

SPOĽAHLIVOSŤ A PRESNOSŤ HODNOTENIA KOGNITÍVNEHO VÝKONU ŠTUDENTMI ŠTUDIJNÉHO ODBORU UČITEĽSTVO A PEDAGOGICKÉ VEDY V PREDMETE MATEMATICKÁ GRAMOTNOSŤ

Jana HNATOVÁ, Marek MOKRIŠ
Prešovská univerzita v Prešove, Pedagogická fakulta (Slovensko)
jana.hnatova@unipo.sk, marek.mokris@unipo.sk

Abstrakt

Meranie spoľahlivosti a presnosti hodnotenia kognitívneho matematického výkonu edukantmi dovoľuje kvantifikovane uchopiť konštrukt metakognitívneho monitorovania v edukačnej praxi. V článku sa zameriavame na špecifickú skupinu vysokoškolských študentov študijného odboru Učiteľstvo a pedagogické vedy ($N = 654$), v ktorej zisťujeme dosahovanú mieru spoľahlivosti a presnosti ich hodnotenia matematického výkonu pomocou kalibračných indexov BIAS a AAI. Ich porovnaním s ideálnymi hodnotami a následnou dekompozíciou efektívnych hypotéz vzhľadom na dosiahnutý klasifikačný stupeň výstupného hodnotenia v predmete Matematická gramotnosť spresňujeme nami dosiahnuté výsledky v tejto špecifickej výskumnej skupine.

Kľúčové slová: presnosť hodnotenia, skreslenie, BIAS, AAI

BIAS AND ACCURACY OF EVALUATION OF COGNITIVE PERFORMANCE BY STUDENTS OF THE STUDY DEPARTMENT TEACHING AND PEDAGOGICAL SCIENCES IN THE SUBJECT MATHEMATICAL LITERACY

Abstract

The measuring the bias and accuracy of evaluating cognitive mathematical performance by educators allows a quantified grasp of the construct of metacognitive monitoring in educational practice. In this article we focus on a specific group of university students in the field of study Teaching and Pedagogical Sciences ($N = 654$), in which we determine the achieved degree of bias and accuracy of their evaluation of mathematical performance using calibration indices BIAS and AAI. By comparing them with ideal values and subsequent decomposition of effective hypotheses with respect to the achieved classification level of the final evaluation in the subject Mathematical Literacy, we specify the results achieved by us in this specific research group.

Keywords: evaluation accuracy, BIAS, IAA

1. Úvod

Výučba matematiky v predmete Matematická gramotnosť patrí k úvodným kurzom vysokoškolského matematického vzdelávania na Pedagogickej fakulte Prešovskej univerzity v Prešove. Bez ohľadu na realizovanú formu vzdelávania, vyžaduje aktívnu účasť študentov, v rámci ktorej sú nútení kontrolovať, ako interagujú s podpornými študijnými materiálmi

svojho kurzu. Táto kontrola je zrejmá najmä zvýšenou aktivitou na seminároch i v počte prístupov do elektronického podporného systému LMS Moodle, a to predovšetkým v období prípravy študentov na priebežné hodnotenia. Úspech v rámci hodnotených aktivít predmetu je totiž aj podľa Zimmermana (2008) do značnej miery ovplyvnený schopnosťou študenta zapojiť sa do samoregulačného učenia, ktorý chápe ako proces vyžadujúci efektívne zvládanie metakognitívnych monitorovacích a kontrolných procesov.

Študent prvého ročníka, ktorý má vo svojom prvom skúškovom období na vysokej škole absolvovať niekoľko skúšok, musí začať rozumne manažovať čas prípravy na ne, pričom zohľadňuje zameranie a ním subjektívne posúdenú náročnosť prípravy na jednotlivé skúšky. Učí sa odhadnúť aktuálny stav svojich vedomostí a dosiahnutú úroveň schopností v porovnaní s očakávaniami, ktoré sú naňho kladené v rámci úspešného absolvovania konkrétneho predmetu. Tento spôsob metakognitívneho monitorovania sa opiera o presnosť a spoľahlivosť, s akou študent pri monitoringu pracuje a často je jedným z faktorov ovplyvňujúcich termínový plán skúšok a časový plán prípravy na ne.

2. Presnosť a spoľahlivosť metakognitívneho monitorovania

Úvodné kurzy vysokoškolského štúdia sa zvyčajne zameriavajú na pokrytie širokej vzorky tém, preto nie je jednoduché nájsť v nich priestor pre ďalší rozvoj metakognitívnych schopností (Morphew, 2020). Metakognitívne schopnosti odkazujúce na spôsoby, akými sa študenti zapájajú do samoregulačných procesov učenia sa, zahŕňajú okrem procesov kontroly aj procesy monitorovania. O zapojení študentov do metakognitívneho monitorovania hovoríme vtedy, ak študenti hodnotia svoj súčasný stav učenia sa na základe daného kritéria (Merritt, 2019).

