byl zcela zanedbatelný. Za takové můžeme považovat položky QD2, QD5, QD4, QD12, QD14 a QD15 a QD17.

Dále se u některých položek zdá, jako by existoval pouze jeden distraktor, zatímco druhý jako by pro většinu ani nepřipadal v úvahu (QD1, QD3, QD4, QD12, QD14).

V této fázi je jasnější problém u položek QD19 a QD20, u kterých se zdá, že je rozložení odpovědí způsobeno spíše špatným dabingem. U QD19 byl výslednou emocí překvapení, distraktory byly neutrál a radost. Scénka vedla k takové situaci, ve které by připadal v úvahu také vztek či strach, ale druhou nejvíce volenou možností byla radost, která v nahrávce zní relativně podobně jako překvapení, ale do kontextu scénky nezapadá (výslednou emocí je překvapení v negativním smyslu). Položka QD20 měla výslednou emoci radost s distraktory strach a znechucení. Ačkoliv scénka jako taková by strach vzbuzovat neměla, výsledná intonace strachu je opět velmi podobná nahrávce intonace radost.

### 3.5.3 Dílčí závěry CTT analýzy

Z CTT analýzy subtestu plyne, že se jedná o nástroj s relativně akceptovatelnou vnitřní konzistencí, na kterém je však nutné provádět další úpravy. Konkrétně se jedná o některé problematické položky, u kterých bych doporučovala přezkoumání či volbu nových distraktorů.

Pouze málo položek obsahuje tři možnosti, mezi kterými by se dalo volit, obvykle jsou možnosti dvě (což vede k riziku toho, že při tipování bude mít respondent 50% šanci odpovědět správně). V tomto ohledu nelze opomenout položku QD2, kde jeden z distraktorů nezvolil nikdo a druhý pouhých 2,60 % respondentů. Test by měl obsahoval lehčí i těžší položky, ale v určitém bodě se položka na úkor své jednoduchosti stává nepříliš užitečnou.

Problematická je položka QD19, a to kvůli záporné korelaci se zbytkem škály. Je možné, že za zápornou korelaci může nevhodný dabing distraktorů.

Nabízí se otázka, zda je smysluplné hodnotit v testu celkově 7 položek, u nichž všichni respondenti volili velmi podobně (QD2, QD5, QD4, QD12, QD14 a QD15 a QD17), či jestli se celkově nezamyslet nad zatraktivněním distraktorů tak, aby respondenti měli u více otázek před sebou tři možnosti, nad kterými by bylo možné uvažovat.

## 3.6 IRT analýza subtestu Líná Džina

Teorie odpovědi na položku (IRT) je přístup k vývoji a hodnocení testů, který vznikl v padesátých a šedesátých letech zejména přispěním autorů M.F. Lorda a M. R. Novicka (Zanon et al., 2016). Jejich cílem bylo vytvořit přístup, který bude méně závislý na konkrétních položkách v testu a konkrétním vzorku, jako je tomu v případě CTT. V IRT je každá položka hodnocena parametry, které by měly být neměnné napříč vzorky a položkami, a každý respondent dostává svou hodnotu θ (*théta*), která odpovídá úrovni jeho schopnosti – latentní proměnné, jíž by test měl měřit (DeMars, 2010). Výhodou IRT je zároveň i možnost se hlouběji zaměřit na jednotlivé položky a hodnotit fungování testu pro respondenty s různou úrovní θ (De Ayala, 2009).

### 3.6.1 Ověření předpokladů pro IRT model

Použití tradičních IRT modelů vyžaduje dodržení tří předpokladů. Jedná se o předpoklad jednodimenzionality, lokální nezávislosti a správné specifikace modelu (DeMars, 2010).

Předpoklad jednodimenzionality jsem ověřovala za pomoci exploračního 2PL modelu. Zdá se, že jednofaktorové řešení dobře popisuje data a je smysluplné z teoretického hlediska. Nenašla jsem smysluplně interpretovatelnou vícedimenzionální strukturu testu.

Lokální nezávislost znamená, že spolu odpovědi na jakýkoliv pár dvou položek nesouvisí jinak než vlivem modelované schopnosti (de Ayala, 2009, *s.* 131-134). Posuzovala jsem za pomoci metody Yenovo Q3 (Yen, 1984) na 2PL modelu subtestu. V datech se nenachází žádný pár položek, u kterého by reziduální korelace překročila 0,20, proto i tento předpoklad považuji za splněný.

Poslednímu z předpokladů, kterým je správná specifikace modelu, se věnuji v následující podkapitole *Volba modelu*.

### 3.6.2 Volba modelu

Rozhodla jsem se vytvořit dva modely, 1PL a 2PL pro porovnání, který z nich lépe popisuje data.

**Tabulka 7.** Srovnání fitu 1PL a 2PL modelu.

| | *M2* | *df* | *p* | RMSEA | SRMSR | TLI |
|---|---|---|---|---|---|---|
| **1PL** | 340,89 | 189 | <0,001 | 0,04 | 0,07 | 0,43 |
| **2PL** | 222,94 | 152 | <0,001 | 0,03 | 0,06 | 0,70 |

Z hodnot ukazatelů shody modelu s daty se zdá, že data lépe popisuje 2PL model. Některé ukazatele poukazují na dobrou shodu 2PL modelu s daty. Konkrétně se jedná o hodnotu RMSEA = 0,03, 90 %CI [0,02; 0,04], která i s intervalem spolehlivosti poukazuje na dobrou shodu. SRMSR vyjadřuje průměrnou chybu odhadu korelace, kterou náš model implikuje oproti korelacím zjištěným v datech. Hu & Bentler (1999) označují za přijatelné hodnoty od 0 do 0,08, což opět poukazuje na uspokojivý fit 2PL modelu, v kontextu dalších ukazatelů shody se jedná o dobrou hodnotu. Naopak hodnota TLI, ačkoliv vyšší u 2PL modelu, se zdá být neuspokojivá pro oba modely. Za dobrou hodnotu bych považovala hodnotu vyšší než 0,95 (Kenny, 2020). Příčinou nízkého TLI mohou být relativně nízké vzájemné korelace mezi položkami v testu. TLI je inkrementální index fitu, což znamená, že hodnotí nárůst informace proti nulovému modelu (nulový model je takový, který předpokládá, že za odpověďmi na položky nestojí žádná společná příčina). Kenny (2020) uvádí, že informativnost TLI lze posuzovat spočítáním hodnoty RMSEA nulového modelu. Pokud je hodnota RMSEA nulového modelu nižší než 0,158, není TLI podle autorů příliš informativní, protože nulový model není ve srovnání s naším modelem tolikrát horší. V našem případě je hodnota RMSEA nulového modelu 0,05, což může vysvětlovat nízkou hodnotu TLI (matici tetrachorických korelací lze případně najít v příloze B).

### 3.6.3 Položková analýza dle IRT

V dalších analýzách pracuji s 2PL modelem.

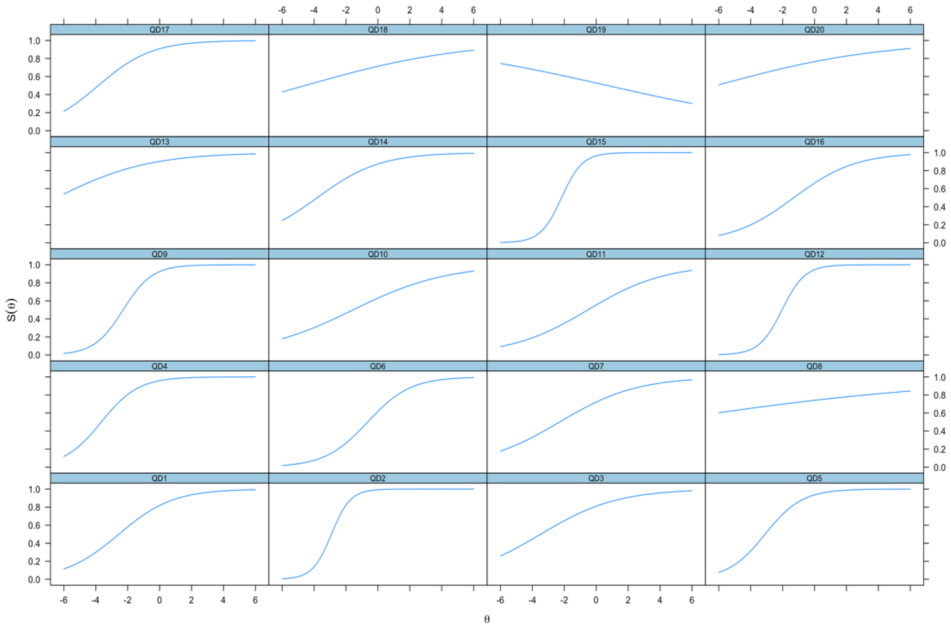
**Tabulka 8.** Vlastnosti položek v 2PL modelu.

