

SKOLĒNU KOGNITĪVĀS DARBĪBAS DZILUMS VIDĒJĀS IZGLĪTĪBAS VALSTS PĀRBAUDES DARBOS MATEMĀTIKĀ

Students' Depth of Cognitive Performance on Secondary School State Mathematics Tests

Marta Mikite

Latvijas Universitāte, Latvija

Ilze France

Latvijas Universitāte, Latvija

Abstract. Latvia began curriculum reform implementation in 2020. Not only has the learning content changed, but students are also now able to choose from one of three levels to study mathematics in secondary school. This requires new exams at the end of each of these courses. The aim of this study is to investigate student performance and to what extent students can use higher cognitive skills on the new exams. Results were analyzed using Classical Test Theory and Item Response Theory. Comparing the results of the new exams using Wright maps shows a trend that, although there are flaws, they are more representative of the relevant population than the previous exam. A group of experts determined the level of cognition for each of the tasks and analyzed students' performance. The results show that there are tasks where students can demonstrate the highest level of cognitive abilities. However, such tasks are relatively few, and they are mathematical rather than problems in a real context. Performance analysis shows that students have difficulty expressing their thoughts and this is one of the reasons why only a small number of students solve challenging problems.

Keywords: cognitive abilities, item analysis, mathematics assessment, national large-scale assessment, performance analysis, secondary school exams.

Ievads

Introduction

Kopš 2020. gada Latvijā notiek jaunā mācību standarta matemātikā ieviešana. Mainās ne tikai mācību saturs, bet izmaiņas ir arī vidusskolas matemātikas kursa apguves formā. Skolēni var izvēlēties vienu no trīs mācību satura apguves līmeņiem, kādā apgūt matemātikas kursu - vispārīgajā, optimālajā vai augstākajā līmenī (Ministru kabinets [MK], 2019). Sākot ar 2021./22. mācību gadu vispārizglītojošo

skolu skolēniem, kuri apgūst matemātiku pēc jaunā mācību priekšmeta standarta, jākārtu centralizētais eksāmens optimālajā satura apguves līmenī (CEOAL), vai augstākā (MK, 2019). Savukārt, profesionālās izglītības iestāžu audzēkņiem, kuri apgūst vispārējās izglītības saturu, jākārtu matemātikas eksāmens vispārīgajā satura apguves līmenī (CEVAL) vai augstākā (MK, 2020). Pētījumā analizēti 2022. gada CEVAL un CEOAL.

Eksāmeni un citi vērtēšanas darbi vienmēr ir bijuši visas izglītības sistēmas dzinulis. Tie mēra un ietekmē mācīšanas praksi un formas - ja eksāmenā ir šaura, uz rezultātu balstīta nevis uz reālo dzīvi, procesu un attīstību orientēta pieeja, tas atstāj ietekmi arī uz mācīšanu un mācīšanos (Cimen, 2010), tādejādi valsts pārbaudes darbu izpēte ir viens no instrumentiem, kas dod informāciju par jaunā satura ieviešanu (Pestovs, Saleniece, & Namsone, 2019). Ņemot vērā vidējās izglītības eksāmenu rezultātu augsto ietekmi uz skolēnu turpmāko dzīvi, jāpārlicinās par testu kvalitāti. To var raksturot novērtējot eksāmenu ticamību (ang. *validity*) un drošumu (ang. *reliability*). Informācija par testētajiem, kuri iespējams nemēra, to, kas plānots, ir noderīga gan eksāmenu veidotājiem, gan skolotājiem, gan visiem, kuri tieši vai netieši izmanto eksāmenu rezultātus lēmumu pieņemšanā (Latifi, Bulut, Gierl, Christie, & Jeeva, 2016).

Jaunā satura ieviešana paredz kompetences kā galveno mācību rezultātu (Oliņa, Namsone & France, 2018). Līdzšinējo vidusskolas centralizēto eksāmenu matemātikā šobrīd aizstāj trīs dažādi eksāmeni, kas ir atbilstoši izvēlētajam satura apguves līmenim. Būtiski ir noskaidrot, cik lielā mērā šie eksāmeni mēra augstākos izziņas darbības līmeņus un kādu sniegumu skolēni šajos uzdevumos demonstrē. Pētījumam izvirzīti jautājumi: (1) Ko rāda eksāmenu rezultāti, par eksāmenu ticamību, drošumu un piemērotību skolēnu kopai, kas to kārtu? (2) Kā tiek mērīta skolēnu izziņas darbība augstākajos līmeņos šajos eksāmenos un kāds ir skolēnu sniegums?

Literatūras apskats *Literature Review*

Centralizēto eksāmenu mērķis ir novērtēt skolēnu sniegumu matemātikā atbilstoši izvēlētajam satura apguves līmenim un iegūt datus skolēnu snieguma un mācību satura izvērtēšanai, metodisko ieteikumu izstrādei un profesionālās pilnveides plānošanai izglītības iestādes, dibinātāja un valsts līmenī (VISC, 2023). Pasaulē arvien lielāku lomu iegūst datu par skolēnu sniegumu izmantošana, izglītības sistēmas uzlabošanā (Gneezy et al., 2019, Shiel & Cartwright, 2015, Eacott & Holmes, 2010). To ļauj izdarīt prasmīga datu apstrāde un palūkošanās uz eksāmena darbiem no dažādiem rakursiem. Datu analīze ļauj izdarīt secinājumus par

konstruktu kopumā un identificēt problemātiskos testelementus, taču, lai izdarītu precīzākus secinājumus nepieciešama arī kvalitatīva rezultātu analīze. Šādu kvantitatīvu un kvalitatīvu novērojumu kopsavilkums ļauj noderīgi, jēgpilni un derīgi skaidrot datu būtību (Bond & Fox, 2013).

