

TESTOVÁNÍ ZNALOSTÍ STŘEDOŠKOLSKÉ FYZIKY U STUDENTŮ 1. ROČNÍKŮ LÉKAŘSKÝCH FAKULT V ČR – PRAKTICKÉ VYUŽITÍ METODIKY ANALÝZY VÝSLEDKŮ

**Běláček Jaromír, Komarc Martin, Kymplová Jaroslava, Kvašňák
Eugen, Mornstein Vojtěch, Kubeš Zdeněk, Sochorová Hana**

Anotace

Pedagogičtí pracovníci biofyzikálních ústavů lékařských fakult v ČR organizují již ve druhém školním roce prezenční anketu/testování znalostí středoškolské fyziky u studentů 1. ročníků. Cílem testu je zmapovat souhrnně i specifické znalosti studentů v rámci 8 tématických oblastí (Optika, Elektřina, Teplo, Akustika, Radioaktivita, Magnetismus, Tlak a Fyzikální veličiny). Testování proběhlo ve školním roce 2012/13 na 1. a 3. LF UK v Praze, na LF UK v Plzni a na MU v Brně, v metodicky vylepšené verzi i v r. 2013/14 (připojila se LF OU v Ostravě a 2.LF UK v Praze).

Test sestává z 30 resp. (letos) z 32 otázek s možností zatržení právě jediné správné ze čtyř promptovaných odpovědí. Součástí testu jsou rovněž základní geodemografické charakteristiky studentů použité jako třídící pro následnou ANOVA analýzu diferenciací zjištěných znalostí. Kromě standardních třídění jsme provedli rovněž analýzu baterie všech otázek/položek testu prostřednictvím Cronbachova alfa (CA) a několika IRT (Item Response Theory) modelů.

Nejlepších souhrnných výsledků v r. 2012/13 bylo dosaženo na MU Brno (v průměru 19,0 správných odpovědí), následovala 3. LF UK (18,0), 1. LF UK (17,2) a LF UK Plzeň (16,8). Všeobecně lepších výsledků dosahovali absolventi víceletých resp. čtyřletých gymnázií (18,3 resp. 17,6) oproti absolventům ostatních SŠ (15,2), vyšší znalostí základů fyziky měli muži (19,4) než ženy (17,1). Hlavní závěry byly potvrzeny i ve školním roce 2013/14, kde jsme použili již „vylepšenou“ verzi testu zohledňující výsledky analýzy reliability baterie 30 otázek a IRT modelů provedených oddělením BioStat při ÚBI 1. LF UK v předchozím roce.

Budoucnost testování středoškolských znalostí studentů lze vidět např. v efektivnějším využití internetových (jmenovitě i dvou již provedených Moodle) aplikací anebo v redukci počtu testovaných osob na základě kvalifikované připravených výběrových šetření.

Klíčová slova:

testování znalostí z fyziky, ANOVA, analýza reliability testu, Cronbachovo alfa, optimální výběr otázek/položek testu (IRT modely)

1. Úvod

Prověřování znalostí studentů by mělo být nezbytnou součástí jakéhokoli pedagogického procesu ([1]). Testování prostřednictvím standardizované sady (baterie) otázek je vhodným nástrojem zejména v případech, kdy potřebujeme ověřit nabyté znalosti ve větším měřítku (větší počet studentů nebo tématických oblastí). Tento příspěvek je zaměřen na znalosti studentů 1. ročníků lékařských fakult v ČR, ať již z důvodů porovnání organizace a efektů přijímacího řízení

na jednotlivých VŠ anebo z pohledu porovnávání znalostí studentů z hlediska přípravy na různých typech středních škol, v odlišných regionech, podle pohlaví apod.

V rámci této studie jsou stručně shrnuty výsledky formální statistické analýzy testů z fyziky, provedených pedagogickými pracovníky biofyzikálních ústavů několika lékařských fakult v ČR, kteří se dohodli na přípravě a realizaci prezenční ankety znalostí středoškolské fyziky u studentů 1. ročníků. Cílem testu bylo zmapovat souhrnné i specifické znalosti studentů v rámci 8 standardně vyučovaných oblastí: Optika, Elektřina, Teplo, Akustika, Radioaktivita, Magnetismus, Tlak a Fyzikální veličiny.

2. Materiál a metody

Testování proběhlo ve školním roce 2012/13 na 1. a 3. LF UK v Praze, na LF UK v Plzni a na MU v Brně, v metodicky vylepšené verzi i v r. 2013/14 (připojila se LF OU v Ostravě a 2.LF UK v Praze). Test sestával z 30 resp. (letos) z 32 otázek s možností zatržení právě jediné správné ze čtyř promptovaných odpovědí (single response). Výběr otázek, reprezentujících (až na oblast radioaktivity) vždy po čtyřech každou z výše jmenovaných oblastí, návrh jejich přesného znění a verifikace správnosti proběhla na půdě všech zúčastněných fakult. Pilotáž, včetně stanovení maximální povolené doby vyplňování testu, znáhodnění pořadí otázek do několika Excelovských verzí testu a formální statistická analýza výsledků byla provedena na oddělení BioStat při ÚBI na 1. LF UK Praha.

