



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

DIPLOMOVÁ PRÁCE

Martin Landa

Psychometrická analýza Lawsonova testu vědeckého uvažování

Katedra didaktiky fyziky

Vedoucí diplomové práce: PhDr. Chvál Martin, Ph.D.

Studijní program: Učitelství fyziky pro střední školy

Studijní obor: Učitelství fyziky pro střední školy se sdruženým studiem

Učitelství matematiky pro střední školy

Praha 2023

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona.

V Praze dne 17.7.2023

Martin Landa

Velké poděkování patří vedoucímu této práce PhDr. Martinu Chválovi, Ph.D., za nekonečnou ochotu a čas, bez jeho odborného vedení a cenných rad by tato práce nikdy nevznikla. Další poděkování patří RNDr. Ireně Dvořákové, Ph.D. za poskytnutí české verze testu a dat z testování na českých školách.

Název práce: Psychometrická analýza Lawsonova testu vědeckého uvažování

Autor: Martin Landa

Katedra: Katedra didaktiky fyziky

Vedoucí bakalářské práce: PhDr. Martin Chvál, Ph.D., Katedra didaktiky fyziky

Abstrakt: V předložené diplomové práci se zabýváme pojmem vědecké myšlení a problematikou určení úrovně vědeckého myšlení. Náplní práce bylo zpracování dat z České republiky, která byla shromážděná Lawsonovým testem vědeckého myšlení. Byla provedena psychometrická analýza úloh i celého testu, použita byla klasická teorie testů a Item Response Theory. Mimo jiné došlo k podrobnému analyzování distraktorů včetně komentářů nabízejících vysvětlení jejich volby respondenty. Následovalo určení úrovně vědeckého myšlení respondentů, tedy žáků českých základních škol a gymnázií. Na závěr bylo provedeno srovnání našich výsledků s výsledky českých i zahraničních studií.

Klíčová slova: vědecké myšlení, Lawsonův test, psychometrická analýza úloh a testu, interpretace výsledků

Title: Psychometric analysis of Lawson's test of formal reasoning

Author: Martin Landa

Department: Department of Physics Education

Supervisor: PhDr. Martin Chval, Ph.D., Department of Physics Education

Abstract: The thesis discusses the concept of formal reasoning and determination of its level. The content of the thesis was processing of data from the Czech republic, which was collected via Lawson's test of formal reasoning. Psychometric analysis of the questions and the test as a whole was made using the Classical Test Theory and Item Response Theory. Also detailed distractor analysis, including commentary offering a possible reason respondent might have chosen said distractor, was made. Next was determination of the level of formal reasoning of our respondents, who were Czech elementary level students. Lastly, comparison of our results with the results of Czech and foreign studies was made.

Keywords: formal reasoning, Lawson's test, psychometric analysis the test and its questions, interpretation of result

Obsah

Úvod	7
1 Vědecké myšlení a Lawsonův test	8
1.1 Vědecké myšlení	8
1.2 Lawsonův test vědeckého myšlení	9
1.3 Charakteristika zkoumaného Lawsonova testu	10
2 Metodologie analýzy Lawsonova testu	12
2.1 Zpracování dat	12
2.2 Analýza Lawsonova testu dle klasické testové teorie	13
2.3 Analýza Lawsonova testu dle teorie odpovědi na položku	14
2.4 Zavedení jednotné struktury	15
3 Výsledky analýzy dvojúloh Lawsonova testu	17
3.1 Úlohy 1 a 2	17
3.2 Úlohy 3 a 4	20
3.3 Úlohy 5 a 6	23
3.4 Úlohy 7 a 8	26
3.5 Úlohy 9 a 10	29
3.6 Úlohy 11 a 12	32
3.7 Úlohy 13 a 14	35
3.8 Úlohy 15 a 16	38
3.9 Úlohy 17 a 18	41
3.10 Úlohy 19 a 20	44
3.11 Úlohy 21 a 22	47
3.12 Úlohy 23 a 24	50
4 Výsledky analýzy Lawsonova testu jako celku	55
4.1 Výsledky s ohledem na kognitivní úroveň	55
4.2 Výsledky s ohledem na typ uvažování	56
4.3 Porovnání našich výsledků s výsledky jiných domácích a zahraničních studií	57
Závěr	62
Příloha	64
Literatura	65

Úvod

Tato diplomová práce byla vytvořena se záměrem vypracování psychometrické analýzy nejpoužívanější verze Lawsonova testu vědeckého uvažování. Vědecké uvažování je v dnešní době rychle se rozvíjející vědy a techniky důležitá kompetence pro úspěch právě v těchto odvětvích. Lawsonův test nabízí validní a snadno prakticky realizovatelný prostředek zjišťující úroveň vědeckého myšlení respondenta. Určení této úrovně může pomoci vzdělávacím institucím či přímo učitelům v přizpůsobení výuky za účelem lepšího rozvoje vědeckého myšlení žáků či studentů.

Součástí této práce byla příprava a úprava dat z České republiky shromážděná Lawsonovým testem v průběhu několika let. Následně byla provedena psychometrická analýza úloh i celého testu a to s využitím jak klasické teorie testů, tak Item Response Theory.

První kapitola práce se zabývá zavedením pojmu vědecké myšlení a stručně vysvětluje jeho podstatu a význam. Dále se krátce zabývá různými metodami určování úrovně vědeckého myšlení a propojuje vědecké myšlení s Lawsonovým testem. Nakonec podrobně popisuje verzi Lawsonova testu, kterou se tato práce zabývá.

V druhé kapitole je popsán způsob přípravy a úpravy poskytnutých dat, následně se práce zabývá metodami analyzování jednotlivých úloh i celého testu. Dále je uvedena jednotná forma prezentování výsledků analýz.

Ve třetí kapitole jsou prezentovány výsledky analýzy jednotlivých úloh formou dvanácti menších kapitol. Každá z těchto kapitol obsahuje výčet podstatných parametrů a komentář, který výsledky interpretuje.

Poslední kapitola prezentuje výsledky analýzy celého testu, a to dvěma různými pohledy na test. Na závěr je uvedeno srovnání našich výsledků s výsledky dalších českých i zahraničních analýz.

Motivací autora k vytvoření této práce byla kombinace jeho zájmů o pedagogiku, matematickou statistiku a hodnocení didaktických testů. Zároveň chtěl přispět k problematice vědeckého myšlení ve vzdělávání a doufá, že výsledky jeho práce budou k užitku mnoha učitelům, kteří mají zájem o zlepšení vědeckého myšlení svých žáků a studentů.

1 Vědecké myšlení a Lawsonův test

1.1 Vědecké myšlení

Pojem vědecké myšlení můžeme chápat jako soubor myšlenkových a úvahových procesů, které využíváme při vyšetřování, experimentování, vyhodnocování důkazů, odvozování a argumentování konceptů a teorií našeho světa [1]. Tyto dovednosti jsou považovány za zásadní pro co největší přispívání jedince ve společnosti 21. století [2]. Piaget [3] zavádí hypotézu kognitivního vývoje člověka, ve které rozlišuje čtyři hlavní stádia úrovně poznání. Konečné stádium, stádium formálních operací, charakterizuje výčet schopností jako je abstraktní myšlení, metodický postup, či vytváření hypotéz. Vědecké myšlení tedy úzce souvisí s pojmem formálního uvažování [4], v české literatuře se pojem vědeckého myšlení často substituuje pojmem kognitivního myšlení. Tato práce bude tyto dva pojmy volně zaměňovat.

Bao [5] ve své práci uvádí velmi malý vliv tradičního vyučování na přírodovědných a technických vysokých školách zaměřeného na memorování na úroveň vědeckého myšlení. Zároveň Cavallo [6] zmiňuje pozitivní korelaci mezi vědeckým uvažováním a lepšími studijními výsledky vysokoškolských studentů přírodovědných oborů. Obdobné výsledky poukazující na souvislost vědeckého uvažování a úspěchu v akademickém prostředí uvádí například Ates [7] a Jensen [8].

Tyto výsledky ukazují na význam vědeckého myšlení, který by měl být zohledněn ve vzdělávacích systémech. Vzdělávací instituce i učitelé potřebovali kvalitní, jednoduše použitelné metody určení úrovně vědeckého myšlení svých žáků a studentů. Díky určení této úrovně je možné použít vhodné výukové metody a formy, aby mohlo dojít k co nejlepšímu rozvoji vědeckého myšlení. Dále můžou instituce i učitelé zjistit vliv svých dosavadně používaných metod a forem výuky na úroveň vědeckého myšlení žáků a studentů.

Historicky byla jedna z prvních metod určování kognitivní úrovně Piagetova metoda klinického rozhovoru. Piaget, případně jiný tazatel, a respondent vedli rozhovor, který byl tvořen pokládáním otázek nebo představením nějakého problému. Tazatel mohl své otázky měnit na základě odpovědí respondenta za

účelem dosáhnouti co nejlepšího určení úrovně myšlení. Velkou výhodou metody rozhovoru bylo dobré pochopení důvodů odpovědi respondenta. Tato metoda byla ovšem velmi časově i finančně náročná, tudíž nebyla vhodná pro testování velkého množství žáků zároveň [9].

1.2 Lawsonův test vědeckého myšlení

Výzkumníci inspirovaní Piagetovými dotazy tedy začali vytvářet další metody testování, které by byly více vhodné pro použití ve škole. Nejúspěšnějšími metodami se staly tři různé testy, a to *Group Assessment of Logical Thinking* (GALT), *Test of Logical Thinking* (TOLT) a *Lawson's Classroom Test of Scientific Reasoning* (CTFR-78) [10], všechny umožňující určit úroveň vědeckého myšlení, ovšem různými způsoby a s jinou přesností. Původní Lawsonův test CTFR-78 vyvinutý Lawsonem v roce 1978 byl v roce 2000 předělán na novější verzi *Lawson classroom test of scientific reasoning* (LCTSR) [11], se kterou pracuje i tato práce, kdykoliv bude zmíněn Lawsonův test, budeme mít na mysli právě verzi LCTSR. Novější verze používá větší počet úloh, které formátu uzavřených otázek, zatímco původní verze používá otázky otevřené. Validita původní verze CTFR-78 byla několikrát ověřována (např. studie Pratt [12] a Stefanich [13]), u verze LCTSR nedošlo k dostatečnému ověření validity. Bao [14] píše o horší validitě nové verze u celkem pěti úloh, díky které jen relativně malý počet řešitelů dosahuje nejvyššího skóre. Po jejich úpravě by tento problém měl zaniknout.

Lawson ve svém původním testu vytvořil bodovací systém, který měl vzdělávacím institucím a učitelům pomoci interpretovat úroveň myšlení svých žáků a studentů vzhledem k Piagetovu třídění. Lawson definoval tři úrovně vědeckého myšlení: *Konkrétně operační kognitivní úroveň* (0-5 bodů), *Přechodně kognitivní úroveň* (6-11 bodů) a *Formálně operační úroveň* (12-15 bodů). Stadium konkrétních operací Piaget [3] definuje jako „období, které znamenají provázanost myšlení na konkrétní obsah. V tomto věku většina dětí dospívá k principu konzervace – když se mění tvar, nemění se množství, lépe si uvědomují příčinnost, jsou schopny lépe pojmenovat vnější příčiny jevů, lépe uvažují i o tom, jak prožívají a přemýšlejí ostatní lidé. Nejsou však ještě schopny abstrakce.“. Stadium formálních operací je charakterizováno „schopností abstraktního myšlení, metodického postupu, vytváření hypotéz apod.“.

Přechodové stádium Piaget nezavádí, jedná se ale o stádium mezi předchozími dvěma stádii. Novější verze testu používá stejné úrovně vědeckého myšlení, vzhledem k jinému počtu úloh má ale jiné rozdělení těchto úrovní podle bodů.

1.3 Charakteristika zkoumaného Lawsonova testu

Anglickou verzi zkoumaného Lawsonova testu můžeme nalézt v odkazu [11]¹. Zkoumaný test vědeckého myšlení obsahuje 24 úloh. Všechny úlohy jsou uzavřené s výběrem ze tří až pěti odpovědí, správná je vždy právě jedna odpověď. Úlohy jsou dále sestaveny do dvojic. První úloha z dvojice představí problém a většinou se ptá na nějaký jev, druhá úloha zase předkládá různá vysvětlení onoho jevu.

Při standardním vyhodnocení řešitel získá dva body, pouze pokud odpoví na obě otázky ze dvojice správně, v ostatních případech nedostává bod žádný. Výjimkou je poslední dvojice úloh, které se bodují nezávisle, každá po jednom bodu. Maximálně může řešitel získat 24 bodů. Existují i alternativní verze bodování, jako např. bodování každé úlohy zvlášť jedním bodem, které vedou k alternativní interpretaci výsledků. Bodování v naší verzi vede k výraznému snížení uhodnutí správného řešení, jelikož by musel řešitel pro získání bodů uhodnout vždy obě části dvojúlohy.

¹ Tuto verzi přeložil do češtiny L. Dvořák. Ve snaze zabránit všeobecnému zveřejnění české verze (neboť chceme zachovat vypovídající hodnotu testu) nebude tato práce na českou verzi odkazovat.

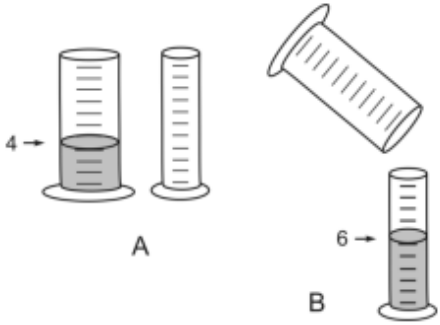
Dále je možné úlohy rozdělit do sedmi různých kategorií zabývajících se různými typy uvažování.

Rozdělení je následující [15]:

- Zachování hmotnosti - úlohy 1+2
- Zachování vytlačeného objemu - úlohy 3+4
- Poměrové myšlení - úlohy 5+6, 7+8
- Identifikace a kontrola změny - úlohy 9+10, 11+12, 13+14
- Pravděpodobnostní myšlení - úlohy 15+16, 17+18
- Korelační myšlení - úlohy 19+20
- Kombinační myšlení zahrnující prvky z předchozích oblastí - úlohy 21+22, 23+24

Následující obrázek ukazuje typický formát úlohy Lawsonova testu:

5. Na obrázku vpravo jsou zobrazeny široký a úzký válec. Na válcích jsou značky ve stejných vzdálenostech od sebe. Do širšího válce je nalita voda po 4. značku (viz A). Když je voda přelita do úzkého válce, vystoupí po 6. značku (viz B).



Oba válce jsou vylity (není znázorněno) a pak je do širšího válce nalita voda po 6. značku.
Jak vysoko vystoupí voda, když ji přelijeme do prázdného úzkého válce?

a. asi k 8. značce
b. asi k 9. značce
c. asi k 10. značce
d. asi k 12. značce
e. žádná z předchozích odpovědí není správná

6. *protože*

a. odpověď nelze určit ze zadaných informací.
b. v prvním případě stoupla o 2 značky, takže opět vystoupí o 2 značky.
c. pro každé 2 značky v širokém válci stoupne o 3 značky v úzkém.
d. druhý válec je užší.
e. vodu musíme do válců skutečně nalít a pozorovat, co se stane.

Obrázek. 1.1: Příklad dvojice otázek v Lawsonově testu [11]

2 Metodologie analýzy Lawsonova testu

2.1 Zpracování dat

Původní data, se kterými bylo pracováno, poskytla RNDr. Irena Dvořáková Ph.D. z MFF UK. Jednalo se o data v souboru Excel z testování celkem 1422 žáků českých základních škol a gymnázií ve věku od 13 do 16 let shromážděná v období od roku 2011 do roku 2018. Průměrný věk řešitele byl 15,0 let. V poskytnutých datech je u každého jednotlivého řešitele uvedeno jeho pohlaví, celkový počet získaných bodů, jeho řešení každé z 24 úloh, škola a věk. Dále poskytnutý materiál uváděl některé výsledky základní analýzy těchto dat jako úspěšnost v jednotlivých kategoriích typů uvažování, četnost správných odpovědí a rozložení celkového skóre.

Z původního množství 1422 řešitelů jich bylo 186 vyřazeno. 8 řešitelů zadalo více než jednu odpověď na úlohu, 167 na úlohu neodpovědělo. Dále bylo ve vyhodnocovaných datech 11 chyb formátu neplatné odpovědi. Výsledný počet řešitelů, kteří byli použiti pro analýzu, je 1236. Původní data pracují se všemi 1422 řešiteli, s ohledem na to nebude v této práci použit žádný z původních výsledků.

Data byla následně upravena do různých formátů vhodných pro další analýzu. Zatímco řešení každého řešitele v původních datech byla uvedena ve formátu:

úloha - odpověď

(například: úloha 3 - odpověď b),

v novém podání je formát podaný stylem:

úloha - odpověď - ano/ne

(například: úloha 3 - odpověď a - ne, odpověď b - ano, odpověď c - ne atd.)

díky kterému bylo možné provést hlubší analýzu.

Vzhledem ke způsobu standardního vyhodnocování byl následně vytvořen formát stejného stylu, který se ale zajímá o řešení celých dvojic úloh.

S ohledem na vysoký počet možných kombinací odpovědí, z nichž některé mají u dané dvojúlohy velmi nízké zastoupení, by velké množství distraktorů

mělo příliš malou hodnotu atraktivity určenou jako relativní četnost. Byl tedy zaveden pojem relevantní dvojice odpovědí. Odpověď je považována za relevantní, pokud ji zvolilo alespoň 68 řešitelů ($\approx 5,5\%$) ze všech. Všechny zbylé distraktory, které měly příliš malou hodnotu atraktivity, byly sečteny a označeny jako relevantní dvojice *Ostatní*. U všech relevantních distraktorů byl určen korelační koeficient r s celkovým skóre v testu. Pro záporné hodnoty r můžeme distraktor považovat za dobrý, jelikož jeho zvolení vede k horšímu výsledku v testu. Pro hodnoty blízké nule či kladné je distraktor problémový. Oba případy jsou vždy diskutovány v komentáři u příslušné dvojúlohy.

2.2 Analýza Lawsonova testu dle klasické testové teorie

Po zpracování dat proběhla jejich analýza pomocí klasické testové teorie (dále jen CTT). Prvním výstupem CTT je úspěšnost při řešení dvojúlohy, která byla počítána jako procentuální zastoupení řešitelů, kteří danou dvojúlohu vyřešili správně.

Dalším výstupem je diskriminace položky respektive item discrimination dané dvojúlohy. K určení diskriminace položky byli rozděleni respondenti na dvě skupiny: skupinu s horšími výsledky a skupiny s lepšími výsledky. Respondenti byly na základě jejich úspěšnosti v testu seřazeny dle jejich celkových výsledků. Následně bylo vybráno vrchních 27 % řešitelů (UG) a spodních 27 % řešitelů (LG), kteří tvoří tyto dvě skupiny. Index diskriminace položky (D) byl počítán jako koeficient ULI daný vztahem [16]:

$$D = \frac{UG-LG}{N}, \quad (1)$$

kde N je celkový počet řešitelů. Čím vyšší je hodnota indexu D , tím lépe dokáže úloha rozlišovat mezi řešiteli s vysokým a nízkým celkovým výsledkem. Na základě Ebelových [17] pokynů pro vypracování CTT jsou položky rozlišeny dle hodnoty indexu D . Položky s hodnotou $0 < D < 0,19$ jsou považovány za špatné, $0,20 < D < 0,29$ za přijatelné, $0,30 < D < 0,39$ za dobré a $D > 0,40$ výborné. Položky se zápornou hodnotou D jsou vyřazeny.

Dále byla určena reliabilita testu. Reliabilita byla vypočtena jako Cronbachovo alfa pomocí vztahu [16]:

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k \sigma_i^2}{\sigma_t^2}\right), \quad (2)$$

kde k je počet položek, σ_i^2 je rozptyl i -té položky a σ_t^2 je rozptyl celého testu.