| | *a* | $SE_a$ | *b* | $SE_b$ | χ2 | *df* | RMSEA | *p* |
|---|---|---|---|---|---|---|---|---|
| **QD1** | 0,59 | 0,19 | -2,55 | 0,76 | 4,94 | 7,00 | 0,00 | 0,67 |
| **QD2** | 1,68 | 0,54 | -2,91 | 0,57 | 1,64 | 2,00 | 0,00 | 0,44 |
| **QD3** | 0,42 | 0,20 | -3,48 | 1,56 | 6,48 | 7,00 | 0,00 | 0,49 |
| **QD4** | 0,86 | 0,30 | -3,66 | 1,07 | 3,02 | 6,00 | 0,00 | 0,81 |
| **QD5** | 0,86 | 0,27 | -3,10 | 0,79 | 4,92 | 7,00 | 0,00 | 0,67 |
| **QD6** | 0,75 | 0,20 | -0,65 | 0,20 | 9,91 | 8,00 | 0,02 | 0,27 |
| **QD7** | 0,41 | 0,18 | -2,30 | 0,94 | 10,90 | 8,00 | 0,03 | 0,21 |
| **QD8** | 0,11 | 0,16 | -9,95 | 15,17 | 7,31 | 8,00 | 0,00 | 0,50 |
| **QD9** | 1,10 | 0,29 | -2,29 | 0,45 | 6,93 | 7,00 | 0,00 | 0,44 |
| **QD10** | 0,34 | 0,16 | -1,56 | 0,74 | 2,85 | 8,00 | 0,00 | 0,94 |
| **QD11** | 0,42 | 0,18 | -0,52 | 0,31 | 10,88 | 8,00 | 0,03 | 0,21 |
| **QD12** | 1,40 | 0,34 | -2,04 | 0,33 | 5,10 | 7,00 | 0,00 | 0,65 |
| **QD13** | 0,34 | 0,24 | -6,48 | 4,32 | 7,60 | 7,00 | 0,01 | 0,37 |
| **QD14** | 0,51 | 0,21 | -3,81 | 1,47 | 6,89 | 7,00 | 0,00 | 0,44 |
| **QD15** | 1,51 | 0,38 | -2,17 | 0,35 | 3,12 | 6,00 | 0,00 | 0,79 |
| **QD16** | 0,52 | 0,17 | -1,30 | 0,43 | 5,70 | 8,00 | 0,00 | 0,68 |
| **QD17** | 0,60 | 0,23 | -3,84 | 1,34 | 3,83 | 7,00 | 0,00 | 0,80 |
| **QD18** | 0,20 | 0,16 | -4,58 | 3,62 | 6,44 | 8,00 | 0,00 | 0,60 |
| **QD19** | -0,16 | 0,14 | 0,71 | 0,86 | 10,68 | 8,00 | 0,03 | 0,22 |
| **QD20** | 0,19 | 0,17 | -6,17 | 5,32 | 7,14 | 8,00 | 0,00 | 0,52 |

Podle hodnot RMSEA se zdá, že model všechny položky popisuje uspokojivě. Co se týče parametrů obtížnosti (b), u všech položek mimo QD19 pozorujeme zápornou obtížnost, což znamená, že jsou vhodné spíše pro participanty s podprůměrnou schopností rozpoznávat emoci z intonace, protože hodnotu 0 by měla položka s průměrnou obtížností a *SD* ($\hat{\theta}$) je rovna 1. De Ayala (2009) tvrdí, že by se obtížnosti položek měly ideálně pohybovat v rozmezí od -2 do 2, k čemuž se některé položky blíží. Velmi nízkou obtížnost naopak vykazují položky QD8, QD13, QD18 a QD20. U těchto položek (a zejména pak u QD8) pozorujeme velké výběrové chyby u odhadů obtížností, což poukazuje na to, že jsou vzorce odpovědí na ně málo konzistentní s odpověďmi na jiné položky. Z CTT

analýzy vycházela jako velmi jednoduchá položka QD2, ale v IRT analýze se extrémně lehká nezdá. Naopak má nejvyšší diskriminační index ze všech položek. Je možné, že v ní opravdu chybovali pouze participanti s velmi nízkou úrovní θ a tak spolehlivě rozlišuje mezi průměrnými a silně podprůměrnými participanty.

**Obrázek 5.** Charakteristické funkce položek subtestu Líná Džina.

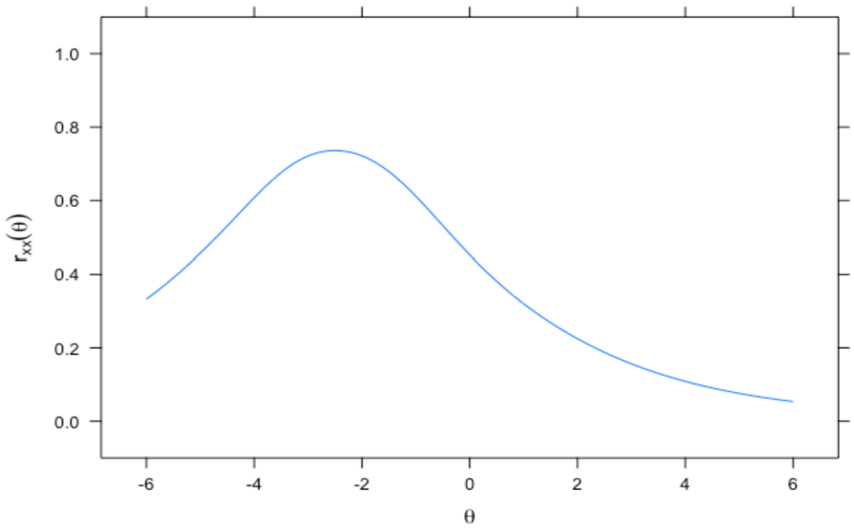



Záporný diskriminační index má položka QD19, což je patrné i z její charakteristické funkce. Mezi velmi málo diskriminující bychom mohli zařadit položky QD8, QD18 a QD20. Charakteristické křivky zbylých položek však také rostou velmi pozvolna a dalo by se celkově říct, že test příliš nediskriminuje. Nižší hodnoty zároveň poukazují na spíše nižší souvislosti položek s latentní proměnnou schopnosti rozpoznání emoci z intonace.

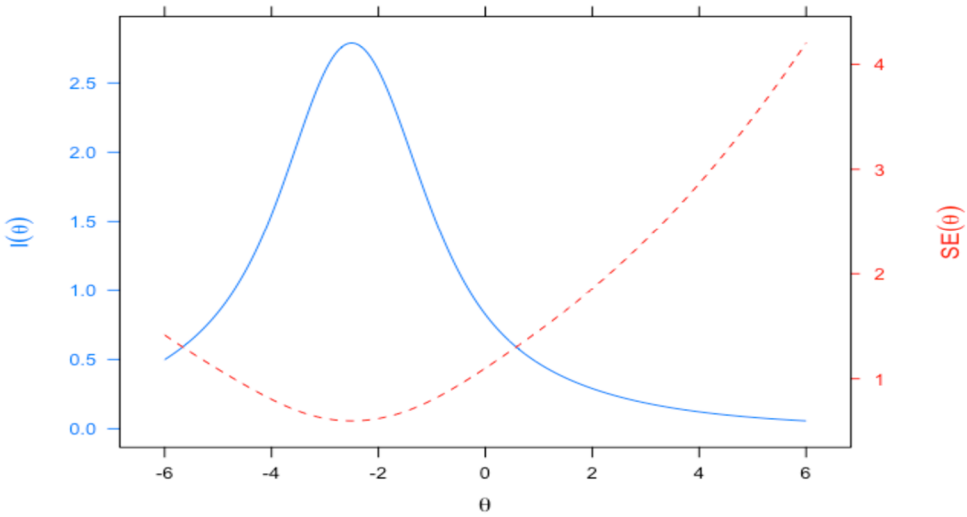
Co se týče IRT odhadu reliability, odhad marginální reliability je 0,46, což je výrazně nižší odhad než hodnota Cronbachovy alfy a McDonaldovy omegy. Výhodou IRT modelování je však to, že lze zjistit, při jaké úrovni schopnosti (θ) je reliabilita testu nejvyšší. Na obrázku číslo 6 je možné vidět, že nejvyšší reliability dosahuje subtest při θ = -2, a to lehce nad 0,7, což přibližně odpovídá hodnotám CTT ukazatelů vnitřní konzistence. Podíváme-li se pak na informační křivku testu na obrázku 7, je z ní patrné, že test nám dává nejvíce informací o lidech s podprůměrnou

schopností rozpoznat emoci z intonace, a to konkrétně o lidech, jejichž théta je okolo -2 (zhruba od -4 do 0). Naopak není téměř vůbec informativní pro lidi s průměrnou či vyšší schopností, což ilustruje i zvyšující se standardní chyba měření.