Ve školním roce 2012/13 byl test vyplněný prezenčně všemi studenty, kteří se dostavili na 1. hodinu praktických cvičení z biofyziky. Jednotně modifikovaná verze testu byla předložena studentům 1. ročníku na týchž fakultách o rok později, na některých fakultách ale při pozměněné organizaci vyplňování testu (termíny realizace terénu, počty zúčastněných studentů a rozdíly v organizaci vyplňování formulářů viz Tabulka 1).

Výsledky testu reprezentované jednak průměry souhrnných/kompozitních skóre (= součet všech správných odpovědí) a průměry dílčích součtů v rámci každé z osmi testovaných oblastí fyziky byly vyhodnoceny prostřednictvím ANOVA podle fakult, místa trvalého bydliště respondentů (kraje ČR, stát), typu střední školy (gymnázium, víceleté gymnázium, ostatní SŠ), právního statutu absolvované střední školy (státní, soukromá) a podle pohlaví.

V rámci analýzy výsledků po prvním roce testování znalostí byla provedena analýza reliability vstupní baterie 30 otázek a validity odpovědí na jednotlivé položky testu s pomocí explorační faktorové analýzy (EFA), několikánásobnými přepočty Cronbachova alfa a dvou IRT (Item Response Theory) modelů (PL1 a PL2). Stejně statistické modely byly použity pro analýzu výsledků ve školním roce 2013/14, kdy byla použita již modifikovaná verze testu (opravená a rozšířená na více konzistentní baterii 32 otázek) a výsledky porovnány s předchozími (po odpovídajících přepočtech).

Pro analýzu dat byl použit statistický SW SPSS [2], STATISTICA [3] a balíček LTM v programu R [4].

2. Výsledky

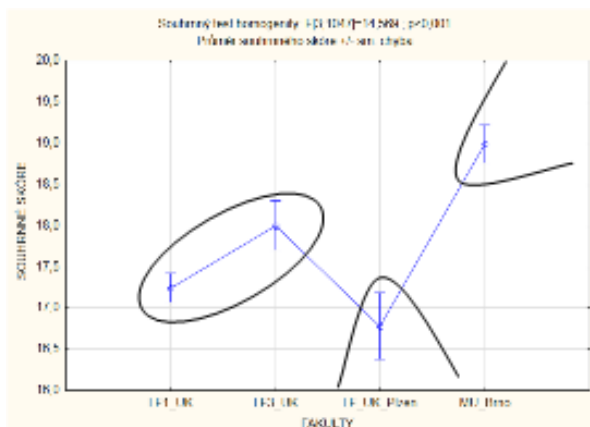
2.1 Souhrnná skóre a tématické oblasti

Průměr souhrnných skóre znalostí středoškolské fyziky za všechny čtyři lékařské fakulty ve školním roce 2012/13 (tzn. z 30 vybraných single response otázek) činil 17,8 (+/- SEM = 0,123). Jak ovšem vyplývá z údajů uvedených v Tabulce 1, nejlepších výsledků bylo dosaženo na LF MU v Brně (18,99 +/- 0,283), nejhorších výsledků na LF UK v Plzni (16,78 +/- 0,368), průměrná souhrnná skóre na 1. LF UK Praha (17,24 +/- 0,169) a na 3. LF UK Praha (18,00 +/- 0,286) se mezi sebou statisticky příliš neliší. Na Grafu 1 jsou tyto hodnoty a jejich statistické významnosti vyjádřeny ještě doplněnými parabolami a elipsami.

Fakulta:	Měsíc/Rok	Formulář	N	Souhrnné skóre	Směrodatná chyba
1. LF UK Praha	10/2012	v Excelu	502	17,24	,169
3. LF UK Praha	10/2012	v Excelu	169	18,00	,286
LF UK v Plzni	2/2013	v Excelu	95	16,78	,368
LF MU v Brně	10/2012	na papíře	285	18,99	,253
Celkem:			1051	17,80	,123

Tabulka 1 — Termíny realizace testu znalostí SŠ fyziky ve školním roce 2012/13 a vztahové specifikace

Legenda: N — počet řádně vyplněných testů (= počet respondentů)



Graf 1 — Průměry souhrnného skóre [0-30] z testu znalostí SŠ fyziky u studentů 1. ročníku 4 lékařských fakult ve školním roce 2012/13

Legenda: Souhrnné výsledky testu z fyziky se v rámci 4 lékařských fakult atestovaných ve školním roce 2012/13 liší statisticky významně ($F[3;1047] = 14,569$; $p < 0,001$). Statisticky významně nejlepších výsledků bylo dosaženo na LF MU Brno (papírové verze testu), statisticky významně nejhorších výsledků na LF UK v Plzni, průměrná souhrnná skóre na 1. a 3. LF UK Praha se statisticky významně neliší. Paraboly resp. elipsy značí průměry statisticky významně oddělitelné resp. nerozlišitelné od jiných.