Reliabilita testu vyjadřuje, zda při opakovaném použití testu dostaneme podobné výsledky. Reliabilita nabývá hodnot 0 až 1 a lze ji chápat jako procentuální nepřítomnost chyby měření. Z analyzovaných dat bylo Cronbachovo alfa Lawsonova testu určeno jako $\alpha = 0,72$. Dále bylo určeno, jak by se tato hodnota změnila v závislosti na odstranění určité dvojúlohy. Tím bylo určeno, zda by se při odstranění některé dvojúlohy nezlepšila reliabilita testu.

Nakonec byla spočtena korelace mezi každou konkrétní relevantní odpovědí na dvojúlohu a celkovým počtem bodů v testu, čímž můžeme rozpoznat odpovědi, které měli velký či malý vliv na celkový výsledek v testu.

Všechny tři tyto koeficienty přinášejí obdobnou interpretaci, cílem určení koeficientů je rozlišení lepších a horších úloh testu.

2.3 Analýza Lawsonova testu dle teorie odpovědi na položku

Dále byla provedena analýza dle teorie odpovědi na položku (dále jen IRT). Hlavním výstupem jsou IRT grafy charakteristických křivek úlohy a parametry těchto křivek logistické regrese.

K analýze byl využit program ShinyItemAnalysis [18]. Vzhledem k formátu dat byly využity dva modely IRT.

Dichotomní dvouparametrový model (2PL) předpokládá položky lišící se v parametru obtížnosti (b_i), což je reprezentováno pozicí inflexních bodů na charakteristické křivce úlohy. Dále předpokládá, že se položky liší parametrem diskriminace (a_i), což je reprezentováno strmostí charakteristické křivky úlohy kolem jejího inflexního bodu. IRT parametr diskriminace nám podává stejný typ informace jako CTT index diskriminace z (2), můžeme tedy očekávat konzistentní výsledky diskriminace položky. Pravděpodobnost správné odpovědi respondenta p na položku i je:

$$P(\theta_p) = \frac{e^{a_i(\theta_p - b_i)}}{1 + e^{a_i(\theta_p - b_i)}}, \quad (3)$$

kde θ_p je parametr schopnosti řešitele, kterou si můžeme představit jako souhrnný výsledek v testu.

Druhým používaným modelem je Nominal response model s parametrizací Baseline-category Logit Intercept-Slope (BLIRT). BLIRT využívá stejné dva parametry jako model předchozí, pouze pracuje s nominálním tvarem dat a zároveň fixuje hodnoty (b_i) a (a_i) pro správnou odpověď na hodnotu 0 a ostatní hodnoty (b_i) a (a_i) dopočítává vzhledem k této fixaci.

2.4 Zavedení jednotné struktury

Výsledky analýzy Lawsonova testu jsou nejprve prezentovány formou dvanácti kapitol, každá zabývající se příslušnou dvojúlohou. Struktura těchto kapitol je následující:

- Zadání obou úloh se všemi možnostmi odpovědí společně se správným řešením.
- Kategorie typu uvažování, který dvojúloha testuje.
- Hodnota indexu diskriminace položky.
- Změna hodnoty Cronbachova alfa při odstranění dvojúlohy (dále jen Změna reliability).
- Tabulka se zastoupením všech jednotlivých odpovědí.
- Tabulka se zastoupením relevantních odpovědí a jejich hodnoty parametrů obtížnosti b a diskriminace a .
- IRT graf závislosti pravděpodobnosti zvolení správné odpovědi na schopnostech řešitele.
- IRT graf závislosti pravděpodobnosti zvolení dané relevantní odpovědi na schopnostech řešitele.
- Komentář k výsledkům.

Číslo úlohy	Možné odpovědi		Relevantní odpovědi		Parametry diskriminace a obtížnosti			
	5	6	Zastoupení	Kor. koef	a	b	Správná odpověď	
a	26,0%	2,0%	ab	23,22%	-0,47	-17,36	-0,69	
b	69,0%	24,0%	bc	57,61%	0,59	1,49	-0,29	
c	3,0%	61,0%	bd	9,55%	-0,07	-1,11	-1,34	
d	1,0%	12,0%	Ostatní	9,63%	-0,24	-3,48	-0,59	
e	1,0%	1,0%						

Tabulka zastoupení všech odpovědí Tabulka relevantních odpovědí

Obrázek. 2.2: Ukázka tabulek zastoupení

V tabulkách jsou zeleně označeny správné odpovědi. V tabulce relevantních odpovědí byly hodnoty parametrů obtížnosti b a diskriminace a určeny pomocí modelu BLIRT. Vzhledem k tomu, že model fixuje tyto hodnoty pro správnou odpověď na 0, byly hodnoty těchto parametrů v případě správné odpovědi určeny z modelu 2PL. Díky tomu můžeme provádět srovnání s jinými dvojúlohami. Pokud by nebyl použit tento druhý model, tak přestože budou mít různé správné odpovědi různé hodnoty obtížnosti a diskriminace, model BLIRT by je všechny fixoval na hodnotu 0 a srovnání by nebylo možné.

První IRT graf charakteristických křivek úlohy byl získán pomocí modelu 2PL. Z grafu můžeme vyčíst, s jakou pravděpodobností zvolí řešitel s určitou schopností správnou odpověď.

Druhý IRT graf charakteristických křivek úlohy byl získán pomocí modelu BLIRT. Z grafu můžeme vyčíst, s jakou pravděpodobností zvolí řešitel s určitou schopností danou relevantní odpověď. Každá křivka odpovídá dané relevantní odpovědi.

Komentář interpretuje výše zmíněné výsledky a porovnává je s jinými dvojúlohami. Dále se zabývá analýzou konkrétních relevantních odpovědí a pokouší se podávat vysvětlení pro jejich zvolení řešitelem.²

Po této analýze dvojúloh následuje analýza Lawsonova testu jako celku, kde jsou nejprve prezentovány naše výsledky analýzy zaměřené na kognitivní úroveň řešitelů, které jsou následně srovnávány s výsledky dalších českých a zahraničních studií. Poté jsou prezentovány naše výsledky analýzy zaměřené na různé kategorie uvažování, do kterých jsou dvojúlohy rozdělené. Výsledky jsou opět srovnány s výsledky dalších studií.

² Celkové shrnutí všech relevantních parametrů lze dohledat v příloze.

3 Výsledky analýzy dvojúloh Lawsonova testu

3.1 Úlohy 1 a 2

Zadání

1. Máte dvě hliněné kuličky stejné velikosti a tvaru. Obě kuličky mají také stejnou váhu. Jedna kulička je pak splácnuta do tvaru koláčku. *Které z následujících tvrzení je správné?*
 - a. Splácnutý kousek hlíny váží víc než kulatý
 - b. Oba kousky váží stejně
 - c. Kulička váží víc než splácnutý kousek

2. *protože*
 - a. spláclý kousek pokrývá větší plochu.
 - b. kulička tlačí dolů víc na jednom místě.
 - c. když se něco splácne, ztratí to na váze.
 - d. hlína nebyla přidána ani odebrána.
 - e. když se něco splácne, získá to na váze.

Obrázek 3.1.1: Zadání úloh 1. a 2. [11]

Správné řešení b. + d.

Kategorie Zachování hmotnosti

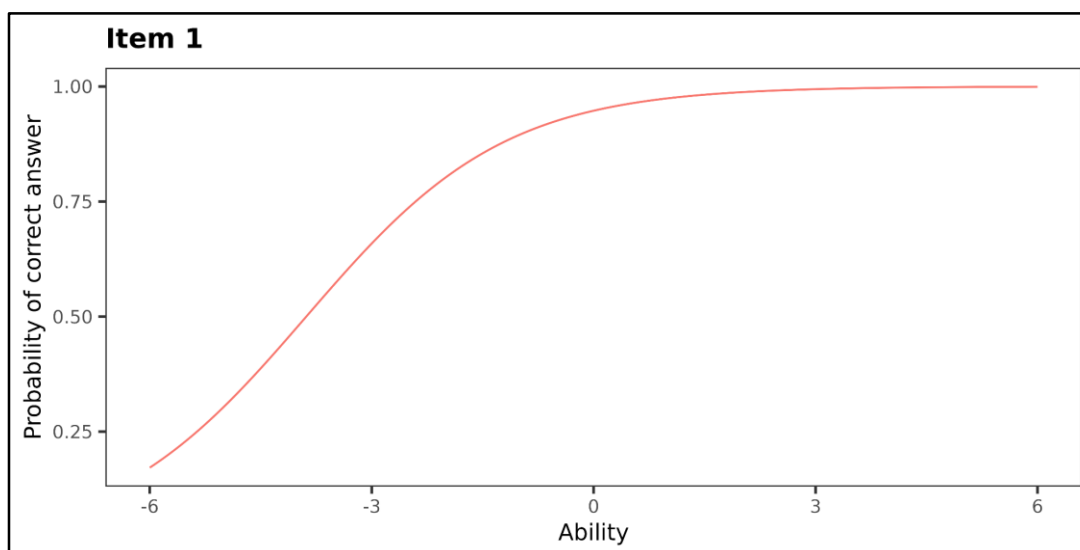
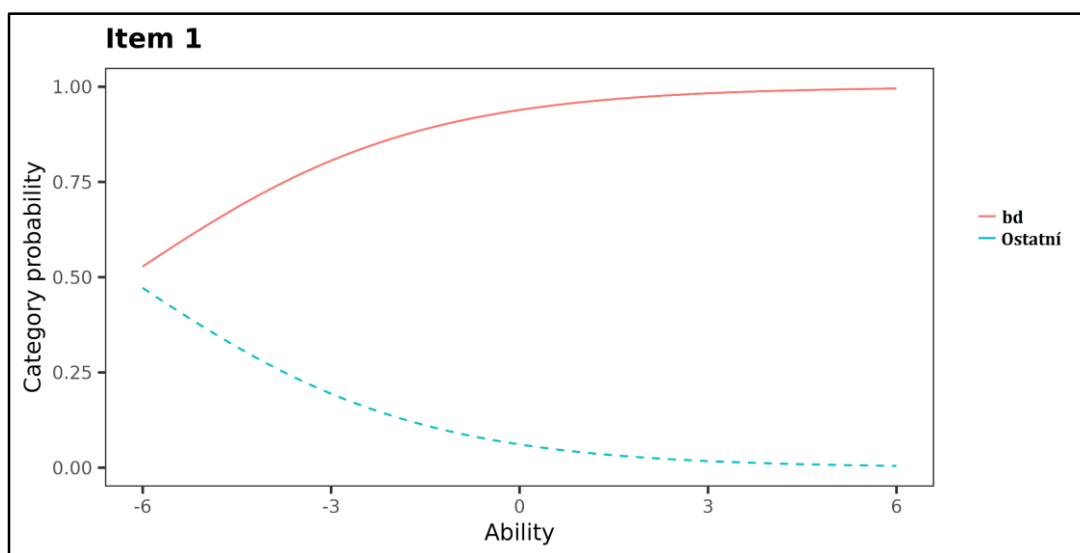
Úspěšnost 93 %

Index diskriminace položky ULI 0,13

Změna reliability Při odstranění dvojice úloh 1. a 2. nedojde ke změně hodnoty Cronbachovo alfa, hodnota tedy zůstává 0,72.

Tabulky 3.1.1 a 3.1.2 : Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	1	2		Zastoupení	Kor. koef	a	b
a	1,6%	2,3%	bd	93,45%	0,24	0,74	-3,88
b	93,9%	2,7%	Ostatní	6,55%	-0,24	-0,44	-6,26
c	4,5%	0,7%					
d	-	93,8%					
e	-	0,5%					



Obrázky 3.1.2 a 3.1.3: Grafy křivek předmětové charakteristiky

Úloha 1;2 má ze všech úloh výrazně nejvyšší úspěšnost, jde tedy o úlohu velmi jednoduchou. Jedná se o jedinou úlohu kategorie *Zachování hmotnosti*, výsledky v této kategorii nekorelují s výsledky v žádné jiné kategorii, což může být dáno právě i její vysokou úspěšností.

Úlohy 1. i 2. jednotlivě nemají žádný relevantní distraktor, což odpovídá nízké obtížnosti úlohy. Úloha by se dala považovat za příliš jednoduchou, jedná se o jednu ze dvou úloh, u kterých by nedošlo ke změně Cronbachovo alfa při jejím odstranění, nedošlo by tedy ke zhoršení kvality testu. Zároveň má také nejmenší hodnotu indexu diskriminace položky. Vzhledem k tomu, že se jedná o první úlohu, a je vhodné mít lehčí první úlohu za účelem uklidnění řešitele, je nízká obtížnost úlohy ospravedlněna. Zároveň úloha dobře rozlišuje mezi nejslabšími řešiteli.

Vzhledem k nízkému počtu možných odpovědí na úlohu 1. je větší pravděpodobnost uhádnutí správné odpovědi, ovšem téměř všichni řešitelé, kteří zvolili správnou odpověď *b*, zvolili i správné vysvětlení *d*.

Nejběžnější distraktor úlohy 1. *c* byl nejčastěji pojen s vysvětleními *a* a *b*, což poukazuje na špatné řešitelovo pochopení fyzikálních konceptů tlaku a hmotnosti. Tato kombinace měla s celkovým skóre téměř nulovou korelaci. Z grafu vidíme, že tuto možnost volili i celkově úspěšnější studenti, kteří sice mohli chápat zachování poměru, nedokázali ho už ovšem správně odůvodnit.

Ze všech úloh testu má správná odpověď *bd* nejslabší korelaci s celkovým skóre $r = 0,24$.

3.2 Úlohy 3 a 4

Zadání

3. Na pravé straně jsou obrázky dvou válců naplněných vodou do stejné výšky. Válce mají stejnou velikost a tvar.

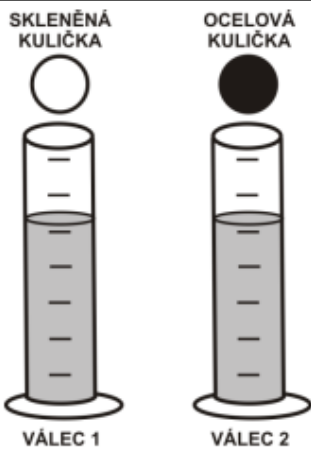
Na obrázku jsou ukázány také dvě kuličky, jedna skleněná a jedna ocelová. Kuličky jsou stejně velké, ale ocelová je mnohem těžší než skleněná.

Když je skleněná kulička vložena do válce 1, potopí se na dno a voda stoupne k šesté značce. Jestliže vložíme ocelovou kuličku do válce 2, voda stoupne:

a. do stejné výšky jako stoupla ve válci 1
b. do větší výšky než do níž stoupla ve válci 1
c. do menší výšky než do níž stoupla ve válci 1

4. protože

a. ocelová kulička se potopí rychleji.
b. kuličky jsou z různých materiálů.
c. ocelová kulička je těžší než skleněná.
d. skleněná kulička vyvolá menší tlak.
e. kuličky mají stejnou velikost.



Obrázek 3.2.1: Zadání úloh 3. a 4. [11]

Správné řešení a. + e.

Kategorie Zachování vytlačeného objemu

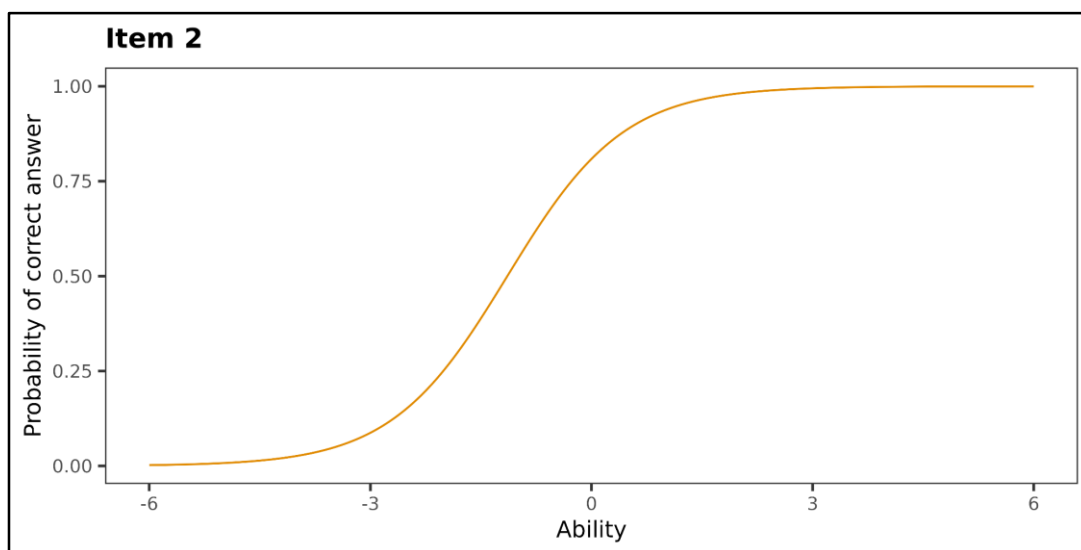
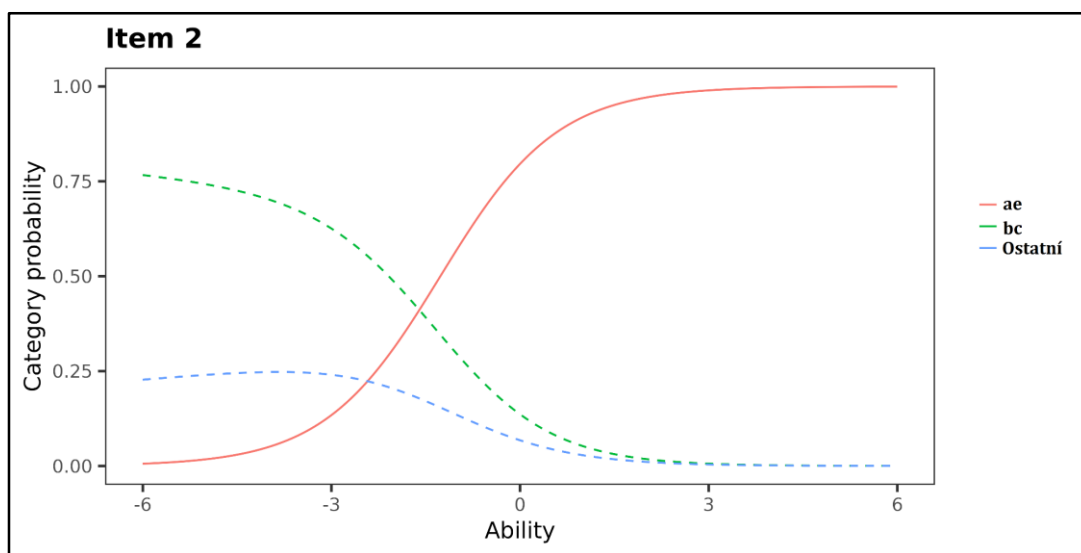
Úspěšnost 75 %

Index diskriminace položky ULI 0,54

Změna reliability Při odstranění dvojice úloh 3. a 4. dojde ke zhoršení hodnoty Cronbachovo alfa z 0,72 na 0,70.

Tabulky 3.2.1 a 3.2.2 : Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	3	4		Zastoupení	Kor. koef	a	b
a	76,1%	0,6%	ae	75,24%	0,50	1,26	-1,14
b	22,9%	2,5%	bc	16,83%	-0,39	-1,10	-1,60
c	1,0%	17,8%	Ostatní	7,93%	-0,26	-1,02	-2,43
d	-	3,2%					
e	-	75,8%					



Obrázky 3.2.2 a 3.2.3: Grafy křivek předmětové charakteristiky

Úloha 3;4 má ze všech úloh druhou nejvyšší úspěšnost, můžeme ji tedy považovat za úlohu snadnou. Jedná se o jedinou úlohu kategorie *Zachování vytlačeného objemu*, výsledky v této kategorii mají slabou korelaci s výsledky v kategorii *Pravděpodobnostní myšlení* $r = 0,25$.

