

Konštrukcia a pilotné psychometrické overenie testu porozumenia výrokovej logiky

Construction and Pilot Psychometric Validation of a Propositional Logic Understanding Test

Anita Molnárová^a – Kitti Páleníková^b – Valéria Švecová^b

^aAnita Molnárová, Cabajská 10, 949 01 Nitra, Slovakia

^bDepartment of Mathematics, Faculty of Natural Sciences and Informatics, Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, 949 01 Nitra, Slovakia

Received February 2, 2026 ; received in revised form May 13, 2026; accepted May 27, 2026

Abstract

This article presents the development and pilot psychometric evaluation of a test assessing secondary school students' understanding of propositional logic. The test was constructed in two parallel forms, A and B, and focused on basic concepts of propositional logic, logical connectives, negation, implication, propositional formulas, truth tables, and word problems. The pilot study included 368 students, of whom 190 completed form A and 178 completed form B. The equivalence of the forms was examined by comparing total scores. Because the score distributions deviated from normality and a slight ceiling effect was observed, the Mann–Whitney U test was used. No statistically significant difference between the forms was found, and the effect size was negligible, indicating that the forms can be considered comparable in terms of overall performance. Item difficulty analysis showed that students were more successful in straightforward tasks involving terminology and simple symbolic representation. More demanding items involved quantified statements, truth tables, and implication in word problems. Reliability analysis indicated acceptable internal consistency for a short pilot test and identified items requiring revision. The results suggest that the test may serve as a useful diagnostic tool and provide a basis for further refinement and validation.

Keywords: propositional logic; test; version; difficulty; reliability.

Classification: 97E30

Úvod

Výroková logika predstavuje jednu zo základných oblastí matematického vzdelávania, ktorá vytvára predpoklady pre porozumenie matematickému jazyku, formulovanie tvrdení, posudzovanie ich pravdivosti a rozlišovanie správnych a nesprávnych úsudkov. V stredoškolskej matematike sa často zaraďuje medzi úvodné tematické celky, pretože na rozdiel od mnohých ďalších oblastí matematiky nevyžaduje rozsiahle špecifické vedomostné predpoklady zo základnej školy. Pracuje predovšetkým s prirodzeným jazykom, pravdivostnou hodnotou tvrdení a vzťahmi medzi výrokmi, čo umožňuje nadviazať na bežné jazykové a argumentačné skúsenosti študentov. Zároveň však už na začiatku stredoškolského štúdia

*Corresponding author; email: kpalenikova@ukf.sk

otvára cestu k presnejšiemu matematickému vyjadrovaniu a k uvedomelému používaniu pojmov ako výrok, negácia, konjunkcia, disjunkcia, implikácia a ekvivalencia.

Význam výrokovej logiky nespočíva iba v osvojení si formálnych pravidiel alebo pravdivostných tabuliek. Jej dôležitosť vyplýva najmä z toho, že poskytuje študentom nástroje na presné čítanie a interpretáciu matematických tvrdení. Porozumenie logickým spojkám je nevyhnutné pri práci s definíciami, vetami, dôkazmi, podmienkami, ekvivalentnými úpravami rovníc či pri posudzovaní správnosti argumentácie. Štátny vzdelávací program pre gymnáziá zahŕňa v tematickom okruhu Logika, dôvodenie, dôkazy požiadavky, aby študent vedel určiť, či je daná vetná konštrukcia výrokom, správne vnímať logické spojky v rôznych prostrediach, rozlišovať implikáciu a ekvivalenciu, tvoriť zložené výroky, určovať ich štruktúru a pravdivosť, negovať výroky a hľadať chyby v argumentácii a usudzovaní (ŠPÚ, 2015).

Osobitne významné miesto má výroková logika vo vzťahu k matematickému dôkazu. Dôkaz nie je len formálnym zápisom postupu, ale predstavuje zdôvodnenú líniu argumentácie, prostredníctvom ktorej sa overuje pravdivosť matematického tvrdenia. Aristidou (2020) upozorňuje, že hoci sa v bežnej matematickej praxi nepoužívajú všetky formálne techniky matematickej logiky v plnej symbolickej podobe, logika je pre matematické dokazovanie a analytické myslenie dôležitá. Podobne Fehér a kol. (2023) zdôrazňujú, že výroková logika sa uplatňuje pri rozhodovaní, riešení problémov, rozvíjaní kritického myslenia, komunikácii a vyhodnocovaní pravdivosti informácií.

Z didaktického hľadiska je dôležité, aby sa výroková logika nevyučovala iba ako súbor pravidiel, ktoré majú študenti mechanicky aplikovať. Viaceré výskumy v oblasti matematického vzdelávania upozorňujú na riziko mechanického riešenia úloh bez hlbšieho porozumenia. Lithner (2008) upozorňuje, že ak sa študenti pri riešení matematických úloh príliš spoliehajú na zapamätané postupy a vzorové riešenia, môže to obmedzovať rozvoj ich hlbšieho porozumenia. Za dôležité preto považuje aj také úlohy, pri ktorých musia študenti samostatne premýšľať, hľadať súvislosti a zdôvodňovať svoj postup. Jonsson a kol. (2020) dopĺňajú, že úlohy vyžadujúce konštrukciu vlastného postupu môžu viesť k lepšiemu porozumeniu než úlohy založené iba na nasledovaní predpísaného algoritmu. V kontexte výrokovej logiky to znamená, že nestačí, aby študent vedel formálne vyplniť pravdivostnú tabuľku; dôležité je aj to, či rozumie významu logických spojok a dokáže ich aplikovať v matematických aj slovných situáciách.