Vidusskolas noslēguma eksāmeņiem matemātikā ir liela ietekme uz skolēnu turpmāko dzīvi, tāpēc ir nepieciešams pārliecināties, ka iegūtie dati ir ticami un godīgi (Betels, 2003). Tas ir nozīmīgi (1) lai būtu pārliecība, ka skolēns tiešām ir sasniedzis paredzēto satura apguves līmeni, (2) lai varētu taisnīgi sarindot skolēnus pēc viņu snieguma eksāmenā un (3) eksāmens sniedz empīriskus pierādījumus, kas varētu veicināt izglītības mērķu, eksāmenu politikas un mācību priekšmetu satura kritisku izvērtēšanu un, plašākā nozīmē, varētu veicināt uzlabojumus (Latifi et al., 2016).

Pārbaudes darba ticamība raksturo, kādā mērā šis darbs ļauj novērtēt tieši tās skolēna zināšanas un prasmes, kura pārbaudei tas paredzēts (Betels, 2003). Lai tiktu mērītas standartā noteiktās prasības, eksāmena vērtēšanas saturu veido trīs kategorijas: (1) satura modulis, (2) sasniedzamo rezultātu grupas, (3) izziņas darbības līmenis (VISC, 2023). Viens no eksāmena drošuma faktoriem ir testelementu grūtības pakāpe, jo labāk testelementu grūtība atbilst skolēnu spēju līmenim, jo augstāks ir testa drošums (Betels, 2003).

Raita karte ir Raša analīzes vizuāls attēlojums, kurā personas (kartes kreisajā pusē) un testelementi (kartes labajā pusē) tiek attēloti uz vienas skalas atbilstoši to spēju un grūtības pakāpes novērtējumam (Bond & Fox, 2013). Raša analīze balstās apgalvojumu, ka: “personai ar lielākām spējām kā citai vajadzētu būt lielākai varbūtībai atrisināt jebkuru uzdevumu, un līdzīgi grūtāku uzdevumu kā citu jebkurai personai atrisināt ir mazāka varbūtība” (Rasch, 1960). Tiek aprēķināta varbūtība, personai ar tās spēju līmeni (konkrēta vērtība, balstoties uz kopējo punktu skaitu testā), pareizi atbildēt uz jautājumu (ar konkrētu grūtības pakāpi, kas atkarīga no pareizi atbildējušo personu skaita). Šī varbūtība ir atkarīga no personas spēju un testelementa grūtības pakāpes starpības (Bond & Fox, 2013). Svarīgākais ir princips, ka personas novietojums attiecībā pret testelementu attēlo, kāda ir varbūtība, ka šī persona pareizi atbildēs uz konkrēto jautājumu (Boone, 2016). Raita karte atspoguļo tikai saistību starp skolēnu sniegumu un testelementu grūtības pakāpi, kas ir tikai daļa no visas Raša analīzes, taču tā ļauj izdarīt secinājumus par eksāmenu piemērotību skolēnu kopai, kas tos kārtos (Bond & Fox, 2013).

Matemātikas mācību jomas plānotie skolēnam sasniedzamie mācību rezultāti paredz, ka skolēns izprot matemātiku kā zināšanu un prasmju sistēmisku kopumu, kas ļauj kvantitatīvi aprakstīt un izzināt apkārtējo pasauli, lietojot algoritmus, matemātisko modelēšanu un citus matemātikai raksturīgus paņēmienus dažādos kontekstos, spriež inductīvi un deduktīvi, veidojot risinājumus un skaidrojot savu

darbību un rezultātu (MK, 2019). Labs pārbaudes darbs mēra ne tikai skolēnu prasmes reproducēt zināšanas un prasmes, tajā ir jāiekļauj jautājumi kuros skolēns saprot un lieto informāciju jaunās situācijās un kontekstos. Darbā jāiekļauj arī uzdevumi, kuros skolēnam zināšanas un prasmes būtu jālieto jaunā situācijā, lai demonstrētu savu kompetenci (Cimen, 2010). Zināšanās balstītos uzdevumos vajadzētu iekļaut arī jautājumus *Kāpēc?* un *Kā?*, ne tikai atbildes uz jautājumiem *Kas? Kur? Kad?*, uzdevumos jāiekļauj konteksti, kas saistīti ar reālo dzīvi (Heyneman & Ransom, 1990).