Úloha 3. má jediný relevantní distraktor *b*, úloha 4. jediný relevantní distraktor *c*. Řešitel volí tuto kombinaci distraktorů kvůli prisouzení většího stoupnutí hladiny větší hmotnosti ocelové kuličky.

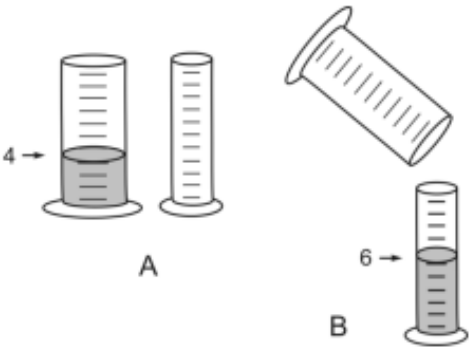
Přestože se jedná o úlohu na zachování, má tato úloha na rozdíl od úlohy předchozí mnohem lepší hodnotu indexu diskriminace. To může být způsobeno větším množstvím fyzikálních konceptů, které musí vzít řešitel v potaz při řešení této úlohy oproti úloze předchozí.

Ze všech úloh testu má správná odpověď *ae* pátou nejsilnější pozitivní korelaci s celkovým skóre $r = 0,50$. Distraktor *bc* má s celkovým skóre korelaci $r = -0,39$ což je třetí nejsilnější negativní korelace. Porozumění či neporozumění zachování vytlačeného objemu tedy hraje velkou roli v určení úrovně kognitivního vývoje řešitele.

3.3 Úlohy 5 a 6

Zadání

5. Na obrázku vpravo jsou zobrazeny široký a úzký válec. Na válcích jsou značky ve stejných vzdálenostech od sebe. Do širšího válce je nalita voda po 4. značku (viz A). Když je voda přelita do úzkého válce, vystoupí po 6. značku (viz B).



Oba válce jsou vylity (není znázorněno) a pak je do širšího válce nalita voda po 6. značku.
Jak vysoko vystoupí voda, když ji přelijeme do prázdného úzkého válce?

a. asi k 8. značce
b. asi k 9. značce
c. asi k 10. značce
d. asi k 12. značce
e. žádná z předchozích odpovědí není správná

6. *protože*

a. odpověď nelze určit ze zadaných informací.
b. v prvním případě stoupla o 2 značky, takže opět vystoupí o 2 značky.
c. pro každé 2 značky v širokém válci stoupne o 3 značky v úzkém.
d. druhý válec je užší.
e. vodu musíme do válců skutečně nalít a pozorovat, co se stane.

Obrázek 3.3.1: Zadání úloh 5. a 6. [11]

Správné řešení b. + c.

Kategorie Poměřové myšlení

Úspěšnost 58 %

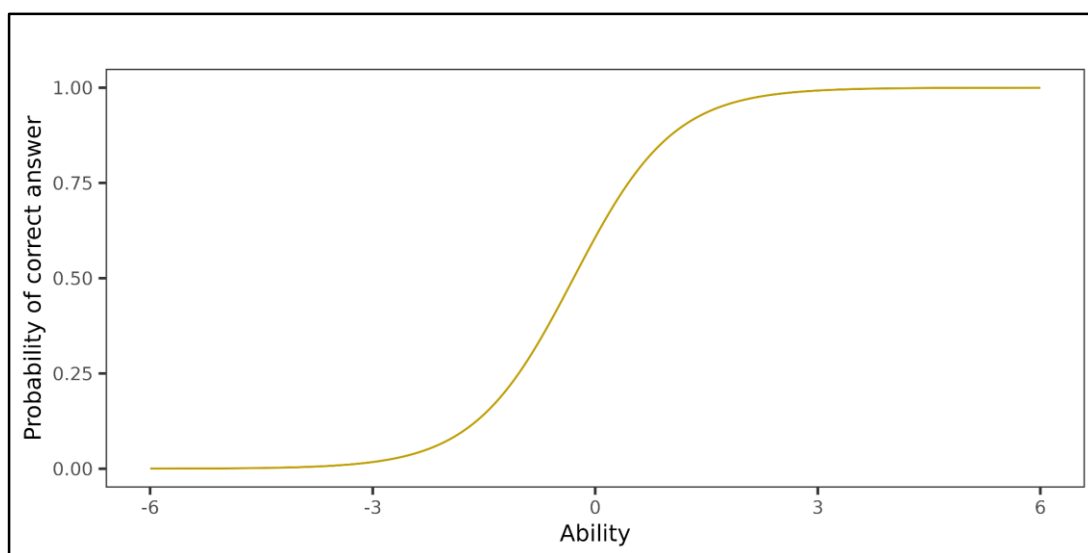
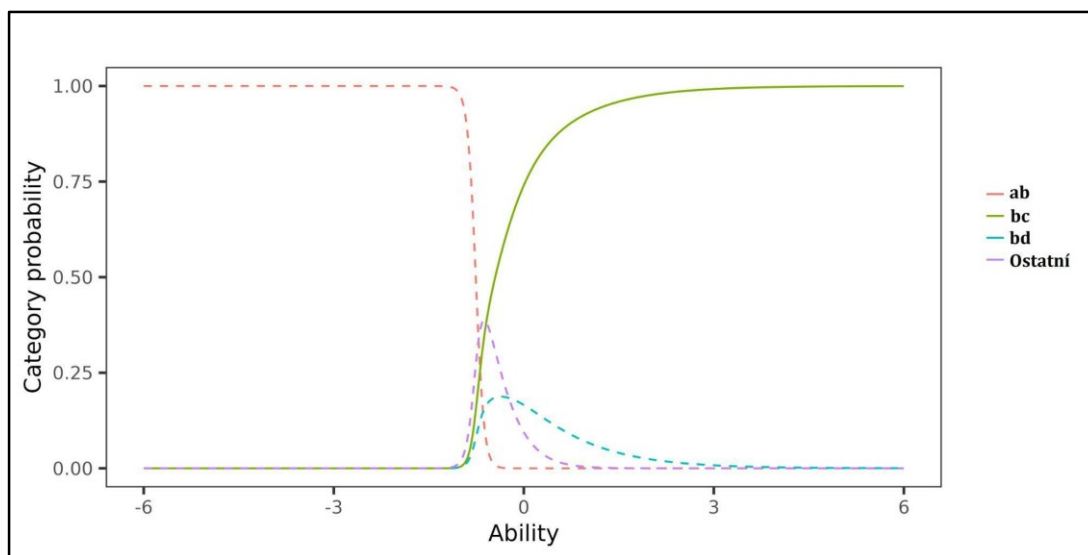
Index diskriminace položky ULI 0,73

Reliabilita Při odstranění dvojice úloh 5. a 6. dojde ke zhoršení hodnoty

Cronbachovo alfa z 0,72 na 0,69.

Tabulky 3.3.1 a 3.3.2 : Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	5	6		Zastoupení	Kor. koef	a	b
a	25,8%	1,8%	ab	23,22%	-0,47	-17,36	-0,69
b	68,5%	24,1%	bc	57,61%	0,59	1,49	-0,29
c	3,1%	60,5%	bd	9,55%	-0,07	-1,11	-1,34
d	1,2%	12,3%	Ostatní	9,63%	-0,24	-3,48	-0,59
e	1,4%	1,3%					



Obrázky 3.3.2 a 3.3.2: Grafy křivek předmětové charakteristiky

Úloha 5;6 má ze všech úloh pátou nejvyšší úspěšnost, můžeme ji tedy považovat za úlohu snazší. Jedná se o první ze dvou úloh kategorie *Poměrové myšlení*, přesto je mezi nimi pouze slabá pozitivní korelace $r = 0,30$.

Úloha 5. má jediný relevantní distraktor *a*. Řešitel tento distraktor pravděpodobně volí kvůli neporozumění vlivu rozdílných průměrů válců na poměr, což můžeme vidět na častém chybném zdůvodnění *b*. Správná odpověď *b* je nejčastěji spojená se správným vysvětlením *c*, které je v podstatě popsáním zachování konstantního poměru. Řešitel chápající např. přímou úměrnost by tedy měl být při řešení této úlohy zvýhodněn. Dále byla odpověď *b* často pojena s vysvětlením *d*. Vysvětlení *d* by se dalo považovat za správné, ale nedostatečně přesné. Vysvětlení *d* je totiž správné pro jakoukoliv z odpovědí *a*, *b*, *c*, *d*. Nevysvětluje ovšem, proč jediná správná odpověď právě *b*.

Tato kombinace měla s celkovým skóre téměř nulovou korelaci. Z grafu vidíme, že tuto možnost volili i celkově úspěšnější studenti, kteří sice mohli chápat zachování poměru, nedokázali ho už ovšem správně odůvodnit.

Ze všech úloh testu má správná odpověď *bc* třetí nejsilnější korelaci s celkovým skóre $r = 0,59$. Oproti tomu distraktor *ab* má ze všech úloh testu s celkovým skóre nejsilnější zápornou korelaci $r = -0,47$. Neporozumění zachování poměru má tedy značnou souvislost s úrovní kognitivního vývoje řešitele.

3.4 Úlohy 7 a 8

Zadání

7. Nyní nalijeme vodu do úzkého válce (popsaného výše v bodě 5) k 11. značce. *Jak vysoko voda vystoupí, když ji přelijeme do prázdného širokého válce?*
- a. asi na 7 ½ (polovinu nad 7 .značku)
 - b. asi k 9. značce
 - c. asi k 8. značce
 - d. asi k 7 1/3 (třetinu nad 7. značku)
 - e. žádná z předchozích odpovědí není správná
8. *protože*
- a. poměry musí zůstat stejné.
 - b. vodu musíme do válců skutečně nalít a pozorovat, co se stane.
 - c. odpověď nelze určit ze zadaných informací.
 - d. předtím byla výška o 2 nižší, tak musí být zas o 2 nižší.
 - e. pro každé 3 značky v úzkém válci odečtete 2 v širším.

Obrázek 3.4.1: Zadání úloh 7. a 8. [11]

Správné řešení d. + a.

Kategorie Poměřové myšlení

Úspěšnost 30 %

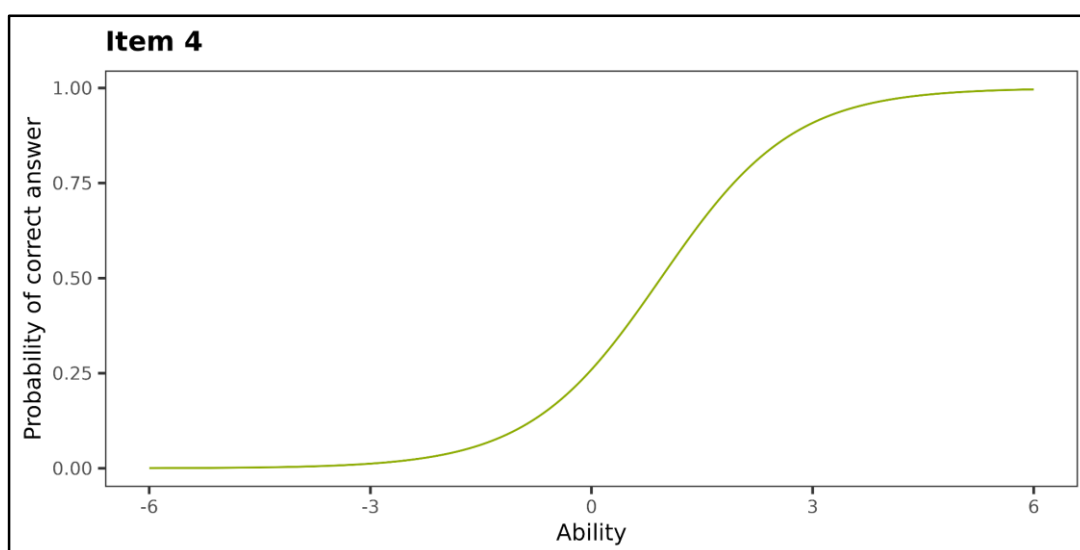
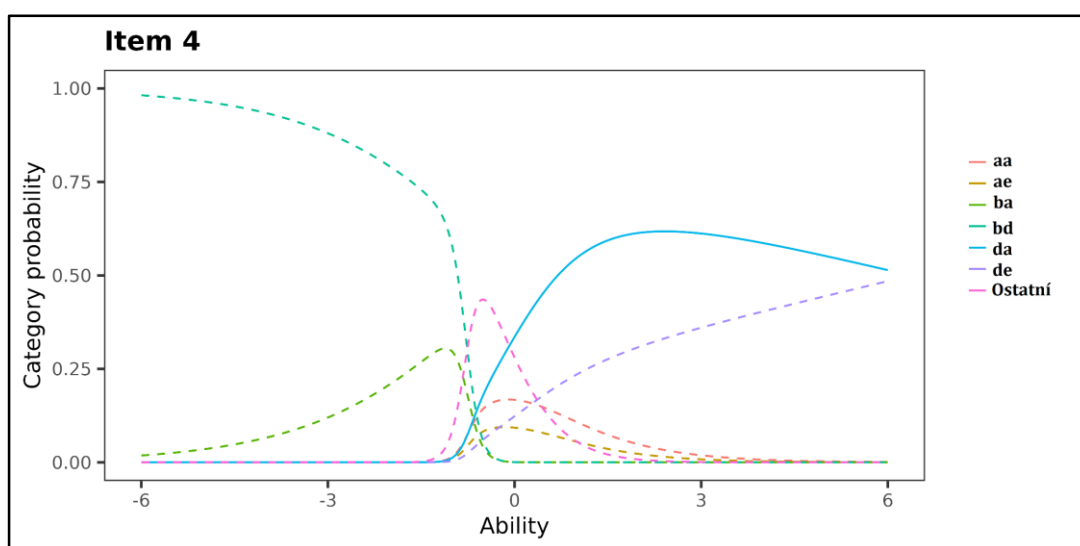
Index diskriminace položky ULI 0,58

Reliabilita Při odstranění dvojice úloh 7. a 8. dojde ke zhoršení hodnoty

Cronbachovo alfa z 0,72 na 0,70.

Tabulky 3.4.1 a 3.4.2 : Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	7	8		Zastoupení	Kor. koef	a	b
a	17,6%	55,1%	aa	10,36%	-0,20	-0,93	-0,75
b	25,6%	1,5%	ae	5,66%	-0,03	-1,01	-1,27
c	4,9%	2,5%	ba	7,28%	-0,20	-9,38	-0,66
d	43,0%	17,3%	bd	15,86%	-0,37	-10,04	-0,62
e	9,0%	23,5%	da	30,10%	0,50	1,11	0,94
			de	12,38%	0,16	0,16	6,37
			Ostatní	18,37%	-0,20	-2,11	-0,09



Obrázky 3.4.2 a 3.4.3: Grafy křivek předmětové charakteristiky

Úloha 7;8 má ze všech úloh čtvrtou nejnižší úspěšnost, můžeme ji tedy považovat za úlohu obtížnější. Jedná se o druhou ze dvou úloh kategorie *Poměrové myšlení*.

Úloha 7. má čtyři relevantní distraktory, úloha 8. má tři. Kombinace *bd* byla nejčastěji zvolená řešiteli s celkově horšími výsledky, zároveň se jedná o nejčastější distraktor. Řešitelé tuto možnost pravděpodobně vybírají kvůli přímo nabídnuté souvislosti v předchozí úlohou ve vysvětlení *d*.

Překvapivé je, že četnost výběru distraktoru *de* roste s celkovým skóre řešitelů stejně jako správné řešení *da*, ovšem neúspěšnější řešitelé volili správné řešení méně často, než méně úspěšní řešitelé. Možným vysvětlením je přílišná jednoduchost vysvětlení *a* oproti vysvětlení *e*. Zároveň je formát vysvětlení *e* stejný, jako je správné vysvětlení v předchozí dvojúloze, řešitel tedy může být sváděn k použití stejného formátu. Tvrzení *e* je sice pravdivé ale nevysvětluje odpověď *d*. Han [10] ve své práci uvádí myšlenku, že i řešení *e* by se dalo považovat za správné a navrhuje tento distraktor předělat, s čímž, vzhledem ke kladné korelaci distraktoru s celkovým skóre, souhlasíme.

Ostatní kombinace byly voleny řešiteli s průměrným celkovým výsledkem. Jejich výběr můžeme přisoudit špatnému počítání v první úloze a následné snaze odůvodnit špatný výpočet. Většina řešitelů ovšem i tak pravděpodobně intuitivně věděla, že poměr by se měl zachovat a volili správné vysvětlení *a*.

Ze všech úloh testu má správná odpověď *da* šestou nejsilnější korelaci s celkovým skóre $r = 0,50$. Distraktor *bd* má s celkovým skóre nejsilnější zápornou korelaci $r = -0,39$ z ostatních distraktorů. Distraktor *de* má s celkovým skóre slabou pozitivní korelaci $r = 0,15$. Přestože řešitel vyřešil úlohu špatně, schopnost správně spočítat vystoupení hladiny a zvolit odpověď *d* i tak předpovídá větší úspěšnost v testu. Můžeme tedy tuto schopnost považovat za podstatnou v kognitivním vývoji řešitele.

Distraktory *aa* a *ae* mají velmi slabou negativní korelaci s celkovým výsledkem, distraktor *ba* a *Ostatní* mají slabou negativní korelaci, uchýlovali se k nim tedy celkově spíše slabší žáci.

3.5 Úlohy 9 a 10

Zadání

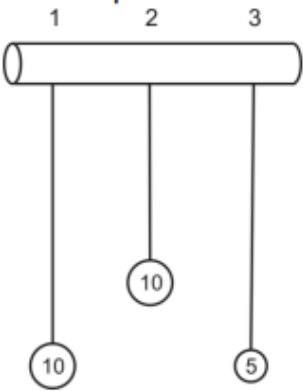
9. Obrázek vpravo ukazuje tři provázky visící z tyčky. Na koncích všech tří provázků jsou připevněna kovová závaží. Provázky 1 a 3 mají stejnou délku. Provázek 2 je kratší. Na konci provázku 1 je závaží o tíze 10 jednotek. Závaží o tíze 10 jednotek je připevněno i na konci provázku 2. Závaží o tíze 5 jednotek je připevněno na konci provázku 3. Provázky a k nim připevněná závaží se mohou kývat a můžeme měřit dobu jejich kmitu.

Řekněme, že byste chtěli určit, zda má délka provázku vliv na dobu, za níž se kývá tam a zpět. Které provázky byste si vybrali, abyste to zjistili?

a. jen jeden provázek
b. všechny tři provázky
c. 2 a 3
d. 1 a 3
e. 1 a 2

10. protože

a. musíme užít nejdelší provázek.
b. musíme porovnat provázky s lehkým a s těžkým závažím.
c. se liší jen délka.
d. potřebujeme provést všechna možná srovnání.
e. se liší hmotností.



Obrázek 3.5.1: Zadání úloh 9. a 10. [11]

Správné řešení e. + c.