Z uvedených dôvodov je potrebné mať k dispozícii nástroj, ktorý umožní zachytiť nielen celkovú úspešnosť študentov, ale aj konkrétne problematické oblasti ich porozumenia výrokovej logike. Diagnostické hodnotenie môže učiteľovi poskytnúť informácie o pretrvávajúcich miskoncepciách a podporiť rozhodovanie o ďalších krokoch vo výučbe (Ketterlin-Geller a Yovanoff, 2009; Schildkamp a kol., 2020). Cieľom predkladaného príspevku je preto predstaviť konštrukciu testu porozumenia výrokovej logike, porovnať jeho paralelné varianty a pilotne overiť vybrané psychometrické vlastnosti testu, najmä obťažnosť položiek a reliabilitu.

1 Metodológia

Pilotné overenie testu bolo zamerané na prvotné posúdenie jeho použiteľnosti pri diagnostike porozumenia výrokovej logike u žiakov stredných škôl, ako zvládajú vybrané pojmy a vzťahy výrokovej logiky. Z tohto dôvodu bola pozornosť venovaná najmä základným psychometrickým ukazovateľom, ktoré môžu slúžiť ako podklad na ďalšie úpravy testu.

Overenie sa sústredilo na tri oblasti. Prvou bolo porovnanie dvoch paralelných variantov testu, ktoré mali rovnakú štruktúru a analogické obsahové zameranie. Druhou bola analýza obťažnosti jednotlivých položiek, ktorá umožnila identifikovať položky ľahké, primerane náročné a ťažké. Treťou oblasťou bolo posúdenie reliability testu ako ukazovateľa jeho vnútornej konzistencie. Výsledky týchto analýz boli následne využité na zhodnotenie kvality testových položiek a na formulovanie odporúčaní pre ďalšiu revíziu testu.

1.1 Výskumná vzorka

Pilotné overenie testu bolo realizované na vzorke 368 študentov stredných škôl. Podrobnejšiu štruktúru výskumnej vzorky podľa školy uvádzame v tabuľke 1.

Tabuľka 1: Charakteristika výskumnej vzorky podľa zapojenej školy

Škola	Počet tried	Počet žiakov
Gymnázium (S1)	4	104
Gymnázium (S2)	2	53
Stredná odborná škola (S3)	5	132
Stredná odborná škola (S4)	4	79
Spolu	15	368

Výber škôl mal charakter dostupného výberu, pričom zámerom bolo zahrnúť do pilotného overenia študentov z oboch typov stredných škôl, teda z gymnázií aj zo stredných odborných škôl.

Z hľadiska ročníka boli do pilotného overenia zaradení študenti prvého ročníka štvorročného gymnázia, študenti piateho ročníka osemročného gymnázia, študenti prvého ročníka strednej odbornej školy a študenti druhého ročníka strednej odbornej školy. Zaradenie študentov druhého ročníka jednej zo stredných odborných škôl súviselo s organizáciou školského vzdelávacieho programu, keďže učivo výrokovej logiky sa na danej škole preberá až na začiatku druhého ročníka.

1.2 Test porozumenia výrokovej logike a jeho administrácia

Test porozumenia výrokovej logike bol zostavený ako didaktický test určený na overenie porozumenia učiva výrokovej logiky po jeho prebratí vo vyučovaní. Jeho cieľom bolo zistiť, do akej miery študenti rozumejú základným pojmom výrokovej logiky, logickým spojkám, práci s pravdivosťnými hodnotami a vybraným aplikáciám výrokovej logiky v slovných úlohách.

Test bol vytvorený v dvoch paralelných variantoch označených ako variant A a variant B. Oba varianty mali rovnakú štruktúru, rovnaký počet položiek, rovnaký maximálny počet bodov a analogické obsahové zameranie. Položky v jednotlivých variantoch overovali rovnaké typy poznatkov a zručností, líšili sa však konkrétnym zadáním a ponúkanými možnosťami odpovedí. Cieľom vytvorenia dvoch paralelných variantov bolo umožniť administráciu testu vo viacerých skupinách študenti a zároveň posúdiť ich porovnateľnosť z hľadiska dosiahnutých výsledkov.

Každý variant testu obsahoval 10 položiek. Položky mali formu úloh s výberom jednej správnej odpovede z troch alebo štyroch ponúkaných možností. Test bol hodnotený dichotomicky: za správnu odpoveď získal študenti 1 bod, za nesprávnu odpoveď 0 bodov. Maximálny počet bodov, ktorý tak mohol študent v teste dosiahnuť, bol 10 bodov.

Z hľadiska obsahu test postupne prechádzal od jednoduchších položiek zameraných na základné pojmy výrokovej logiky k náročnejším položkám vyžadujúcim interpretáciu zložených

výrokov a aplikáciu poznatkov v kontexte slovných úloh. Prvé položky boli orientované na identifikáciu výroku, negáciu a negáciu výrokovej formy. Ďalšie položky sa sústreďovali na zložené výroky, používanie logických spojok, ich terminológiu, určovanie pravdivostnej hodnoty a symbolický zápis. Súčasťou testu bola aj položka zameraná na prácu s pravdivostnou tabuľkou výrokovej formuly. Záverečné položky overovali schopnosť aplikovať poznatky z výrokovej logiky v slovných úlohách, pričom jedna z nich bola zameraná na negáciu implikácie.

Takto zostavený test umožňoval získať informáciu nielen o celkovom výsledku študenta, ale aj o úspešnosti riešenia jednotlivých typov položiek. To bolo dôležité najmä z hľadiska pilotného psychometrického overenia testu, v rámci ktorého sa sledovala porovnateľnosť paralelných variantov, obťažnosť položiek a reliabilita testu.

Test bol administrovaný na jeseň 2025, počas vyučovacej hodiny matematiky po prebratí príslušného učiva výrokovej logiky. Test mal papierovú formu a študenti ho riešili samostatne v priebehu jednej vyučovacej hodiny, teda počas 45 minút. Varianty A a B boli študentom prideľované náhodne tak, aby študenti sediaci vedľa seba nemali rovnaké zadanie. Tento spôsob administrácie bol zvolený s cieľom obmedziť možnosť vzájomného ovplyvňovania odpovedí a zároveň umožnil následné porovnanie výsledkov dosiahnutých v oboch variantoch testu.