Literatūrā tiek aprakstītas dažādas pieejas un definīcijas, ko saprot ar augstākā līmeņa izziņas darbību, piemēram, tās ir prasmes, kas veicina dziļākas, konceptuāli virzītas izpratnes veidošanu (Schraw & Robinson, 2011). Daudzi autori, veidojot augstāko izziņas darbības līmeņu definīciju, izmanto Blūma taksonomiju vai tās precizētu versiju (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956; Anderson & Krathwohl, 2001), kas piedāvā skolēnam sasniegumus rezultātus klasificēt sešos izziņas darbības līmeņos - zināšanas (atcerēšanas), saprašana, lietošana, analīze, sintēze un izvērtēšana. Aprakstot plašāk modeli kā vērtēt augstāka līmeņa izziņas darbību (Paul & Nosich, 1992) min 21 būtiskākos mērķus, lai novērtētu skolēnu izziņas darbību augstā līmenī. Veidojot uzdevumus mācību procesam un pārbaudes darbiem, ir būtiski novērtēt, cik lielā mērā uzdevums dod iespēju skolēnam domāt “dziļi”. Uzdevumus var klasificēt četros līmeņos: (1) iegaumēšana, (2) procedūru izpilde, neveicinot to saistību ar izpratni, (3) procedūru izpilde, veicinot to saistību ar izpratni, (4) matemātikas lietošana dažādās situācijās, kur trešais un ceturtais līmenis raksturo skolēnu iespēju izmantot augstākās izziņas darbības prasmes (Smith & Stein, 1998). Domājot par skolēna iespējamo sniegumu, Austrālijas pētnieki (Biggs & Collis, 1982) izveidojuši SOLO (ang. *structure of observed learning outcomes*) taksonomiju, ko var izmantot, lai analizētu skolēna demonstrētās izziņas darbības dziļumu (Namsone & Oliņa, 2018). Sākot ar 2021./22.mācību gadu centralizēto matemātikas eksāmenu programmas nosaka to, ka eksāmena vērtēšanas saturu raksturo trīs kategorijas: (1) sasniegamo rezultātu veids un grupa; (2) satura modulis; (3) izziņas darbības līmenis (VISC, 2023), kur izziņas darbības līmenis tiek noteikts izmantojot SOLO taksonomiju. Arī iepriekš dažādos valsts pārbaudes darbos tika piedāvāts skolēnu izziņas darbību aplūkot trīs līmeņos, kuros nepieciešama: (1) zema; (2) vidēja vai (3) augsta izziņas līmeņa darbība. Pētot dažādus valsts pārbaudes darbus matemātikā (France, Cakane, Namsone & Cirulis, 2017), ir vērojams, ka pārbaudes darbos ir neliels uzdevumu skaits, kuros skolēniem jādemonstrē augsts izziņas darbības līmenis.

Metodoloģija Research Methodology

Pētījumā izmantots paralēlais jauktais pētījuma dizains (1. att). Vienlaicīgi eksāmena testelementi un rezultāti tiek analizēti izmantojot kvantitatīvas un kvalitatīvas pētījuma metodes.



1.attēls. *Pētījuma dizaina shematiskais attēlojums (autoru veidots)*
Figure 1 *Schematic representation of the research design (created by authors)*

Pētījumā tika izmantoti 2021./2022. mācību gada dati par CEVAL un CEOAL eksāmenu kārtotajiem skolēniem. CEVAL kārtoja 860 skolēni, bet CEOAL kārtoja 4338 skolēni. Eksāmeni tika vērtēti centralizēti. Katra eksāmena otro daļu vērtēja divi vērtētāji. Skolēns saņēma vidējo vērtējumu.

Eksāmenu rezultāti analizēti izmantojot klasiskās testu teorijas elementus un Raša analīzes Raita kartes. Lai pārliecinātos par eksāmenu konstruktīvu validitāti, katrā eksāmenā tika aprēķināti Pīrsona korelācijas koeficienti starp katru no eksāmena testelementu kopām (Swank & Mullen, 2017), grupējot tos pēc (1) satura moduļa, (2) sasniedzamo rezultātu grupas, (3) izziņas darbības līmeņa. Lai pārliecinātos par eksāmenu drošumu un noturību eksāmenu daļām tika aprēķināti Kronbaha alfa koeficienti (Tavakol & Dennick, 2011). Tika izmantotas Raša analīzes Raita kartes, lai analizētu eksāmenus un to atbilstību skolēnu kopai, kas kārtot katru no eksāmeniem (Bond & Fox, 2011).

Ekspertu grupa, ko veido 3 izglītības eksperti, analizēja eksāmena testelementus izmantojot SOLO taksonomiju, kas izmantota aprakstot izziņas darbības līmeņus un CEVAL un OEVAL programmas (VISC, 2023). Eksāmenu izziņas darbības līmeņu apraksti: I Atceras, lieto faktus, īsas procedūras vai atsevišķas idejas., II Veic tipiskus algoritmus, lieto formulas, paņēmienus vai prasmes pazīstamās situācijās., III Saista, skaidro, lieto zināšanas vai prasmes jaunās situācijās, demonstrējot patiesu izpratni.,

IV Veido un pierāda vispārinājumus, lieto zināšanas un prasmes situācijās ar augstu kompleksuma pakāpi (VISC, 2023) Kvalitatīvai skolēnu snieguma analīzei testelementos ar augstu izzināšanas līmeni (trešo vai ceturto) atlasītās izlases kopas sastāv no 100 skolēnu darbiem (CEOAL) un 70 darbiem (CEVAL). Veicot šo atlasu tika nejauši izvēlēti skolēnu darbi, kuru sniegums nebija zemāks par 20%, tos sagrupējot: A grupa skolēni, kuru sniegums ir intervālā no 80% līdz 100%, B grupa - no 60% līdz 80% (neieskaitot), C grupa - no 40% līdz 60% (neieskaitot) un D grupa no 20% līdz 40% (neieskaitot).