**Obrázek 6.** Reliabilita subtestu.




**Obrázek 7.** Informační křivka subtestu a křivka chyby měření.

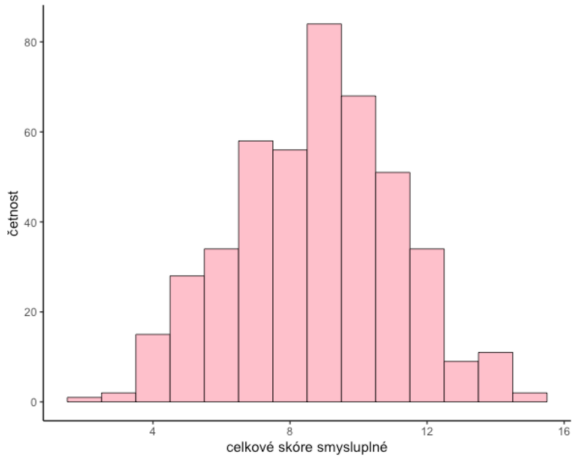


## 3.7 Analýza testu Věřte, nevěřte!

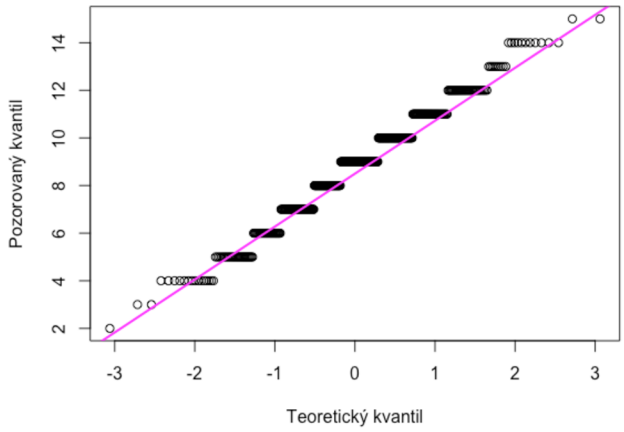
Testu *Věřte, nevěřte!*, založeného na principu verbálních absurdit, se zúčastnilo méně participantů než první části (N = 454), protože 8 dětem zabralo většinu času vyhrazeného na testování vyplňování první části testu (subtestu *Lína Džina*).

Z důvodu zjištěné dvoufaktorové struktury testu reportuji deskriptivní statistiky zvlášť pro položky náležející ke každému faktoru zvlášť (smysluplné a nesmyslné položky).

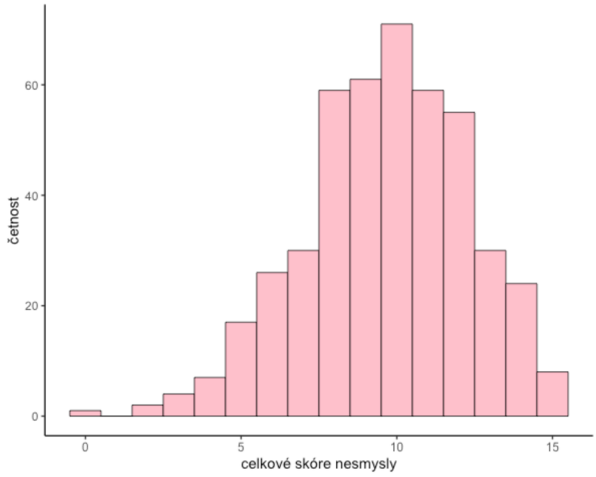
**Obrázek 8.** Rozložení celkových skórů smysluplných položek.



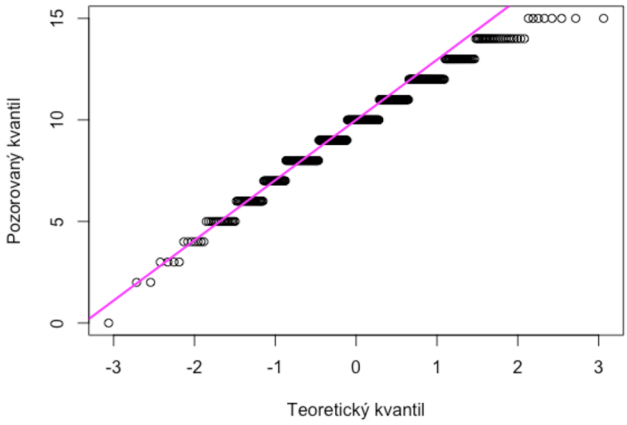

**Obrázek 9.** Q-Q plot smysluplných položek testu.



**Obrázek 10.** Rozložení celkových skórů nesmyslných položek.




**Obrázek 11.** Q-Q plot nesmyslných položek.




Průměrný počet správně zodpovězených nesmyslných položek je *M* = 9,65, *SD* = 2,64. Jeden respondent nezodpověděl ani jednu otázku správně, někteří respondenti naopak získali plný počet bodů. Rozložení je mírně zešikmené zleva (šikmost = -0,34; špičatost = -0,03). Odhady vnitřní konzistence škály jsou α = 0,79, 95 %CI [0,58; 0,92] a ω = 0,82.

Průměrný výsledek správně zodpovězených smysluplných položek je *M* = 8,77; *SD* = 2,39, rozložení je přibližně normální (šikmost = -0,05; špičatost = -0,32). Minimální počet správných odpovědí je 2, maximální 15. Odhady reliability škály jsou nižší než u nesmyslných výroků, kon-

krétně α = 0,60, $_{95\ \%}$CI [0,23; 0,84], interval spolehlivosti pro Cronbachovu alfu je tedy značně široký, a ω = 0,69. Po bližším přezkoumání položek se ukázalo, že některé položky se škálou korelují velmi slabě, a to konkrétně položky Q3, Q40, Q49, Q69 a Q75, které by všechny měly sytit faktor smysluplných položek. Ukázalo se, že tyto položky společný faktor sytí velmi slabě či téměř vůbec (v 2PL modelu měly nízké hodnoty diskriminačního parametru *a*, který je analogický faktorovému náboji ve faktorové analýze). Obsahově mají tyto položky společné to, že se řadí mezi zamýšleně jednodušší položky (cílová slova, která obsahovaly, byla například *podvlíkačky* či *rýč,* tj. obecně známější slova). Je možné, že u těchto položek, které jsou smysluplné a relativně jednoduché, probíhá proces odpovědi jinak (např. v nich respondenti mohou hledat další složitosti, které je u již tak obtížných položek nenapadají), což by bylo třeba zhodnotit například při kognitivní pilotáži. Z analýz jsem se každopádně rozhodla tyto položky vyřadit.

Finální verzi tedy tvoří dva faktory, smysluplné výroky (10 položek) a nesmyslné výroky (15 položek). Faktory spolu korelují velmi slabě, v rámci IRT analýzy jsem vytvořila dvoufaktorový 2PL model, ve kterém spolu faktor smysluplný a faktor nesmyslů korelovaly *r* = 0,086, $_{95\%}$CI [-0,08; 0,25]. Je navíc potřeba přihlédnout k tomu, že se jedná o korelaci latentních proměnných, která je očištěná od chyby měření. Nízká korelace mezi oběma faktory vybízí k otázce, proč se odpovědi na smysluplné a nesmyslné výroky natolik liší.

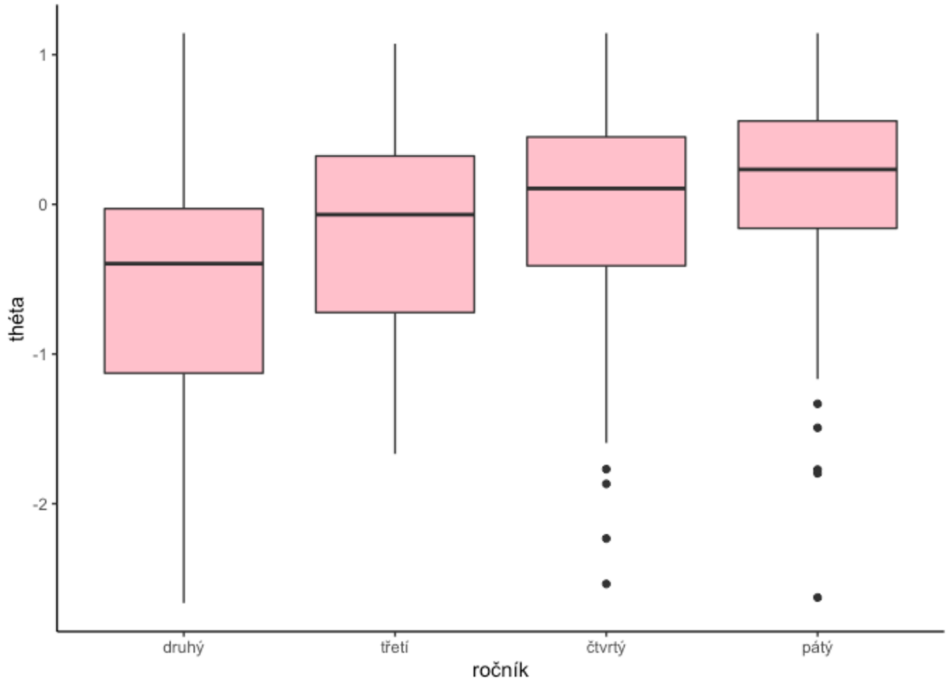
Dvoufaktorový 2PL model testu *Věřte, nevěřte!* celkově vykazuje uspokojivou shodu modelu s daty, M2(*410*) = 532,8; *p* <0,001; RMSEA = 0,03; $_{90\ \%}$CI$_{RMSEA}$ = [0,02; 0,03], SRMSR = 0,05, TLI = 0,84. Hodnota TLI je patrně opět nižší z toho důvodu, že i po vyřazení velmi slabě korelujících položek nejsou korelace mezi položkami příliš silné, hodnota RMSEA nulového modelu je 0,05.