Kategorie Identifikace a kontrola změny

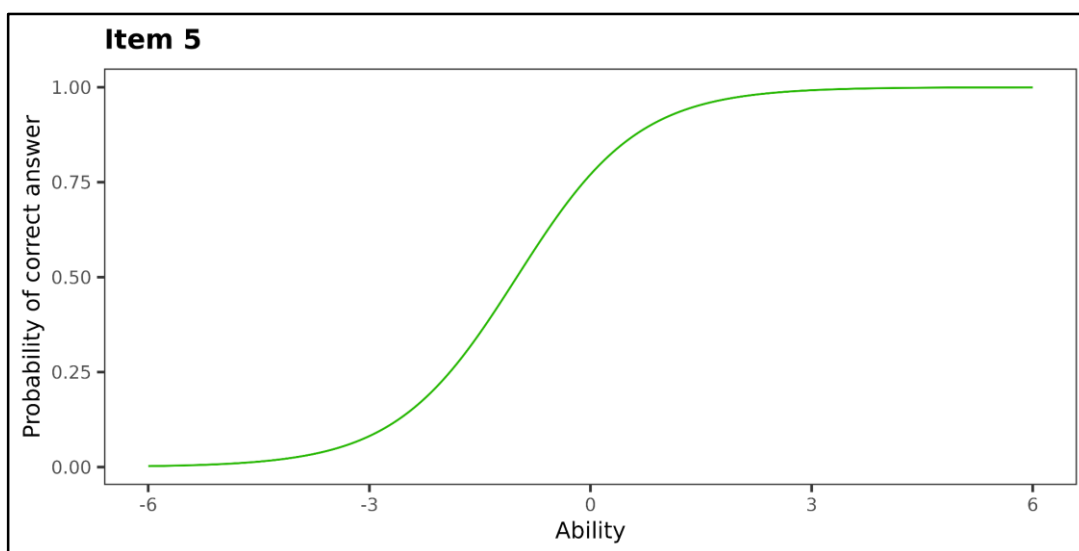
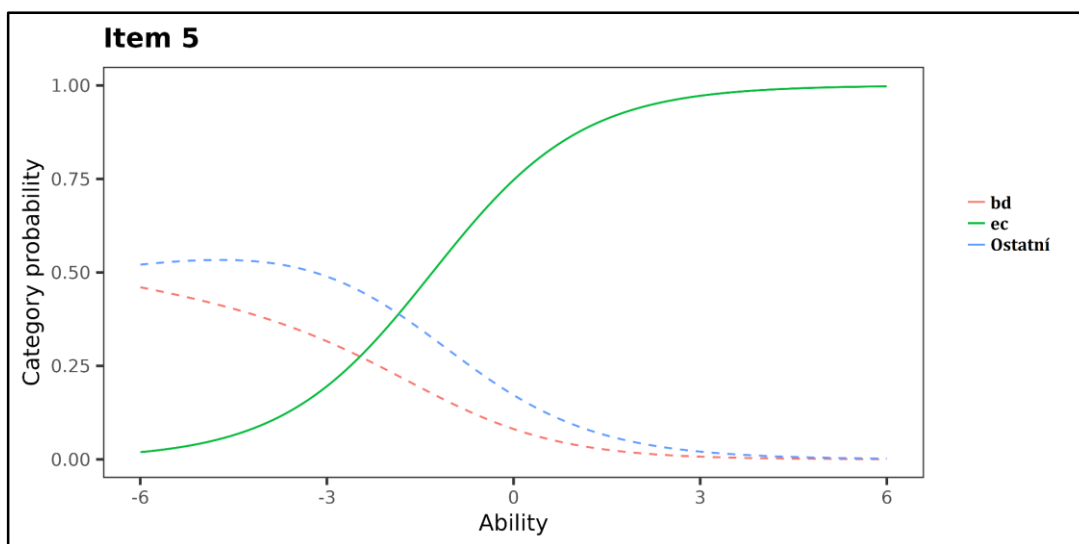
Úspěšnost 72 %

Index diskriminace položky ULI 0,59

Reliabilita Při odstranění dvojice úloh 9. a 10. dojde ke zhoršení hodnoty Cronbachovo alfa z 0,72 na 0,70.

Tabulky 3.5.1 a 3.5.2 : Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	9	10		Zastoupení	Kor. koef	a	b
a	1,2%	1,9%	bd	9,30%	-0,27	-0,90	-2,47
b	11,2%	9,3%	ec	72,01%	0,51	1,21	-1,00
c	4,1%	73,9%	Ostatní	18,69%	-0,39	-0,79	-1,85
d	7,6%	11,5%					
e	75,8%	3,4%					



Obrázky 3.5.2 a 3.5.3: Grafy křivek předmětové charakteristiky

Úloha 9;10 má ze všech úloh třetí nejvyšší úspěšnost, můžeme ji tedy považovat za úlohu snazší. Jedná se o první ze tří úloh kategorie *Identifikace a kontrola změny*. Přesto mají výsledky v této úloze s ostatními úlohami této kategorie pouze slabou pozitivní korelaci $r = 0,14$ respektive $r = 0,16$.

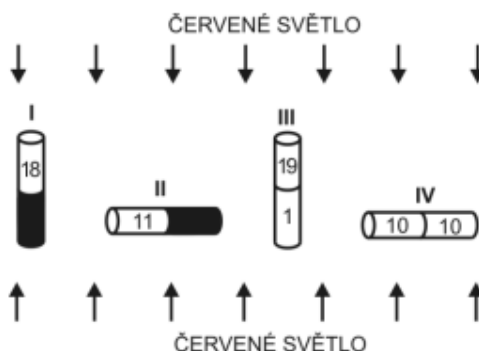
Úloha 9. i 10. mají každá dva relevantní distraktory. Kombinace *bd* byla jediným kombinovaným relevantním distraktorem. Řešitelé, kteří zvolili tento distraktor mohli považovat tíhu závaží za relevantní faktor, přestože se úloha ptá pouze na závislost doby kyvu na délce provázků.

Ze všech úloh testu má správná odpověď *ec* čtvrtou nejsilnější korelaci s celkovým skóre $r = 0,51$. Distraktor *bd* má s celkovým skóre zápornou korelaci $r = -0,27$, distraktor *Ostatní* má s celkovým skóre zápornou korelaci $r = -0,39$. Ostatní kombinace odpovědí tedy implikují horší celkové skóre než kombinace *bd*.

3.6 Úlohy 11 a 12

Zadání

11. Do každé ze čtyř skleněných trubiček dáme dvacet ovocných mušek. Trubičky těsně uzavřeme. Trubičky I a II jsou zčásti pokryty černým papírem. Trubičky III a IV nejsou vůbec zakryté. Trubičky jsou umístěny tak, jak to ukazuje obrázek. Pak je vystavíme červenému světlu po dobu pěti minut. Na obrázku jsou uvedeny počty mušek v nezakrytých částech trubiček.



Tento experiment ukazuje, že mušky reagují (to znamená, že se posunou blíž nebo dál) na:

- červené světlo, ale ne na gravitaci
 - gravitaci, ale ne na červené světlo
 - červené světlo i na gravitaci
 - nereagují ani na červené světlo ani na gravitaci
12. *protože*
- většina mušek je v horní části trubičky III, ale jsou rozmístěny zhruba rovnoměrně v trubičce II.
 - většina mušek nejde ke dnu trubiček I a III.
 - mušky potřebují světlo, aby viděly a musí letět proti gravitaci.
 - většina mušek je v horních koncích a v osvětlených koncích trubiček.
 - nějaké mušky jsou na obou koncích každé trubičky.

Obrázek 3.6.1: Zadání úloh 11. a 12. [11]

Správné řešení b. + a.

Kategorie Identifikace a kontrola změny

Úspěšnost 17 %

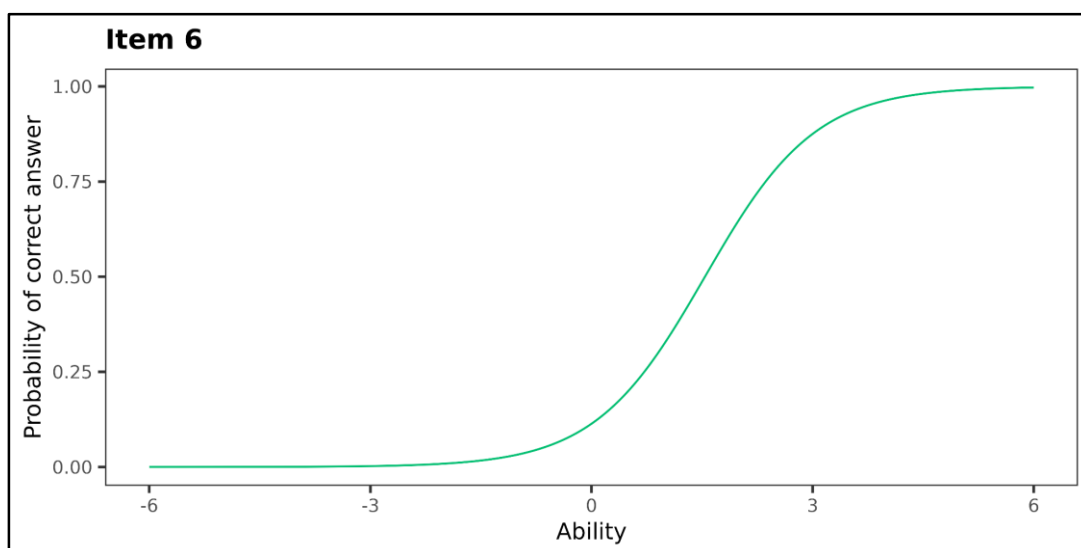
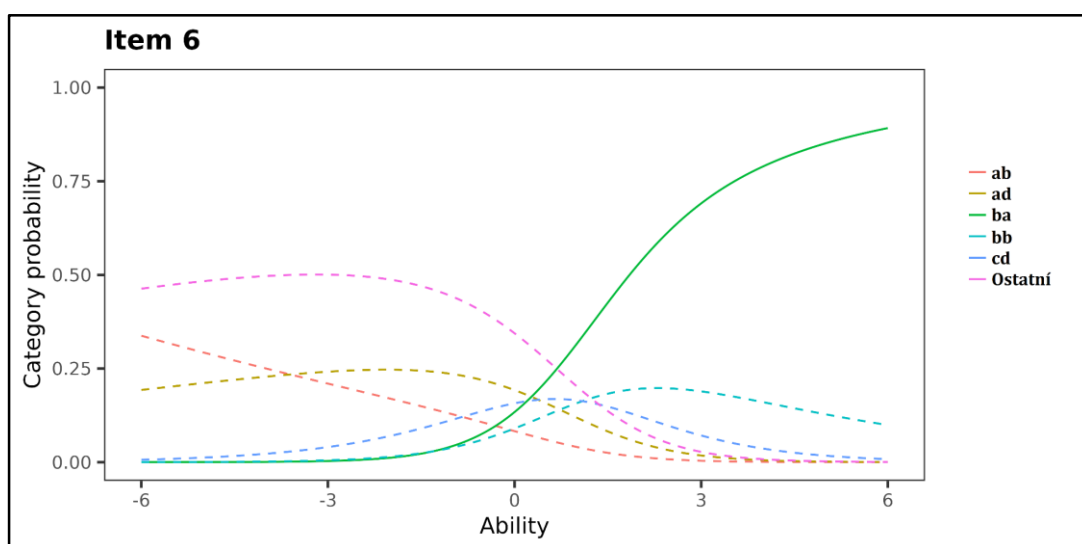
Index diskriminace položky ULI 0,42

Reliabilita Při odstranění dvojice úloh 11. a 12. dojde ke zhoršení hodnoty

Cronbachovo alfa z 0,72 na 0,70.

Tabulky 3.6.1 a 3.6.2 : Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	11	12		Zastoupení	Kor. koef	a	b
a	33,7%	27,6%	ab	8,50%	-0,16	-1,57	-0,30
b	29,9%	22,2%	ad	18,04%	-0,15	-1,34	0,27
c	28,6%	9,0%	ba	17,23%	0,47	1,33	1,54
d	7,8%	33,8%	bb	9,47%	0,14	-0,30	-1,32
e	-	7,4%	cd	14,00%	0,04	-0,81	0,20
			Ostatní	32,77%	-0,28	-1,39	0,68



Obrázky 3.6.2 a 3.6.3: Grafy křivek předmětové charakteristiky

Úloha 11;12 má ze všech úloh nejnižší úspěšnost, jedná se tedy o nejtěžší úlohu celého testu. Jedná se o druhou ze tří úloh kategorie *Identifikace a kontrola změny*. Výsledky v této úloze mají s výsledky třetí úlohy této kategorie slabou pozitivní korelaci $r = 0,32$ přestože se jedná o úlohy velmi podobné.

Úloha 11. i 12. mají všechny své distraktory relevantní, přesto existuje pouze pět relevantních kombinací distraktorů. Jedná se o jednu ze dvou úloh, ve které má kombinace *Ostatní* nejvyšší četnost, což poukazuje na vysoký počet dalších kombinací odpovědí, které ovšem nevyhovovaly kritériu pro relevantní kombinaci. Zároveň se jedná o jedinou úlohu, ve které měl jiný konkrétní distraktor vyšší četnost, než správné řešení.

Pokud vyřadíme řešení *d* úlohy 11., které většina řešitelů považovala za nesprávné, mají zbylá řešení téměř stejnou četnost. Jelikož se jedná o úvahově rozdílná řešení, může být podobná četnost způsobena vysokým počtem hádání odpovědi.

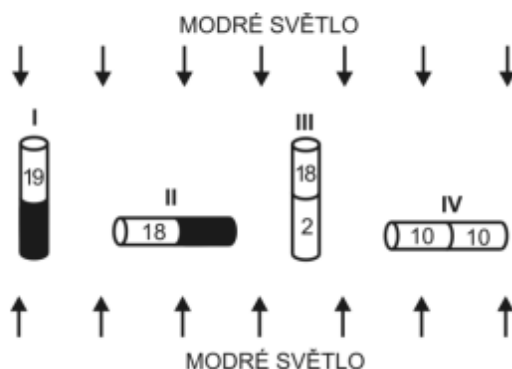
Četnost správného řešení úlohy 11. a úlohy 12. zvláště je výrazně menší než četnost správné kombinace řešení, což by potvrzovalo vysokou míru hádání odpovědi i v případě zvolení správného řešení pouze v úloze 12.. Zároveň ovšem správnou kombinaci volili převážně celkově úspěšnější řešitelé, volba správného řešení tedy většinou nebyla způsobena hádáním. Úloha tedy rozlišuje dobře mezi nejlepšími řešiteli.

Distraktory volené celkově úspěšnějšími řešiteli jsou distraktory *bb* a *cd*. Vysvětlení *d* v kombinaci *cd* by správně vysvětlovalo řešení *c*, to ovšem neodpovídá situaci popsané v zadání. Vysvětlení *b* v kombinaci *bb* naopak vysvětluje pouze část správného řešení *b*. V obou těchto případech tedy řešitel provedl částečně správnou úvahu. Řešitel tedy projevil dobrou ale nedostatečnou úroveň kognitivního myšlení, což potvrzuje i velmi slabě pozitivní korelace s celkovým skóre $r = 0,14$ pro kombinace *bb* a $r = 0,04$ pro kombinaci *cd*. Ostatní kombinace distraktorů mají s celkovým skóre pouze slabou negativní korelaci. Správná odpověď *ba* má korelaci s celkovým skóre $r = 0,47$.

3.7 Úlohy 13 a 14

Zadání

13. V druhém pokusu byl užít jiný druh mušek a modré světlo. Obrázek ukazuje výsledky.



Tato data ukazují, že mušky reagují (to znamená, že se posunou blíže nebo dál) na:

- a. modré světlo, ale ne na gravitaci
- b. gravitaci, ale ne na modré světlo
- c. modré světlo i na gravitaci
- d. nereagují ani na modré světlo ani na gravitaci

14. protože

- a. nějaké mušky jsou na obou koncích každé trubičky.
- b. mušky potřebují světlo, aby viděly, a musí letět proti gravitaci.
- c. mušky jsou rozděleny zhruba stejně v trubičce IV a v horní části trubičky III.
- d. většina mušek je v osvětlené části trubičky II, ale nejdou dolů v trubičkách I a III.
- e. většina mušek je v horní části trubičky I a v osvětlené části trubičky II.

Obrázek 3.7.1: Zadání úloh 13. a 14. [11]

Správné řešení c. + d.

Kategorie Identifikace a kontrola změny

Úspěšnost 29 %

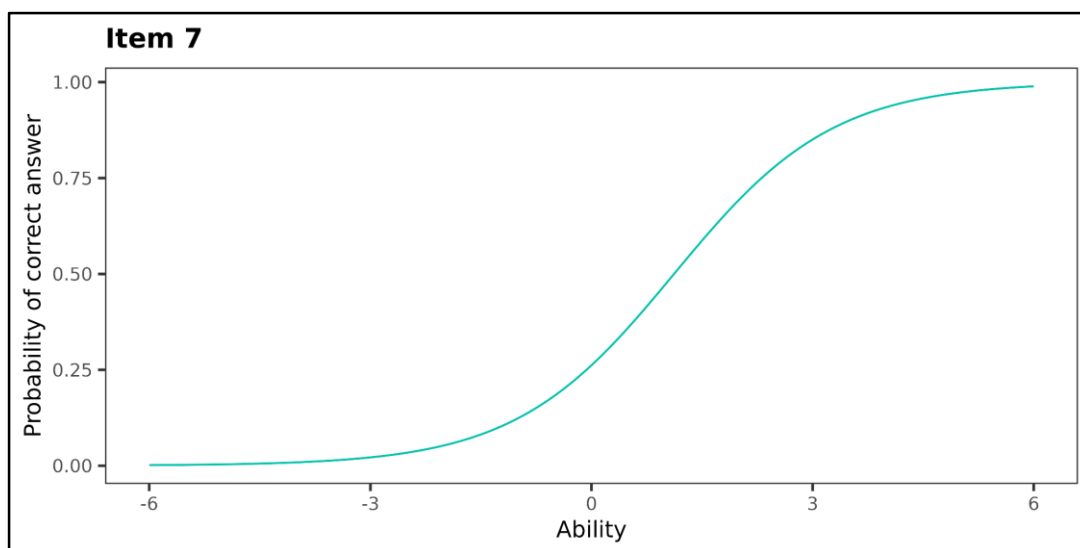
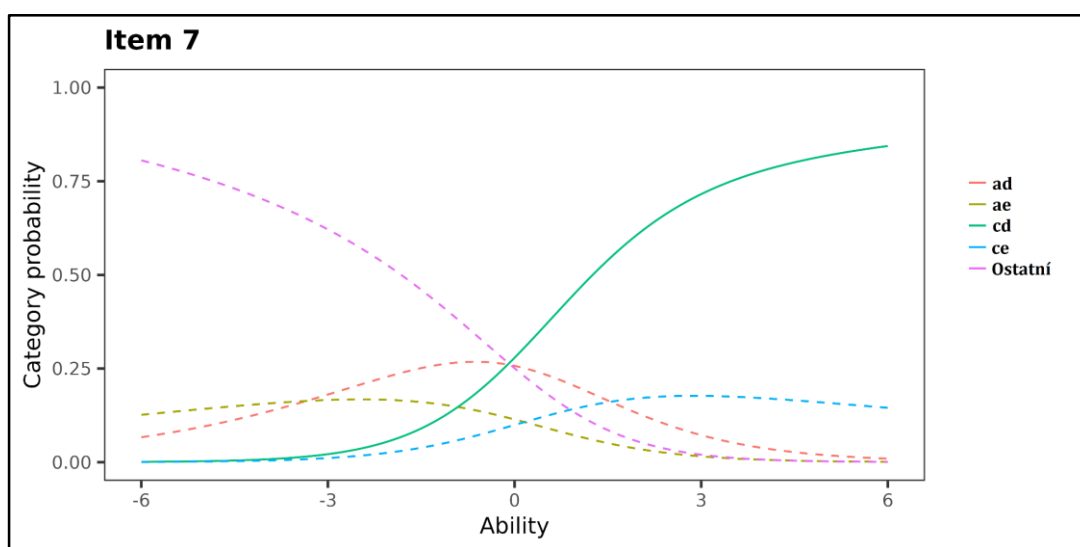
Index diskriminace položky ULI 0,50

Reliabilita Při odstranění dvojice úloh 13. a 14. dojde ke zhoršení hodnoty

Cronbachovo alfa z 0,72 na 0,71.

Tabulky 3.7.1 a 3.7.2 : Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	13	14		Zastoupení	Kor. koef	a	b
a	41,7%	6,5%	ad	23,62%	-0,11	-0,74	-0,11
b	5,9%	10,4%	ae	11,08%	-0,11	-0,98	-0,91
c	47,7%	4,6%	cd	29,29%	0,46	0,93	1,12
d	4,8%	56,3%	ce	9,87%	0,05	-0,12	-8,49
e	-	22,2%	Ostatní	26,13%	-0,33	-1,16	-0,09



Obrázky 3.7.2 a 3.7.3: Grafy křivek předmětové charakteristiky

Úloha 13;14 má ze všech úloh třetí nejnižší úspěšnost, jedná se tedy o náročnou úlohu. Jedná se o poslední ze tří úloh kategorie *Identifikace a kontrola změny*.