Rezultāti

Results

Nosakot CEOAL ticamību pēc vērtēšanas satura kategorijām, aprēķinātie Pīrsona koeficienti ir lielāki kā 0,5, kas liecina par vidēji ciešu un ciešu korelāciju starp eksāmena testelementu grupām un eksāmena satura ticamību. CEVAL vairums aprēķināto Pīrsona koeficientu ir virs 0,5. Nebūtiski vājāka korelācija novērota IV izzināšanas līmeņa testelementiem ar atsevišķām satura grupām. Kopumā arī CEVAL saturs ir ticams.

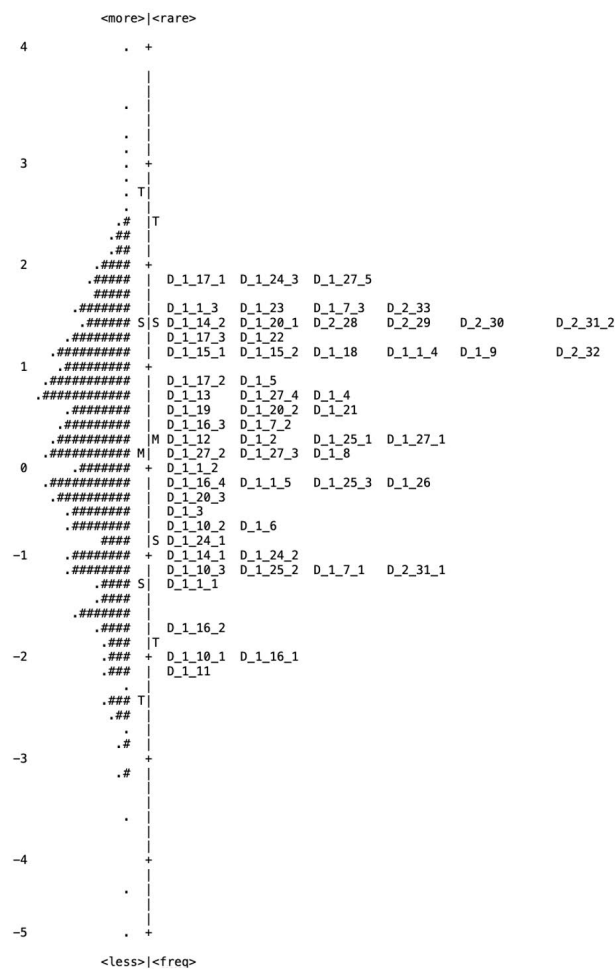
CEVAL un CEOAL tika aprēķināts Kronbaha alfa koeficients atsevišķām eksāmena daļām, lai pārlicinātos par eksāmenu noturību un drošumu. CEOAL Pirmās daļas Kronbaha alfa koeficients ir 0,95, otrajā daļā – 0,85. CEVAL Kronbaha alfa koeficients ir 0,90, otrajā daļā – 0,80. Visās daļās Kronbaha alfa ir lielāks kā 0,8, tāpēc abus eksāmenus var uzskatīt par drošiem (Ekulu & Quainoo, 2019, Krishnan, 2013).

Spriežot par eksāmenu ticamību, CEOAL Raita kartē (3. att.) novērojams, ka skolēnu snieguma un testelementu grūtības vidējās (M) un vienas standartnovirzes (S) vērtības ir ļoti tuvas. Būtu sagaidāms, testelementu izkārtojums kartē simetrisks skolēnu sadalījumam, taču redzams, ka tā nav. Ir daudz testelementu ar līdzīgu grūtības pakāpi. Raita kartē redzams, ka trūkst uzdevumu ar augstu izzināšanas līmeni, jo skolēni izkārtojas krietni virs testelementiem, tātad šiem skolēniem eksāmens nav bijis pietiekoši izaicinošs un eksāmens neļauj precīzi izmērīt šo skolēnu sniegumu.

CEVAL Raita kartē (2. att.) novērojams, ka testelementu grūtības vidējā vērtība (M) atrodas kartē krietni virs skolēnu snieguma vidējās vērtības (M). Tātad zem skolēnu snieguma vidējās vērtības ir mazāk testelementu, kā virs. Tas liecina, ka šis pārbaudes darbs šai skolēnu kopai bijis pārāk grūts. Starp skolēnu snieguma pirmo apakšējo (S) un otro apakšējo (T) standartnovirzi Raita kartē atbilstoši ir tikai 2 testelementi. Skatoties virs testelementu vidējās vērtības līdz pirmajai augšējai standartnovirzei (S) ir daudz testelementu, ar līdzīgu grūtības pakāpi. Lai tests kā mērisinstrumentu būtu precīzs, testelementiem būtu jāizkārtojas simetriski pret