## 3.8 Testy hypotéz

První z předpokladů byl ten, že skóre v subtestu *Líná Džina* poroste s věkem participanta. Ověřovala jsem ho za pomoci Welchovy ANOVY, která nepředpokládá stejnost rozptylů. V našem případě nelze stejnost rozptylů předpokládat s ohledem na to, jak velké jsou rozdíly v počtu respondentů z jednotlivých ročníků. Výsledky Welchovy ANOVA jsou

$F(3;92{,}67) = 6{,}34$; $p < 0{,}001$; $\eta^2 = 0{,}05$. Zdá se tedy, že se mezi jednotlivými skupinami vyskytují rozdíly, což svědčí o tom, že se výsledek v testu zlepšuje s rostoucím věkem. Rozdíly mezi jednotlivými třídami lze pozorovat také v boxplotech na obrázku 12. Lineární trend mezi jednotlivými třídami je statisticky významný, s každou navštěvovanou třídou roste schopnost rozpoznat emoci z intonace průměrně o 0,21 θ, což je relativně velké zlepšení, největší pozorujeme mezi druhým a třetím ročníkem.

**Obrázek 12.** Odhadnuté théty pro participanty v jednotlivých ročnících.




Druhou testovanou hypotézou je korelace mezi oběma testy. Je třeba brát v potaz, že hodnota Pearsonova korelačního koeficientu by mohla být podhodnocená vlivem toho, že žádná z metod není vysoce relabilní. Rozhodla jsem se proto faktory korelovat na úrovni latentních proměnných jako korelované faktory v 2PL modelu. Pro subtest *Líná Džina* a subtest se smysluplnými položkami byla hodnota korelace $r = 0{,}33$; $_{95\%}$CI [0,13; 0,52] a pro subtest s absurdními položkami $r = 0{,}32$, $_{95\%}$CI [0,15; 0,49].

# 4 Diskuze

## 4.1 Shrnutí výzkumných zjištění

Cílem práce bylo pilotní ověření vybraných psychometrických vlastností nově vzniklé metody určené k měření schopnosti rozpoznání emoce z intonace u dětí mladšího školního věku. Subtest je součástí vznikajícího testu *Frída a podivuhodná galerie*, určeného pro měření socio-emočních schopností. Z pilotáže vyplývá, že se jedná o nástroj, který je pro většinu dětí jednoduchý a relativně spolehlivým měřítkem by mohlo být pouze pro děti, které jsou v této schopnosti podprůměrné. Data však podporují jednodimenzionální strukturu testu a ukazatele relativně konzistentně poukazují na položky, které by bylo vhodné dále upravovat.

### 4.1.1 Konkrétní návrhy na úpravu položek

Několik položek v testu by na základě analýz bylo vhodné upravit. Jedná se o položku QD8, která v IRT analýze neměla příliš dobré hodnoty parametrů a v kognitivních pilotážích se ukázalo, že příčinou toho může být dvojznačnost výsledné situace, ve které záleží, zda se děti přikloní spíše k tomu, že se hlavní postava cítí vystrašeně, či spíše naštvaně (obě možnosti byly mezi možnými odpověďmi). Podobně je na tom pak položka QD18, ve které je sice výsledná emoce, kterou měla scénka navozovat jednoznačná, ale výpovědi respondentů z kognitivních pilotáží se shodují na tom, že dvě z možností odpovědí zní velmi klidně a nevyjadřují silně ani jednu z emocí. Dalším faktorem, který může vstupovat do rozhodování, je fakt, že položka obsahuje ironii (hlavní postava říká větu *„Ne, to vůbec není nebezpečný“*, zatímco druhým vytýká, že provedli něco nebezpečného). Zejména u mladších dětí by mohl tento rozpor mezi lexikálním a emocionálním obsahem vyvolat rozpor, což je v souladu s přítomností lexikálního zkreslení v mladším školním věku (např. Covic et al., 2021). Dále je u položky problémem výraz postavy v závěru scénky, ve kterém by se postava měla tvářit vážně, ale do videa se nezáměrně dostala verze, kdy se postava usmívá, což také mohlo pro participanty být matoucí. Položka QD19, jak již bylo zmíněno, obsahuje nahrávku s intonací *radost,* která však připomíná *překvapení*, jež by mohlo být legitimním výstupem scénky (stejně jako *vztek* či *strach*). Tato nejednoznačnost je dalším z fak-

torů, které mohou přispívat k horšímu fungování položky. Další položkou, kterou bych doporučovala k úpravě, je QD20. V jejím případě byl velmi často volen distraktor strach, který ale scénka neměla evokovat. Danou intonaci však lze interpretovat také jako radost vyřčenou s určitou naléhavostí, která je kontextu scénky představitelná.

## 4.2 Celkové zhodnocení testu

Ze způsobu konceptualizace testu plyne mnoho výhod. Oproti stávajícím měřícím nástrojům nabízí řešení, které je možné administrovat skupinově pomocí počítače. Dětem předkládá ekologicky validní stimuly v podobě scének, které jsou pro ně většinou srozumitelné (jak vyplývá z kognitivních pilotáží) a pro děti není obtížné pochopit, co je v subtestu úkolem. Není zároveň nutné nic číst, což přináší možnost zkusit testovat také předškolní děti. V možnostech odpovědi není cílem eliminovat obsah toho, co je řečeno (jako např. u některých testů, používajících elektronicky upravené či nesrozumitelné hlasy), nabízí emočně zabarvené intonace v takové formě, ve které se s nimi potkáváme v běžném životě. Od relativně velkého počtu dětí, které si subtest měly možnost vyzkoušet, se k autorům nedostala žádná negativní zpětná vazba. Při hromadné i individuální zpětné vazbě se participanti vyjadřovali, že pro ně byl subtest zábavný, což je u testu určeného pro dětskou populaci zvlášť klíčové.

### 4.2.1 Validita metody

Co se týče očekávání o validitě nové metody, naše předpoklady byly z větší části podpořeny. Platí, že úspěšnost v testu narůstá se zvyšujícím se navštěvovaným ročníkem. Na relativně uspokojivý výsledek co do diferenciální validity poukazuje také zjištěná korelace s metodou měřící verbální schopnost. Pro obě části metody *Věřte, nevěřte!* je korelace latentních proměnných rovná asi $r = 0{,}33$. Korelace latentních faktorů však nebere v úvahu vliv věku, který by mohl vysvětlovat určitou část sdíleného rozptylu výsledku v testech. Původně jsme očekávali korelaci rovnou přibližně $r = 0{,}20$, odhad výsledné korelace však není vzhledem k dříve řečenému diametrálně odlišný (i s přihlédnutím k tomu, jak široké intervaly spolehlivosti odhady odhady mají) a může sloužit jako podpora pro relativní nezávislost (oproti jiným testům) výkonu v testech, tedy pro to, že úspěch v subtestu *Líná Džina* není silně závislý na verbální schopnosti. Také z kognitivních pilotáží vyplývá, že název emoce (a to

zvláště u sekundárních emocí, jako je třeba *pohrdání*) nemusí být pro děti vždy známý, byť emoci zvládnou identifikovat na implicitní úrovni (Zeidner et al., 2001). Z toho důvodu se jedná o cenný poznatek o validitě metody.

Verbální schopnost však není jedinou schopností, která by mohla výkon v subtestu ovlivňovat. Vzhledem k omezené kapacitě této práce jsme se nezaměřovali i na jiné schopnosti, které by úspěch v testu mohly ovlivňovat. Do procesu odpovědi může vstupovat například krystalická nebo fluidní inteligence (např. Maul et al., 2012) či schopnost sluchové diferenciace. V rámci standardizace metody by bylo vhodné zjišťovat korelace mezi měřítky těchto schopností a subtestem *Líná Džina.*

Co se týče souběžné validity, je výběr testů, které by k jejímu ověření bylo možné administrovat, relativně omezený. V českém prostředí totiž nemáme (pokud je nám známo) jiný test na měření rozpoznání emoce z intonace v hlase, což je zapříčiněno patrně také závislostí takových testů na jazyku. Pro ověření souběžné validity by se tak dal použít některý ze zahraničních testů, který nepoužívá srozumitelnou řeč, či některý z dalších subtestů vyvíjených v rámci testu *Frída a podivuhodná galerie.* Neměří sice totožnou schopnost, ale můžeme mezi subtesty očekávat silnější souvislosti s ohledem na to, že by se mělo jednat o různé úzké schopnosti kandidátské široké schopnosti *Gei* v rámci CHC teorie.