Jak vyplývá dále z údajů shrnutých v rámci Tabulky 2, statisticky významně lepších výsledků z testu dosahovali absolventi gymnázií (u víceletých: 18,33 +/- 0,162; u čtyřletých: 17,61 +/- 0,196), absolventi ostatních středních škol (zdravotních, SOŠ ev. jiných SŠ) dosahovali vesměs horších výsledků (v průměru 15,18 +/- 0,429). Posledně jmenovaných jsme však v souboru znalostně testovaných osob shledali pouze 6,9%. Absolventi státních vs. soukromých středních škol se ve výše uvedeném smyslu svými výsledky neodlišovali statisticky významně.

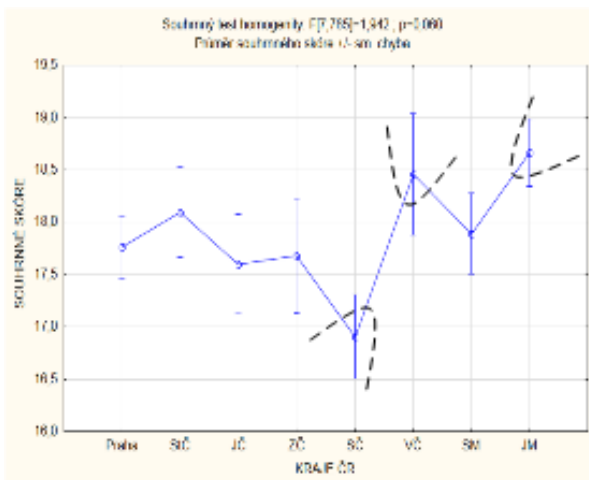
Typ SŠ:	N	Souhrnné skóre	Směrodatná chyba	p-hodnota (LSD post-hoc)		
1: gymnázium čtyřleté	437	17,61	,196	x	0,004	< 0,001
2: gymnázium víceleté	538	18,33	,162	0,004	x	< 0,001
3: ostatní SŠ	72	15,18	,426	< 0,001	< 0,001	x
Celkem:	1047	17,81	,123	$F[2;1044] = 21,878$; $p < 0,001$		

Tabulka 2 — Průměrná souhrnná skóre a vztažené specifikace podle typu absolvované střední školy ve školním roce 2012/13

Legenda: N — počet řádně vyplněných testů (= počet respondentů)

Metodicky stejně provedená analýza nerozlišila studenty 1. ročníku lékařských fakult nijak zásadně ani podle místa jejich trvalého bydliště, ačkoli souhrnný One-Way ANOVA test nulové hypotézy o homogenitě na úrovni 8 krajů ČR byl ve školním roce 2012/13 téměř na hranici statistické významnosti ($F[7;785] = 1,940$; $p = 0,060$) a následně provedený LSD post-hoc test by indikoval jako statisticky významně nejhorší výsledky pro 97 studentů 1. ročníku LF příchozích ze Severočeského kraje (s průměrem 16,91) vůči 43 absolventům SŠ z Východočeského kraje (18,45), a zejména vůči 153 z kraje Jihomoravského (18,67) – viz na Grafu 2. Statisticky významné rozdíly jsme ale shledali mezi výsledky studentům české nebo slovenské národnosti vůči zbývajícím 24 studentům ze zahraničí (s průměrem 15,33 +/- 0,719) – viz Tabulka 3.

Jako statisticky významný se ukázal rovněž rozdíl ve znalostech SŠ fyziky mezi pohlavími: zjištěné souhrnné skóre u 331 atestovaných mužů činilo 19,44 +/- 0,215, zatímco u 715 žen 17,06 +/- 0,141 ($F[1;1044] = 87,937$; $p < 0,001$).