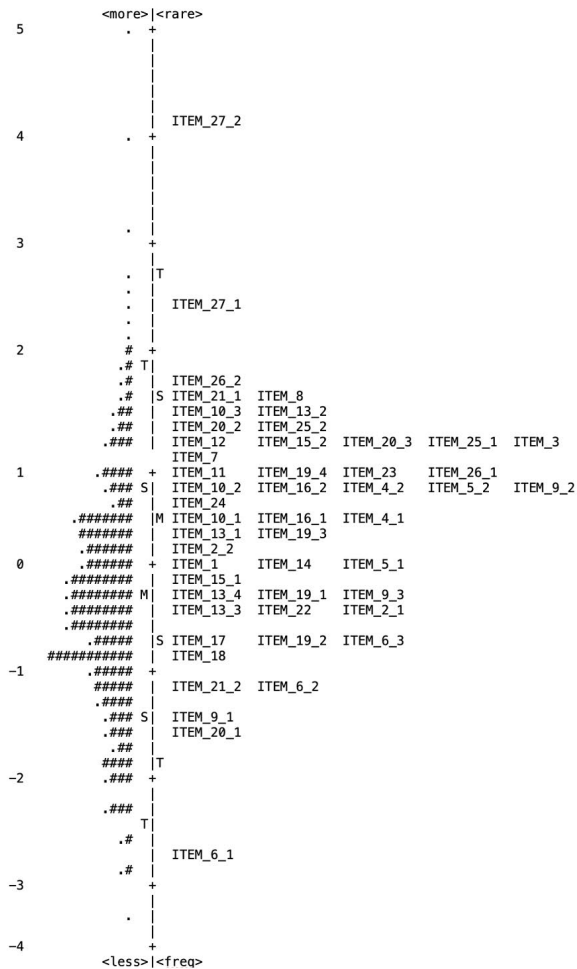
skolēniem, veidojot normālu sadalījumu, šajā gadījumā redzams, ka šis princips nav ievērots. Redzams, ka ir divi testelementi, kas ir bijuši skolēniem izaicinoši 27.1 un 27.2. Savukārt, trūkst testelementu starp 20.1. un 6.1., jo starp tiem veidojas atstarpe, kurai pretī ir liels skaits skolēnu.

CEOAL



apzīmē 18 skolēnus
 . apzīmē 1 līdz 17 skolēnus

CEVAL



apzīmē 6 skolēnus
 . apzīmē 1 līdz 5 skolēnus

2.attēls. CEOAL un CEVAL Raita karte (autoru veidots)
 Figure 2 Wright-map of CEOAL and CEVAL (created by authors)

CEOAL un CEVAL pirmo reizi tika atsevišķi vērtēta matemātiskās valodas lietojums un, kā organizēts un skaidrots risinājums. Ņemot vērā esošo darbu vērtēšanas kārtību, nav iespējams iegūt detalizētāku informāciju par to, kuros uzdevumos skolēniem ir labāk veicies ar valodas un risinājumu organizēšanu, kuru

simbolu vai pieraksta lietošana skolēniem nesagādā grūtības un, kuros gadījumos tā nav ierasta prakse.

1. tabulā redzams, ka, veicot uzdevumu analīzi un nosakot to izziņas darbības līmeni, tā atbilst plānotajam, atšķirība ir vērojama I un II līmeņa uzdevumu īpatsvarā CEVAL eksāmenā.

1. Tabula Uzdevumu izziņas darbības līmeņi un skolēnu sniegums (autoru veidota)
Table 1 Cognitive levels of items and students' performance (created by authors)

Izziņas darbības līmenis	Plānotais īpatsvars (%)	CEOAL		CEVAL	
		Īpatsvars eksāmenā, (%)	Skolēnu vidējais sniegums, (%)	Īpatsvars eksāmenā, (%)	Skolēnu vidējais sniegums, (%)
I	23 – 27	28	52	36	40
II	58 – 62	53	41	40	29
III	11 – 13	15	26	15	22
IV	2 – 4	4	23	4	6

Uzdevumi ar augstu izziņas darbības līmeni (III un IV līmenis) ir CEOAL eksāmenā 31.2, 33. (algebra), 29. (analītiskā ģeometrija), 32. (ģeometrija) un 30. (trigonometrija), nav iekļauti uzdevumi ar augstu izziņas darbības līmeni satura modulī varbūtības (t.sk.kombinatorika) un statistika. CEVAL eksāmenā uzdevumi ar augstu izziņas darbības līmeni ir 23., 26.2 (algebras modulis), 25.1., 27. (ģeometrijas modulis), 24. (varbūtības). Raita kartēs, redzams, ka testelementi ar augstu izziņas līmeni (III un IV) izkārtojas virs skolēnu augšējās standartnovirzes (S), izņemot CEOAL 32. uzdevumu un CEVAL 24. uzdevumu.

CEOAL 29. uzdev. skolēniem bija jāizmanto informāciju, kas dota tekstā un attēlā, jāpieņem lēmums par taisnes vienādojuma veidu, kuru izmantos, jāizmanto trijstūra laukuma aprēķināšana un trijstūru vienādības pierādīšana. 9% no skolēniem demonstrēja spriešanu III izziņas darbības līmenī. I līmenī savu sniegumu demonstrēja 17% skolēni, veicot kādu konkrētu spriedumu vai aprēķinus. Analizējot darbus, vērojams, ka skolēni pareizi atrisina uzdevumu, tomēr ir ļoti lakoniski un nepamato/neskaidro veiktās darbības, kļūdās arī aprēķinos.