Tabuľka 1. Prehľad konštruktov a možností ich merania

	Konštrukt	Index	Výpočet	Interpretácia: Index vyhodnocuje...
1.	Absolútna presnosť	Index absolútnej presnosti (AAI)	$p_i = \frac{p'_i}{m_i}$ $AAI = \frac{1}{N} \sum_{i=1}^N (c_i - p_i)^2$... presnosť úsudku spoľahlivosti v porovnaní s objektívnym hodnotením výkonu t. j. hodnotením podľa daného kritéria
2.	Relatívna presnosť	Index relatívnej presnosti (r)	$r = \frac{\sum_{i=1}^N (c_i - \bar{c})(p_i - \bar{p})}{\sqrt{\sum_{i=1}^N (c_i - \bar{c})^2 \sum_{i=1}^N (p_i - \bar{p})^2}}$... hodnotí (lineárny) vzťah medzi úsudkami spoľahlivosti a objektívnym hodnotením výkonu podľa daného kritéria
3.	Skreslenie	Index skreslenia (Bias)	$Bias = \frac{1}{N} \sum_{i=1}^N (c_i - p_i)$... skreslenie vlastného úsudku určením hodnoty nesúladu a jeho smeru k možnému podhodnocovaniu alebo nadhodnocovaniu vlastného výkonu
4.	Rozptyl	Bodový index (SI)	$N = N_+ + N_-$ $SI = \frac{1}{N} [N_+ \sigma_+^2 - N_- \sigma_-^2]$... rozdiely v rozptyle spoľahlivo (+) a nespoľahlivo (-) vyhodnotených položiek
5.	Diskriminácia	Diskriminačný index (DI)	$DI = \frac{1}{N} \left[\sum_{i=1}^{N_+} c_{i+} - \sum_{i=1}^{N_-} c_{i-} \right]$... diskrimináciu medzi spoľahlivo (+) a nespoľahlivo (-) vyhodnotenými položkami

Vysvetlivky:

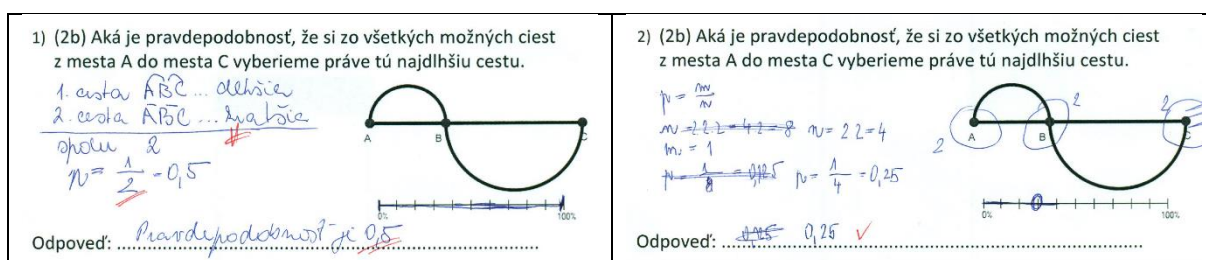
- c_i ... miera subjektívnej istoty správnosti odpovede vyjadrená relatívnou dĺžkou úsečky istoty,
 c_{i+} ... miera subjektívnej istoty správnosti odpovede v spoľahlivo vyhodnotenej položke,
 c_{i-} ... miera subjektívnej istoty správnosti odpovede v nespoľahlivo vyhodnotenej položke,
 p'_i ... objektívne hodnotenie matematického výkonu bodovým skóre v i -tej položke testu,
 m_i ... maximálne bodové skóre v i -tej položke testu,
 p_i ... relatívne bodové hodnotenie matematického výkonu v i -tej položke testu,
 N_+ ... početnosť spoľahlivo vyhodnotených položiek,
 N_- ... početnosť nespoľahlivo vyhodnotených položiek,
 σ_+^2 ... rozptyl spoľahlivo vyhodnotených položiek,
 σ_-^2 ... rozptyl nespoľahlivo vyhodnotených položiek.

Zdroj: (1) Maki a kol. 2005; (2) Nelson 1996; (3,4,5) Schaw 2009; (4) Yates 1998 modifikované podľa potreby výskumného šetrenia

Viacerí autori (Maki a kol., 2005; Nelson, 1996; Schaw, 2009; Yates 1998) uvádzajú vo svojich prácach rôzne konštrukty umožňujúce kvantifikáciu procesov metakognitívneho monitorovania (tab. 1). Vychádzajúc z potrieb štúdie je možné (aj opakovane) realizovať merania vstupných premenných príslušných indexov:

- pred podaním výkonu študenta alebo po jeho výkone (preddikčné alebo postdikčné merania),
- na úrovni každého dosiahnutého stupňa, časti resp. položky výkonu alebo na úrovni celého posudzovaného výkonu (globálne alebo lokálne merania).