Přestože je úloha 13;14 velmi podobná úloze předchozí, výsledky v této úloze ukazují na nižší náročnost.

Úloha 13 má dva relevantní distraktory, distraktor *a* má ovšem výrazně vyšší zastoupení než distraktor *b*. Vyšší úspěšnost oproti předchozí úloze je pravděpodobně způsobena vysokou četností u pouze jednoho distraktoru v první části úlohy oproti dvěma distraktorům v úloze předchozí.

Můžeme si všimnout, že distraktor *a* i správná odpověď *c* se liší v názoru, zda mušky reagují na gravitaci či ne. Přestože zadání přímo ukazuje reakci mušek na gravitaci v trubičkách I. a II., 41,7 % řešitelů tuto možnost nezvolilo. Je tedy možné, že si velké množství řešitelů nedokázalo představit, jak by se reakce mušek na gravitaci v obrázku vůbec projevila, a proto volilo distraktor *a*, ke kterému hledalo jim srozumitelné vysvětlení. Je ovšem překvapivé, že nejčastější kombinací konkrétních distraktorů je kombinace *ad*, jejíž vysvětlení *d* přímo vylučuje řešení *a*. Tuto kombinaci volili celkově průměrní řešitelé.

Distraktor *ce* byl volen celkově úspěšnějšími řešiteli. Vysvětlení *e* se řešiteli mohlo na první pohled zdát správné, jelikož poukazuje jak na polohu vůči gravitaci, tak na polohu vůči světlu, ovšem ve skutečnosti by vedlo k nejednoznačnosti. Tento distraktor má velmi slabou pozitivní korelaci s celkovým skóre $r = 0,05$, podobně jako v předchozí úloze řešitel projevili dobrou, ale nedostatečnou úroveň kognitivního myšlení.

Kombinace *Ostatní* má s celkovým skóre negativní korelaci $r = -0,33$, ostatní relevantní kombinace distraktorů mají s celkovým skóre pouze slabou negativní korelaci. Správná odpověď *cd* má korelaci s celkovým skóre $r = 0,46$.

3.8 Úlohy 15 a 16

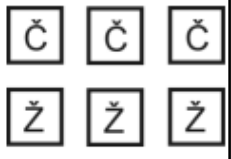
Zadání

15. Šest čtvercových kousků dřeva dáme do látkového sáčku a promícháme. Všech šest kousků má stejnou velikost a tvar, ale tři kousky jsou červené a tři žluté. Předpokládejme, že někdo sáhne do sáčku (aniž by se díval) a vytáhne jeden kousek. *Jaká je šance, že vytažený kousek je červený?*

a. 1 ze 6
b. 1 ze 3
c. 1 ze 2
d. 1 z 1
e. nelze určit

16. *protože*

a. 3 ze 6 kusů jsou červené.
b. nelze nijak říci, který kousek bude vytažen.
c. je vytažen jen 1 kus ze 6 kusů v sáčku.
d. všech 6 kusů má stejnou velikost a tvar.
e. může být vytažen jen 1 červený kus ze 3 červených kusů.



The diagram consists of two rows of three squares each. The top row contains three squares, each with the letter 'Č' (red) inside. The bottom row contains three squares, each with the letter 'Ž' (yellow) inside.

Obrázek 3.8.1: Zadání úloh 15. a 16. [11]

Správné řešení c. + a.

Kategorie Pravděpodobnostní myšlení

Úspěšnost 53 %

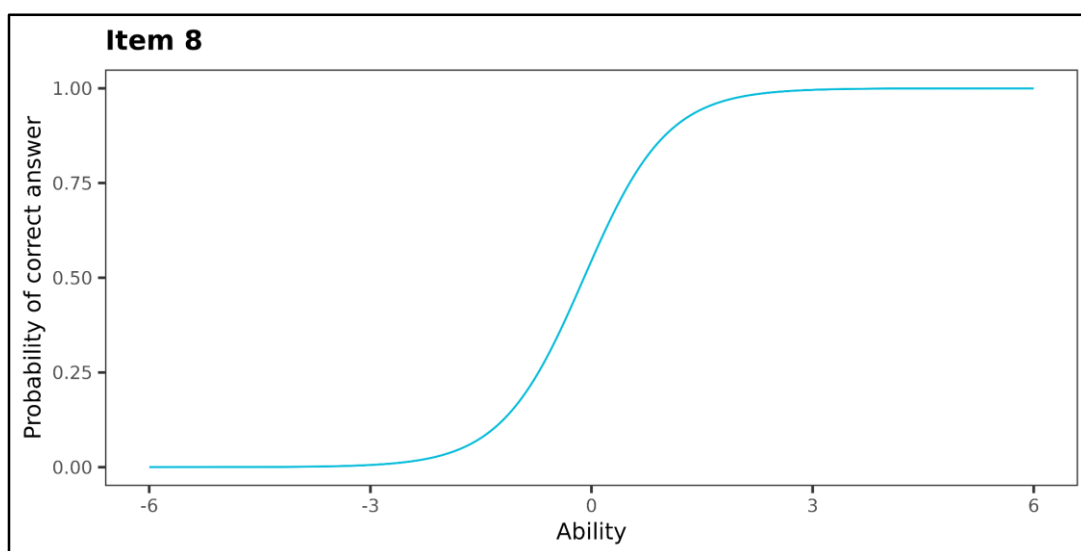
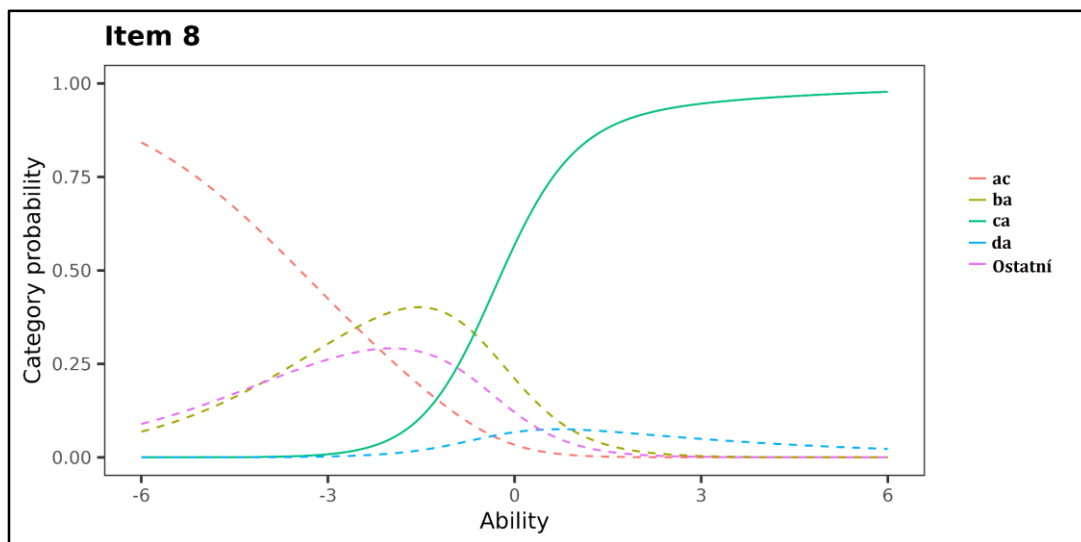
Index diskriminace položky ULI 0,80

Reliabilita Při odstranění dvojice úloh 15. a 16. dojde ke zhoršení hodnoty

Cronbachovo alfa z 0,72 na 0,68.

Tabulky 3.8.1 a 3.8.2 : Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	15	16		Zastoupení	Kor. koef	a	b
a	9,3%	81,7%	ac	5,99%	-0,23	-2,26	-1,25
b	27,7%	3,9%	ba	21,93%	-0,35	-1,53	-0,65
c	54,1%	6,6%	ca	52,83%	0,63	1,78	1,12
d	6,6%	1,9%	da	5,66%	0,00	-0,27	-7,79
e	2,3%	5,9%	Ostatní	13,59%	-0,33	-1,67	-0,93



Obrázky 3.8.2 a 3.8.3: Grafy křivek předmětové charakteristiky

Úloha 15;16 tvoří medián obtížností úloh testu. Jedná se o první ze dvou úloh kategorie *Pravděpodobnostní myšlení*. Výsledky v této úloze mají s výsledky z následující úlohy korelaci $r = 0,37$.

Úloha 15 má tři relevantní distraktory. V úloze 16 byla nejčastěji volená správná odpověď *a*, pravděpodobně protože přímo odpovídá obrázku v zadání. Kombinace *ac* byla nejčastěji volená nejméně úspěšnými řešiteli. Byla pravděpodobně volena kvůli nedostatečné úrovni pravděpodobnostního myšlení. Kombinace *ba* a *Ostatní* již byly voleny více úspěšnými řešiteli, stále ovšem celkově méně úspěšnými.

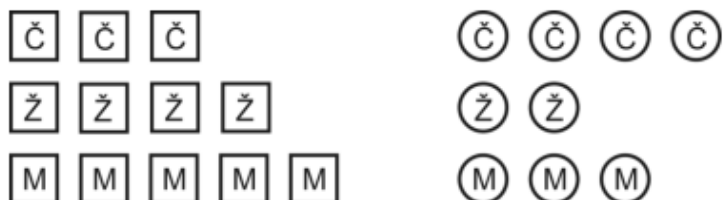
Kombinace *da* byla volena celkově mírně úspěšnějšími řešiteli. Vzhledem k této kombinaci tito řešitelé pravděpodobně interpretovali odpověď *d* chybně jako *jedna ku jedné*, což by odpovídalo jejich následnému správnému vysvětlení *a*. Tato kombinace má s celkovým skóre nulovou korelaci.

Kombinace *ba* má s celkovým skóre negativní korelaci $r = -0,35$, kombinace *Ostatní* $r = -0,33$. Kombinace *ac* má z ostatních distraktorů nejslabší negativní korelaci $r = -0,23$. Správné řešení *ca* má ze všech řešení v testu nejvyšší pozitivní korelaci s celkovým skóre $r = 0,63$. Zároveň má nejvyšší hodnotu indexu diskriminace položky. Můžeme tedy pochopení základních konceptů pravděpodobnostního myšlení považovat za nejlivnější faktor v určení vysoké úrovně kognitivního vývoje řešitele.

3.9 Úlohy 17 a 18

Zadání

17. Do látkového sáčku dáme tři červené dřevěné čtverečky, čtyři žluté dřevěné čtverečky a pět modrých dřevěných čtverečků. Do sáčku dáme také čtyři červená kolečka, dvě žlutá kolečka a tři modrá kolečka. Všechny kousky pak v sáčku promícháme. Předpokládejme, že někdo sáhne do sáčku (aniž by se díval a aniž by se staral o tvar, na který sahá) a vytáhne jeden kousek.



Jaká je šance, že vytažený kousek je červené kolečko nebo modré kolečko?

- a. nelze je určit
- b. 1 ze 3
- c. 1 z 21
- d. 15 z 21
- e. 1 ze 2

18. *protože*

- a. 1 ze 2 tvarů je kulatý.
- b. 15 z 21 kousků je červených nebo modrých.
- c. nelze určit, který kousek bude vytažen.
- d. ze sáčku je vytažen jen 1 kousek z 21.
- e. 1 ze 3 kousků je červené nebo modré kolečko.

Obrázek 3.9.1: Zadání úloh 17. a 18. [11]

Správné řešení b. + e.

Kategorie Pravděpodobnostní myšlení

Úspěšnost 55 %

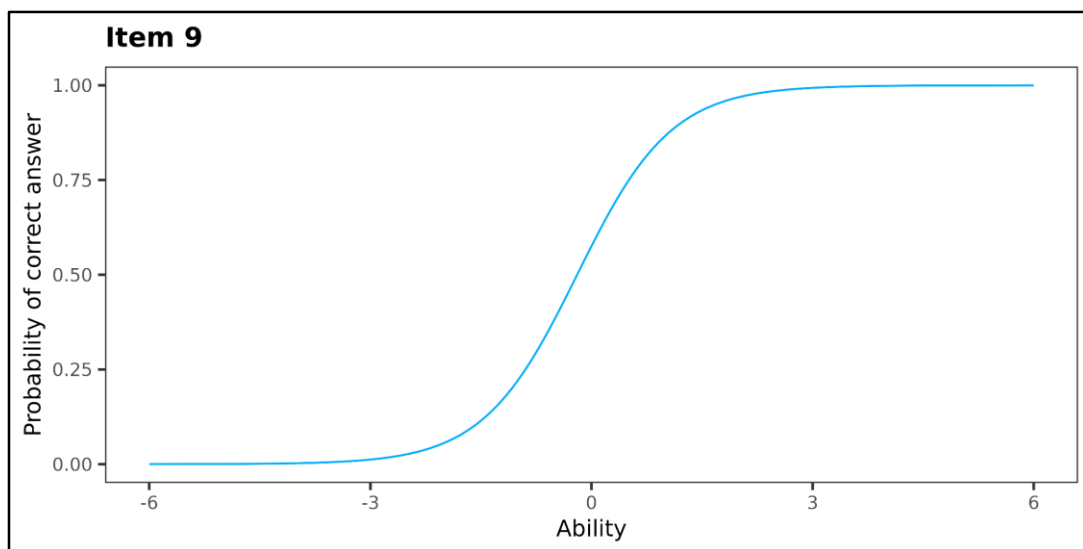
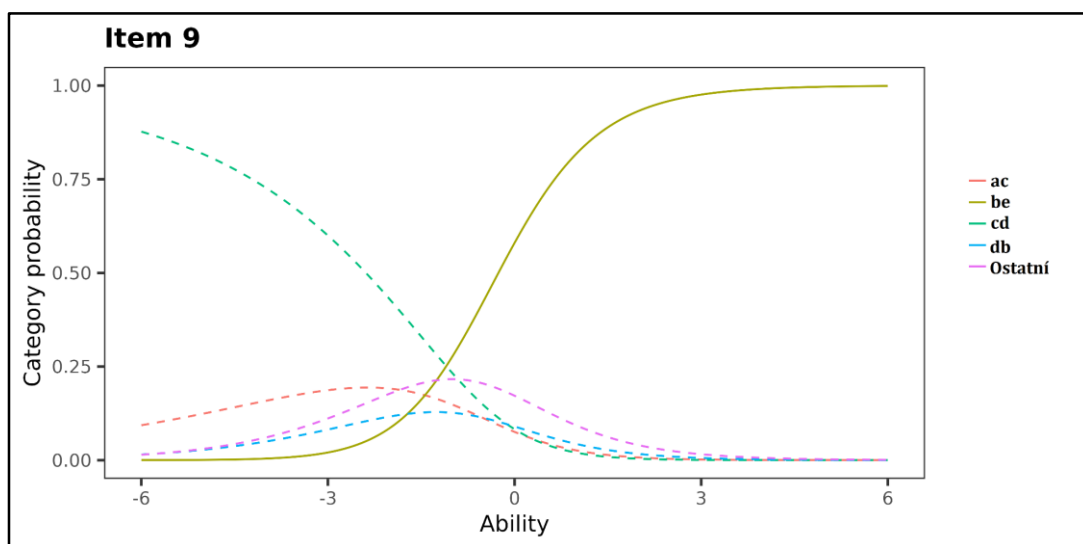
Index diskriminace položky ULI 0,75

Reliabilita Při odstranění dvojice úloh 17. a 18. dojde ke zhoršení hodnoty

Cronbachovo alfa z 0,72 na 0,69.

Tabulky 3.9.1 a 3.9.2 : Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	17	18		Zastoupení	Kor. koef	a	b
a	10,4%	4,4%	ac	8,50%	-0,25	-1,42	-1,44
b	59,1%	9,9%	be	55,18%	0,60	1,56	-0,19
c	13,2%	10,1%	cd	11,97%	-0,29	-1,77	-1,10
d	9,5%	15,3%	db	8,58%	-0,20	-1,08	-1,72
e	7,8%	60,3%	Ostatní	15,78%	-0,21	-0,97	-1,25



Obrázky 3.9.2 a 3.9.3: Grafy křivek předmětové charakteristiky

Úloha 17;18 má ze všech úloh šestou nejvyšší úspěšnost, jedná se tedy o úlohu mírně snazší. Jedná se o druhou ze dvou úloh kategorie *Pravděpodobnostní myšlení*.

Úlohy 17 i 18 mají všechny své distraktory relevantní, kromě vysvětlení *a*. Přesto má úloha pouze 4 relevantní kombinace distraktorů. To je způsobeno vhodným vysvětlením k nesprávnému řešení první části úlohy.

Kombinace *cd* byla nejčastěji volená nejméně úspěšnými řešiteli. Porovnáním této kombinace s kombinací *ac* v předchozí úloze vidíme, že se jedná o stejný úvahový proces. Opět byla pravděpodobně volena kvůli nedostatečné úrovni pravděpodobnostního myšlení.

Ostatní relevantní kombinace distraktorů byly stejně jako v předchozí úloze voleny více úspěšnými řešiteli, kteří ovšem byli celkově méně úspěšní.

Kombinace *cd* má s celkovým skóre negativní korelaci $r = -0,29$, ostatní relevantní dvojice distraktorů mají slabou negativní korelaci. Správné řešení *be* má ze všech řešení v testu druhou nejvyšší pozitivní korelaci s celkovým skóre $r = 0,60$. Zároveň má druhou nejvyšší hodnotu indexu diskriminace položky. Tento výsledek tedy podporuje myšlenku, že pochopení základních konceptů pravděpodobnostního myšlení je nejvlivnějším faktorem v určení vysoké úrovně kognitivního vývoje řešitele.

3.10 Úlohy 19 a 20

Zadání

19. Farmář Brown pozoroval myši žijící na jeho poli. Zjistil, že každá je buď tlustá nebo hubená. Také měla každá z nich buď černý nebo bílý ocas. Napadlo ho, jestli velikost myši nějak souvisí s barvou jejího ocasu. Pochytil proto všechny myši v určité části pole a prohlédl si je. Obrázek ukazuje všechny myši, které pochytal.



Myslíte, že je nějaká souvislost mezi velikostí myší a barvou jejich ocasu?

- a. zdá se, že souvislost existuje
- b. zdá se, že žádná souvislost neexistuje
- c. nelze učinit rozumný závěr

20. *protože*

- a. pro každou možnost je tu několik myší.
- b. souvislost mezi velikostí myší a barvou jejich ocasu může být dána geneticky.
- c. farmář nepochytil dost myší.
- d. většina tlustých myší má černé ocasy a většina hubených myší má bílé ocasy.
- e. když myši tloustnou, jejich ocasy tmavnou.

Obrázek 3.10.1: Zadání úloh 19. a 20. [11]

Správné řešení a. + d.