Graf 2 — Průměry souhrnného skóre [0-30] z testu znalostí SŠ fyziky 4 lékařských fakult podle kraje místa trvalého bydliště respondentů (školní rok 2012/13)

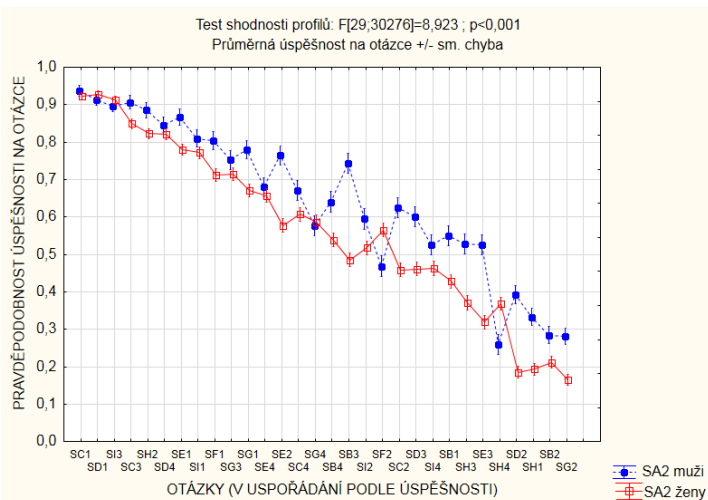
Legenda: Souhrnné výsledky testu z fyziky podle kraje místa trvalého bydliště respondentů se ve školním roce 2012/13 statisticky významně nelišily, ačkoli byly téměř na hranici statistické významnosti $p = 0,05$ ($F[8;785] = 1,942$; $p = 0,060$). Čárkovane vyznačené paraboly značí kraje, z nichž přichází studenti do 1. ročníku studia dosahovali vůči sobě statisticky významně nižšího (Stř) resp. vyššího (VČ a JM) průměru souhrnného skóre (na základě Fischerova LSD post-hoc testu).

Stát, kde absolvoval SŠ:	N	Souhrnné skóre	Směrodatná chyba	p-hodnota (LSD post-hoc)		
1: Česko	806	17,87	,141	x	0,852	0,002
2: Slovensko	221	17,81	,256	0,852	x	0,004
3: jiný stát	24	15,33	,719	0,002	0,004	x
Celkem:	1051	17,80	,123	F[2;1048] = 4,767 ; p = 0,009		

Tabulka 3 — Průměrná souhrnná skóre a vztážené specifikace podle státu absolvované střední školy ve školním roce 2012/13

Legenda: N — počet řádně vyplněných testů (= počet respondentů)

Pro studenty 1. ročníků lékařských fakult je tato skutečnost známa již ze známých výsledků přijímacích testů. Proto za zajímavější považujeme spíše naše zjištění (viz na Grafu 3), že u mužů je evidentní dokonce statisticky významně „pomalejší trend poklesu úspěšnosti směrem k obtížnějším otázkám“ nežli u žen (pro Two-Way ANOVA model s jedním faktorem „opakování“ /% úspěšnosti baterie 30 otázek/ a druhým faktorem fixním

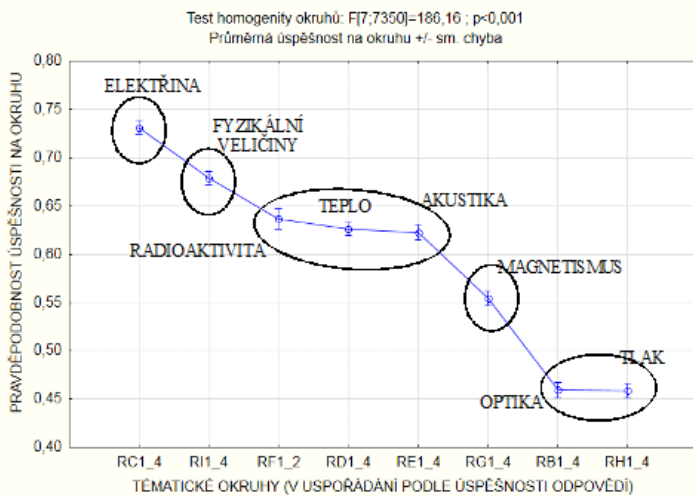


Graf 3 — Příspěvky souhrnnému skóre testu znalostí SŠ fyziky podle jednotlivých otázek a podle pohlaví (školní rok 2012/13)

Legenda: Otázky jsou uvedeny v sestupném uspořádání od nejlehčích po nejtěžnější. Muži dosáhli statisticky významně lepších výsledků než ženy (un block: $F[1;1044] = 87,937$; $p < 0,001$). Na grafu je patrný rovněž statisticky významně „rychlejší trend poklesu úspěšnosti směrem k obtížnějším otázkám“ u žen než u mužů ($F[29;30276] = 8,923$; $p < 0,001$). U nejlehčích otázek rozdíl mezi pohlavími patrný není.

/pohlaví/ jsme zamítli nulovou hypotézu o shodě profilů /resp. „trendů po uspořádání otázek podle % úspěšnosti“/ prostřednictvím statistiky $F[29;30276] = 8,923$; $p < 0,001$.