1.3 Metódy analýzy údajov

Výsledky testu boli spracované kvantitatívne. Keďže test pozostával z desiatich dichotomicky hodnotených položiek, každá položka bola kódovaná hodnotou 1 v prípade správnej odpovede a hodnotou 0 v prípade nesprávnej odpovede. Celkové skóre študenta bolo určené ako súčet bodov za jednotlivé položky, maximálny možný výsledok tak bol 10 bodov.

V prvej fáze bola spracovaná deskriptívna štatistika celkovej úspešnosti študentov. Výsledky boli vyhodnotené za celý súbor respondentov a následne aj samostatne podľa zapojených škôl. Sledované boli najmä priemerné skóre, smerodajná odchýlka, medián, minimálna a maximálna dosiahnutá hodnota. Cieľom tejto časti analýzy bolo získať základný obraz o celkovej úrovni zvládnutia testu a o rozdieloch medzi jednotlivými školami.

V druhom kroku analýzy bola venovaná pozornosť porovnaniu paralelných variantov testu. Pre variant A a variant B boli najskôr vypočítané základné deskriptívne ukazovatele celkového skóre. Následne boli výsledky študentov riešiacich jednotlivé varianty porovnané pomocou Mann-Whitneyho U testu (Markechová a kol., 2011). Tento neparametrický test bol zvolený vzhľadom na charakter údajov, obmedzený rozsah škály celkového skóre a cieľ porovnať dve nezávislé skupiny respondentov. Okrem štatistickej významnosti rozdielu bola vypočítaná aj veľkosť efektu prostredníctvom rank-biseriálnej korelácie, označovanej ako RBC (Kerby, 2014). Tá umožňuje vyjadriť veľkosť rozdielu medzi dvoma nezávislými skupinami pri použití Mann-Whitneyho U testu; Kerby uvádza jej vzťah k Mann-Whitneyho testu a interpretuje ju ako rozdiel medzi podielom priaznivých a nepriaznivých párov. Rank-biseriálna korelácia bola použitá preto, aby sa porovnanie variantov neopieralo iba o hodnotu štatistickej významnosti. Pri väčšom počte respondentov totiž môže byť aj malý rozdiel štatisticky významný, hoci z praktického hľadiska nemusí byť podstatný. Hodnota RBC blízka nule naznačuje malý rozdiel medzi variantmi, zatiaľ čo vyššia absolútna hodnota poukazuje na výraznejšie odlišnosti medzi výsledkami študentov riešiacich variant A a variant B. V kontexte pilotného overenia bola preto rovnocennosť variantov posudzovaná na základe kombinácie deskriptívnych ukazovateľov, výsledku Mann-Whitneyho U testu a veľkosti efektu RBC.

Tretiu časť analýzy tvorilo určenie obťažnosti jednotlivých položiek. Obťažnosť položky bola vyjadrená ako index P - podiel počtu žiakov, ktorí danú položku vyriešili správne, z celkového počtu žiakov, ktorí danú položku riešili. V tomto chápaní vyššia hodnota indexu P znamená ľahšiu položku, keďže vyjadruje vyšší podiel správnych odpovedí. V klasickej testovej teórii sa hodnota P bežne používa ako index obťažnosti položky pri úlohách s výberom odpovede; vysoká hodnota znamená, že položka bola pre respondentov ľahká, nízka hodnota naopak poukazuje na náročnú položku (Thompson, 2023). Pri interpretácii obťažnosti boli položky orientačne rozdelené podľa Thompsona do štyroch skupín: položky s hodnotou nižšou ako 0,40 boli považované za veľmi ťažké, položky s hodnotou medzi 0,40 a 0,60 za náročné, položky s hodnotou od 0,60 do 0,95 za typické položky a položky s hodnotou nad 0,95 za veľmi ľahké. Uvedené hranice boli použité ako orientačné kritérium, nie ako absolútne pravidlo, keďže vhodná obťažnosť položky závisí od účelu testu, cieľovej skupiny a fázy výučby.

V poslednom kroku analýzy bola vypočítaná reliabilita testu na posúdenie vnútornej konzistencie testu. Keďže všetky položky boli hodnotené dichotomicky, teda hodnotami 0 a 1, reliabilita bola interpretovaná ako koeficient KR-20 (Kuder a Richardson, 1937). Vzhľadom na pilotný charakter overenia bola hodnota reliability interpretovaná opatrne. Cieľom nebolo potvrdiť definitívnu podobu testu ako štandardizovaného nástroja, ale získať podklad na ďalšiu úpravu testu a prípadnú revíziu vybraných položiek.

Na štatistické spracovanie a vyhodnotenie získaných údajov boli použité programy Microsoft Excel s doplnkom Real Statistics ([www1](#)) a R ([www2](#)).

2 Výsledky pilotného overenia testu

V rámci tejto časti uvádzame výsledky pilotného psychometrického overenia testu porozumenia výrokovej logiky. Analýza bola zameraná na celkové výsledky študentov v teste: deskriptívne charakteristiky celkového skóre za celý súbor respondentov a podľa jednotlivých škôl, porovnanie dvoch paralelných variantov testu a posúdenie vybraných psychometrických vlastností testových položiek - analýza obťažnosti jednotlivých položiek a reliabilita testu.