Kategorie Korelační myšlení

Úspěšnost 40 %

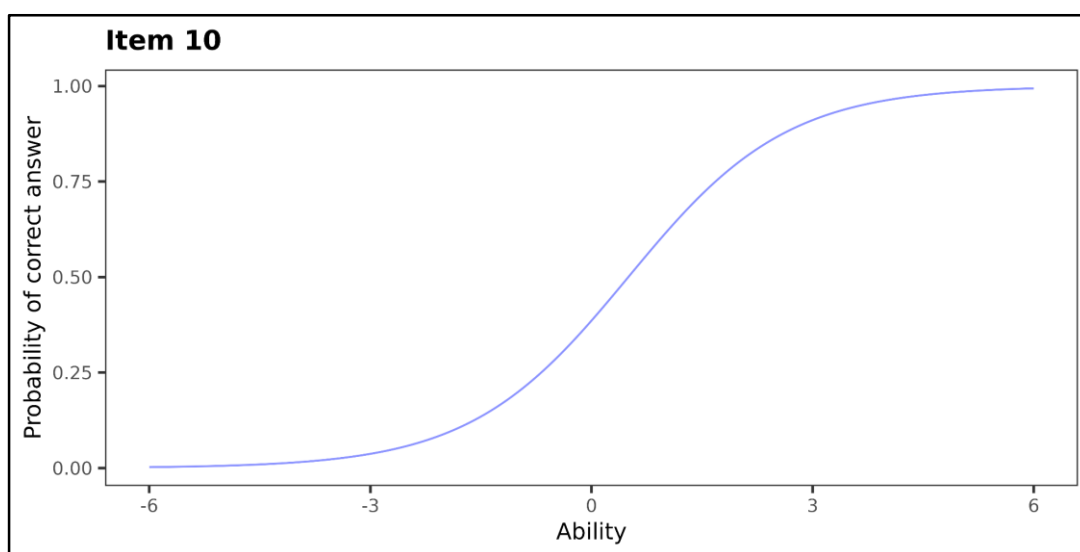
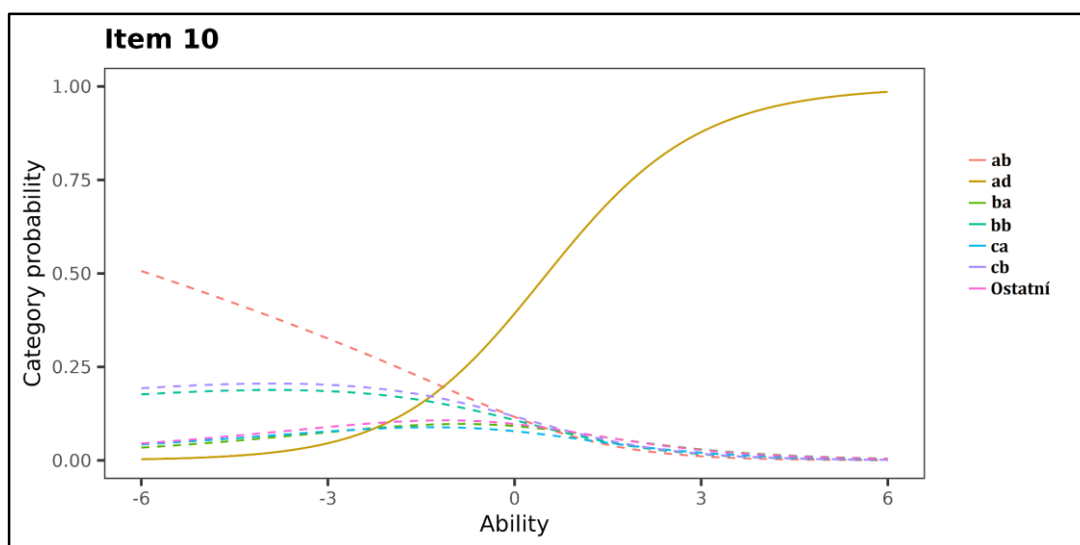
Index diskriminace položky ULI 0,63

Reliabilita Při odstranění dvojice úloh 19. a 20. dojde ke zhoršení hodnoty

Cronbachovo alfa z 0,72 na 0,70.

Tabulky 3.10.1 a 3.10.2: Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	19	20		Zastoupení	Kor. koef	a	b
a	54,4%	16,5%	ab	12,30%	-0,23	-1,06	-1,15
b	21,9%	34,6%	ad	40,29%	0,50	0,93	0,50
c	23,7%	3,7%	ba	8,58%	-0,09	-0,65	-2,24
d	-	43,9%	bb	10,68%	-0,15	-0,90	-1,44
e	-	1,2%	ca	7,36%	-0,09	-0,71	-2,27
			cb	11,65%	-0,12	-0,90	-1,34
			Ostatní	9,14%	-0,12	-0,69	-2,03



Obrázky 3.10.2 a 3.10.3: Grafy křivek předmětové charakteristiky

Úloha 19;20 má ze všech úloh pátou nejnižší úspěšnost, jedná se tedy o úlohu náročnější. Jedná se o jedinou úlohu kategorie *Korelační myšlení*. Výsledky v této úloze nejvíce korelují s výsledky v úloze 5;6, $r = 0,23$.

Úloha 19 má oba distraktory relevantní. Přestože jsou distraktory pouze dva, je celkem šest kombinací relevantních distraktorů. Všechny kombinace distraktorů byly voleny celkově méně úspěšnými řešiteli. Většina řešitelů správně objevila souvislost mezi velikostí myši a barvou jejího ocasu. Například kombinací distraktorů s nejvyšší četností je kombinace *ab*. Řešitelé, kteří volili tuto kombinaci, objevení této souvislosti přisoudili genetice myši, o které se zadání vůbec nezmiňuje. Ovšem i řešitelé, kteří závislost neobjevili či usoudili, že závislost nelze určit nejčastěji argumentovali vysvětlením *b*. Možným vysvětlením může být velká důležitost, kterou část řešitelů přikládá vlivu genetiky, jiná formulace této možnosti která, nezmiňuje genetiku, by mohla vést k jiným výsledkům.

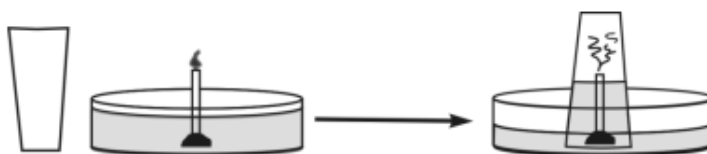
Kombinace *ab* má s celkovým skóre nejsilnější negativní korelaci $r = -0,23$, ostatní relevantní dvojice distraktorů mají slabou negativní korelaci. Správné řešení *ad* má ze všech řešení v testu šestou nejvyšší pozitivní korelaci s celkovým skóre $r = 0,60$

Převážně slabé negativní korelace naznačují menší vliv nevyvinutého korelačního myšlení na celkový kognitivní vývoj řešitele.

3.11 Úlohy 21 a 22

Zadání

21. Levý obrázek níže ukazuje skleničku a hořící svíčku stojící v nádobce s vodou. Svíčka je na dno přichycena malým kouskem plastelíny. Když skleničku obrátíme, přiklopíme jí svíčku a postavíme do vody, svíčka rychle zhasne a voda se nahně do sklenice (jak ukazuje obrázek vpravo).



Když pozorujeme, co se stalo, může nás napadnout otázka: Proč se voda nahně do sklenice?

Jedno možné vysvětlení: Plamen mění kyslík na oxid uhličitý. Protože kyslík se rozpouští ve vodě jen pomalu, ale oxid uhličitý velmi rychle, vytvořený oxid uhličitý se rychle rozpustí ve vodě a tím se pod sklenicí sníží tlak.

Předpokládejme, že máte výše uvedené věci a navíc zápalky a nějaký suchý led (suchý led je zmrzlý oxid uhličitý). *Jak byste s využitím těchto materiálů mohli otestovat výše uvedené možné vysvětlení?*

- Nasytit vodu oxidem uhličitým a opakovat experiment. Všimnout si přítom, o kolik stoupne voda.
 - Voda stoupne, protože se spotřebovává kyslík, takže je potřeba zopakovat pokus přesně stejným způsobem, abychom ukázali, že voda stoupá díky úbytku kyslíku.
 - Provést kontrolovaný pokus, při němž budeme měnit pouze počet svíček, abychom viděli, zda nastane nějaký rozdíl.
 - Za vzestup hladiny může sání, takže je třeba vzít na obou koncích otevřený válec, přes jeden jeho konec navléknout gumový pouťový balónek a tímto válcem pak přiklopit svíčku.
 - Opakovat experiment, ale ujistit se, že všechny nezávisle proměnné jsou konstantní, potom měřit výšku, do níž voda vystoupí.
22. Který výsledek vašeho testování (uvedeného výše v bodě 21) by ukázal, že uvedené vysvětlení je pravděpodobně špatné?
- Voda stoupne do téže výše jako v předešlém pokusu.
 - Voda stoupne méně než v předešlém pokusu.
 - Balónek se nafoukne ven z válce.
 - Balónek je vsáván do válce.

Obrázek 3.11.1: Zadání úloh 21. a 22. [11]

Správné řešení a. + a.

Kategorie Kombinační myšlení zahrnující prvky z předchozích oblastí

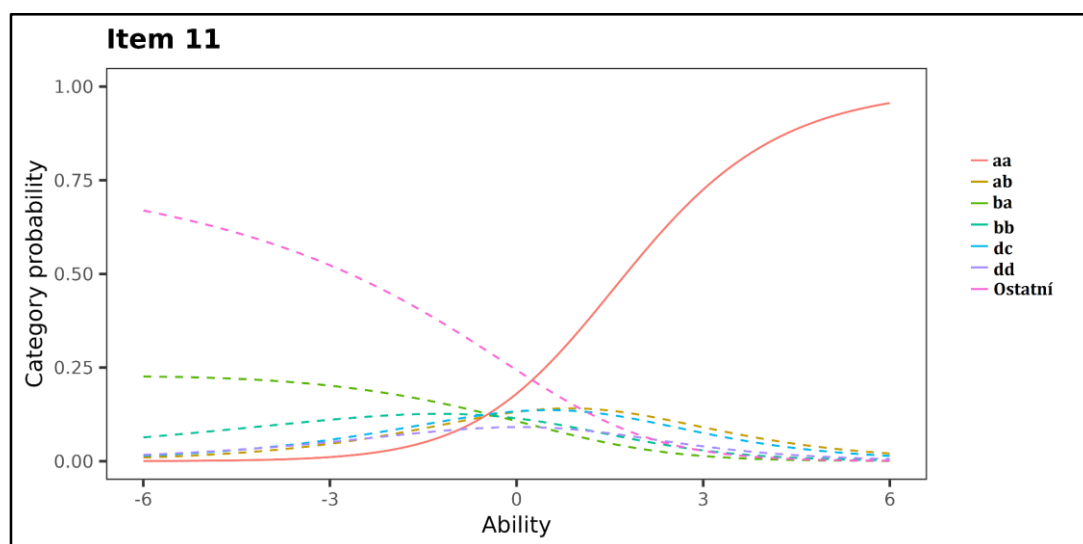
Úspěšnost 21 %

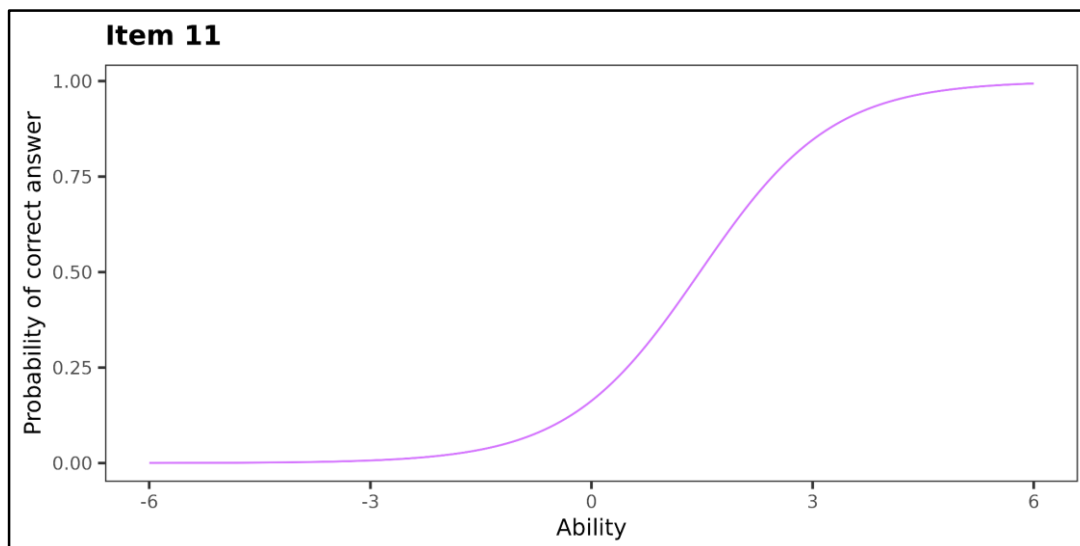
Index diskriminace položky ULI 0,47

Reliabilita Při odstranění dvojice úloh 21. a 22. dojde ke zhoršení hodnoty Cronbachovo alfa z 0,72 na 0,70.

Tabulky 3.11.1 a 3.11.2: Zastoupení jednotlivých odpovědí a dvojic relevantních odpovědí

	21	22		Zastoupení	Kor. koef	a	b
a	35,4%	39,1%	aa	20,95%	0,46	1,11	1,47
b	27,3%	29,4%	ab	12,22%	0,01	-0,59	-0,54
c	6,8%	18,4%	ba	10,68%	-0,14	-1,15	-0,45
d	23,5%	13,1%	bb	10,76%	-0,10	-0,92	-0,49
e	7,0%	-	dc	12,30%	0,00	-0,65	-0,47
			dd	8,41%	-0,02	-0,74	-0,92
			Ostatní	24,68%	-0,26	-1,19	0,25





Obrázky 3.11.2 a 3.11.3: Grafy křivek předmětové charakteristiky

Úloha 21;22 má ze všech úloh druhou nejnižší úspěšnost, jedná se tedy o úlohu obtížnou. Jedná se o jednu ze tří úloh kategorie *Kombinační myšlení zahrnující prvky z předchozích oblastí*. Výsledky v této úloze mají pouze slabé korelace s výsledky v ostatních úlohách.

Úlohy 21. i 22. mají všechny distraktory relevantní. Jedná se o jednu ze dvou úloh, kde má kombinace *Ostatní* vyšší četnost než správná kombinace. Kombinace volené celkově méně úspěšnými řešiteli jsou *Ostatní*, *ba*, *bb*. Kombinace *ba* i *bb* ve svých vysvětleních obě předpokládají změnu ve výsledku testování i přesto, že odpověď na první část *b* požaduje provedení naprosto stejného testu. Tyto distraktory i tak mají pouze slabou negativní korelaci s celkovým výsledkem.

Ostatní distraktory byly voleny celkově mírně úspěšnějšími řešiteli, jejich korelace s celkovým skóre je téměř nulová. Stejně jako v případě úlohy 11;12 je četnost správné kombinace výrazně menší, než četnost jednotlivých správných odpovědí. Jako možné vysvětlení se opět nabízí vyšší míra hádání odpovědi.

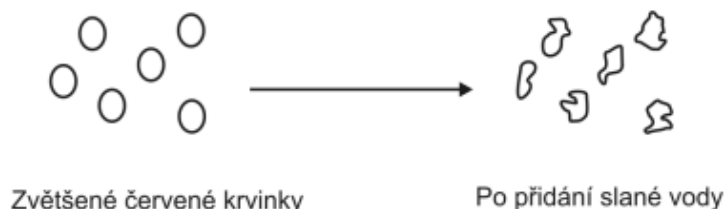
Správná kombinace *aa* má s celkovým skóre pozitivní korelaci $r = 0,46$, což je čtvrtá nejnižší pozitivní korelace. S ohledem na to a na velmi slabé negativní korelace v případě zvolení distraktorů má schopnost resp. neschopnost vyřešení této úlohy relativně malý vliv na kognitivní vývoj řešitele.

3.12 Úlohy 23 a 24

Úlohy 23 a 24 jsou bodovány jednotlivě, je tedy uvedena analýza každé úlohy zvlášť.

Zadání

23. Student kápl kapku krve na podložní sklíčko a podíval se na krev mikroskopem. Jak ukazuje obrázek níže, zvětšené červené krvinky vypadají jako kulaté balóčky. Po přidání několika kapek slané vody si student všiml, že se buňky zmenšují.



Pokus nás přivádí k zajímavé otázce: Proč se zdá, že se červené krvinky zmenšují?

Zde jsou dvě možná vysvětlení: I. Ionty soli (Na^+ a Cl^-) tlačí na membránu buňky a způsobují tak, že se buňky zmenšují. II. Molekuly vody se přitahují k iontům soli, takže se molekuly vody pohybují ven z buněk a buňky se zmenšují.

Aby otestoval tato vysvětlení, student použil slanou vodu, velmi přesné váhy a plastové sáčky naplněné vodou. Předpokládal přitom, že daný plast se chová stejně jako membrána červených krvinek. Pokus zahrnoval přesné zvážení sáčku naplněného vodou, jeho následné umístění do roztoku soli na dobu deset minut a opětovné zvážení.

Jaký výsledek experimentu by nejlépe ukázal, že vysvětlení I je patrně špatné?

- a. sáček je při druhém zvážení lehčí
 - b. sáček váží stejně
 - c. sáček vypadá menší
24. *Jaký výsledek experimentu by nejlépe ukázal, že vysvětlení II je patrně špatné?*
- a. sáček je při druhém zvážení lehčí
 - b. sáček váží stejně
 - c. sáček vypadá menší

Obrázek 3.12.1: Zadání úloh 23. a 24. [11]

Úloha 23

Správné řešení a.

Kategorie Kombinační myšlení zahrnující prvky z předchozích oblastí

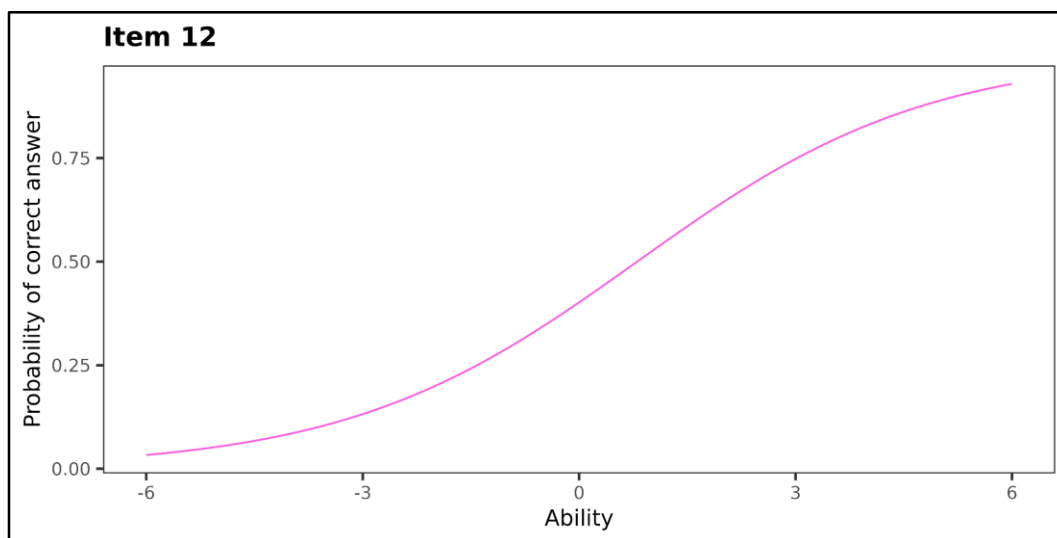
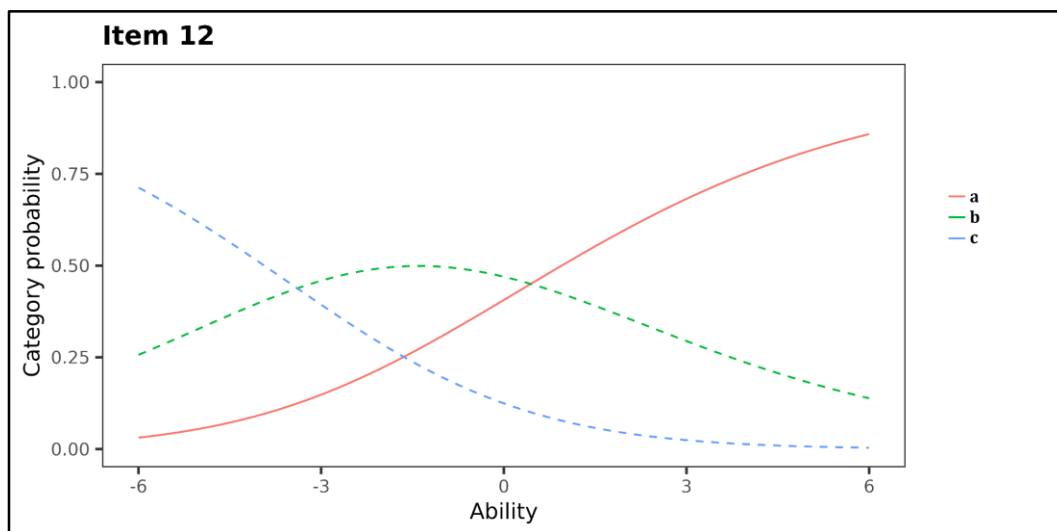
Úspěšnost 41 %

Index diskriminace položky ULI 0,37

Reliabilita Při odstranění úlohy 23. dojde ke zhoršení hodnoty Cronbachova alfa z 0,72 na 0,72.