30. uzdev. dota situācija ar reālu kontekstu. Risinājumā bija jāaskata sakarības vairākos trijstūros, jāizmanto sinusu teorēma, jāveic trigonometrisko vērtību aprēķināšana, izmantojot kalkulatoru un noapaļošana. 12% skolēnu demonstrēja spriešanu III izziņas darbības līmenī, savukārt I līmenī 14%, iesākot risinājumu un izdarot spriedumus, piemēram, par trijstūru leņķiem. Līdzīgi kā iepriekšējā uzdevumā skolēni mēdz būt ļoti lakoniski, piemēram, raksta uzreiz sinusa teorēmu,

nenorādot, kurš trijstūris tiek izmantots, nekorekti lieto vienādības zīmi. 15% skolēnu no izlases kopas nespēja tikt galā ar skaitļošanu, lietojot sinusu teorēmu. Grūtības sagādā starprezultātu iegūšana un noapaļošana, kaut arī skolēni lietoja kalkulatorus.

31.2. uzd. dota situācija ar reālu kontekstu. Risinājumā bija izvēle veidot algebrisku modeli - eksponentvienādojumu vai spriest, jāveic identiski pārveidojumi eksponentvienādojumā un ar logaritmiem, jāizmanto logaritmi saknes noteikšanai. 10% skolēnu demonstrēja spriešanu III izziņas darbības līmenī, savukārt, I līmenī sprieda 25% rakstot konkrētus piemērus un vērtības, savukārt, tikai 9% izveidoja vienādojumu, bet tālākus risinājuma soļus neveica - daļai skolēnu grūtības sagādāja darbošanās ar logaritmiem.

32.uzd. jāveic ģeometrisks pierādījums, spriežot par trijstūriem un telpas elementiem. 16% skolēnu demonstrēja spriešanu III izziņas darbības līmenī. Arī te daļa skolēnu spēja veikt atsevišķus spriedumus, grūtības sagādāja spriešanas gaitas korekta pierakstīšana.

33. uzd. dota situācija ar ģeometrisku kontekstu, kuras risinājumā jāizmanto algebriski modeļi. 5% skolēnu demonstrēja spriešanu IV izziņas darbības līmenī, savukārt, I līmenī sprieda 21% , apskatot zīmējumos dotās situācijas vai spriežot par sānu virsmu diagonāļu skaitu, neveicot tālākus vispārinājumus.

CEVAL 23.uzd. dota situācija ar reālu kontekstu, kura risināšanā jāizmanto procenti. Tālākajā risinājumā skolēni varēja veidot spriedumus, algebrisku vienādojumu vai nevienādību, izdarot spriedumus par rezultātu. 4% no skolēniem demonstrēja spriešanu III izziņas darbības līmenī. I līmenī savu sniegumu demonstrēja 55%, veicot skaitliskus aprēķinus, algebrisku modeli izveido tikai 7% skolēni.

24.uzd. varēja veikt sistemātisku pilno pārlassi vai izmantot kombinatorikas likumus izlašu skaita noteikšanai, bija jāaprēķina klasiskā varbūtība. 14% no skolēniem demonstrēja spriešanu III izziņas darbības līmenī. Tā kā konteksts - skaitļu veidošana, dod iespēju demonstrēt atsevišķus spriedumus, to veica 14% skolēnu, bet tālāk ar risinājumu netika.

25.1.uzd. dota situācija ar reālu kontekstu, tajā jāskaidro iespēja izmantot prizmas tipluma aprēķināšanu. 12% no skolēniem demonstrēja spriešanu III izziņas darbības līmenī, izvēloties sev piemērotāko risināšanas paņēmienu.

26.2. uzd. dota situācija ar reālu kontekstu, jāveido eksponentvienādojums. Uzdevuma pirmajā solī jau ir nepieciešams spriest vispārīgi, to veic 9% skolēnu, bet pilnībā spriešanu III līmenī veic tikai 5% skolēnu.

27.2. uzd. jāveic pierādīšana, ko III izziņas līmenī demonstrē tikai 1% skolēnu. Šāda veida uzdevumi visticamāk nav ierasta prakse šajā mācību kursā.

Skolēnu risinājumos kopumā vērojams, ka kļūdas ir īpašību un pārveidojumu lietojumā, parādot to, ka skolēniem nav izpratnes par to ko viņi dara, tiek veikti pavirši un neprecīzi spriedumi. Skolēnu daļa, kuri iesāk risināt uzdevumus ar augstu izziņas darbību, tajos bieži demonstrē tikai zemākā darbības līmeņa sniegumu.

Secinājumi un diskusija *Conclusions and discussion*

CEVAL un CEOAL konstruktīvi uzskatāmi par ticamiem un drošiem. CEOAL Raša analīzes Raita kartē redzams, ka, lai arī skolēnu snieguma un testelementu vidējās vērtības ir ļoti tuvas, testelementi neizkārtojas simetriski pret skolēniem. Lai eksāmenu varētu uzskatīt par piemērotu populācijai būtu jāsamazina testelementu skaits, kuriem ir līdzīga grūtības pakāpe un jāpievieno augstas izziņas līmeņa testelementi. Pētnieki (Planinic, Boone, Susac & Ivanjek, 2019) iesaka, lai uzlabotu eksāmenu piemērotību skolēnu kopai, kas tos kārtot, eksāmena izstrādes procesā, veidot uzdevumu banku, apbēt to, tad veikt Raša analīzi (ne tikai izmantojot Raita karti) un izmantot iegūtos datus, lai atlasītu piemērotākos uzdevumus.