2.1 Deskriptívna analýza výsledkov testu

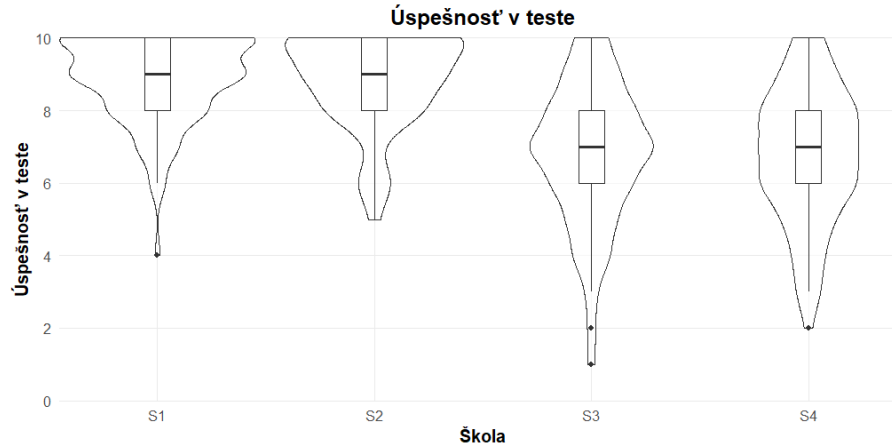
Sledovanou premennou bola úspešnosť študentov v teste porozumenia výrokovej logiky, vyjadrená počtom získaných bodov. Maximálne skóre predstavovalo 10 bodov, pričom za správnu odpoveď v každej z desiatich úloh získal študent 1 bod. V tabuľke 2 uvádzame deskriptívne charakteristiky úspešnosti v teste podľa jednotlivých škôl a zároveň za celý výskumný súbor.

Tabuľka 2: Deskriptívne charakteristiky úspešnosti v teste podľa škôl.

Škola	Počet	Priemer	SD	Medián	IQR	min	max
S1	104	8,904	1,170	9	2	4	10
S2	53	8,849	1,292	9	2	5	10
S3	132	6,765	1,933	7	2	1	10
S4	79	6,759	1,876	7	2	2	10
Spolu	368	7,668	1,951	8	3	1	10

Z údajov v tabuľke vyplýva, že najvyššiu priemernú úspešnosť dosiahli študenti škôl S1 a S2 – obe gymnáziá. Priemery a mediány naznačujú, že typický výkon študentov v týchto školách sa

pohyboval na pomerne vysokej úrovni. Zároveň sa v týchto školách ukázala aj nižšia variabilita výsledkov. Naopak, nižšiu priemernú úspešnosť zaznamenali školy S3 a S4 (stredné odborné školy) v porovnaní so školami S1 a S2. Vyššie hodnoty smerodajnej odchýlky ukazujú vyššiu variabilitu výsledkov. Tieto charakteristiky umožňuje vizuálne lepšie posúdiť aj graf na obr. 1, z ktorého vidíme polohu mediánu, variabilitu výsledkov aj tvar rozdelenia testového skóre.



Obr. 1: Rozdelenie celkovej úspešnosti študentov v teste podľa školy

Z grafu je zrejmé, že rozdelenie úspešnosti na gymnáziách je posunuté do vyšších hodnôt testového skóre. V oboch prípadoch sa väčšia časť výsledkov sústreďuje v hornej časti škály, najmä medzi 8 až 10 bodmi, čo je v súlade s vyššími priemerami aj mediánmi uvedenými v tabuľke. Box-ploty zároveň naznačujú menší rozptyl výsledkov, teda väčšiu homogénnosť výkonu študentov na týchto dvoch školách. Naopak, SOŠ majú rozdelenie výsledkov rozptýlenejšie a posunuté do nižších hodnôt. Výsledky sú viac rozložené v strednej časti škály, pričom sa objavujú aj nižšie skóre, čo zodpovedá nižším priemerom, nižším mediánom a širšiemu rozsahu hodnôt.

Celkovo možno konštatovať, že pri premennej úspešnosť v teste sa medzi školami ukazujú pomerne výrazné rozdiely už na úrovni deskriptívnych ukazovateľov. Vzhľadom na zaradenie dvoch gymnázií a dvoch stredných odborných škôl do pilotného overenia bol takýto výsledok do istej miery očakávaný. Vyššie hodnoty dosiahnuté študentmi gymnázií môžu súvisieť s odlišnou matematickou prípravou študentov, vyššími nárokmi na výkon, ako aj väčším rozsahom vyučovania matematiky na gymnáziách.

2.2 Porovnanie variantov testu

Pearsonovým chí-kvadrát testom nezávislosti sme overili, že priradenie variantov testu (A a B) nebolo štatisticky významne asociované so školou ($\chi^2(3, 368) = 4,39, p = 0,223$). Varianty boli teda na jednotlivých školách zastúpené približne rovnomerne. Pri ďalších analýzach preto nemožno predpokladať skreslenie výsledkov v dôsledku toho, že by sa jednotlivé varianty koncentrovali na konkrétnych školách.

Rovnocennosť paralelných foriem testu bola overovaná porovnaním celkového skóre študentov vo variante A a vo variante B. V tabuľke 3 uvádzame deskriptívne charakteristiky skóre v jednotlivých variantoch testu. Keďže rozdelenie sledovanej premennej sa v oboch variantoch odchyľovalo od normálneho rozdelenia, čo potvrdil Shapiro-Wilkov test normality ($p < 0,001$), a zároveň sa prejavil mierny stropový efekt, na porovnanie výsledkov medzi variantmi bol použitý Mann-Whitneyho U test.

Mann-Whitneyho U test nepreukázal štatisticky významný rozdiel v rozdelení celkového skóre medzi variantom A a variantom B ($U = 15\,251,0$; $p = 0,099$). Keďže tento test pracuje s poradím hodnôt, výsledok neinterpretujeme ako priame porovnanie priemerov, ale ako porovnanie rozdelenia, resp. poradí dosiahnutých skóre v oboch skupinách.