Tabulka 3.12.1: Zastoupení jednotlivých odpovědí

	Zastoupení	Kor. koef	a	b
a	40,70%	0,29	0,50	0,80
b	45,87%	-0,16	-0,33	0,44
c	13,43%	-0,18	-0,72	-1,64



Obrázky. 3.12.2 a 3.12.3: Grafy křivek předmětové charakteristiky

Úloha 23 má ze všech úloh šestou nejnižší úspěšnost, jedná se tedy o úlohu mírně obtížnější. Jedná se o druhou ze tří úloh kategorie *Kombinační myšlení zahrnující prvky z předchozích oblastí*. Výsledky v této úloze nejvíce korelují s výsledky v úloze následující, $r = 0,21$. Zároveň se jedná o jednu ze dvou úloh, která se hodnotí samostatně, nikoliv jako dvojúloha.

Vzhledem k malému počtu možností jsou oba distraktory relevantní. Distraktor *c* byl častěji volený celkově horšími řešiteli, úspěšnější řešitelé nepovažovali vizuální velikost sáčku za relevantní. Druhý distraktor *b* byl volený úspěšnějšími řešiteli, dokonce častěji, než správné řešení. Řešitelé, kteří zvolili tento distraktor pravděpodobně předpokládali, že zmenšením plastových sáčků by došlo ke snížení hmotnosti sáčku, a tedy nesnížení hmotnosti sáčku by znamenalo neplatnost vysvětlení. Zadání ovšem nic takového nepředpokládá. Možným vysvětlením je chybná řešitelova představa, že při zmenšení objemu vždy dojde ke zmenšení hmotnosti.

Oba distraktory mají slabou negativní korelaci s celkovým skóre. Správná odpověď *a* má s celkovým skóre kladnou korelaci $r = 0,29$, což je druhá nejmenší kladná korelace v testu. Zároveň má nejnižší hodnotu indexu diskriminace položky, což může být způsobeno i vyšší šancí na uhádnutí odpovědi. Přestože se tedy nejedná o snadnou úlohu, schopnost jejího správného vyřešení má relativně malý vliv na kognitivní úroveň řešitele.

Úloha 24

Správné řešení b.

Kategorie Kombinační myšlení zahrnující prvky z předchozích oblastí

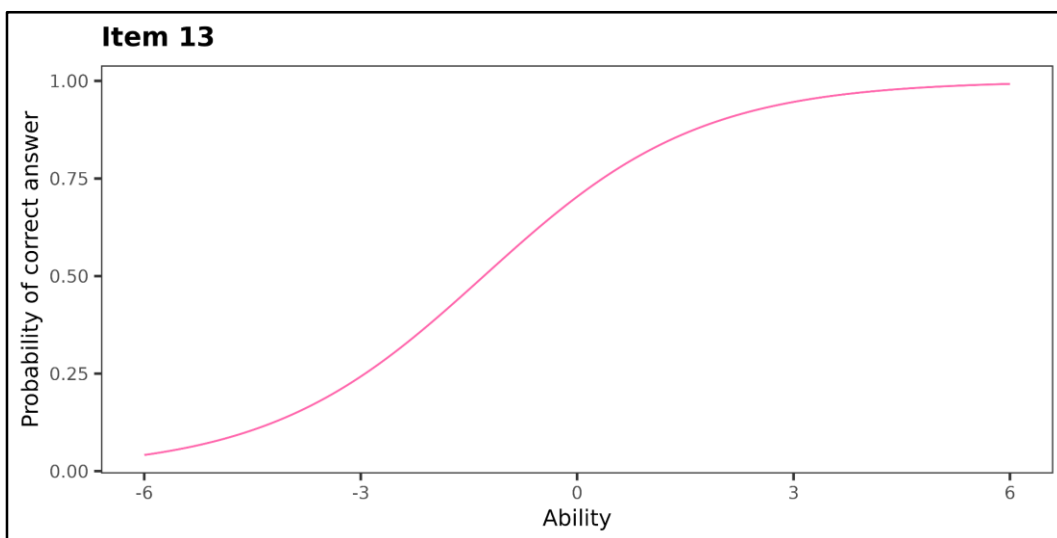
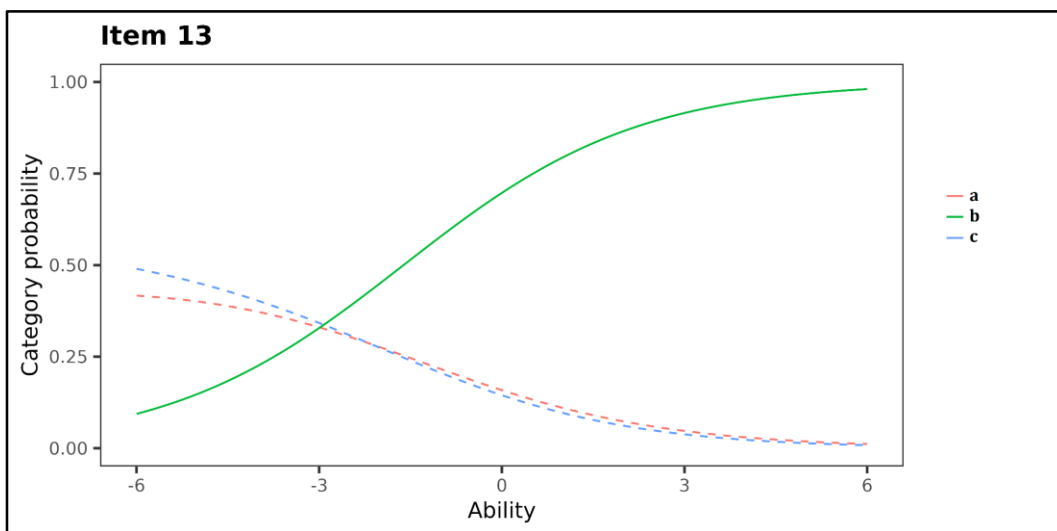
Úspěšnost 69 %

Index diskriminace položky ULI 0,38

Reliabilita Při odstranění dvojice úloh 1. a 2. nedojde ke změně hodnoty Cronbachovo alfa, hodnota tedy zůstává 0,72.

Tabulka 3.12.2: Zastoupení jednotlivých odpovědí

	Zastoupení	Kor. koef	a	b
a	16,26%	-0,21	-0,50	-2,99
b	68,69%	0,33	0,67	-1,29
c	15,05%	-0,22	-0,54	-2,92



Obrázky 3.12.4 a 3.12.5: Grafy křivek předmětové charakteristiky

Úloha 24 má ze všech úloh čtvrtou nejvyšší úspěšnost, jedná se tedy o úlohu snazší. Jedná se o třetí ze tří úloh kategorie *Kombinační myšlení zahrnující prvky z předchozích oblastí*. Zároveň se jedná o druhou ze dvou úloh, která se hodnotí samostatně, nikoliv jako dvojúloha.

Stejně jako v předchozí úloze, vzhledem k malému počtu možností jsou oba distraktory relevantní. Tentokrát byly ovšem oba voleny výrazně méně často, než správné řešení. Správné řešení *b* je stejně formulované jako předchozí distraktor *b*. Mnoha předchozím neúspěšným řešitelům pravděpodobně přišel tento druhý test principiálně stejný jako první, tentokrát ovšem při zmenšení sáčku dochází k úniku vody a tedy se jeho hmotnost skutečně zmenšuje. Tito řešitelé tedy mohli použít stejnou úvahu jako v předchozí úloze, tentokrát už ovšem správnou. Proto má v této úloze správná odpověď mnohem vyšší četnost.

Správná odpověď *b* má s celkovým skóre kladnou korelaci $r = 0,33$, což je třetí nejmenší kladná korelace v testu. Podobně jako předchozí úloha má druhou nejnižší hodnotu indexu diskriminace položky, což může být opět způsobeno vyšší šancí na uhádnutí odpovědi. Přestože se jedná o velmi podobnou úlohu úloze předchozí, je korelace mezi předchozím a tímto výsledkem pouze slabá ($r = 0,21$). S ohledem na to je relativně vysoká úspěšnost této úlohy pravděpodobně způsobena neporozuměním rozdílu mezi předchozí a touto úlohou a následným nezměněním úvahy, která ovšem tentokrát vede ke správnému řešení. Schopnost jejího správného vyřešení má tedy opět relativně malý vliv na kognitivní úroveň řešitele.

4 Výsledky analýzy Lawsonova testu jako celku

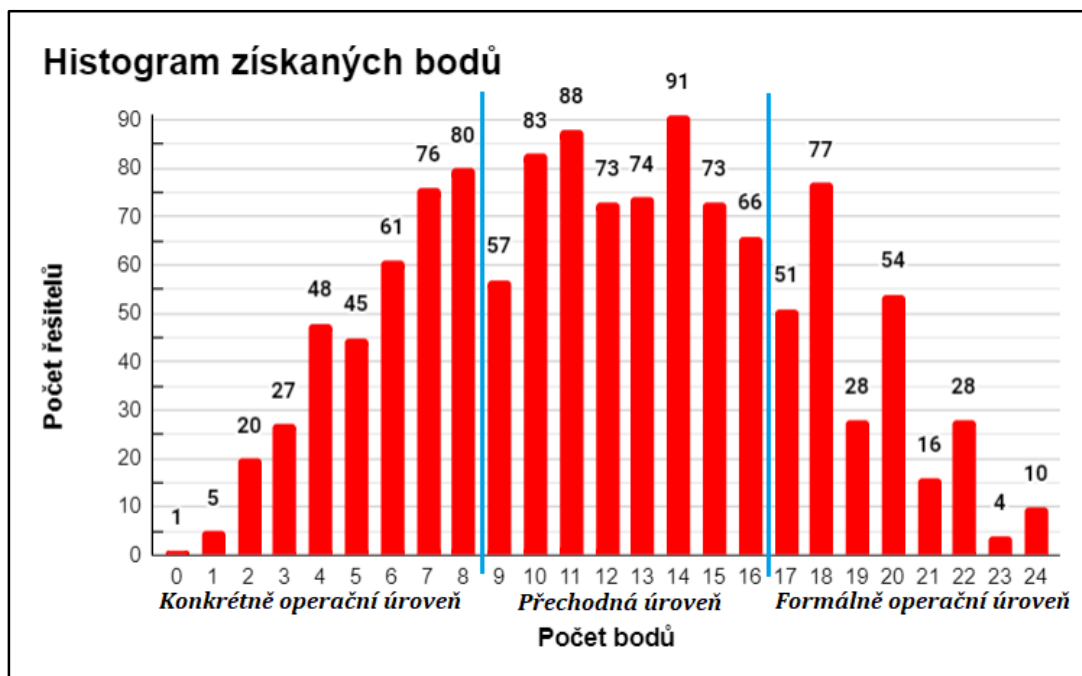
4.1 Výsledky s ohledem na kognitivní úroveň

V histogramu 4.1 níže můžeme vidět, že nejnižšího skóre 0 bodů dosáhl pouze jeden řešitel, nejvyššího skóre 24 bodů dosáhlo 10 z 1236ti řešitelů. Průměrné dosažené skóre je 12,0 bodů, což odpovídá průměrné úspěšnosti 50 %. Hodnota směrodatné odchylky σ je 5,2.

Šikmost hustoty rozdělení bodů je levostranná ($\gamma_1 = -0,31$), většina řešitelů tedy dosáhla lepšího výsledku, než byl průměr.

Koeficient špičatosti je záporný ($\gamma_2 = -1,30$), rozdělení je tedy rovnoměrnější a křivka hustoty je plošší než u normálního rozdělení.

Celkem 363 (29,4 %) řešitelů dosáhlo konkrétně operační kognitivní úrovně (0-8 bodů), 605 (48,9 %) řešitelů dosáhlo přechodné kognitivní úrovně (9-16 bodů) a 268 (21,7 %) řešitelů dosáhlo formálně operační úrovně (17-24 bodů).



Obrázek. 4.1: Histogram získaných bodů

4.2 Výsledky s ohledem na typ uvažování

V tabulce 4.1 níže můžeme vidět výsledky analýzy dle dané kategorie typu uvažování.

Tabulka 4.1: Výsledky testu s ohledem na kategorie

Kategorie	Počet dvojúloh v kategorii	Průměrná úspěšnost	Úspěšnost jednotlivých dvojúloh	Průměrná korelace správné odpovědi s celkovým skóre	Korelace s celkovým skóre jednotlivých dvojúloh
Zachování hmotnosti	1	93 %	93 %	0,24	0,24
Zachování vytlačeného objemu	1	75 %	75 %	0,50	0,50
Poměrové myšlení	2	44 %	58 %; 30 %	0,55	0,59; 0,50
Identifikace a kontrola změny	3	39 %	72 %; 17 %; 29 %	0,48	0,51; 0,47; 0,46
Pravděpodobnostní myšlení	2	54 %	53 %; 55 %	0,62	0,63; 0,60
Korelační myšlení	1	40 %	40 %	0,50	0,50
Kombinační myšlení	3	44 %	21 %; 41 %; 69 %	0,43	0,46; 0,50; 0,33

Kategorie *Zachování hmotnosti* má výrazně nejvyšší úspěšnost, což je pravděpodobně úmysl viz. komentář k úloze 1 a 2. v kapitole 3. Nejnižší průměrnou úspěšnost má kategorie *Identifikace a kontrola změny*. Přestože má jedna ze tří dvojúloh této kategorie třetí nejnižší obtížnost, zbylé dvě mají naopak nejvyšší, respektive třetí nejvyšší obtížnost. Přes tyto velké rozdíly v úspěšnosti se příliš neliší v hodnotách korelace správné odpovědi s celkovým skóre. Možné vysvětlení je k nalezení v komentáři k příslušným dvojúlohám v kapitole 3.

Kategorie s nejmenším rozdílem úspěšnosti mezi dvojúlohami byla kategorie *Pravděpodobnostního myšlení*. Všechny ostatní kategorie s více dvojúlohami vyžadovaly pro správné vyřešení snazší i obtížnější úvahy, které se vzájemně mírně lišily. Oproti tomu úvaha pro vyřešení dvojúloh kategorie *Pravděpodobnostního myšlení* je v obou případech v podstatě totožná, řešitelé, kteří správně vyřešili jednu z dvojúloh, většinou vyřešili správně i druhou viz. komentáře k příslušným dvojúlohám v kapitole 3.

4.3 Porovnání našich výsledků s výsledky jiných domácích a zahraničních studií

Průměrný věk našeho řešitele byl 15,0 let. Tyto výsledky můžeme srovnat s výsledky Lawsona [9], který prováděl testování na amerických školách s jinou verzí svého testu, ovšem na podobně starých řešitelích (15,2 let). Američtí řešitelé dosáhli průměrné úspěšnosti 49,7 %, což je srovnatelné s naší průměrnou úspěšností. 35,3 % dosáhlo úrovně konkrétně operační, 49,5 % úrovně přechodné a 15,2 % úrovně formálně operační, čeští řešitelé tedy měli menší zastoupení v nejhorší kognitivní úrovni a vyšší zastoupení v nejvyšší kognitivní úrovni. Ve srovnání s našimi výsledky byly tedy výsledky amerických řešitelů stejného věku horší.

Dále můžeme naše výsledky srovnat s výsledky Hejnové [19], která provedla v roce 2017 testování na 165 žácích českých základních škol a gymnázií ve věku 14-15 let. Průměrné dosažené skóre bylo 8,0 bodů, což odpovídá úspěšnosti 33,3 %. Tito žáci tedy dosáhli horších výsledků, než naši řešitelé. Z těchto testovaných žáků patřila část do skupiny, kterou učila fyziku Dvořáková, a ti dosáhli průměrného skóre 14,8 bodů, což odpovídá úspěšnosti 61,6 %. Dvořáková toto přisuzuje „způsobu vedení výuky, ve které hraje velmi podstatnou roli experimentování, jež žákům dává dostatek podnětů k rozvoji vědeckého uvažování.“ Han [10] ve svém výzkumu uvádí, že na kognitivní úroveň řešitele má větší vliv úroveň myšlení učitele, než jeho obsahové znalosti. Tvzení Dvořákové by tomuto odpovídalo.

Bao [20] porovnává výsledky mezi americkými a čínskými studenty prvních ročníků STEM oborů (science, technology, engineering, and mathematics) s ohledem na různé způsoby a úrovně výuky fyziky a matematiky na amerických a čínských středních a základních školách. Přestože se čínské školy více zaměřovali na klasické trénování fyzikálních znalostí jako je například memorizace faktů, nedosahovali čínští řešitelé lepších výsledků než řešitelé američtí. Přestože některé úlohy Lawsonova testu mohou mít fyzikální charakter, řešitelovo dobré znalosti oborů STEM nevedou nutně k lepší kognitivní úrovni. Zároveň ovšem Bao zmiňuje lepší výsledky u řešitelů, kteří absolvovali výuku, jejíž součástí bylo explicitní zlepšení kognitivního vnímání studentů.

Moore a Rubbo [21] ve svém výzkumu uvádějí, že studenti prvních ročníků STEM oborů mají výrazně lepší výsledky než studenti prvních ročníků jiných oborů. Přestože se tedy vysoká úroveň kognitivního myšlení vyskytuje převážně u studentů STEM oborů, jejich znalosti v těchto oborech nevedou k vyšší úrovni kognitivního myšlení.

Piaget [22] předpokládá, že se do finálního kognitivního stádia, stádia formálních operací, dostává člověk ve věku od 12 let výše. Vzhledem k tomu, že nejnovější verze lawsonova testu, kterou tato práce analyzuje, je určena zejména pro studenty, kteří dokončují středoškolské studium či začínají vysokoškolské studium [20], nenabízí se mnoho vhodných srovnání s řešiteli podobného věku. S ohledem na Piagetův předpoklad a na výše uvedené výsledky se můžeme pokusit o srovnání našich řešitelů s řešiteli výrazně staršími, jelikož jejich lepší znalosti fyziky a matematiky by neměli mít na výsledek velký vliv.

Bao [20] uvádí hodnotu průměrné úspěšnosti jeho řešitelů kolem 70 %. Jeho výzkum se ale zaměřuje zejména na studenty STEM oborů, u kterých se dle Moora a Rubbeho vyskytuje vyšší úroveň uvažování. Ti ve své práci [21] uvádějí hodnotu průměrné úspěšnosti u studentů prvních ročníků jiných, než STEM oborů 54 %. Tito řešitelé tedy dosahují mírně lepších výsledků, než naši řešitelé.

Následně můžeme naše řešitele srovnat s výsledky Hrouzkové a Richterka [23], kteří testovali 446 studentů prvních ročníků oborů fyzika, chemie a učitelství fyziky a chemie na Palackého Univerzitě v Olomouci. Průměrný počet získaných bodů těchto řešitelů byl 12,7, což odpovídá průměrné úspěšnosti 52,8 %. 24,7 % dosáhlo úrovně konkrétně operační, 50,4 % úrovně přechodné a 24,9 % úrovně formálně operační. Ve výsledku byli tedy tito řešitelé také mírně úspěšnější, než naši řešitelé.