Spoľahlivosť úsudkov študentov porovnaním subjektívneho a objektívneho hodnotenia výkonu je možné vyjadriť indexom skreslenia (pozri tab. 1). Index skreslenia (ďalej Bias) nadobúda hodnoty z intervalu $[-1, 1]$ a poskytuje informácie o smere a závažnosti nesúladu medzi deklarovanou subjektívnou istotou a objektívne dosiahnutou výkonnosťou. Ak študent v položke prezentuje vysokú mieru istoty pri nízkej úspešnosti podaného výkonu, index nadobúda kladné hodnoty, ktoré signalizujú nadhodnocovanie vlastného výkonu samotným študentom (obr. 1a). Ak je prezentovaná miera istoty nízka, avšak úspešnosť výkonu je vysoká, index nadobúda záporné hodnoty signalizujúce podhodnocovanie vlastného výkonu študentom (obr. 1b). Navyše, absolútna hodnota Bias indexu vychádzajúca z geometrickej definície vzdialenosti poskytuje taktiež informácie o závažnosti chyby v úsudku. Čím je vzdialenosť hodnoty indexu od nuly väčšia, tým väčší nesúlad medzi subjektívnym a objektívnym hodnotením výkonu index signalizuje.



Obrázok 1. Testová položka s nadhodnoteným výkonom študenta (a) vľavo a s podhodnoteným výkonom študenta (b) vpravo

Konštrukt presnosti úsudku je možné merať pomocou relatívnych alebo absolútnych indexov presnosti (tab. 1). Absolútna presnosť sa vzťahuje na veľkosť odchýlky medzi premennými vyjadrujúcimi (subjektívny) úsudok o istote a (objektívne) hodnotený výkon t. j. podľa jednoznačne daného kritéria. Relatívna presnosť sa zvyčajne meria pomocou korelačného koeficienta (Pearsonovo r) alebo pohotovostného koeficienta (d'Gamma)

a predstavuje schopnosť jednotlivca rozlišovať medzi dvoma premennými, napríklad medzi spoľahlivo a nespoľahlivo riešenými položkami (Lindel, Lenhart, Schneider, 2019). Za spoľahlivo riešené položky sú považované tie, v ktorých sa študentov subjektívny úsudok o istote zhodne s objektívnym hodnotením položky. Ostatné položky, v ktorých zhoda nenastane, sú považované za nespoľahlivo riešené. Podobne aj ďalšie konštrukty (diskriminácia a rozptyl, pozri tab. 1) reprezentované diskriminačným a bodovým indexom pracujú s porovnávaním premenných súvisiacich so spoľahlivo a nespoľahlivo riešenými položkami.

V tomto príspevku budeme vyhodnocovať miery dvoch z uvedených konštruktov pomocou indexov absolútnej presnosti (AAI) a skreslenia (Bias), ktoré korešpondujú s predpokladaným cieľom študentov - úspešne absolvovať testy priebežného hodnotenia, ktoré im pomáhajú určiť, či sa ich aktuálna úroveň poznania zhoduje s cieľom úspešne absolvovať priebežné hodnotenie predmetu podľa daných kritérií. Naším zámerom je prezentovať vybrané sumárne zistenia týkajúce sa danej problematiky, ktoré vychádzajú a dopĺňujú publikované čiastkové výsledky (Hnatová, Mokriš, Lipták 2019; Hnatová, Mokriš 2019).

3. Základné údaje a otázky výskumného šetrenia

Výskum bol realizovaný počas rokov 2018-2019 na skupinách študentov prvého ročníka bakalárskeho štúdia študijného odboru Učiteľstvo a pedagogické vedy ($N = 654$) na Pedagogickej fakulte Prešovskej univerzity v Prešove. Výskumná vzorka sa kreovala ako dostupná a zámerná, vytváraná študentmi na báze dobrovoľnosti.

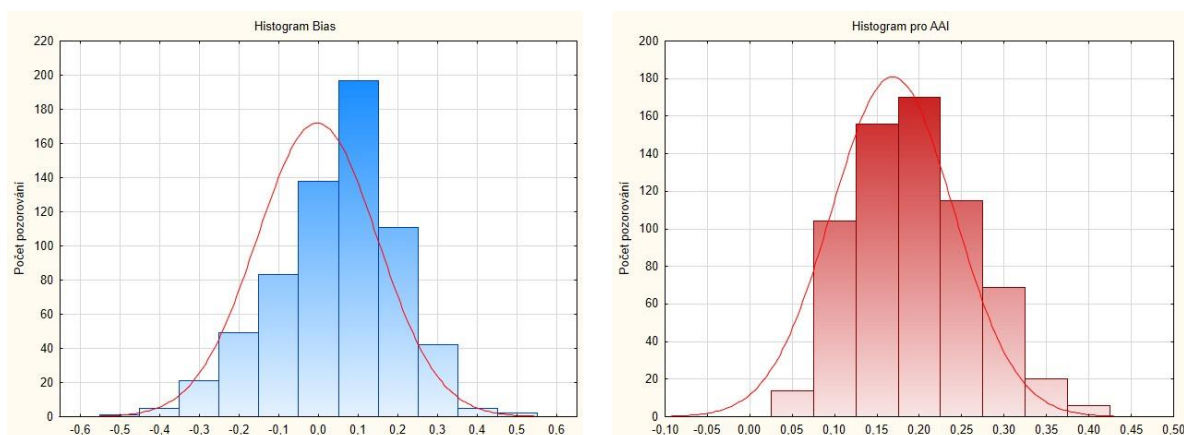
Hľadali sme odpovede na nasledujúce výskumné otázky:

- VO1: Aké skreslenie dosahujú študenti študijného odboru Učiteľstvo a pedagogické vedy medzi subjektívnym a objektívnym hodnotením matematického výkonu?
- VO2: Existuje závislosť medzi skreslím hodnotenia matematického výkonu študentmi študijného odboru Učiteľstvo a pedagogické vedy a výsledným hodnotením študentov v predmete Matematická gramotnosť?
- VO3: Akú presnosť hodnotenia matematického výkonu dosahujú študenti študijného odboru Učiteľstvo a pedagogické vedy?
- VO4: Existuje závislosť medzi presnosťou hodnotenia matematického výkonu študentmi študijného odboru Učiteľstvo a pedagogické vedy a výsledným hodnotením študentov v predmete Matematická gramotnosť?

4. Zber a spracovanie údajov

V priebehu výučby predmetu Matematická gramotnosť, ktorý patrí do korpusu povinných predmetov uvedeného študijného odboru vo všetkých jeho študijných programoch, sa študenti zúčastnili dvoch priebežných hodnotení. Obe boli realizované prezenčne s využitím didaktických testov ako meracích nástrojov (Cronbachovo $\alpha_{PH1} = 0,8368$, Cronbachovo $\alpha_{PH2} = 0,7493$). Testy sa sumárne skladali z 56 položiek, ktoré svojim obsahovým zameraním a početnosťou zastúpenia v 12 tematických oblastiach vychádzali z obsahového a rozsahového zamerania predmetu.

Rozloženia početností údajov oboch indexov vypočítaných podľa vzťahov (tab. 1- riadky (1) a (3)) sú znázornené histogramami (graf 1a a 1b), ktoré naznačujú potrebu pomocou Kolmogorovho-Smirnovovho testu overiť ich zhodu s normálnym rozdelením údajov, keďže Gaussovú krivku (v grafe vyznačenú červenou čiarou) nekopírujú presne.



Graf 1. Histogram Bias indexu (a) a AAI (b)

Na základe výsledkov pre Bias index ($K-S d = 0,06005$, $p < 0,05$) konštatujeme na hladine významnosti $\alpha = 0,05$, že údaje o Bias indexe nepochádzajú z normálneho rozdelenia. Jeho absenciu musíme zohľadniť vo výbere štatistických testov použitých pri následnom spracovaní údajov. Dosahované p -hodnoty pre AAI ($K-S (d) = 0,03878$, $p > 0,20$) nám dovoľujú rozdelenie údajov pre AAI považovať za normálne.

Pri hľadaní odpovede na VO1 bolo potrebné pre každého študenta v každej položke spárovať merania súvisiace s jeho subjektívnou mierou istoty a objektívnym hodnotením výkonu, vypočítať jeho Bias index a vybranú charakteristiku polohy indexu celého výskumného súboru následne porovnať s referenčnou nulovou hodnotou. Pre Bias index predstavuje referenčná hodnota ideálny stav, v ktorom dochádza k absolútnej zhode medzi subjektívnym a objektívnym hodnotením študenta, a teda rozdiel medzi nimi je nulový.

Rovnakým postupom sme zisťovali aj odpoveď na otázku týkajúcu sa dosahovania čo najväčšej presnosti hodnotenia matematického výkonu študentom (VO3). Referenčnou hodnotou pre AAI je ideál absolútnej presnosti, a teda nulovej druhej mocniny rozdielov oboch hodnotení.

V tejto súvislosti formulujeme dve hypotézy testujúce stredné hodnoty sledovaných indexov (mediány \tilde{x} resp. priemery \bar{x}) oproti ich nulovým referenčným hodnotám:

$$H1_1: \tilde{x}(Bias) \neq 0 \quad \text{vs.} \quad H0_1: \tilde{x}(Bias) = 0.$$

$$H1_2: \bar{x}(AAI) \neq 0 \quad \text{vs.} \quad H0_2: \bar{x}(AAI) = 0.$$

Pri overovaní platnosti nulovej hypotézy $H0_1$, s ohľadom na absenciu normálneho rozdelenia súboru údajov, využijeme neparametrický Wilcoxonov test, pri nulovej hypotéze $H0_2$ použijeme, vzhľadom na dosiahnuté normálne rozdelenie, Studentov t -test. Na hladine významnosti $\alpha = 0,01$ podľa p -hodnôt týchto testov (tab. 2) nezamietame nulovú hypotézu $H0_1$ a naopak, zamietame nulovú hypotézu $H0_2$ v prospech jej alternatívnej hypotézy $H1_2$. Konštatujeme teda, že kým stredné hodnoty Bias indexu vo výskumnom súbore študentov nadobúdajú ideálne hodnoty, stredné hodnoty AAI sa od ideálneho stavu signifikantne líšia.