## 4.3 Možné způsoby dalšího vývoje testu

Ukazuje se, že pro děti z běžné populace základních škol je test jednoduchý. Nad tím, co to znamená pro další vývoj metody, se dá uvažovat více způsoby. Vezmeme-li v úvahu to, že celý test *Frída a podivuhodná galerie* je primárně určen k identifikaci deficitů v socio-emoční oblasti, mohl by subtest najít využití právě k tomuto účelu. Tím by se odlišoval od zbylých testů ze systému Invenio, které jsou plošně administrovány za účelem identifikace nadaných žáků. V kontextu subtestu *Líná Džina* je smysluplnost plošné administrace otázkou, protože pro osoby s průměrnou či nadprůměrnou schopností není test v současné podobě příliš informativní. Mohlo by být přínosné zkusit test administrovat dětem, u kterých na základě minulých studií očekáváme ve schopnosti rozpoznání emoce z intonace nějaké deficity, jedná se například děti s autismem (Fridenson-Hayo et al., 2016), ADHD (Ludlow et al., 2014), ale i opožděným jazykovým vývojem (Löytömäki, 2019).

Zároveň je možné, že by test v současné podobě mohl být vhodnější spíš pro mladší děti, než na kterých byl ověřován v této práci. Zajímavé výsledky by mohlo přinést například srovnání dětí v předškolním a školním věku. Výsledky pilotáže naznačují, že i pro mladší děti by mohl subtest být zvládnutelný.

Další možností je pak změna znění distraktorů v testu. Při tvorbě testu (i v rámci této práce) bylo snahou, aby bylo vyznění scének jednoznačné, jasně poukazovalo na jednu emoci. U většiny položek se toho povedlo dosáhnout, u některých se v rámci pilotáže ukázalo, že zcela jednoznačné nejsou. Dále pak bylo snahou, aby byla správná odpověď jednoznačná pro dospělého člověka. Znamená to snahu vyhnout se vyhodnocování způsobem, kterým se to dělá u některých dalších testů na měření emoční inteligence (například MSCEIT), tedy určením správné odpovědi za pomocí shody expertů nebo konsenzuálně (správná odpověď je ta, na které se shodne nejvíce participantů). Pilotáž však poukazuje na to, že se jedná o obtížný úkol. Subtest je sice jednoznačný pro dospělého člověka, ale zdá se, že tomu je tak do značné míry i pro dětské participanty. Bylo by možné test ztížit, a to tím, že by se k dané scénce vybraly jako možnosti emoce, které jsou si navzájem významově bližší (například smutek-strach-vztek). Správnou odpověď by se pak bylo nutné určovat jedním z výše popsaných způsobů. Alternativou by bylo také navýšení počtu distraktorů.

## 4.4 Limity

Ačkoliv se subtest *Líná Džina* jeví jako slibné měřítko rozpoznání emoce z intonace, při tvorbě testu i samotném provedení pilotáže se vyskytly určité limity.

### 4.4.1 Limity teorie

Jak už vyplývá z části Úvod, teoretické ukotvení je celkově neuralgickým bodem celého konstruktu emoční inteligence. Ve výkonovém přístupu k měření EI dominuje Salovey-Mayerova teorie emoční inteligence (Mayer et al., 1999), která je však značně obtížně přenositelná do praxe. Dokladem o tom je fakt, že test MSCEIT, který vytvořili samotní autoři teorie tak, aby kopíroval zamýšlenou strukturu teorie, jí faktorově neodpovídá. Místo očekávaných čtyř faktorů různí autoři nalézají většinou jeden

obecný faktor EI, nebo jeden faktor prvního řádu a v jeho rámci tři faktory druhého řádu (Fiori et al., 2014; Palmer et al., 2005). Salovey et al. (2016) upřesňují, že se faktorovou strukturu nedaří empiricky podpořit, a původní faktory nazývají *oblastmi zájmu.* Přesto trvají na tom, že by čtyři větve jejich modelu měly stát úzkými schopnostmi *Gei* v rámci CHC. Salovey-Mayerova teorie tak zatím zůstává spíše v rovině přání autorů, kteří věří, že se lidé liší ve schopnosti nakládání s emocemi, ale zatím není zcela jasné, v čem přesně se liší a jak. Na takovém základě se těžko staví testovatelné hypotézy. Maul et al. (2012) tvrdí, že v oblasti emoční inteligence by výzkumníci udělali nejlépe, kdyby teorii začali budovat znovu a odspodu, zjišťováním souvislostí mezi měřítky jasně definovaných, specifických schopností, spíše než stavěním na předpokladu, že se jednou podaří ověřit Salovey-Mayerovu teorii.

### 4.4.2 Limity metody

Limity teoretického ukotvení se promítají do limitů metody. Jedná se zejména o fakt, že nelze úplně jasně říci, kterou konkrétní schopnost subtest *Líná Džina* měří. Teoretické ukotvení v tom příliš nepomůže, protože je samotné značně vágní a abstraktní. V momentě, kdy v testu bude mít participant nízké skóre, nemůžeme vědět, zda u něj problém nastává v identifikaci emoce ze scénky, identifikaci emoce z možnosti odpovědí či například nesouladu mezi tím, jaký je výraz v obličeji kreslené postavy a tím, jak zní její hlas. Ze zpětných vazeb vyplývá, že poslední jmenovaný problém se relativně běžně vyskytoval. Při hromadné administraci nejsme schopní určit příčinu problému a neznalost těchto skutečností podrývá validitu metody. Lze říci, že bychom subtest mohli považovat za validní v Messickově (Messick, 1995) pojetí validity v momentě, kdy by spolehlivě odlišoval děti, které mají v socio-emoční oblasti nějaké limity (resp. by byly validní jeho závěry), což ale v současné chvíli říct nelze.

Druhým limitem je samozřejmě už dříve popsaná jednoduchost subtestu a otázka jeho použitelnosti pro normativně se vyvíjející populaci, jak je popsáno v podkapitole *Možné způsoby dalšího vývoje testu*.

### 4.4.3 Limity designu a sběru dat

Použitý výzkumný design je specifický tím, že jeho cílem bylo pouze provedení pilotáže (subtest totiž doposud nikdy administrován nebyl), nejedná se o validizační studii. V rámci ověření diferenciální validity jsme administrovali nestandardizovaný test *Věřte, nevěřte!*, který vznikl

v rámci diplomové práce Mgr. Kláry Hečlové. Původním záměrem bylo použití subtestu Verbální konceptualizace z Inteligenční vývojové škály (IDS), její použití pro naše účely nám však bylo vydavatelem testu, Hogrefe Testcentrem, zakázáno. Přistoupili jsme proto k použití metody, o jejímž fungování jsme neměli mnoho důkazů (kromě pilotáže provedené autorkou). Na opačnou stranu však věříme, že tato práce přinesla zároveň další vhled do fungování metody *Věřte, nevěřte!,* kterou může tým Invenio dále rozvíjet.

Nedostatky nacházíme také ve způsobu sběru dat. Nejvíce očividnou je absence informace o pohlaví participantů. Srovnání mezi tím, jak si v subtestu vedly dívky a chlapci, by mohlo přinést další informace o fungování subtestu. Dalším limitem je pak nevyrovnanost v zastoupení jednotlivých ročníků, která souvisí s rozdílnou ochotou škol nabízet testování jednotlivým ročníkům. Autorka byla přítomná pouze u jedné administrace, u zbylých administrací nezbývá než předpokládat, že proběhly standardním způsobem. Do budoucna by bylo vhodné se více zaměřit na identifikaci dětí se speciálními potřebami. U některých dětí z našeho vzorku jsme sice informaci měli, ale pouze u části z nich známe charakter deficitu a zároveň nevíme, zda se jedná o diagnostikovaný problém, či postřeh pedagoga. Při dalších administracích bude proto důležité se zaměřit na to, jak jsou v subtestu úspěšné děti s konkrétními formami deficitů oproti běžné populaci.

# Použité zdroje

de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Press.

Aguert, M., Laval, V., Lacroix, A., Gil, S., & Le Bigot, L. (2013). Inferring emotions from speech prosody: Not so easy at age five. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0083657

Barnes, S. (2015). Computer-based testing and young children. In O. N. Saracho, (Eds.), *Contemporary Perspectives in research in assessment and evaluation in early childhood education* (pp. 373-395). Information Age Publishing, Inc.

Bar-On R. (2006). The Bar-On model of emotional-social intelligence (ESI). *Psicothema*, *18*, 13–25.