CEVAL Raita kartē redzams, ka trūkst vieglu, zemas izziņas darbības testelementu, taču nepieciešama sīkāka izpēte, kāpēc skolēni ir demonstrējuši tik zemu sniegumu. Par zemu sniegumu liecina skolēnu skaits, kuri nerasniedz 10% robežu. Tie ir 9,9% no visiem kārtotajiem. 2022./23. mācību gadā skolēni ar šādu snief nebūs nokārtojuši gala pārbaudījumu, CEOAL 10% robežu nerasniedza 8,1 % skolēni. Lai eksāmenus nokārtotu, plānots paaugstināt snieguma “sliksni” līdz 15% 2023./24.mācību gadā, līdzīgi tas ir jau šobrīd Lietuvā (Jakaitiene, Želvys, Vaitekaitis, Raižiene & Dukynaite, 2021). CEVAL 15% robežu nerasniedz 19,4 % skolēnu. Tas liek spriest, ka skolēni vidējās izglītības saturu kopumā šajā programmā apgūst vāji.

Eksāmenos uzdevumu ar augstu izziņas darbības līmeni apjoms ir iekļauts plānotajā apjomā. Viens no pētījuma ierobežojumiem bija COVID-19 ietekme uz mācību procesu, novērtējot vai uzdevums ir jauna, vai jau zināma situācija skolēniem. Skolēnu vidējais sniegums šāda veida uzdevumos ir zems abos eksāmenos. Tomēr, gala spriedumus par visiem vidējās izglītības beidzēju rezultātiem varēs izdarīt tad, kad beigu eksāmenu kārtos arī skolēni, kuri izvēlējušies mācīties matemātiku augstākajā apguves līmenī. Šis ir īpašs signāls, lai veiktu padziļinātu izpēti par mācību procesu un skolēnu mācību motivāciju. Pētnieki (Eklöf and Nyroos, 2013) secinājuši, ka 30% no skolēniem nav motivēti veikt nacionāla līmeņa valsts pārbaudes darbus. Būtiski ir veikt pētījumus par mācību procesu stundās, lai meklētu iemeslus kopumā zemajam skolēnu sniegumam, un rastu risinājumus tā uzlabošanai. Būtu jāuzlabo datu ieguve par skolēnu prasmi lietot

matemātisko valodu un risinājuma organizēšanu un skaidrošanu, ko konkrēti skolēni prot, bet ko neprot. Pētot skolēnu darbus, vērojams, ka ir uzdevumi, kuros skolēni ne vienmēr skaidri saprot uzdevumu nosacījumus vai arī tie ir saprotami dažādi, kas rosina izvērtēt uzdevumu aprobācijas procesu. Lielais skolēnu skaits abos eksāmenos, kuri nesāk risināt eksāmenu 2.daļas uzdevumus, rosina meklēt iemeslus tam. Viens no iemesliem varētu būt eksāmenu apjoms un izpildes laiks. Risinājums tam būtu precīzāka eksāmenu sasniedzamo rezultātu atlase, samazinot tā apjomu un skolēnu anketēšana par eksāmena izpildi.

Kopsavilkums *Summary*

Latvia began curriculum reform implementation in 2020. The form of the national examination in mathematics at the end of secondary school changed from one exam for all students to three different exams depending on the chosen curriculum level. For the first time students took the exam in optimal (CEOAL) or general (CEVAL) level depending on their chosen curriculum. The aim of this study is to evaluate validity, reliability and appropriateness of the exams, to investigate how high-order thinking skills are measured and to assess student performance. To estimate the validity of the exams, Pearson correlation coefficients were calculated between the results of test item groups. The results show that the constructs of both exams are valid. Cronbach's alpha coefficients were calculated to test the reliability of the exams. They exceed the 0.8 threshold, indicated that the exams are reliable. The next step in the evaluation was the Wright map analysis. Results showed that there should be more items that measure high-order thinking skills in the CEOAL exam. The Wright map of this exam shows a very disbalanced distribution of items pointing to the necessity to improve the procedures for developing the exams. The Wright map of the CEVAL exam shows very low student performance. This may indicate a need for less challenging items or a focus on how the curriculum is taught in this program. Reasons for low performance in tasks requiring students to demonstrate High-Order Thinking Skills were investigated and explained. One finding was that students do not even start solving these items. Most students made inaccurate justifications. Results show that students have difficulty reasoning at the third and fourth cognitive levels. Further research is needed on the motivation of students and its impact on test results.

Pateicības Acknowledgements

Pētījums noticis ar projekta “LU doktorantūras kapacitātes stiprināšana jaunā doktorantūras modeļa ietvarā” nr. 8.2.2.0/20/I/006 atbalstu.