Tabuľka 2. Prehľad výsledkov testovania hypotéz $H0_1$ vs. $H1_1$ a $H0_2$ vs. $H1_2$

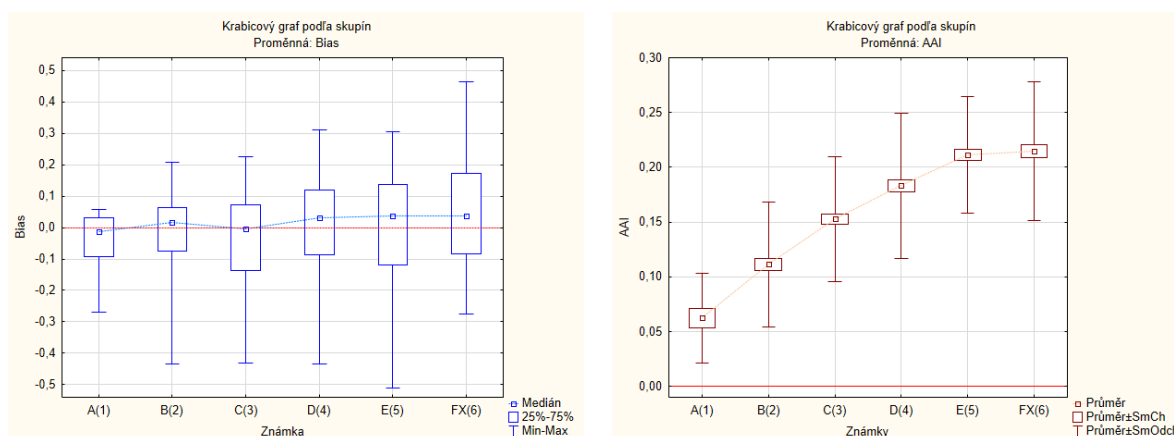
Testovanie hypotéz	Test	p -hodnota
$H0_1: \tilde{x}(Bias) = 0$ vs. $H1_1: \tilde{x}(Bias) \neq 0$	Wilcoxonov test	0,519338
$H0_2: \bar{x}(AAI) = 0$ vs. $H1_2: \bar{x}(AAI) \neq 0$	Studentov t -test	0,000000

Zdroj: vlastné spracovanie

Zodpovedanie otázok VO2 a VO4 vyžaduje dekompozíciu uvedených hypotéz, ktorou dokážeme popísať nielen stav, ale aj vývoj hodnôt indexov v smere zavedenej postupnosti

klasifikačných stupňov celkového hodnotenia študentov v predmete. V grafickom znázornení vizualizujeme polohu všetkých stredných hodnôt Bias indexov (graf 2a) a AAI (graf 2b) vo všetkých dosiahnutých klasifikačných stupňoch celkového hodnotenia predmetu a môžeme porovnať ich polohu vzhľadom:

- k nulovým referenčným hodnotám popísanými konštantnou nulovou funkciou, ktorá je na grafe (graf 2) vyznačená priamkami červenej farby,
- k vzájomným polohám stredných hodnôt indexov faktorizovaných dosiahnutým klasifikačným stupňom výsledného hodnotenia predmetu.



Graf 2. Krabicové grafy stredných hodnôt indexov Bias indexu (a) a AAI (b) v jednotlivých klasifikačných stupňoch výsledného hodnotenia predmetu Matematika gramotnosť

Stredné hodnoty Bias indexu skupín študentov hodnotených rôznymi klasifikačnými stupňami sa nachádzajú v blízkosti referenčnej priamky (graf 2a). Testovanie vzájomných zhôd týchto stredných hodnôt je v tomto prípade potrebné realizovať pomocou neparametrického mediánového testu. Pre test vzájomných zhôd mediánov (resp. distribučných funkcií) Bias indexov v jednotlivých klasifikačných stupňoch formulujeme nulovú hypotézu:

$$H_{03}: \tilde{x}(Bias_{A(1)}) = \tilde{x}(Bias_{B(2)}) = \tilde{x}(Bias_{C(3)}) = \dots = \tilde{x}(Bias_{FX(6)})$$

oproti alternatívnej hypotéze H_{13} , že aspoň jeden medián (resp. aspoň jedna distribučná funkcia) sa od iného (resp. inej) navzájom odlišujú. Výsledky mediánového testu ($\chi^2 = 8,942808$ $sv = 5$ $p = 0,1114$) platnosť nulovej hypotézy nezamietajú, môžeme teda skonštatovať existenciu zhody mediánov Bias indexu naprieč celým spektrom výsledného hodnotenia predmetu.

Stredné hodnoty AAI skupín študentov hodnotených rôznymi klasifikačnými stupňami vizualizujeme krabicovým grafom (graf 2b). Aj v tomto prípade otestujeme platnosť nulovej hypotézy o zhode priemerov indexov AAI v daných klasifikačných stupňoch. Nulová hypotéza má tvar:

$$H_{04}: \bar{x}(AAI_{A(1)}) = \bar{x}(AAI_{B(2)}) = \bar{x}(AAI_{C(3)}) = \dots = \bar{x}(AAI_{FX(6)})$$

oproti alternatívnej hypotéze H_{14} dovoľujúcej nezgodu aspoň jednej dvojici priemerov z uvedených skupín výsledného hodnotenia študentov.