Baron-Cohen, S., Golan, O., Wheelwright, S., Granader, Y., & Hill, J. (2010). Emotion word comprehension from 4 to 16 years old: A developmental survey. *Frontiers in Evolutionary Neuroscience*, *2*. https://doi.org/10.3389/fnevo.2010.0010

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in The eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251. https://doi.org/10.1111/1469-7610.00715

Bänziger, T., Grandjean, D., & Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: The multimodal emotion recognition test (MERT). *Emotion*, *9*(5), 691–704. https://doi.org/10.1037/a0017088

Baum, K. M., & Nowicki, S., Jr. (1998). Perception of emotion: Measuring decoding accuracy of adult prosodic cues varying in intensity. *Journal of Nonverbal Behavior, 22*(2), 89–107. https://doi.org/10.1023/A:1022954014365

Bernieri, F. J. (2001). Toward a taxonomy of interpersonal sensitivity. In J. A. Hall & F. J. Bernieri (Eds*.), Interpersonal sensitivity: Theory and measurement* (pp. 3-20). Lawrence Erlbaum Associates Publishers.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments* (2nd ed). University of California Press.

Buchanan, R. D., & Finch, S. J. (2005). History of Psychometrics. *Encyclopedia of Statistics in Behavioral Science* (pp. 1–4). https://doi:10.1002/0470013192.bsa282

Carpenter, R., & Alloway, T. (2018). Computer versus paper-based testing: Are they equivalent when it comes to working memory? *Journal of Psychoeducational Assessment*, *37*(3), 382–394. https://doi.org/10.1177/0734282918761496

Chalmers. R. P. (2012). "mirt: A Multidimensional Item Response Theory Package for the R Environment." *Journal of Statistical Software*, *48*(6), 1–29. https://doi:10.18637/jss.v048.i06.

Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., & Sonuga-Barke, E. J. (2014). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *British Journal of Developmental Psychology*, *33*(2), 218–236. https://doi.org/10.1111/bjdp.12075

Covic, A., von Steinbüchel, N., & Kiese-Himmel, C. (2019). Emotion recognition in kindergarten children. *Folia Phoniatrica Et Logopaedica*, *72*(4), 273–281. https://doi.org/10.1159/000500589

Conzelmann, K., Weis, S., & Süß, H. (2013). New findings about social intelligence: Development and application of the Magdeburg test of social intelligence (MTSI). *Journal of Individual Differences*, *34*(3), 119-137. https://doi:10.1027/1614-0001/a000106

Dai, S., Wang, X., Svetina, D. (2022). subscore: Computing Subscores in Classical Test Theory and Item Response Theory. R package version 3.3. dostupné z https://CRAN.R-project.org/package=subscore

Davis, S. K., Morningstar, M., Dirks, M. A., & Qualter, P. (2020). Ability emotional intelligence: What about recognition of emotion in voices? *Personality and Individual Differences*, *160*. https://doi.org/10.1016/j.paid.2020.109938

DeMars, C. (2010). Item Response Theory: Understanding Statistics Measurement. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195377033.001.0001

Ekman, P., & Davidson, R. J. (1994). The nature of emotion: Fundamental questions. Oxford University Press.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191. https://doi.org/10.3758/BF03193146

Fiori, M., Antonietti, J.-P., Mikolajczak, M., Luminet, O., Hansenne, M., & Rossier, J. (2014). What is the ability emotional intelligence test (MSCEIT) good for? an evaluation using item response theory. *PLoS ONE*, *9*(6). https://doi.org/10.1371/journal.pone.0098827

Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., Baron-Cohen, S., & Golan, O. (2016). Basic and complex emotion recognition in children with autism: Cross-cultural findings. *Molecular Autism*, *7*(1). https://doi.org/10.1186/s13229-016-0113-9

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences.* Basic Books.

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2006). The 'Reading the Mind in the Voice' test-revised: A study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *37*(6), 1096–1106. https://doi.org/10.1007/s10803-006-0252-5

Golan, O., Sinai-Gavrilov, Y., & Baron-Cohen, S. (2015). The Cambridge mindreading face-voice battery for children (CAM-C): Complex emotion recognition in children with and without autism spectrum conditions. *Molecular Autism*, *6*(1). https://doi.org/10.1186/s13229-015-0018-z

Goleman, D. (1995). *Emotional intelligence.* Bantam Books, Inc.

Grosse, G., Streubel, B., Gunzenhauser, C., & Saalbach, H. (2021). Let's talk about emotions: The development of children's emotion vocabulary

from 4 to 11 years of age. *Affective Science*, *2*(2), 150–162. https://doi.org/10.1007/s42761-021-00040-2

Grossmann, T., Oberecker, R., Koch, S. P., & Friederici, A. D. (2010). The developmental origins of voice processing in the human brain. *Neuron*, *65*(6), 852–858. https://doi.org/10.1016/j.neuron.2010.03.001

Hečlová, K. (2020). Vývoj a pilotní ověření metody k měření verbálních schopností u dětí pomocí počítačového testování [Diplomová práce, Masarykova univerzita]. Institucionální repozitář MU. https://is.muni.cz/th/dyffp/

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ikeda, S., Sudo, M., Matsui, T., & Haryu, E. (2021). Developmental changes in understanding emotion in speech in children in Japan and the United States. *Cognitive Development*, *60*(12). https://doi.org/10.1016/j.cogdev.2021.101110

Izard, C. E., & Harris, P. (1995). Emotional development and developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol. 1. Theory and methods* (pp. 467–503). John Wiley & Sons.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). semTools: Useful tools for structural equation modeling. R package version 0.5-6. Dostupné z https://CRAN.R-project.org/package=semTools

Kan, K.-J., Kievit, R. A., Dolan, C., & der Maas, H. van. (2011). On the interpretation of the CHC factor GC. *Intelligence*, *39*(5), 292–302. https://doi.org/10.1016/j.intell.2011.05.003

Kenny, D. (2020). *SEM: Fit*. SEM: Fit (David A. Kenny). Dostupné z http://www.davidakenny.net/cm/fit.htm

Lewis, M. (1993). Self-conscious emotions: Embarrassment, pride, shame, and guilt. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 563–573). The Guilford Press.

Löytömäki, J., Ohtonen, P., Laakso, M. L., & Huttunen, K. (2019). The role of linguistic and cognitive factors in emotion recognition difficulties in children with ASD, ADHD or DLD. *International Journal of Language & Communication Disorders*, *55*(2), 231–242. https://doi.org/10.1111/1460-6984.12514

Ludlow, A. K., Garrood, A., Lawrence, K., & Gutierrez, R. (2014). Emotion recognition from dynamic emotional displays in children with ADHD. *Journal of Social and Clinical Psychology*, *33*(5), 413–427. https://doi.org/10.1521/jscp.2014.33.5.413

MacCann, C., Jiang, Y., Brown, L. E., Double, K. S., Bucich, M., & Minbashian, A. (2020). Emotional intelligence predicts academic performance: A meta-analysis. *Psychological Bulletin*, *146*(2), 150–186. https://doi.org/10.1037/bul0000219

Maul, A. (2012). The validity of the Mayer–Salovey–Caruso emotional intelligence test (MSCEIT) as a measure of emotional intelligence. *Emotion Review*, *4*(4), 394–402. https://doi.org/10.1177/1754073912445811

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, *27*(4), 267–298. https://doi.org/10.1016/s0160-2896(99)00016-1

Mayer, J. D., Caruso, D. R., & Salovey, P. (2016). The ability model of emotional intelligence: Principles and updates. *Emotion Review*, *8*(4), 290–300. https://doi.org/10.1177/1754073916639667

Mayer, J. D., Salovey, P., & Caruso, D. R. (2012). The validity of the MSCEIT: Additional analyses and evidence. *Emotion Review*, *4*(4), 403–408. https://doi.org/10.1177/1754073912445815

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT v2.0. *Emotion*, *3*(1), 97–105. https://doi.org/10.1037/1528-3542.3.1.97

Mavilidi, M.-F., Hoogerheide, V., & Paas, F. (2014). A quick and easy strategy to reduce test anxiety and enhance test performance. *Applied Cognitive Psychology*, *28*(5), 720–726. https://doi.org/10.1002/acp.3058

McGill, R. J., & Dombrowski, S. C. (2019). Critically reflecting on the origins, evolution, and impact of the cattell-horn-carroll (CHC) model. *Applied Measurement in Education*, *32*(3), 216–231. https://doi.org/10.1080/08957347.2019.1619561

McGrew, K. S. (2009). CHC theory and the Human Cognitive Abilities Project: Standing on the shoulders of the Giants of Psychometric Intelligence Research. *Intelligence*, *37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741–749. https://doi.org/10.1037/0003-066x.50.9.741

Morningstar, M., Nelson, E. E., & Dirks, M. A. (2018). Maturation of vocal emotion recognition: Insights from the developmental and neuroimaging literature. *Neuroscience & Biobehavioral Reviews*, *90*, 221–230. https://doi.org/10.1016/j.neubiorev.2018.04.019

Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development*, *72*(3), 834–843. https://doi.org/10.1111/1467-8624.00318

Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Somerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion*, *20*(5), 773–792. https://doi.org/10.1037/emo0000609

Nonverbal receptivity: The diagnostic analysis of nonverbal accuracy (DANVA). (2001). *Interpersonal Sensitivity*, 199–216. https://doi.org/10.4324/9781410600424-19

O'Connor, P. J., Hill, A., Kaya, M., & Martin, B. (2019). The measurement of Emotional Intelligence: A Critical Review of the literature and recommendations for researchers and Practitioners. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01116

Odukoya, A. J. (2020). Review of studies on the Psychometric Properties of mayer–salovey–caruso emotional intelligence test. *Journal of Advanced Research in Dynamical and Control Systems*, *12*(SP7), 1961–1968. https://doi.org/10.5373/jardcs/v12sp7/20202311

Palmer, B. R., Gignac, G., Manocha, R., & Stough, C. (2005). A psychometric evaluation of the Mayer–Salovey–Caruso Emotional Intelligence Test version 2.0. *Intelligence*, *33*(3), 285–305. https://doi.org/10.1016/j.intell.2004.11.003

Portešová, Š., Straka, O., Jabůrek, M., Palíšek, P., Jajcaj, F., Šromová, V., Macek, D., Hlavica, M., Jirásek, P., Doubravová, L., Gregor, L., Novák, Z., Kottman, J., Dostálek, I., Kolín, O., Bíza, M., Beňuš, A., Burian, J., Cupáková, D., . . . Zatloukal, P. (2022). *Subtest pro identifikaci socio-emočních deficitů*.