Literatūras saraksts References

- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition*. New York: Longman.
- Betels, Dž. (2003). *Rokasgrāmata pārbaudes darbu veidotājiem*. Rīga: Izglītības un zinātnes ministrija.
- Biggs, J., Collis, K. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.
- Bloom, B., Englehart, M. Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York, Toronto: Longmans, Green.
- Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Boone, W. J. (2016). Rasch analysis for instrument development: Why, when, and how? *CBE—Life Sciences Education*, 15(4), rm4.
- Cimen, E. E. (2010). How compatible are the 9th grade mathematics written exams with mathematical power assessment criteria. *Procedia-Social and Behavioral Sciences*, 2(2), 4462-4467.
- Eacott, S., & Holmes, K. (2010). Leading Reform in Mathematics Education: Solving a Complex Equation. *Mathematics Teacher Education and Development, Vol. 12.2*, 84–97.
- Eklöf, H., & Nyroos, M. (2013). Pupil perceptions of national tests in science: perceived importance, invested effort, and test anxiety. *European journal of psychology of education*, 28(2), 497-510
- Ekolu, S. O., & Quainoo, H. (2019). Reliability of assessments in engineering education using Cronbach's alpha, KR and split-half methods. *Global journal of engineering education*, 21(1), 24-29.
- France, I., Cakane, L., Namsone, D., & Cirulis, A. (2017). Cognitive Demand in Observed Lessons and National Testing Compared to PISA Mathematics Results in Latvia. *EDULEARN Proceedings*. DOI:10.21125/edulearn.2017.1102
- Gneezy, U., List, J.A., Livingston, J.A., Qin, X., Sadoff, S., & Xu, Y. (2019). Measuring success in education: the role of effort on the test itself. *American Economic Review: Insights*, 1(3), 291-308.
- Heyneman, S. P., & Ransom, A. W. (1990). Using examinations and testing to improve educational quality. *Educational Policy*, 4(3), 177-192.
- Jakaitiene, A., Želvys, R., Vaitekaitis, J., Raižiene, S., & Dukynaite, R. (2021). Centralised Mathematics Assessments of Lithuanian Secondary School Students: Population Analysis. *Informatics in Education, Vol. 20, No. 3*, 439–462. DOI: 10.15388/infedu.2021.18

- Krishnan V., (2013). *The Early Child Development Instrument (EDI): An item analysis using Classical Test Theory (CTT) on Alberta's data.*
- Latifi, S., Bulut, O., Gierl, M., Christie, T., & Jeeva, S. (2016). Differential performance on national exams: Evaluating item and bundle functioning methods using English, mathematics, and science assessments. *Sage Open*, 6(2). DOI: <https://doi.org/10.1177/2158244016653791>
- Ministru kabinets (2019). *Noteikumi par valsts vispārējās vidējās izglītības standartu un vispārējās vidējās izglītības programmu paraugiem*, Nr. 416. Pieejams: <https://likumi.lv/ta/id/309597-noteikumi-par-valsts-visparejas-videjas-izglitibas-standartu-un-visparejas-videjas-izglitibas-programmu-paraugiem>
- Ministru kabinets (2020). *Noteikumi par valsts profesionālās vidējās izglītības standartu un valsts arodizglītības standartu*, Nr. 332. Pieejams: <https://likumi.lv/ta/id/315146-noteikumi-par-valsts-profionalas-videjas-izglitibas-standartu-un-valsts-arodizglitibas-standartu>
- Namsone, D., & Oliņa, Z. (2018). Kā vērtē kompleksu sniegumu, No (sast. Namsone), *Mācīšanās lietpratībai /Learning for Proficiency/ (44-65)*. Rīga: LU Akadēmiskais apgāds. DOI: <https://doi.org/10.22364/ml.2018>
- Oliņa Z., Namsone, D., & France, I. (2018). Kompetence kā komplekss skolēna mācīšanās rezultāts. No (sast. Namsone), *Mācīšanās lietpratībai /Learning for Proficiency/ (18-43)*. Rīga: LU Akadēmiskais apgāds. DOI: <https://doi.org/10.22364/ml.2018.1>
- Paul, R., & Nosich, G. M. (1992). A Model for the National Assessment of Higher Order Thinking. In R. Paul (Ed.), *Critical Thinking: What Every Student Needs to Survive in a Rapidly Changing World* (pp. 78-123). Dillon Beach, CA: Foundation for Critical Thinking.
- Pestovs, P., Saleniece I., & Namsone, D. (2019). Science Large-Scale Assessment to the Revised Science Curriculum. *Proceedings of the 3rd International Baltic Symposium on Science and Technology Education, BalticSTE2019*, Šiauliai, Lithuania.
- Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, 15(2), 020111.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Paedagogiske Institut.
- Schraw, G., & Robinson, D. R. (2011). Conceptualizing and assessing higher order thinking skills. In G. Schraw & D. R. Robinson (Eds.), *Assessment of higher order thinking skills* (pp. 1–15). IAP Information Age Publishing.
- Shiel, G., & Cartwright, F. (2015). *National Assessments of Educational Achievement, Volume 4: Analyzing Data from a National Assessment of Educational Achievement*. World Bank Publications.
- Smith, M. S., & Stein, M. K. (1998). Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School*, 3(5), 344–350.
- Swank, J. M., & Mullen, P. R. (2017). Evaluating evidence for conceptually related constructs using bivariate correlations. *Measurement and Evaluation in Counseling and Development*, 50(4), 270-274.
- Tavakol, M. & Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011 Jun 27;2:53-55. DOI: 10.5116/ijme.4dfb.8dfd. PMID: 28029643; PMCID: PMC4205511.
- VISC. (2023). *Valsts pārbaudes darbu programmas*. Pieejams: <https://www.visc.gov.lv/lv/valsts-parbaudes-darbu-programmas>