Tabuľka 3: Deskriptívne charakteristiky celkového skóre v teste podľa variantu a výsledok Mann-Whitneyho U testu.

charakteristika	variant A (n = 190)	variant B (n = 178)	porovnanie
priemer	7,53	7,81	U = 15 251,0
SD	1,94	1,96	
medián	8	8	p = 0,099
IQR	6 – 9	7 – 9	RBC = -0,099
Min - max	2 – 10	1 – 10	

Poznámka: Mann-Whitneyho U test porovnáva poradie dosiahnutých skóre v oboch skupinách, nie priamo ich priemery. RBC (rank-biseriálna korelácia) predstavuje veľkosť efektu; hodnoty blízke nule indikujú zanedbateľný rozdiel medzi variantmi.

Deskriptívne charakteristiky zároveň ukazujú veľmi podobné výsledky v oboch variantoch, najmä zhodný medián a porovnateľnú variabilitu skóre. Hodnota veľkosti efektu bola zanedbateľná ($RBC = -0,099$), čo naznačuje iba veľmi malý rozdiel medzi skupinami. Vzhľadom na tieto výsledky možno varianty A a B z hľadiska celkového skóre v pilotnom overení považovať za porovnateľné.

2.3 Obťažnosť položiek

V tejto časti uvádzame analýzu obťažnosti položiek vyjadrenú indexom P (tab. 4), pričom vyššia hodnota indexu znamená vyšší podiel správnych odpovedí.

Tabuľka 4: Náročnosť jednotlivých položiek testu pre oba varianty A a B.

položka	obsahové zameranie	P (variant A)	P (variant B)
U1	výrok	0,847	0,742
U2	negácia	0,784	0,758
U3	negácia výrokovej formy	0,547	0,725
U4	zložené výroky - logické spojky - terminológia	0,895	0,921
U5	zložené výroky - logické spojky	0,821	0,860
U6	Zložené výroky - pravdivostná hodnota	0,705	0,815
U7	zložené výroky - logické spojky – symbolický zápis	0,884	0,944
U8	výroková formula – pravdivostná tabuľka	0,542	0,663
U9	slovná úloha – negácia implikácie	0,868	0,669
U10	slovná úloha - výroková formula – pravdivostná tabuľka	0,637	0,719

Poznámka. Index P vyjadruje podiel správnych odpovedí na danú položku; čím je hodnota P vyššia, tým bola položka pre respondentov ľahšia.

Oba varianty testu sú ako celok primerané. Priemerná hodnota indexu P vo variante A bola 0,753 a vo variante B 0,782, z čoho vyplýva, že variant B bol mierne ľahší. Podľa Thompsonovej stupnice väčšina položiek patrí medzi typické položky. Ako náročnejšie sa ukázali najmä položky U3 a U8 vo variante A. Žiadna položka nespadá do kategórie príliš náročných položiek a žiadna neprekročila hranicu príliš ľahkej položky, hoci položka U7 vo variante B sa tejto hranici výrazne priblížila.

Pri úprave alebo výmene položiek bude vhodné zamerať sa najmä na položky U1, U3, U6, U8 a U9, pri ktorých bol rozdiel v obťažnosti medzi variantmi väčší ako jedna desatina. Pri položke U1B mohla byť pre časť respondentov mätúca možnosť „Čokoláda.“, keďže formálne pripomína oznamovaciu vetu, ale nespĺňa podmienku výroku s určitelnou pravdivostnou hodnotou. Pri položke U3 je rozdiel medzi variantmi pomerne výrazný, preto bude potrebné zvážiť úpravu distraktorov tak, aby boli paralelné položky vyrovnannejšie. Pri položke U8 je vhodné preformulovať možnosti odpovede, keďže kategória splniteľnej formuly sa môže prekrývať s tautológiou. Pri položke U9B mohla byť vyššia náročnosť spôsobená skôr formuláciou slovnej úlohy než samotným logickým obsahom. Položky U4 a U7, najmä vo variante B, patria medzi ľahšie položky; v prípade potreby by bolo možné ich sťažiť alebo nahradiť aplikačne zameranými úlohami. Ďalšie rozhodnutia o úprave alebo výmene položiek budú posúdené aj na základe ukazovateľov reliability v nasledujúcej podkapitole.

2.4 Reliabilita testu

Reliabilita testu bola posudzovaná prostredníctvom vnútorných konzistenčných ukazovateľov, a to osobitne pre variant A, variant B aj pre celý súbor respondentov (tab. 5). Keďže test pozostával z dichotomicky skórovaných položiek, ako hlavný ukazovateľ bol použitý koeficient KR-20, doplnený o McDonalduvu omegu (ω) a položkové korelačné charakteristiky.

Tabuľka 5: Ukazovatele vnútornej konzistencie testu podľa variantu

Ukazovateľ	Variant A	Variant B	Celkovo
počet respondentov (N)	190	178	368
KR-20	0,609	0,640	0,618
McDonalduva ω	0,838	0,875	0,836
Medzipoložková korelácia r (min – max; priemer)	-0,121 – 0,402; 0,134	0,020 – 0,422; 0,164	-0,007 – 0,382; 0,145
Korigovaná korelácia položka–celok r	0,101 – 0,471	0,164 – 0,417	0,118 – 0,414
Priemerná náročnosť položiek (P)	0,753	0,782	0,767

Poznámka. Koeficient KR-20 je ekvivalent Cronbachovej alfy (α) pre dichotomické položky (0/1). P vyjadruje priemernú úspešnosť (obťažnosť) položiek. Korigovaná korelácia položka–celok (r_{drop}) je korelácia položky s celkovým skóre po vylúčení danej položky. Varianty A a B obsahovali rovnaký počet položiek (10 úloh).