Qualter, P., Urquijo, I., Henzi, P., Barrett, L., & Humphrey, N. (2018). Ability emotional intelligence and children's behaviour in the Playground. *Social Development*, *28*(2), 430–448. https://doi.org/10.1111/sode.12340

Reeve, J. (2018). *Understanding motivation and emotion*. John Wiley & Sons, Inc.

Revelle, W. (2023). psych: Procedures for Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois. R package version 2.3.3, dostupné z https://CRAN.R-project.org/package=psych.

Riggio, R. E. (2010). Before emotional intelligence: Research on nonverbal, emotional, and social competences. *Industrial and Organizational Psychology*, *3*(2), 178–182. https://doi.org/10.1111/j.1754-9434.2010.01221.x

Rivers, S. E., Brackett, M. A., Reyes, M. R., Mayer, J. D., Caruso, D. R., & Salovey, P. (2012). Measuring emotional intelligence in early adolescence with the MSCEIT-YV. *Journal of Psychoeducational Assessment*, *30*(4), 344–366. https://doi.org/10.1177/0734282912449443

Rosseel, Y. (2012). "lavaan: An R Package for Structural Equation Modeling." *Journal of Statistical Software*, *48*(2), 1-36. doi:10.18637/jss.v048.i02.

Rosenthal, R., Archer, D., Hall, J. A., DiMatteo, M. R., & Rogers, P. L. (1979). Measuring sensitivity to nonverbal communication: The PONS test. *Nonverbal Behavior*, 67–98. https://doi.org/10.1016/b978-0-12-761350-5.50012-4

Rothman, A. D., & Nowicki Jr., S. (2004). A measure of the ability to identify emotion in children's tone of Voice. *Journal of Nonverbal Behavior*, *28*(2), 67–92. https://doi.org/10.1023/b:jonb.0000023653.13943.31

Saarni, C., Campos, J.J., Camras, L.A. and Witherington, D. (2007). Emotional Development: Action, Communication, and Understanding. In Handbook of Child Psychology (eds W. Damon, R.M. Lerner and N. Eisenberg). https://doi.org/10.1002/9780470147658.chpsy0305

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, *9*(3), 185–211. https://doi.org/10.2190/dugg-p24e-52wk-6cdg

Sauter, D. A., Panattoni, C., & Happé, F. (2012). Children's recognition of emotions from vocal cues. *British Journal of Developmental Psychology*, *31*(1), 97–113. https://doi.org/10.1111/j.2044-835x.2012.02081.x

Scherer, K. R., & Scherer, U. (2011). Assessing the ability to recognize facial and vocal expressions of emotion: Construction and validation of the Emotion Recognition Index. *Journal of Nonverbal Behavior*, *35*(4), 305–326. https://doi.org/10.1007/s10919-011-0115-4

Schlegel, K., & Scherer, K. R. (2015). Introducing a short version of the Geneva Emotion Recognition Test (Gert-S): Psychometric properties and construct validation. *Behavior Research Methods*, *48*(4), 1383–1392. https://doi.org/10.3758/s13428-015-0646-4

Schirmer, A., & Adolphs, R. (2017). Emotion perception from face, voice, and touch: Comparisons and convergence. *Trends in Cognitive Sciences*, *21*(3), 216–228. https://doi.org/10.1016/j.tics.2017.01.001

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* (pp. 73–163). The Guilford Press.

Sternberg, R. (1984). Toward a triarchic theory of human intelligence. Behavioral and Brain Sciences, *7*(2), 269-287. https://doi:10.1017/S0140525X00044629

Širůček, J. (2010). Problémy psychologického testování pomocí počítače a v prostředí internetu. *TESTFÓRUM*, *1*(1), 23–26. https://doi.org/10.5817/tf2010-1-6

Thorndike, E. L. (1920). Intelligence and its use. *Harper's Magazine*, *140*, 227-235.

Wasserman, J. D. (2019). Deconstructing CHC. *Applied Measurement in Education*, *32*(3), 249–268. https://doi.org/10.1080/08957347.2019.1619563

Waterhouse, L. (2006). Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review. *Educational Psychologist*, *41*(4), 207–225. https://doi.org/10.1207/s15326985ep4104_1

Wickham H., Averick M., Bryan J., Chang W., McGowan L. D., François R., Grolemund G., Hayes A., Henry L., Hester J., Kuhn M., Pedersen T. L., Miller E., Bache S. M., Müller K., Ooms J., Robinson D., Seidel D. P., Spinu V., …, Yutani H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *43*(4), 1686. https://doi:10.21105/joss.01686.

Wind, S. A., & Hua, C. (2022). *Rasch measurement theory analysis in R*. CRC Press.

Yen, W. M. (1984). Effects of Local Item Dependence on the Fit and Equating Performance of the Three-Parameter Logistic Model. Applied Psychological Measurement, *8*(2), 125–145. https://doi.org/10.1177/014662168400800201

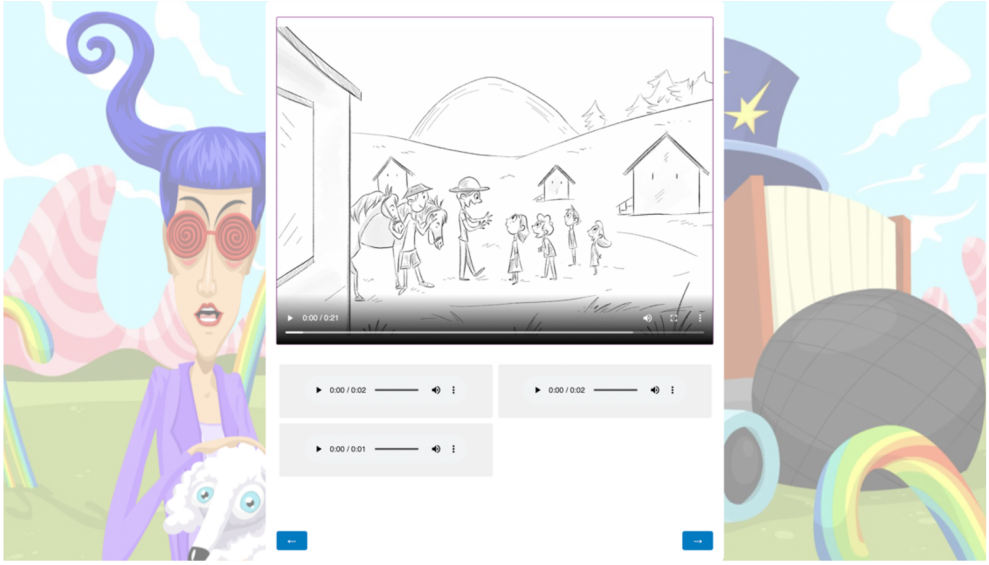
Zanon, C., Hutz, C. S., Yoo, H., & Hambleton, R. K. (2016). An application of item response theory to psychological test development. *Psicologia:*

*Reflexão e Crítica*, *29*(1). https://doi.org/10.1186/s41155-016-0040-x

Zeidner, M., Matthews, G., & Roberts, R. D. (2001). Slow Down, you move too fast: Emotional intelligence remains an "elusive" intelligence. *Emotion*, *1*(3), 265–275. https://doi.org/10.1037/1528-3542.1.3.265

Zijlmans, E. A., Tijmstra, J., van der Ark, L. A., & Sijtsma, K. (2019). Item-score reliability as a selection tool in test construction. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02298