Analýzou rozptylu ANOVA ($F(5,648) = 64,84$; $p = 0,0000$) nulovú hypotézu zamietame a potvrdzujeme existenciu signifikantných rozdielov, ktoré dourčíme Tukeyovým HSD testom (**Chyba! Nenalezen zdroj odkazů.**). Z jeho výsledkov vyplýva, že až na priemery indexov AAI v skupinách študentov hodnotených vo výslednom hodnotení klasifikačnými stupňami E a FX, všetky ostatné vykazujú navzájom štatisticky významné rozdiely.

Tabuľka 3. Prehľad signifikantných výsledkov HSD testu priemerov AAI skupín študentov roztriedených podľa dosiahnutých klasifikačných stupňov

AAI	$\bar{x}(AAI_{B(2)})$	$\bar{x}(AAI_{C(3)})$	$\bar{x}(AAI_{D(4)})$	$\bar{x}(AAI_{E(5)})$	$\bar{x}(AAI_{FX(6)})$
$\bar{x}(AAI_{A(1)})$	** 0,006133	** 0,000020	** 0,000020	** 0,000020	** 0,000020
$\bar{x}(AAI_{B(2)})$		** 0,000020	** 0,000020	** 0,000020	** 0,000020
$\bar{x}(AAI_{C(3)})$			** 0,000213	** 0,000020	** 0,000020
$\bar{x}(AAI_{D(4)})$				** 0,005113	** 0,000349
$\bar{x}(AAI_{E(5)})$					0,998257

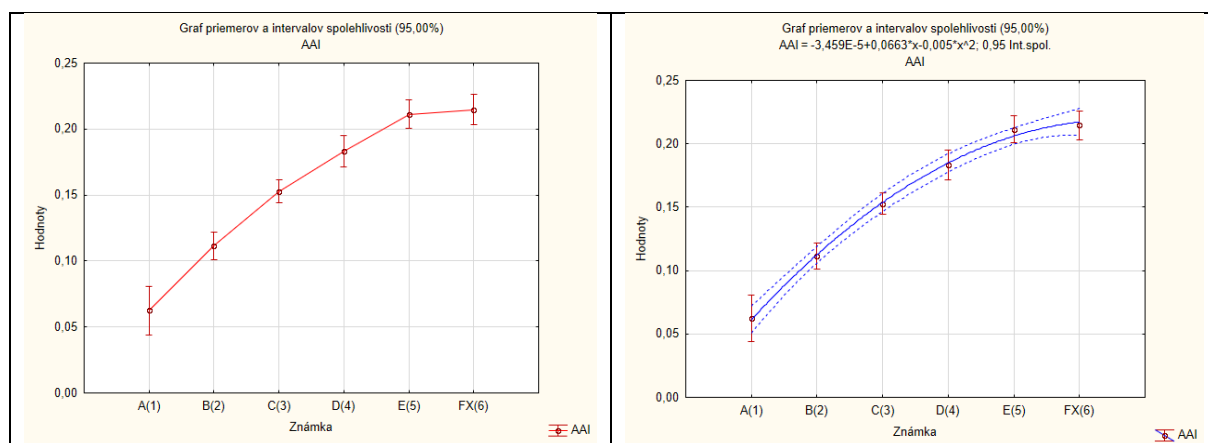
Vysvetlivky:

* signifikantná nezghoda na hladine 0,05,

** signifikantná nezghoda na hladine 0,01.

Zdroj: vlastné spracovanie

Pri modelovaní vývoja stredných hodnôt AAI v smere zavedenej postupnosti klasifikačných stupňov celkového hodnotenia študentov v predmete využijeme GLM (skr. z ang. *Generalized Linear Model*) s polynomicou regresiou, ktorá poskytuje výstup v podobe kvadratického regresného modelu $AAI = -0,005x^2 + 0,0663x - 3,459 \cdot 10^{-5}$ s $R^2 = 0,9877$ (graf 3).



Graf 3. Grafické porovnanie stavu a vývoja AAI v jednotlivých klasifikačných stupňoch výsledného hodnotenia predmetu Matematická gramotnosť

V zhrnutí teda môžeme na hladine významnosti $\alpha = 0,05$ formulovať čiastkové odpovede na položené výskumné otázky (VO1 až VO4).