Obdobné rozdíly mezi pohlavími byly statisticky významné rovněž po rozkladu souhrnného skóre 30 otázek na 7 dílčích skóre s maximem 4 možných správných odpovědí (s výjimkou dvojice otázek týkajících se radioaktivity). S ohledem na cíl této práce považujeme však za důležitější námi atestované diference znalostí v rámci všech 8 tématických okruhů: Simultánní One-Way ANOVA test zamítá nulovou hypotézu „o rovnocennosti znalostí na úrovni oblastí“ prostřednictvím statistiky $F[7;7350] = 186,16$ ($p < 0,001$), tedy vysoce statisticky významně. V tabulce 5 jsou tématické okruhy uspořádány sestupně podle stupně obtížnosti pro studenty: Statisticky významně nejlepších výsledků bylo dosaženo v oblasti elektřiny a znalosti fyzikálních veličin (s průměry po řadě: $2,93 \pm 0,028$ a $2,71 \pm 0,029$), nejhorších u čtveřice otázek z optiky ($1,84 \pm 0,030$) a tlaku ($1,83 \pm 0,032$). Tyto výsledky prezentujeme rovněž na Grafu 4, kde statisticky významně se odlišující skupiny (podle výsledků Fischerova LSD post-hoc testu) diferencujeme prostřednictvím do grafu zakreslených elips.



Graf 4 — Příspěvky souhrnnému skóre testu znalostí SŠ fyziky podle 8 prověřovaných oblastí a podle pohlaví (školní rok 2012/13)

Legenda: Oblasti jsou uvedeny v sestupném uspořádání. Statisticky významně nejlepších výsledků bylo dosaženo v oblasti elektřiny a znalostí fyzikálních veličin (s průměry po řadě: 2,93 +/- 0,028 a 2,71 +/- 0,029), nejhorších u čtveřice otázek z optiky (1,84 +/- 0,030) a tlaku (1,83 +/- 0,032). S výjimkou průměrů na úrovni 2 otázek z radioaktivity dosáhli muži opět statisticky významně lepších výsledků než ženy.

	Tematická oblast:	K	Souhrnné skóre	Směrodatná chyba	CA
1	SC1_4 C: ELEKTŘINA	4	2,93	0,028	0,250
2	SI1_4 I: FYZIKÁLNÍ VELIČINY	4	2,71	0,029	0,232
3	SF1_2 F: RADIOAKTIVITA *)	2	2,54 *)	0,028 *)	-0,042
4	SD1_4 D: TEPLA	4	2,50	0,027	0,221
5	SE1_4 E: AKUSTIKA	4	2,49	0,032	0,282
6	SG1_4 G: MAGNETISMUS	4	2,22	0,031	0,235
7	SB1_4 B: OPTIKA	4	1,84	0,032	0,214
8	SH1_4 H: TLAK	4	1,83	0,030	0,208
	Celkem:	30	19,06 *)	0,132 *)	0,650

Tabulka 5 — Testované oblasti fyziky ve školním roce 2012/13 v uspořádání podle průměru dosažených skóre a Cronbachova alfa

Legenda: *) přepočteno na K = 4 otázky v rámci oblasti radioaktivity

2.2 Reliabilita a validita testových otázek

Výše provedená analýza úspěšnosti studentů 1. ročníků lékařských fakult v testech středoškolských znalostí z fyziky podle různých geodemografických kritérií byla fakticky zcela založena na souhrnném (kompozitním) skóre, kde se jednoduše sčítají všechny správné odpovědi na otázky obsažené v baterii. Otázky vybrané do baterie však nemusí vždy reprezentovat fenomén, který chceme danou baterií hodnotit (v daném případě úroveň znalostí středoškolské fyziky) anebo některá z otázek nemusí být s danou baterií vnitřně konzistentní. V prvním případě by některá z otázek nebyla validní (což může posoudit víceméně pouze odborník v daném tématu či předmětu), ve druhém případě by správnost odpovědi na danou otázku mohla být třeba i kontraproduktivní (pokud by kupř. správná odpověď na otázku de facto snižovala výpovědní schopnost testu jako celku).