Z hľadiska vnútornej konzistencie vykazujú obe formy testu porovnateľnú reliabilitu. Koeficient KR-20, ktorý je ekvivalentom Cronbachovej α pre dichotomické položky, dosiahol hodnoty 0,609 vo variante A, 0,640 vo variante B a 0,618 v celom súbore. Tieto hodnoty naznačujú strednú, resp. hranične prijateľnú vnútornú konzistenciu testu, pričom variant B dosiahol mierne vyššiu hodnotu než variant A. Keďže ide o krátky test s desiatimi

dichotomickými položkami, je prirodzené, že hodnota KR-20 je citlivá na menší počet úloh aj na obsahovú rôznorodosť položiek.

Okrem KR-20 sme na posúdenie reliability použili aj McDonaldovu ω . Tento ukazovateľ dosiahol vyššie hodnoty (A: 0,838, B: 0,875, celkovo: 0,836), čo poukazuje na dobrú spoľahlivosť testu. Rozdiel medzi KR-20 a McDonaldovou ω môže súvisieť s tým, že jednotlivé položky neprispievajú k celkovému skóre úplne rovnako a niektoré rozlišujú výkon študentov lepšie než iné. V takomto prípade môže KR-20 spoľahlivosť testu skôr mierne podhodnocovať, zatiaľ čo McDonaldova ω ju vystihuje presnejšie.

Priemerné medzipoložkové korelácie boli nízke až mierne (A: 0,134, B: 0,164, celkovo: 0,145), čo naznačuje, že test nie je úplne homogénny, ale zachytáva viacero čiastkových stránok výkonu vo výrokovvej logike. Zároveň sa pri niektorých dvojiciach položiek objavili aj záporné korelácie (A: min. -0,121, celkovo: min. -0,007), čo môže znamenať, že tieto položky nesúvisia s ostatnými rovnakým spôsobom. Podrobnejší pohľad do korelačnej matice ukázal záporné korelácie napríklad pri dvojiciach U1 – U3 a U2 – U3. Tieto dvojice vyžadovali odlišný spôsob uvažovania. Kým položka U1 overovala základné rozpoznanie výroku, položka U3 bola viac zameraná na formálne negovanie výrokových foriem; položky U2 a U3 síce obe pracovali s negáciou, avšak jedna bola formulovaná slovne a druhá symbolicky, takže kládli na študentov odlišné nároky.

Rozsahy korigovaných korelácií položka – celok ukazujú, že diskriminačná schopnosť položiek bola celkovo slabšia až stredná (A: 0,101 – 0,471, B: 0,164 – 0,417, celkovo: 0,118 – 0,414). To znamená, že nie všetky úlohy rovnako dobre odlišovali úspešnejších a menej úspešných študentov. Za položky vhodné na revíziu na základe týchto ukazovateľov možno považovať najmä tie, ktorých hodnota bola nižšia ako približne 0,20. Vo variante A ide o úlohy U1 a U9, vo variante B o úlohu U1. Tieto výsledky naznačujú, že uvedené položky mohli k celkovej rozlišovacej schopnosti testu prispievať slabšie, a preto by bolo vhodné pri ich ďalšom použití zvážiť ich úpravu alebo obsahovú revíziu.

Na základe analýzy obťažnosti položiek a ukazovateľov reliability možno konštatovať, že test ako celok vykazuje prijateľné vlastnosti vzhľadom na jeho pilotný charakter. Výsledky však zároveň poukázali na viaceré položky, ktoré si pri ďalšom použití testu vyžadujú úpravu alebo zváženie výmeny. Ide najmä o položky U1, U3 a U9 v oboch variantoch testu, pri ktorých sa prejavili rozdiely v obťažnosti medzi paralelnými formami alebo slabší príspevok k celkovému skóre testu. Tieto položky sú preto nižšie uvedené v pôvodnom znení ako podklad pre ich ďalšiu obsahovú a formálnu revíziu.

Úloha 1

variant A:

Ktorá z nasledujúcich viet je výrok?

- $x - 20 = 8$.
- Pôjdeš dnes do kina?
- Rozdiel dvoch celých čísel je celé číslo.
- Zatvor okno!

variant B:

Ktorá z nasledujúcich viet je výrok?

- Pozor pes!
- $3x - 6 = x + 2$, pre $x = 4$.
- Aké je dnes počasie?
- Čokoláda.

Položka 3

variant A:

Vyber správnu negáciu výrokovvej formy:

$\exists x \in \mathbb{R}$, x je párne číslo.

- $\forall x \in \mathbb{R}$, x je párne číslo.

variant B:

Vyber správnu negáciu výrokovvej formy:

$\forall x \in \mathbb{R}$, $x + 2 = 0$.

- $\forall x \in \mathbb{R}$, $x + 2 \neq 0$.

- | | |
|---|--|
| b) $\exists x \in \mathbb{R}$, x nie je párne číslo. | b) $\exists x \in \mathbb{R}$, $x + 2 = 0$. |
| c) $\forall x \in \mathbb{R}$, x nie je párne číslo. | c) $\exists x \in \mathbb{R}$, $x + 2 \neq 0$. |
| d) $\forall x \notin \mathbb{R}$, x je párne číslo. | d) $\exists x \notin \mathbb{R}$, $x + 2 = 0$. |

Položka 9

variant A:

Rodičia hovoria synovi: „Ak získaš vodičský preukaz, tak dostaneš na narodeniny nové auto.“ Vyber, v ktorej možnosti rodičia nedodržali svoj sľub.

- Syn získal vodičský preukaz, dostal na narodeniny nové auto.
- Syn získal vodičský preukaz, nedostal na narodeniny nové auto.
- Syn nezískal vodičský preukaz, dostal na narodeniny nové auto.
- Syn nezískal vodičský preukaz, nedostal na narodeniny nové auto.

variant B:

Tenisový tréner hovorí mladému tenistovi: „Ak vyhráš zápas, tak získaš medailu.“ Vyber, v ktorej možnosti tréner nevyšlovl pravdu.