Podobné srovnání výsledků našich řešitelů s výsledky Novákové [24], která testovala 125 vysokoškolských studentů oboru učitelství pro 1. stupeň ZŠ, pro změnu vede k výsledkům opačným, než v předchozích příkladech. Tito studenti dosáhli průměrné úspěšnosti pouhých 30 %, což je výrazně horší výsledek než průměrná úspěšnost našich řešitelů. Nováková dále uvádí, že studenti, kteří prošli gymnaziálním studiem, měli výsledky lepší, než studenti z ostatních středních škol. To může být způsobeno rozdílným přístupem k výuce na těchto typech škol.

Piaget sice předpokládá nejvyšší úroveň myšlení již od 12ti let, byl ovšem kritizován za příliš biologicky orientovaný kognitivní vývoj jedince [3]. Zároveň Ding [25] ve svém výzkumu zjišťuje minimální změnu kognitivní úrovně v průběhu vysokoškolského studia, přestože se jedná o přibližně stejný časový rozdíl, jako mezi našimi řešiteli a studenty prvních ročníků vysokých škol. Lepší výsledky vysokoškolských studentů tedy pravděpodobně nebudou způsobeny věkem, ani větší oborovou odborností. Jak uvádí Bao [20] a Han [10], vyšší kognitivní úroveň úzce souvisí s explicitním záměrem učitele zlepšit kognitivní úroveň svých studentů, což potvrzují i výše uvedené výsledky Hejnové [19] a Novákové [24]. Ke stejnému závěru dospěli i Marušić a Sliško [26], kteří ve své práci uvádějí zlepšení kognitivní úrovně u studentů středních škol, jejichž učitelé použili ve své výuce fyziky metody ED (*Experimenting and discussion*) a RPQ (*Reading, Presenting, and Questioning*).

Naši řešitelé tedy dosahují lepších, ale i horších výsledků, než řešitelé zmíněných vysokých škol. Výsledek pravděpodobně závisí na konkrétním stylu výuky, kterou si řešitelé prošli. Možným rozšířením této práce by mohla být analýza výsledků v závislosti na různých typech škol či vzdělávacích systémech, čímž bychom mohli srovnat jejich míru explicitního věnování se rozvoji kognitivního myšlení.

Pro srovnání našich výsledků s ohledem na typ uvažování a výsledků dalších prací je vhodné poznamenat, že část zahraničních prací kombinuje kategorie *Zachování hmotnosti* a *Zachování vytlačeného objemu* do jedné kategorie nazývané obvykle *Zachování*. Výsledky naší analýzy částečně odpovídají výsledkům Hrouzkové a Richterka [23], Piraksa, Srisawasdi a Koul a spol.[27], Moora a Rubbeho [21] i Sriyansyah a Saepuzaman a spol. [28].

Všechny tyto práce uvádí nejvyšší úspěšnost v kategorii *Zachování*, což se shoduje s naším výsledkem. Dále jsou v těchto pracích kategorie *Identifikace a kontrola změny*, *Poměrové myšlení*, *Korelační myšlení* a *Kombinační myšlení* vždy mezi čtyřmi kategoriemi s nejnižší průměrnou úspěšností, mění se pouze jejich pořadí, což také odpovídá našim výsledkům.

Pořadí úloh dle obtížnosti se příliš nemění ani v závislosti na průměrné celkové úspěšnosti. Srovnáme-li například výsledky Moora a Rubbeho [21] s výsledky Novákové [24], pak přestože dosahují řešitelé v první práci průměrné

celkové úspěšnosti o 24 % vyšší, než řešitelé v práci druhé, liší se pořadí úloh dle obtížnosti opět pouze v posledních čtyřech kategoriích. Relativní obtížnosti různých kategorií typů myšlení se tedy vzhledem k různé kognitivní úrovni řešitele mění pouze velmi málo.

Na závěr je v tabulce uvedena *Průměrná korelace správné odpovědi s celkovým skóre*. Kategorie, která má nejsilnější korelaci mezi zvolením správné odpovědi a celkovým skóre je *Pravděpodobnostní myšlení*. Tato kategorie patří, dle našich výsledků, mezi kategorie snazší, přesto dobře rozlišovala mezi nejlepšími řešiteli viz. komentáře k příslušným dvojúlohám v kapitole 3. Zvolení správné odpovědi v náročnějších kategoriích nevede nutně k silné korelaci s celkovým výsledkem, např. přestože je kategorie *Kombinační myšlení* třetí nejnáročnější, má druhou nejslabší korelaci.

Jak již bylo uvedeno, kognitivní úroveň žáka úzce souvisí s explicitní snahou vyučujícího tuto úroveň zlepšit. Marušić a Sliško [26] ve své práci pozorují zlepšení kognitivní úrovně studentů při experimentování a diskuzi v hodinách fyziky. Dle našich výsledků můžeme tedy navrhnout doporučení na zahrnutí experimentů na témata pravděpodobnost a zachování poměru do výuky fyziky. Dobře rozvinuté pravděpodobnostní myšlení silně koreluje s dobrou kognitivní úrovní, špatně rozvinuté porozumění zachování poměru naopak silně koreluje se špatnou kognitivní úrovní viz. komentář k příslušným dvojúlohám v kapitole 3.

Bao [14] diskutuje nad horší validitou u pěti dvojúloh. Jednalo se o dvojúlohy 7;8, 11;12, 13;14, 21;22 a 23;24.

U dvojúlohy 7;8 kritizuje distraktor vysvětlení *e*, který považuje za správný. Uvádí ovšem, že počet řešitelů kteří toto vysvětlení zvolili a zároveň vyřešili předchozí úlohu správně, je velmi blízký počtu řešitelů, kteří předchozí úlohu vyřešili chybně. V našem případě distraktor *e* volili i celkově úspěšnější řešitelé a to v takové míře, že volba vysvětlení *e* spolu se správným řešením úlohy 7 má kladnou korelaci s celkovým skóre. S horší validitou této dvojúlohy tedy souhlasíme.

U dvojúloh 11;12 a 13;14 uvádí zjištění, že obrázková reprezentace uvedených problémů řešitele spíše mátlá, horší a průměrní řešitelé na ně často odkazovali při určování chybných vysvětlení svých chybných odpovědí. Naše testování ukázalo relativně malou hodnotu korelace správné odpovědi s

celkovým skóre, přestože se jedná o dvojúlohy velmi obtížně. S horší validitou těchto dvojúloh tedy také souhlasíme.

U dvojúlohy 21;22 kritizuje použití velkého množství pojmů jako je suchý led a pouťový balónek, díky kterým může být řešitel sváděn k použití zapamatovaných vědomostí, dále kritizuje malou věrohodnost některých z popsaných situací. Naše práce ukazuje velký počet relevantních distraktorů které mají nulovou, či velmi slabou negativní korelaci s celkovým skóre. Zároveň má správné řešení relativně malou korelaci s celkovým skóre i přes vyšší obtížnost stejně jako v případě dvojúloh 11;12 a 13;14. S horší validitou této dvojúlohy tedy, jako u předchozích dvojúloh, souhlasíme.

U úloh 23 a 24 opět kritizuje malou věrohodnost popsaných situací, dále poukazuje na zvýšenou pravděpodobnost nevěnování dostatečné pozornosti zadání, řešitelé mají totiž hypotézy vyvrátit oproti očekávanému potvrzování. Vzhledem k relativně malé korelaci správného vyřešení s celkovým skóre a zvýšenou pravděpodobností uhádnutí správného řešení opět souhlasíme s horší validitou těchto úloh.

Xiao [29] ve své práci diskutuje problematiku analyzování distraktorů dvojúloh. Xiao provádí analýzu distraktorů při použití modelu Rasch a jeho různých modifikací, které pracují s různým tříděním distraktorů. Testuje rozdíly mezi rozdělením distraktorů například na 3 typy - 00 kdy řešitel nevyřešil ani jednu z úloh správně, 10 kdy vyřešil správně pouze první úlohu a 01 kdy vyřešil správně pouze druhou úlohu. Jiné rozdělení zase nerozlišuje distraktory 10 a 01. Ukazuje se, že můžeme získat lepší výsledky za použití různého rozdělení distraktorů pro různé dvojúlohy. Naše práce využila při analýze dle IRT model BLIRT viz. kapitola 2.4, navíc tato práce distraktory třídí ještě podrobněji díky zavedení pojmu relevantní odpovědi viz. kapitola 2.2. Výsledky tedy můžeme srovnat pouze v rovině kvalitativní. Obecně se obě práce nejvíce shodují v určení distraktorů s vysokou hodnotou atraktivity, které dobře rozlišovali nejméně úspěšné od nejvíce úspěšných řešitelů. Dále se výsledky shodovali v případech, kde Xiao použil jemnější třídění distraktorů. V ostatních případech se výsledky nedají dobře srovnat vzhledem k většímu rozdílu v počtu distraktorů.

Závěr

V rámci diplomové práce byla zpracována data z českých základních škol a gymnázií. Jednalo se o řešení úloh celkem 1236 respondentů testu LCTSR. Následně byla provedena analýza úloh i celého testu a to s využitím jak CTT, tak IRT.

Výsledky analýzy jednotlivých dvojúloh testu byly představeny formou dvanácti kapitol. V každé kapitole byly uvedeny relevantní parametry, charakteristické křivky a komentář k příslušné dvojúloze.

Výsledky naší analýzy testu jako celku se shodují se zahraničními studii. Relativní obtížnosti úloh dle kategorie jsou s výsledky ostatních studií, nejobtížnější bývá vždy jedna z kategorií *Identifikace a kontrola změny*, *Poměrové myšlení*, *Korelační myšlení* a *Kombinační myšlení*, nejsnazší je vždy kategorie *Zachování*.

Průměrná úspěšnost našich řešitelů byla vyšší i nižší než u srovnávaných řešitelů, kteří byly starší a měli více fyzikálních odborných znalostí. To se shoduje s výsledky studií poukazující na skutečnost, že na míru úspěšnosti má největší vliv míra učitelovo zaměření na vědecké myšlení, zanedbatelný vliv má míra odborných znalostí a od určitého věku i stáří respondentů.

Vzhledem k většímu množství špatných distraktorů souhlasíme s horší validitou úloh 7;8, 11;12, 13;14, 21;22 a 23;24. Navrhujeme zlepšení validity těchto úloh změněním formulace problémových distraktorů.

Byla určena korelace volby správné či špatné odpovědi s celkovým skórem. Navrhujeme zaměřit výuku zejména na kategorie s vysokými hodnotami korelace. Nejlepší rozlišovací schopnosti mezi lepšími a horšími řešiteli poskytovala kategorie *Pravděpodobnostní myšlení*, zároveň měla tato kategorie nejsilnější pozitivní korelaci správného řešení s celkovým skóre. Dobře rozvinuté pravděpodobností myšlení má tedy velký vliv na dobrou úroveň vědeckého myšlení. Nejsilnější negativní korelaci s celkovým skóre měl jeden z distraktorů v dvojúloze 5;6, která také velmi dobře rozlišovala lepší a horší řešitele. Zvolení tohoto distraktoru naznačuje neporozumění zachování poměru. Porozumění respektive neporozumění těmto dvěma konceptům má, v porovnání s ostatními úlohami, největší význam na rozvoj vědeckého myšlení žáků.

Je třeba upozornit, že přestože tato práce může pomoci vyučujícím s výběrem kategorie myšlení, kterou chtějí ve svých hodinách rozvíjet, aby dosáhli dobrého zlepšení úrovně vědeckého myšlení, neposkytuje žádné konkrétní metody či postupy, které by se daly do výuky implementovat.

Možným rozšířením práce by mohlo být další testování, které by bralo v potaz různé přístupy učitelů k výuce. Bylo by vhodné ověřit, zda zaměření učitele na kategorie myšlení, které mají silnější korelaci s celkovým skóre v testu, skutečně povede k lepším výsledkům, než kdyby se učitel zaměřil na kategorii s korelací slabší.

Příloha

Tabulka 5.1: Shrnutí relevantních parametrů jednotlivých dvojúloh

Úloha	Úspěšnost	Index diskriminace ULI	Kor. koef. pro správné řešení
1 a 2	93 %	0,13	0,24
3 a 4	75 %	0,54	0,50
5 a 6	58 %	0,73	0,59
7 a 8	30 %	0,58	0,50
9 a 10	72 %	0,59	0,51
11 a 12	17 %	0,42	0,47
13 a 14	29 %	0,50	0,46
15 a 16	53 %	0,80	0,63
17 a 18	55 %	0,75	0,60
19 a 20	40 %	0,63	0,50
21 a 22	21 %	0,47	0,46
23	41 %	0,37	0,50
24	69 %	0,38	0,33

Literatura

- [1] C. Zimmerman, The development of scientific reasoning: What psychologists contribute to an understanding of elementary science learning, Paper commissioned by the National Academies of Science (National Research Council's Board of Science Education, Consensus Study on Learning Science, Kindergarten through Eighth Grade) (2005).
- [2] National Research Council, Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century (National Academies Press, Washington, DC, 2012).
- [3] KOHOUTEK, Rudolf. Kognitivní vývoj dětí a školní vzdělávání. Pedagogická orientace 2008, 18(3). ISSN 1211-4969.
- [4] J. Piaget, The stages of the intellectual development of the child, in Readings in Child Development and Personality, edited by P. H. Mussen, J. J. Conger, and J. Kagan (Harper and Row, New York, 1970).
- [5] L. Bao, T. Cai, K. Koenig, K. Fang, J. Han, J. Wang, Q. Liu, L. Ding, L. Cui, Y. Luo et al., Learning and scientific reasoning, Science 323, 586 (2009). [online] [cit. 15.7.2023] Dostupné z: DOI:10.1126/science.1167740
- [6] A. M. Cavallo, M. Rozman, J. Blickenstaff, and N. Walker, Learning, reasoning, motivation, and epistemological beliefs, J. Coll. Sci. Teach. 33, 18 (2003). [online] [cit. 15.7.2023] Dostupné z :
https://www.researchgate.net/publication/285323458_Learning_reasoning_motivation_and_epistemological_beliefs
- [7] S. Ates and E. Cataloglu, The effects of students' reasoning abilities on conceptual understandings and problem-solving skills in introductory mechanics, Eur. J. Phys. 28, 1161 (2007). [online] [cit. 15.7.2023] Dostupné z: DOI 10.1088/0143-0807/28/6/013
- [8] J. L. Jensen and A. Lawson, Effects of collaborative group composition and inquiry instruction on reasoning gains and achievement in undergraduate biology, CBE-Life Sci. Educ. 10, 64 (2011). [online] [cit. 15.7.2023] Dostupné z: <https://doi.org/10.1187/cbe.10-07-0089>

- [9] Lawson, A.E. The Development and Validation of a Classroom Test of Formal Reasoning. JOURNAL OF RESEARCH IN SCIENCE TEACHING, 1978.
- [10] HAN, Jing. Scientific reasoning: research, development, and assessment. Ohio: The Ohio State University 2013. Dissertation.
- [11] A. E. Lawson, Lawson classroom test of scientific reasoning. [online] [cit. 15.7.2023] Dostupné z:
<http://www.public.asu.edu/~anton1/AssessArticles/Assessments/Mathematics%20Assessments/Scientific%20Reasoning%20Test.pdf> (2000).
- [12] C. Pratt and R. G. Hacker, Is Lawson's classroom test of formal reasoning valid?, Educ. Psychol. Meas. 44, 441 (1984). [online] [cit. 15.7.2023]
Dostupné z: <https://doi.org/10.1177/0013164484442025>
- [13] G. P. Stefanich, R. D. Unruh, B. Perry, and G. Phillips, Convergent validity of group tests of cognitive development, J. Res. Sci. Teach. 20, 557 (1983) [online] [cit. 15.7.2023]
Dostupné z: <https://doi.org/10.1002/tea.3660200607>
- [14] BAO, Lei, et. al. Validity evaluation of the Lawson classroom test of scientific reasoning. Physical Review Physics Education Research, 14(2), 020106, 2018. [online] [cit. 15.7.2023]
Dostupné z: DOI: 10.1103/PhysRevPhysEducRes.14.020106
- [15] Dvořáková, I., Sborníky dílen Heuréky 2011 [online] [cit. 9.5.2023].
Dostupné z:
<https://kdf.mff.cuni.cz/heureka/ke-stazeni/sborniky-dilen-heureka/sborniky/SbornikDilenHeureka2011.pdf>
- [16] Zvára, K., Štěpán, J., Pravděpodobnost a matematická statistika, MatfyzPress, 2019.
- [17] Ebel, R. L. (1972). Essentials of Educational Measurement (1st ed.). Upper Saddle River, NJ: Prentice Hall
- [18] Martinkova, P., & Drabinova, A. (2018). ShinyItemAnalysis for teaching psychometrics and to enforce routine analysis of educational tests. The R Journal, 10(2), 503-515, Dostupné z: doi: 10.32614/RJ-2018-074

- [19] HEJNOVA, Eva. Testování vědeckého myšlení. MATEMATIKA-FYZIKA - INFORMATIKA, 27(5), 350-359, 2018.
- [20] Bao. L. (2008) Learning of Content Knowledge and Development of Scientific Reasoning Ability: A Cross Culture Comparison, American Journal of Physics 77(12)
- [21] Moore, J. C., & Rubbo, L. J. (2012). Scientific reasoning abilities of nonscience majors in physicsbased courses. Physical Review Special Topics - Physics Education Research, 8(1), Article 010106. [online] [cit. 1.7.2023]. Dostupné z: <https://doi.org/10.1103/PhysRevSTPER.8.010106>
- [22] Piaget, Jean & Inhelder, Brbel, (1958). The Growth Of Logical Thinking From Childhood To Adolescence: AN ESSAY ON THE CONSTRUCTION OF FORMAL OPERATIONAL STRUCTURES (1st ed.). Routledge. [online] [cit. 1.7.2023] Dostupné z: <https://doi.org/10.4324/9781315009674>
- [23] HROUZKOVÁ T., RICHTEREK L., Proceedings of the 4th International Baltic Symposium on Science and Technology Education, BalticSTE2021, LAWSON CLASSROOM TEST OF SCIENTIFIC REASONING AT ENTRANCE UNIVERSITY LEVEL [online] [cit. 4.7.2023] Dostupné z: <https://files.eric.ed.gov/fulltext/ED620309.pdf>
- [24] NOVÁKOVÁ, Alena, Vlastimil CHYTRÝ a Jaroslav ŘÍČAN. Vědecké myšlení a metakognitivní monitorování studentů učitelství pro 1. stupeň základní školy. Scientia in educatione, 9(1), 66–80, 2018. ISSN 1804-7106.
- [25] Ding, L., Wei, X. & Mollohan, K. Does Higher Education Improve Student Scientific Reasoning Skills?. Int J of Sci and Math Educ 14, (2016). [online] [cit. 1.7.2023] Dostupné z: <https://doi.org/10.1007/s10763-014-9597-y>
- [26] Mirko Marušić & Josip Sliško (2012) Influence of Three Different Methods of Teaching Physics on the Gain in Students' Development of Reasoning, International Journal of Science Education, 34:2, 301-326, [online] [cit. 11.7.2023] Dostupné z: DOI: 10.1080/09500693.2011.582522

[27] Piraksa, C., Srisawasdi, N., & Koul, R. (2014). Effect of gender on student's scientific reasoning ability: A case study in Thailand. *Procedia - Social and Behavioral Sciences*, 116, 486– 491. [online] [cit. 8.7.2023]

Dostupné z: <https://doi.org/10.1016/j.sbspro.2014.01.245>

[28] SRIYANSYAH, Syakti P. a Duden Saepuzaman. Prospective Physic Teachers' Consistency and Scientific Reasoning in the Learning Force Concept. In *International Conference on Mathematics and Science Education*. 57, 21–24.

[29] Yang Xiao, Jing Han, Kathleen Koenig, Jianwen Xiong, and Lei Bao, Multilevel Rasch modeling of two-tier multiple choice test: A case study using Lawson's classroom test of scientific reasoning. *Phys. Rev. Phys. Educ. Res.* 14, 020104 – Published 19 July 2018, [online] [cit. 14.7.2023] Dostupné z: DOI: 10.1103/PhysRevPhysEducRes.14.020104