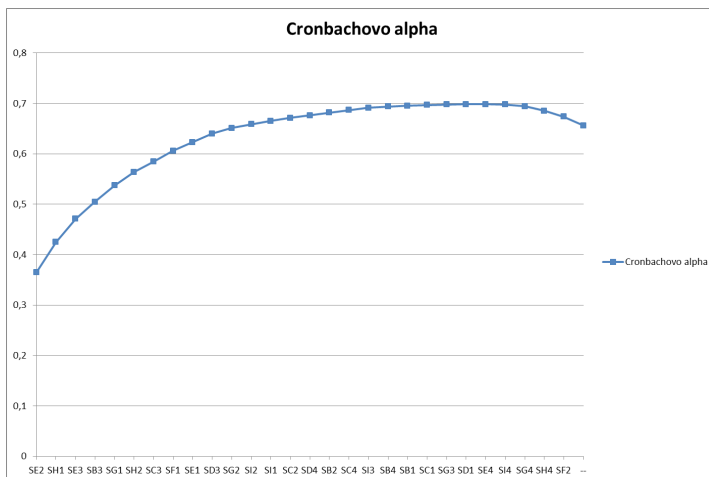
Z těchto a obdobných příčin se po skutečněném terénním testování provádí ještě formální analýza reliability jednotlivých otázek ve vztahu k celé baterii nebo se prověřuje validita položek každé otázky ve vztahu k empiricky zjištěným distribucím odpovědí na každé otázce. Faktorový rozklad Pearsonovy korelační matice indikátorů správnosti odpovědí na každou z 30 otázek testu z fyziky poukázal na existenci jednoho dominantního vlastního čísla, které vysvětlovalo více než 10% variability korelační matice (příslušnou 1. hlavní komponentu považujeme za latentní factor reprezentující pro naše účely žádaný teoretický konstrukt „(vše)obecné znalosti z fyziky“). Výše uvedené skutečnosti legitimovaly naše následné aplikace Cronbachova alfa (zkr. CA) na podsouborech otázek našeho testu z fyziky.

O něco přesněji: CA pro baterii všech 30 otázek testu z roku 2012 nabývala vstupní hodnoty kolem 0,65. Tato hodnota není experty považována za zcela postačující pro verifikaci reliability všech položek baterie. Pro tento případ se doporučuje postupně vylučovat (eliminovat) z baterie nejméně konzistentní položky t.j. v daném případě otázky, jejichž korelace s 1. hlavním faktorem jsou v absolutní hodnotě nejnižší (nebo snad dokonce nabývající záporných hodnot). Tento proces ilustrovaný na Grafu 5 se doporučuje ukončit rámcově v okamžiku, kdy souhrnné CA (pro položky setrvávající v dané baterii) dosáhnou svého přibližného maxima. Na úrovni Grafu 5 vztahujícímu se k baterii našeho testu z roku 2012 se toto týká přibližně 2-4 zprava postupně eliminovaných položek, kdy hodnoty CA dosáhly přibližně 0,7 (což je již všeobecně akceptovatelná hranice doporučená pro souhrnné CA).

Paralelně s výše uvedeným jsme vyčíslili ještě parametry dvou simultánních IRT (Item Response Theory) modelů (viz např. v [5]), které využívají formálně i prezentačně vděčných vlastností logistické křivky tvaru

$$P_{ij}(\Theta_i) = \exp\{b_j(\Theta_i - a_j)\} * (1 + \exp\{b_j(\Theta_i - a_j)\})^{-1}, \quad j=1, \dots, K, \quad i=1, \dots, N,$$

kde a_j resp. b_j (pro $j=1, \dots, K$) jsou parametry obtížnosti resp. diskriminativnosti otázek testu a Θ_i (pro $i=1, \dots, N$) individuální respondenti (studenti 1. ročníku lékařských fakult). V rámci posledních tří sloupců Tabulky 4 jsou parametry



Graf 5 — Hodnoty souhrnného Cronbachova alfa (CA) při postupné eliminaci (vylučování) nejméně konzistentních otázek ze vstupních baterie třiceti (pro školní rok 2012/13)
Legenda: položky jsou postupně eliminovány z původní baterie 30 otázek „zprava doleva“; $N = 1051$ studentů 1. ročníku z 1. a 3. LF, LF UK v Plzni a LF MU v Brně (viz také Tabulka 1)

aj resp. bj vyčísleny pro jedno- a dvouparametrický PL1 a PL2 (IRT) model. V korespondenci s hodnotami parametrů diskriminace v posledním sloupci Tabulky 4 (a výše provedenými experimenty s CA) byly pro školní rok 2013/14 vyloučeny z baterie otázky “SH4: Co měří aneroid(?)” a “SF2 Kde se nepoužívají radioaktivní látky(?)” a substituovány novými.

	Otázka (rámcový název):	Souhrnné skóre	Směrodatná chyba	a_1	a_2	b_2
1	SC1 Vznik kladného iontu	,925	,0081	-2,737	-4,632	0,571
2	SD1 Zahřátí tělesa pohybem částic	,921	,0083	-2,681	-5,901	0,429
3	SI3 Frekvence střídavého proudu	,908	,0089	-2,502	-4,445	0,540
4	SC3 Proč izolanty nevedou proud	,866	,0105	-2,056	-1,846	1,285
5	SH2 Kdy bude těleso plavat	,842	,0113	-1,851	-1,612	1,360
6	SD4 Jednotka měrné tepelné kapacity	,828	,0117	-1,740	-2,644	0,642