# Příloha A Náhled v platformě Qualtrics



# Příloha B Tetrachorické korelace subtestu Líná Džina

| | QD1 | QD2 | QD3 | QD5 | QD4 | QD6 | QD7 | QD8 | QD9 | QD10 | QD11 | QD12 | QD13 | QD14 | QD15 | QD16 | QD17 | QD18 | QD19 | QD20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **QD1** | 1 | | | | | | | | | | | | | | | | | | | |
| **QD2** | 0,32 | 1 | | | | | | | | | | | | | | | | | | |
| **QD3** | 0,16 | 0,22 | 1 | | | | | | | | | | | | | | | | | |
| **QD5** | 0,06 | 0,03 | 0,06 | 1 | | | | | | | | | | | | | | | | |
| **QD4** | 0,33 | 0,12 | 0,3 | 0,09 | 1 | | | | | | | | | | | | | | | |
| **QD6** | 0,04 | 0,47 | 0,18 | 0,3 | 0,18 | 1 | | | | | | | | | | | | | | |
| **QD7** | 0,18 | 0,19 | 0 | 0,26 | 0,03 | 0,08 | 1 | | | | | | | | | | | | | |
| **QD8** | -0,06 | 0,02 | 0,15 | 0,12 | 0,12 | 0,05 | -0,11 | 1 | | | | | | | | | | | | |
| **QD9** | 0,04 | 0,39 | -0,05 | 0,41 | 0,25 | 0,11 | 0,11 | 0,13 | 1 | | | | | | | | | | | |
| **QD10** | -0,08 | 0,29 | 0,04 | 0,09 | 0,18 | 0,08 | -0,1 | 0,04 | 0,09 | 1 | | | | | | | | | | |
| **QD11** | 0,14 | 0,3 | -0,2 | 0,09 | 0,02 | -0,02 | 0,31 | -0,06 | 0,37 | 0,11 | 1 | | | | | | | | | |
| **QD12** | 0,26 | 0,31 | 0,04 | 0,26 | 0,21 | 0,29 | 0,12 | -0,09 | 0,38 | 0,06 | 0,06 | 1 | | | | | | | | |
| **QD13** | 0,11 | -0,03 | 0,32 | 0 | -0,07 | -0,11 | -0,21 | 0,1 | 0,05 | 0,08 | -0,26 | 0,29 | 1 | | | | | | | |
| **QD14** | 0,09 | -0,1 | 0,1 | 0,2 | -0,05 | 0,23 | -0,15 | 0,19 | 0,1 | 0,13 | -0,13 | 0,22 | 0,3 | 1 | | | | | | |
| **QD15** | 0,28 | 0,47 | 0,01 | 0 | 0,21 | 0,25 | 0,23 | -0,11 | 0,31 | 0,11 | 0,26 | 0,52 | 0,18 | 0,13 | 1 | | | | | |
| **QD16** | 0,05 | 0,22 | 0,2 | 0,2 | 0,15 | 0,06 | -0,01 | -0,05 | 0,27 | 0,13 | 0,14 | 0,06 | 0,12 | 0,19 | 0,18 | 1 | | | | |
| **QD17** | 0,18 | -0,03 | 0,22 | 0,13 | 0,2 | 0,18 | 0,24 | 0,06 | 0,04 | 0,14 | 0,05 | 0,23 | 0,01 | -0,03 | 0,17 | 0,04 | 1 | | | |
| **QD18** | 0,13 | 0,09 | 0,21 | -0,05 | 0,28 | 0,12 | -0,04 | -0,03 | 0,01 | -0,04 | -0,11 | 0,12 | -0,06 | -0,06 | 0,07 | 0,05 | 0,17 | 1 | | |
| **QD19** | 0,14 | -0,33 | -0,04 | -0,08 | -0,01 | -0,13 | 0 | -0,09 | 0 | -0,04 | -0,05 | 0,06 | -0,01 | 0,01 | -0,11 | -0,04 | 0,1 | 0,07 | 1 | |
| **QD20** | -0,03 | 0,16 | -0,03 | 0,03 | 0,11 | 0,05 | -0,02 | 0,16 | 0,14 | 0,24 | 0,06 | -0,04 | 0,04 | 0,1 | 0,02 | -0,07 | -0,05 | -0,05 | -0,13 | 1 |

# Příloha C Tetrachorické korelace smysluplných položek

| | Q3 | Q9 | Q11 | Q21 | Q23 | Q26 | Q40 | Q47 | Q55 | Q28 | Q62 | Q69 | Q49 | Q75 | Q82 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Q3** | 1 | | | | | | | | | | | | | | |
| **Q9** | -0,07 | 1 | | | | | | | | | | | | | |
| **Q11** | 0,01 | 0,08 | 1 | | | | | | | | | | | | |
| **Q21** | 0,2 | 0,25 | 0,05 | 1 | | | | | | | | | | | |
| **Q23** | 0,02 | 0,26 | 0,17 | 0,26 | 1 | | | | | | | | | | |
| **Q26** | 0,08 | 0,2 | 0,24 | 0,15 | 0,17 | 1 | | | | | | | | | |
| **Q40** | 0,03 | 0,22 | 0,07 | -0,06 | 0,09 | 0,07 | 1 | | | | | | | | |
| **Q47** | 0,1 | 0,14 | 0,23 | 0,24 | 0,21 | 0,08 | 0,14 | 1 | | | | | | | |
| **Q55** | 0,01 | 0,19 | -0,06 | 0,18 | 0,04 | 0,13 | 0,07 | 0,12 | 1 | | | | | | |
| **Q28** | -0,08 | 0,07 | 0,02 | -0,04 | 0,3 | 0,09 | 0,07 | 0,04 | 0,27 | 1 | | | | | |
| **Q62** | 0,06 | 0,1 | 0,2 | -0,01 | 0,06 | 0,02 | 0,08 | 0,2 | 0,07 | 0,26 | 1 | | | | |
| **Q69** | 0,05 | 0,04 | 0,01 | 0,1 | -0,07 | -0,02 | -0,04 | -0,09 | 0,16 | 0,08 | -0,04 | 1 | | | |
| **Q49** | 0,25 | 0,19 | 0,01 | 0,11 | 0,16 | 0,09 | -0,06 | -0,12 | -0,08 | -0,07 | 0,22 | 0,05 | 1 | | |
| **Q75** | 0,06 | 0,09 | 0,03 | 0,12 | 0,06 | 0,01 | 0,05 | 0,13 | 0,07 | -0,03 | 0,07 | 0,09 | 0,12 | 1 | |
| **Q811** | 0,29 | 0,04 | 0,2 | 0,12 | 0,08 | 0,1 | 0 | 0,05 | 0,18 | 0,13 | 0,27 | 0,22 | 0,02 | 0,07 | 1 |

# Příloha D Tetrachorické korelace nesmyslných položek

| | Q2 | Q6 | Q7 | Q12 | Q14 | Q18 | Q30 | Q31 | Q37 | Q50 | Q53 | Q59 | Q63 | Q68 | Q82 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Q2** | 1 | | | | | | | | | | | | | | |
| **Q6** | 0,1 | 1 | | | | | | | | | | | | | |
| **Q7** | 0,08 | 0,05 | 1 | | | | | | | | | | | | |
| **Q12** | 0,09 | 0,17 | 0,4 | 1 | | | | | | | | | | | |
| **Q14** | 0,37 | 0,13 | 0,07 | 0,2 | 1 | | | | | | | | | | |
| **Q18** | 0,25 | 0,24 | 0,37 | 0,46 | 0,28 | 1 | | | | | | | | | |
| **Q30** | 0,05 | 0,12 | 0,13 | 0,05 | 0,1 | 0,17 | 1 | | | | | | | | |
| **Q31** | 0,04 | 0,21 | 0,12 | 0,11 | 0,16 | 0,26 | 0,14 | 1 | | | | | | | |
| **Q37** | 0,19 | 0,18 | 0,05 | 0,11 | 0,18 | 0,31 | 0,15 | 0,1 | 1 | | | | | | |
| **Q50** | -0,02 | 0,26 | 0,36 | 0,34 | 0,23 | 0,39 | 0 | 0,22 | 0,14 | 1 | | | | | |
| **Q53** | 0,03 | 0,08 | 0,42 | 0,46 | 0,16 | 0,1 | 0,18 | 0,19 | 0,03 | 0,13 | 1 | | | | |
| **Q59** | 0,2 | 0,14 | 0,37 | 0,2 | 0,18 | 0,33 | 0,2 | 0,14 | 0,26 | 0,31 | 0,14 | 1 | | | |
| **Q63** | 0,19 | 0,12 | 0,37 | 0,33 | 0,21 | 0,09 | 0,11 | 0,2 | -0,07 | 0,21 | 0,15 | 0,23 | 1 | | |
| **Q68** | 0,12 | 0,21 | 0,38 | 0,43 | 0,15 | 0,34 | 0,2 | 0,27 | 0,26 | 0,24 | 0,12 | 0,15 | 0,33 | 1 | |
| **Q82** | 0,21 | 0,22 | 0,25 | 0,35 | 0,07 | 0,33 | 0,05 | 0,14 | 0,16 | 0,19 | 0,24 | 0,24 | 0,15 | 0,27 | 1 |