- Tenista vyhral zápas, získal medailu.
- Tenista nevyhral zápas, získal medailu.
- Tenista vyhral zápas, nezískal medailu.
- Tenista nevyhral zápas, nezískal medailu.

3 Diskusia

Výsledky pilotného overenia naznačujú, že varianty A a B sú na úrovni celkového skóre porovnateľné, čo podporuje ich použiteľnosť ako paralelných foriem testu. Zároveň sa však ukázalo, že porovnateľnosť celkového výkonu ešte neznamená úplnú rovnocennosť jednotlivých položiek. Položky s výraznejším rozdielom v obťažnosti medzi variantmi alebo so slabšou koreláciou s celkovým skóre predstavujú dôležitý podklad pre ďalšiu revíziu testu. Vzhľadom na krátky rozsah testu a pilotný charakter overenia je potrebné ukazovatele reliability chápať najmä orientačne, nie ako definitívne potvrdenie kvality nástroja.

Z analýzy obťažnosti položiek vyplýva, že študenti lepšie zvládali úlohy zamerané na základnú terminológiu výrokovej logiky, rozpoznanie logických spojok a jednoduché určovanie pravdivostných hodnôt zložených výrokov. Náročnejšie sa javili položky, ktoré vyžadovali presnejšie chápanie logickej štruktúry výroku. Ťažkosti sa prejavili najmä pri negácii výrokovej formy s kvantifikátorom. Pri tomto type úloh nestačí negovať iba samotné tvrdenie, ale je potrebné správne zmeniť aj kvantifikátor a zachovať význam celej výrokovej formy. Na podobné ťažkosti upozorňuje aj Epp (2003), žiakom spôsobuje problémy najmä práca s výrokmi s implikáciami a s kvantifikovanými výrokmi, osobitne pri ich negovaní, pretože medzi bežným jazykom a matematickou logikou existujú významové rozdiely.

Rovnako Sellers a kol. (2021) poukazujú na to, že študentské chápanie kvantifikovaných premenných nemusí vždy zodpovedať formálnemu matematickému významu kvantifikátorov. Ďalšou problematickejšou oblasťou boli položky pracujúce s výrokovou formulou a pravdivostnou tabuľkou. Pri týchto úlohách museli študenti nielen poznať význam jednotlivých logických spojok, ale zároveň ich správne kombinovať a vyhodnocovať vo viacerých krokoch. Podobné ťažkosti uvádzajú aj Hawthorne a Rasmussen (2015), podľa ktorých študenti pri práci s pravdivostnými tabuľkami a implikáciami často vnímajú jednotlivé časti formuly izolovane, bez dostatočne uceleného pohľadu na celú logickú štruktúru. Osobitne náročné mohli byť aj slovné úlohy, najmä tie, ktoré pracovali s implikáciou. Hoci vety typu „ak..., tak...“ sú bežné aj v každodennom jazyku, ich formálny matematický význam sa nemusí zhodovať s intuitívnym chápaním študentov. Hoyles a Küchemann (2002) vo svojom výskume ukázali, že študenti majú ťažkosti s rozlišovaním implikácie a jej obrátenej implikácie a ich uvažovanie sa postupne posúva od empirického k deduktívnemu. Durand-Guerrier (2008) zároveň upozorňuje na to,

že študenti môžu mať problém rozlíšiť, ako sa výroky chápu v bežnom jazyku a ako sa s nimi pracuje v matematickej logike. Inak povedané, to, čo sa v každodennej komunikácii javí ako prirodzené alebo samozrejmé, nemusí mať rovnaký význam pri formálnom logickom vyhodnocovaní výroku. Tieto zistenia podporujú potrebu venovať vo vyučovaní väčšiu pozornosť prepájaniu prirodzeného jazyka so symbolickým zápisom, práci s kvantifikovanými výrokmi a postupnému nácviku formálnej interpretácie zložených výrokov.

Z didaktického hľadiska môže test slúžiť ako nástroj na identifikáciu oblastí výrokovej logiky, ktoré študenti zvládajú dobre, a oblastí, ktoré si vyžadujú ďalšie vysvetlenie alebo precvičenie. Úspešnosť v jednotlivých položkách môže učiteľovi ukázať, či trieda potrebuje viac pracovať s negáciou, implikáciou, kvantifikátormi, pravdivostnými tabuľkami alebo s prechodom od slovného vyjadrenia k symbolickému zápisu. V súlade s princípmi formatívneho hodnotenia má test význam najmä vtedy, ak sa jeho výsledky využijú pri plánovaní ďalšej výučby, výbere úloh a rozbere typických chýb študentov (Black a Wiliam, 1998; Akveld a Kinnear, 2024). Podobne Sadler (1989) zdôrazňuje, že spätná väzba má význam vtedy, keď pomáha identifikovať rozdiel medzi aktuálnym výkonom študenta a očakávanou úrovňou a zároveň vytvára podmienky na jeho zlepšenie. Hattie a Timperley (2007) takisto poukazujú na to, že spätná väzba patrí medzi významné faktory ovplyvňujúce učenie, pričom jej účinnosť závisí od toho, ako je využitá v ďalšom procese učenia. Test možno v pedagogickej praxi využiť ako krátky vstupný alebo výstupný diagnostický nástroj.

Záver

V predkladanom príspevku sme predstavili konštrukciu a pilotné psychometrické overenie testu porozumenia výrokovej logike v dvoch paralelných variantoch. Výsledky naznačili, že varianty A a B sú na úrovni celkového skóre porovnateľné, no položková analýza zároveň odhalila rozdiely v obťažnosti niektorých paralelných úloh, ktoré si vyžadujú ďalšiu revíziu. Študenti boli úspešnejší v priamočiarych položkách zameraných na základnú terminológiu, logické spojky a jednoduchý symbolický zápis. Väčšie ťažkosti sa objavili pri úlohách vyžadujúcich formálnejšie uvažovanie, najmä pri negácii výrokovej formy s kvantifikátorom, práci s pravdivostnou tabuľkou a interpretácii implikácie v slovnom kontexte.