	Otázka (rámcový název):	Souhrnné skóre	Směrodatná chyba	a_1	a_2	b_2
7	SE1 Co je zvuk	,807	,0122	-1,588	-1,813	0,915
8	SI1 Měření elektroměrem	,783	,0127	-1,429	-2,203	0,631
9	SF1 Určení stáří archeologického nálezů	,739	,0135	-1,165	-1,356	0,896
10	SG3 Feromagnetické látky	,727	,0137	-1,095	-2,342	0,436
11	SG1 Výhody elektromagnetu	,703	,0141	-0,966	-1,018	1,033
12	SE4 Fyzikální interpretace tónu	,664	,0146	-0,766	-1,796	0,394
13	SE2 Šíření zvuku	,636	,0149	-0,626	-0,645	1,076
14	SC4 Co je tužková baterie	,626	,0149	-0,580	-1,104	0,494
15	SG4 Možnost sestavní kompasu	,585	,0152	-0,388	-4,578	0,075
16	SB4 Studené zdroje světla	,572	,0153	-0,327	-0,619	0,496
17	SB3 Duté zrcadlo	,565	,0153	-0,297	-0,331	0,974
18	SI2 Odpor žárovky	,541	,0154	-0,189	-0,283	0,654
19	SF2 Kde se nepoužívají radioaktivní látky *)	,534	,0154	-0,154	0,630	-0,217*
20	SC2 Spotřeba energie přístrojem	,509	,0154	-0,043	-0,072	0,570
21	SD3 Teplota varu vody	,505	,0154	-0,026	-0,038	0,710
22	SI4 Práce jeřábu	,482	,0154	0,077	0,325	0,219
23	SB1 Vzdálenost od Slunce	,467	,0154	0,146	0,320	0,427
24	SH3 Vztlková síla ve vzduchu	,421	,0152	0,358	0,433	0,842

	Otázka (rámcový název):	Souhrnné skóre	Směrodatná chyba	a_1	a_2	b_2
25	SE3 Slyšitelnost harmonického vlnění	,383	,0150	0,531	0,612	0,901
26	SH4 Co měří aneroid *)	,335	,0146	0,767	-54,185	-0,013*
27	SD2 Železo v plynné formě	,250	,0134	1,223	1,321	0,988
28	SH1 Přísavka na skle	,237	,0131	1,303	1,662	0,793
29	SB2 Červené oči na fotografiích	,234	,0131	1,321	2,179	0,583
30	SG2 Kam patří ohebné magnety	,201	,0124	1,535	2,156	0,705
	Celkem:	17,796	,1230			

Tabulka 4 — Testované otázky ve školním roce 2012/13 v uspořádání podle % správných odpovědí (souhrnné skóre) a parametry PL1 a PL2 IRT (logistických) modelů

Legenda: N — a_1 resp. a_2 - parametr obtížnosti v PL1 resp. v PL2 IRT modelu;

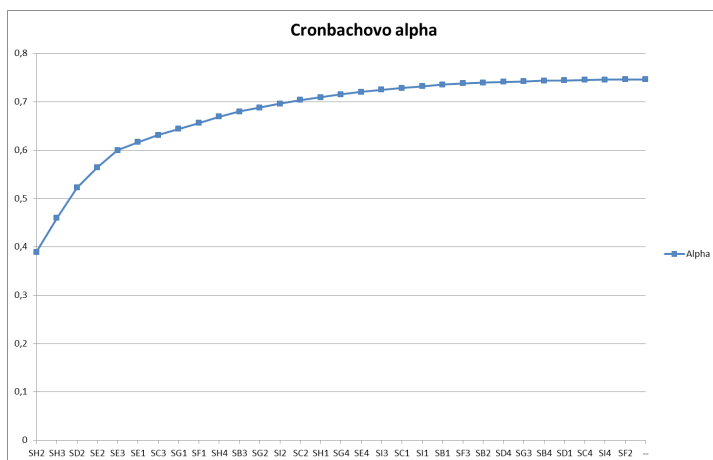
b_2 - diskriminační parametr v PL2 IRT modelu (v PL1 modelu je parametr $b_1 = 1$);

*) otázky vyloučené (substituované) z baterie pro testování ve školním roce 2013/14