Študenti študijného odboru Učiteľstvo a pedagogické vedy po prvom semestri štúdia na vysokej škole dosahujú želaný stav spoľahlivosti hodnotenia svojho matematického výkonu (VO1). Tento nevykazuje žiaden súvis s výsledným hodnotením, ktoré dosiahli v predmete Matematická gramotnosť (VO2). Presnosť, v ponímaní absolútnej presnosti hodnotenia ich matematického výkonu, však po absolvovaní úvodného predmetu z portfólia matematicky orientovaných disciplín ešte dosiahnutá nebola. Pri modelovaní trendu vývoja sa preukazuje kvadratický nárast nepresnosti hodnotenia so zhoršujúcim sa klasifikačným stupňom dosiahnutého výsledného hodnotenia matematického výkonu, pričom záporný kvadratický koeficient $b_2 = -0,005$ spôsobuje znižovanie štatisticky významných rozdielov, avšak iba medzi ostatnou dvojicou skupín študentov, ktorí boli v predmete hodnotení stupňami E a FX. V ostatných skupinách študentov hodnotených rôznymi klasifikačnými stupňami je rozdiel v absolútnej presnosti hodnotenia signifikantne výrazný. Tieto zistenia empiricky dopĺňajú závery metaanalýzy Blacka a Wiliama (1998), ktorí konštatujú silný vplyv formatívneho

hodnotenia na vzdelávanie študentov aj o nami skúmanú špecifickú skupinu študentov – budúcich učiteľov predprimárneho a primárneho stupňa vzdelávania.

5. Limity a obmedzenia

V našom výskumnom súbore nebola v jednom z indexov dosiahnutá normalita rozdelenia údajov aj napriek tomu, že rozdelenie sa v základnom súbore študentov, odvolávajúc na centrálnu limitnú vetu, zvyčajne javí ako normálne (Chráska 2007, s. 67). Je teda dôležité pomenovať skryté faktory, ktoré takémuto rozdeleniu v súbore študentov začínajúcich štúdium na vysokej škole pedagogického zamerania môžu brániť:

- študenti nastupujúci do 1. ročníka bakalárskeho štúdia spravidla neabsolvujú maturitnú skúšku z matematiky, ani žiadnu z jej foriem (externú alebo internú), nakoľko v súčasnosti je výber matematiky ako maturitného predmetu postavený na báze voľiteľnosti (v gymnáziách zo skupiny prírodovedných predmetov) alebo dobrovoľnosti (v ostatných stredných školách),
- prijímacie konanie na VŠ štúdium do týchto študijných programov v súčasnosti nevyžaduje absolvovanie prijímacej skúšky z matematiky alebo jej náhrady,
- predmet Matematická gramotnosť je prvým z predmetov matematickej pregraduálnej prípravy, s ktorým sa študenti v zimnom semestri 1. ročníka svojho štúdia na PF PU stretávajú.

Vedomostná úroveň študentov prichádzajúcich z rôznych typov stredných škôl je rôzna. V prípade existencie (aj závažnejších) problémov v doterajšom štúdiu matematiky si študenti tieto problémy prinášajú so sebou aj na vysokú školu.

Druhým faktorom zatiaľ obmedzujúcim vyhodnotenie ďalších konštruktov metakognitívneho monitorovania je identifikáciu spoľahlivo resp. nespoľahlivo zodpovedanej položky. V prípade použitia dichotomického skórovania je táto identifikácia bezproblémová. Nedichotomické skórovanie však naráža na potrebu stanovenia kritéria správnosti odpovede v prípade položiek s vyšším bodovým skórovaním. Diskusia sa otvára pri úvahách:

- nutnosti jeho fixovania v podobe jednotnej relatívnej miery úspešnosti riešenia každej z položiek (napríklad či bude 80% úspešnosť študenta v riešení každej z položky považovaná za hranicu spoľahlivosti pri posudzovaní odpovede),
- o potrebe jeho dynamizácie zohľadňujúcej výskyt rôznych typov položiek v teste alebo rôznyi počet a váha distraktorov v položke.

6. Záver

Konštatujeme, že u študentov študijného odboru Učiteľstvo a pedagogické vedy, v predmete Matematická gramotnosť, nedochádza ku skresleniu medzi subjektívnym a objektívnym hodnotením matematického výkonu, čo považujeme za potešiteľné. Podľa Hnatovej a Mokriša (2020) ku skresleniu nedochádza ani pri parciálnej časti kurikula predmetu Matematická gramotnosť, ktorá sa zaoberala problematikou nepolyadických číselných sústav. V tejto téme síce študenti dosahujú, v rámci priebežného hodnotenia, požadovanú úspešnosť, avšak s nižšou mierou istoty pri posudzovaní správnosti svojej odpovede. Analýza ostatných subtém kurikula predmetu je obsahom ďalšieho skúmania. Zároveň sa v širšom kontexte nepotvrdil žiaden súvis skreslenia s výsledným hodnotením študentov, ktoré dosiahli v predmete Matematická gramotnosť.

Pri presnosti hodnotenia konštatujeme, že je možné dosiahnutý stav zlepšovať. Z pohľadu učiteľov musíme svoju pozornosť zamerať na študentov s horším záverečným hodnotením matematického výkonu v predmete, nakoľko títo dosahujú signifikantne významnejšiu nepresnosť v hodnotení.

Jedným z ďalších smerovaní výskumu je aplikácia konštruktov založených na porovnaní výsledkov dosiahnutých pri spoľahlivo a nespoľahlivo hodnotených položkách. Našou ambíciou je použiť index relatívnej presnosti (r), ktorý hodnotí (lineárny) vzťah medzi úsudkami spoľahlivosti a objektívnym hodnotením výkonu podľa daného kritéria. V spoločnej interpretácii budú tieto indexy spolu so skreslením komplexnejšie popisovať konštrukty metakognitívneho monitorovania matematického výkonu študentov.