Vzhľadom na pilotný charakter výskumu, krátky rozsah testu a dostupný výber respondentov je potrebné výsledky interpretovať opatrne. Test však predstavuje vhodný základ na ďalšiu revíziu položiek, opätovné psychometrické overenie a možné využitie ako diagnostickej pomôcky v stredoškolskej matematike.

Literatúra

Akveld, M., Kinnear, G. 2024. Improving mathematics diagnostic tests using item analysis. In *International Journal of Mathematical Education in Science and Technology*. ISSN 0020-739X, roč. 55, č. 10, s. 2478–2505. <https://doi.org/10.1080/0020739X.2023.2167132>

Aristidou, M. 2020. Is Mathematical Logic Really Necessary in Teaching Mathematical Proofs? In *Athens Journal of Education*. ISSN 2241-7958, rok 2020, roč. 7, č. 1, s. 99-121. <https://doi.org/10.30958/aje.7-1-5>

Black, P., Wiliam, D. 1998. Assessment and classroom learning. In *Assessment in Education: Principles, Policy & Practice*. ISSN 0969-594X, roč. 5, č. 1, s. 7–74. <https://doi.org/10.1080/0969595980050102>

- Durand-Guerrier, V. 2008. Truth versus validity in mathematical proof. In ZDM Mathematics Education. ISSN 1863-9690, roč. 40, č. 3, s. 373–384. <https://doi.org/10.1007/s11858-008-0098-8>
- Epp, S. S. 2003. The role of logic in teaching proof. In: The American Mathematical Monthly, ISSN 0002-9890, rok 2003, roč. 110, č. 10, s. 886-899. <https://doi.org/10.1080/00029890.2003.11920029>
- Fehér, Z., Jaruska, L., Szarka, K., Tóthová Tarová, E. 2023. Students' propositional logic thinking in higher education from the perspective of disciplines. In Frontiers in Education. ISSN 2504-284X, rok 2023, roč. 8, s. 1 – 10. <https://doi.org/10.3389/feduc.2023.1247653>
- Hattie, J., Timperley, H. 2007. The power of feedback. In Review of Educational Research. ISSN 0034-6543, roč. 77, č. 1, s. 81–112. <https://doi.org/10.3102/00346543029848>
- Hawthorne, C., Rasmussen, C. 2015. A framework for characterizing students' thinking about logical statements and truth tables. In International Journal of Mathematical Education in Science and Technology. ISSN 0020-739X, rok 2015, roč. 46, č. 3, s. 337-353. <https://doi.org/10.1080/0020739X.2014.979895>
- Hoyles, C., Küchemann, D. 2002. Students' understandings of logical implication. In Educational Studies in Mathematics. ISSN 0013-1954, rok 2002, roč. 51, č. 3, s. 193-223. <https://doi.org/10.1023/A:1023629608614>
- Jonsson, B., Granberg, C., Lithner, J. 2020. Gaining mathematical understanding: The effects of creative mathematical reasoning and cognitive proficiency. In Frontiers in Psychology. ISSN 1664-1078, rok 2020, roč. 11, s. 1-16. <https://doi.org/10.3389/fpsyg.2020.574366>
- Kerby, D. S. 2014. The Simple Difference Formula: An Approach to Teaching Nonparametric Correlation. In Comprehensive Psychology. ISSN 2165-2228, rok 2014, roč. 3. <https://doi.org/10.2466/11.IT.3>
- Ketterlin-Geller, L. R., Yovanoff, P. 2009. Diagnostic assessments in mathematics to support instructional decision making. In Practical Assessment, Research, and Evaluation. ISSN 1531-7714, rok 2009, roč. 14, č.16, s. 1-11. <https://doi.org/10.7275/vxrk-3190>
- Kuder, G. F., Richardson, M. W. 1937. The Theory of the Estimation of Test Reliability. In: Psychometrika. ISSN 0033-3123, rok 1937, roč. 2, č. 3, s. 151 – 160. <https://doi.org/10.1007/BF02288391>
- Lithner, J. 2008. A research framework for creative and imitative reasoning. In Educational Studies in Mathematics. ISSN 1573-0816, rok 2008, roč. 67, s. 255-276. DOI 10.1007/s10649-007-9104-2
- Markechová, D., Stehlíková, B., Tirpáková, A. 2011. Štatistické metódy a ich aplikácie. Nitra: UKF v Nitre, 2011. 534 s. ISBN 978-80-8094-807-8
- Sadler, D. R. 1989. Formative assessment and the design of instructional systems. In Instructional Science. ISSN 0020-4277, roč. 18, č. 2, s. 119–144. <https://doi.org/10.1007/BF00117714>
- Sellers, M. E., Roh, K. H., Parr, E. D. 2021. Student quantifications as meanings for quantified variables in complex mathematical statements. In The Journal of Mathematical Behavior. ISSN 0732-3123, roč. 61, s. 1 – 16. <https://doi.org/10.1016/j.jmathb.2020.100802>

Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., Veldkamp, B. P. 2020. Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. In International Journal of Educational Research. ISSN 0883-0355 , rok 2020, roč. 103, s. 1-16. <https://doi.org/10.1016/j.ijer.2020.101602>

Thompson, N. 2023 Classical Test Theory: Item Statistics online. [Cit. 4. 5. 2026]. Dostupné na <https://assess.com/item-statistics-classical-test-theory/>

Štátny pedagogický ústav. 2015. Vzdelávací štandard pre učebný predmet Matematika – gymnázium so štvorročným a päťročným vzdelávacím programom. [cit. 13. 1. 2026] Dostupný online: https://www.statpedu.sk/files/articles/dokumenty/inovovany-statny-vzdelavaci-program/matematika_g_4_5_r.pdf