3. Diskuse a závěry

Jednoduchá analýza CA pro čtveřice (resp. pro oblast radioaktivity dvojice) otázek spadajících do každé z osmi tematických oblastí fyziky přinesla poznatek, že empiricky zjištěné hladiny CA („v rozpětí hodnot od -0,042 do 0,282“ zcela vybočují z běžných norem pro smysluplné úvahy o reliabilitě nebo vnitřní konzistenci položek v bateriích pro námi vybrané tematické oblasti fyziky (viz Tabulka 5). Na úrovni tak malých počtů otázek v baterii, jako jsou 4 nebo dokonce 2, tedy žádné úvahy o diskriminační schopnosti takto koncipovaných baterií otázek pro znalostní testy pravděpodobně nejsou reálné. Ze zkušeností na úrovni i jiných aplikací Cronbachova alfa na položky více sofistikované specifikovaných konstruktů resp. oblastí (viz např. příspěvek [6] ve Sborníku MEDSOFT 2012 zaměřený na “měření křehkosti seniora” prostřednictvím testu SPPB) by snad mohly být již i při nepříliš významném navýšení počtu otázek v konstrukt (tematickém okruhu) baterie otázek pro praxi dobře použitelné. Každopádně by musela být věnována ještě větší pozornost i samotné definici obsahového konceptu každého atestovaného konstruktů a tomu odpovídajícímu ještě pečlivějšímu výběru otázek, včetně detailního vyladění formulace položek, na jejichž základě se má student nebo respondent rozhodovat pro zaškrtnutí “té správné” (tzn. pro něj nejvěrohodnější) odpovědi.

V souladu s výše uvedeným byl pro následující školní rok 2013/14 test z fyziky rozšířen na 32 otázek (tedy na 4 otázky v každé oblasti) a na několika málo položkách týmem spolupracujících expertů (biofyziků) nepatrně formulačně pozměněn. V této verzi byl test opakovaně zadán studentům 1. ročníku na 1., 3. a Plzeňské LF UK v Praze, ke kterým se na podzim 2013 připojila ještě lékařská fakulta Ostravské univerzity, a nově také 2. LF UK v Praze. Předběžné výsledky vypovídají o větší konzistenci modifikované baterie otázek testu. Ukazatel CA (vypočtený stejnou metodikou jako v předešlém školním roce) nabývá na předběžném vzorku $N = 810$ studentů hodnoty téměř 0,75 (viz na Grafu 6), když všech 32 otázek nyní zařazených do baterie můžeme považovat za více reliabilní a vnitřně více konzistentní systém.



Graf 6 — Hodnoty souhrnného Cronbachova alfa (CA) při postupné eliminaci (vylučování) nejméně konzistentních otázek ze vstupních baterie třiceti dvou (pro školní rok 2013/14)

Legenda: položky jsou postupně eliminovány z původní baterie 32 otázek „zprava doleva“; $N = 810$ studentů z 1.LF UK ($N = 554$), z 3. LF UK ($N=159$) Praha a z LF OU v Ostravě ($N=97$)

Na 3. LF UK Praha a na Ostravské univerzitě byly testy provedeny prezenčně v rámci aplikací vytvořených pro systém Moodle. Vzhledem k možnostem automatické permutace individuálních otázek v elektronických aplikacích i nesrovnatelně rychlejší (de facto) automatické přípravě dat pro finální statistické zpracování tento způsob provedení testu považujeme nepochybně za progresivní. S nejbližší perspektivou bude nutné řešit stávající problémy s anonymizací respondentů (např. v aplikaci Moodle) a také (v delším časovém horizontu) přejít od organizace testování založeného „na úplných souborech“ na reprezentativní výběrová šetření – obojí vyžaduje jen pečlivější profesionální (softwarové a statistické) ošetření.

Literatura:

- [1.] Štuka, Č. a kol. (2013). *Testování při výuce medicíny*. Karolinum, Praha
- [2.] SPSS Inc. (2008). *SPSS Statistics for Windows, Version 17.0*. Chicago: SPSS Inc.
- [3.] StatSoft, Inc. (2013). *STATISTICA (data analysis software system), version 12.0*
- [4.] Rizopoulos D. (2006). *Ltm: An R package for Latent Variable Modelling and Item Response Theory Analyses*, *Journal of Statistical Software*, 17(5), 1-25.
- [5.] Hendl, J. (2009). *Přehled statistických metod zpracování dat (3. vydání)*. Portál, Praha
- [6.] Běláček, J., Berková, M., Mádlová, P., Topinková, E. (2012). *K metodice porovnání výsledků testů MMSE, ADL, MNA a SPPB na výběrovém souboru geriatrických pacientů*. In *Sborník příspěvků Medsoft 2012*

Kontakt:

RNDr. Jaromír Běláček, CSc.

Mgr. Komarc Martin

MUDr. Kymplová Jaroslava, Ph.D.

Ústav biofyziky a informatiky 1.LF UK Praha

Salmovská 1, 121 08 Praha 2

tel: +420 22496 5706

e-mail: jaromir.belacek@lf1.cuni.cz

<http://biof.lf1.cuni.cz/biostat.html>

RNDr. Kvašňák Eugen, Ph.D.

Ústav lékařské biofyziky a lékařské informatiky

3. LF UK Praha

Prof. RNDr. Mornstein Vojtěch, CSc.

Biofyzikální ústav, LF MU Brno

Bc. Kubeš Zdeněk

Ústav biofyziky, LF UK v Plzni

RNDr. Sochorová Hana, Ph.D.

Katedra biomedicínských oborů

LF OU v Ostravě