Literatúra

- Black, P. & D. Wiliam. (1998). Inside the Black Box: Raising Standards through Classroom Assessment. The Phi Delta Kappan [online]. 1998, 80(2), 139-147 [cit. 2020-10-26]. ISSN 00317217.
- Boekaerts, M., & Rozendaal, J.S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction* [online]. 2010, 20(5), 372-382. ISSN 09594752. Dostupné z <https://doi.org/10.1016/j.learninstruc.2009.03.002>.
- Hnatová, J. & Mokriš, M. (2019) Sebahodnotenie kognitívneho výkonu študentov predškolskej a elementárnej pedagogiky dosahujúcich rôzne klasifikačné stupne hodnotenia v matematike. *Elementary Mathematics Education Journal*. 2019. Olomouc: Univerzita Palackého v Olomouci, s. 16 – 26, ISSN 2694-813. Dostupné z http://emejournal.upol.cz/Issues/Vol1No2/EMEJ_Vol1No2.pdf.
- Hnatová, J. & Mokriš, M. (2020) Sebahodnotenie kognitívneho výkonu študentov predškolskej a elementárnej pedagogiky v téme nepolyadické číselné sústavy. In Nesmorenko, T. & Bernatová, R. (eds) *Osvita i suspiľstvo: mižnarodnyj zbirnyk naukovych prac*. Opole: Wyższa Szkoła Zarządzania i Administracji w Opolu, 2020. s. 41 – 50, ISBN 978-83-66567-00-9. Dostupné z http://pedagogika.wszia.opole.pl/ebook/1_2020.pdf.
- Hnatová, J., Mokriš, M., & Lipták, J. (2019) Študenti predškolskej pedagogiky a ich sebahodnotenie na začiatku profesijnej matematickej prípravy. In Hnatová, J. & Vužňáková, K. (eds) *Premeny školy a učiteľské vzdelávanie v historickom kontexte a nové perspektívy. Zborník príspevkov z vedeckej konferencie s medzinárodnou účasťou organizovanej Pedagogickou fakultou Prešovskej univerzity v Prešove pri príležitosti 70. výročia založenia Pedagogickej fakulty v Prešove*. Prešov: Prešovská univerzita v Prešove, 2019. s.135-141. ISBN 978-80-555-2267-8.
- Chráška, Miroslav. *Metódy pedagogického výzkumu: Základy kvantitatívneho výzkumu*. Praha: Grada Publishing, 2007. ISBN 978-80-247-1369-4.
- Lindel, K., Lenhart, J., & Schneider W. (2019) Metacognition in Mathematics: Do Different Metacognitive Monitoring Measures Make a Difference? *ZDM: The International Journal on Mathematics Education* [online]. 2019, 51(4), 587-600 [cit. 2020-10-18]. ISSN 18639690. Dostupné z <https://link.springer.com/article/10.1007/s11858-019-01062-8>.
- Maki, R. H., Shields, M., Wheeler A.F. & Zacchilli, T.L. (2005) Individual Differences in Absolute and Relative Metacomprehension Accuracy. *Journal of Educational Psychology* [online]. 2005, 97(4), 723-731 [cit. 2020-10-18]. ISSN 00220663.

- Merritt, R. D. (2019) Teaching Study Skills. Salem Press Encyclopedia [online]. 2019 [cit. 2020-10-18].
- Morphew, J. W. (2020) Changes in metacognitive monitoring accuracy in an introductory physics course. *Metacognition and Learning* [online]. 2020, 1-23 [cit. 2020-10-18]. ISSN 15561623. Dostupné z <https://doi.org/10.1007/s11409-020-09239-3>.
- Nelson, T. O. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not the absolute performance on an individual item: comments on Schraw (1995). *Applied Cognitive Psychology*, 10, 257–260. Dostupné z [https://doi.org/10.1002/\(SICI\)1099-0720\(199606\)10:3<257::AID-ACP400>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1099-0720(199606)10:3<257::AID-ACP400>3.0.CO;2-9).
- Nelson, L. J. & Fyfee, R. (2019) Metacognitive monitoring and help-seeking decisions on mathematical equivalence problems. *Metacognition and Learning* [online]. 2019, 14(2), 167-187 [cit. 2020-10-18]. ISSN 15561623. Dostupné z <https://doi.org/10.1007/s11409-019-09203-w>.
- Schraw, G. (2009) A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning* [online]. 2009, 4(1), 33-45 [cit. 2020-10-18]. ISSN 15561623. Dostupné z <https://doi.org/10.1007/s11409-008-9031-3>.
- Yates, J. F. (1988). Analyzing the accuracy of probability judgments for multiple events: an extension of the covariance decomposition. *Organizational Behavior and Human Decision Processes*, 41, 281–299. Dostupné z [https://doi.org/10.1016/0749-5978\(88\)90031-3](https://doi.org/10.1016/0749-5978(88)90031-3).
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45, 166–183. Dostupné z <https://doi.org/10.3102/0002831207312909>.