

Matematikos gebėjimų testo diagnostinės kokybės charakteristikos

Arkadijus Kiseliovas

Docentas socialinių mokslų
(edukologijos) daktaras
Šiaulių universiteto
Matematikos didaktikos katedra
P. Višinskio g. 25, LT-5400 Šiauliai
Tel. (841) 59 57 15

Danutė Kiseliova

Socialinių mokslų (edukologijos) daktarė
Šiaulių universiteto
Matematikos didaktikos katedra
P. Višinskio g. 25, LT-5400 Šiauliai
Tel. (841) 59 57 15

Straipsnyje analizuojamos ketvirtų klasių moksleivių baigiamojo matematikos gebėjimų testo skaičiavimo užduočių diagnostinės kokybės skaitinės charakteristikos: sunkumo indeksas, standartinis nuokrypis ir t. t. Testo užduočių vidaus struktūra papildomai nagrinėta alfa faktorinės analizės metodu. Pateikiami skaičiavimo užduočių interkoreliacijos ir faktorinės koreliacijos koeficientai bei matematinio faktorinės analizės modelio parametrai.

Tyrimo imtis 1141 ketvirtokas iš įvairių Lietuvos miestų ir rajonų mokyklų.

Viena iš problemų, kurias pedagogika ir psichologija sprendžia kartu, yra mokymosi tikslų, o kartu ir mokymosi rezultatų kognityvioje srityje apibrėžtis. Mokymosi pasiekimų samprata priklauso nuo psichodidaktinės koncepcijos, kuria vadovaujasi vienos ar kitos mokyklos pedagogai, o matavimo procedūros yra derinamos su konkrečios teorijos idėjomis.

Pasirinkta tyrimo tema yra problemiška ir moksliniu teoriniu, ir praktiniu taikomuoju požiūriu.

Pirmąjį problemiško aspekto lemia tai, kad buvusioje SSRS testavimas ilgus dešimtmecius ideologiniais sumetimais buvo uždraustas visai arba buvo esmingai apribotas (M. Depaepe, 1993; G. Merkys, 1997). Šioje srityje atsilikimas iš inercijos vis dar išlieka. Tokiu kontekstu paminėtinas TIMSS (*Third International Mathematics and Science Study*) trečia-

sis projektas, kurį parengė IEA (*International Association for the Evaluation of the Educational Achievement*) organizacija. Tarp 50 valstybių Lietuva (D. Kiseliova, A. Kiseliovas, K. Traškevičius, 1994; Д. Киселёва, А. Киселёв, 1994; V. Čekanavičius, G. Trakas, A. Zabulionis, 1997) pasirodė blogiau, nei buvo tikėtasi, blogiau, nei tikėjosi viešojo nuomonė. Klausimų ir nuomonių, kodėl taip atsitiko, yra įvairių, tačiau TIMSS rezultatai šalyje taip ir netapo nuolatinių ir ilgalaikių mokslinių diskusijų dalimi. Tuo tarpu akivaizdu, kad matematikos ir gamtos dalykų mokymosi pasiekimų tyrimai yra prioritetingi edukacinių mokslų sritis.

Antrąjį (praktinį taikomąjį) mūsų tyrimo problemiško aspekto lemia šalies pedagogų ir švietimo vadybininkų bendruomenės siekis diegti Lietuvoje visų pakopų edukacinį monitoringą, standartizuotą testavimą ir egzaminavi-

ma, atliekama nepriklausomos institucijos – Valstybinio egzaminų centro. Nepaisant pozityvių ir inovacinių siekių, galima konstatuoti, kad realūs žingsniai kol kas yra žengti tik abiturų egzaminų ir imatrikuliacijos srityje. Visa likusi švietimo sistema, ypač žemesniosios jos grandys (pvz., pradinė mokykla), standartizuoto ir nepriklausomo testavimo dar nepatyrė.

Tyrimo objektas – Lietuvos pradinių klasių moksleivių matematikos mokymosi gebėjimai kaip diagnostinis konstruktas.

Tyrimo tikslas – parengti prielaidas (teorinio, metodologinio ir empirinio tyrimo požiūriu) pereiti prie pradinukų matematikos gebėjimų standartizuotos diagnostikos, panaudojant psichometriškai patikimą testą. Taip pat sudaryti darbinį tokio testo variantą, kurį būtų galima taikyti pradinio ugdymo pakopos pabaigoje, t. y. ketvirtoje klasėje.

Tyrimo uždaviniai – atlikti psichometrinę uždavinių atranką, pateikiant tyrimo vartotojui testo uždavinių pavyzdžius ir psichometrines charakteristikas (sunkumo koeficientą, skiriamąją gebą, kai kurias RASCH modelio charakteristikas ir t. t.).

Testo (D. Kiseliuva, 1999) turinio struktūra (pagal bendrojo išsilavinimo standartus (1998) bei Lietuvos bendrojo lavinimo mokyklos bendrąsias programas (1997)) ir struktūrinių dalių proporcijos nusakytos 1 lentelėje.

Matematikos pasiekimų matavimo pagrindu imame tris matematinės veiklos sritis: žinių lygmeniu – supratimą, įgūdžių lygmeniu – algo-

ritmų taikymą, mąstymo lygmeniu – uždavinių sprendimą. Sudarydami matematinės veiklos ir turinio matricą, rėmėmės pažintinės veiklos įgūdžių taksonomija (N. Cole, 1990). Autorių sudarytoje matematikos pasiekimų matavimo metodologinėje schemoje nurodytos mokinių veiklos rūšys, priskirtinos kiekvienai sričiai (S. Balčiūnas, A. Kiseliuva, D. Kiseliuva, 1998).

Siekiant užtikrinti testavimo procedūros situacinį validumą, pasirinktas vidinis tyrimo administravimo būdas. Moksleivius testavo juos mokantys mokytojai, naudodamiesi standartinė testavimo instrukcija. Kadangi moksleivių matematikos pasiekimai kaip psichometrinis konstruktas yra labilus (moksleivių matematikos žinios ir įgūdžiai nuolat kinta, kaitą lemia mokymo procesas ir užmiršimo veiksniai), testo duomenys buvo surinkti per ne ilgesnį kaip 20 dienų laikotarpį. Testą sudaro 14 užduočių blokų (iš viso 40 uždavinių; D. Kiseliuva, 1999), jo trukmė 45 minutės.

Tyrimo rezultatų patikimumui ir tikslumui didelę įtaką turi imties parinkimas. Geografiniu požiūriu miestų kategorijai priskirtos didžiųjų Lietuvos miestų (Vilniaus, Kauno, Klaipėdos, Panevėžio, Šiaulių) mokyklos ir rajonų centrų mokyklos, miestelių kategorijai – kitų miestų ir miestelių mokyklos, kaimo kategorijai – kaimo ir kaimo tipo gyvenviečių mokyklos (2 lentelė).

Duomenų kodavimo ir perkėlimo procedūrų patikimumui įvertinti pakartotinai koduota ir suvesta 5 proc. atsitiktinai atrinktų testo lapų. Klaidų skaičius neviršijo 0,3 proc.

1 lentelė. Pradinukų matematikos pasiekimų testo turinio ir veiklos struktūra %

Turinio kategorijos	Veiklos kategorijos			
	Supratimas	Algoritmų taikymas	Problemų sprendimas	Užduočių skaičius
I. Matematikos taikymas	9,26	0,00	5,56	12,50
II. Skaičiavimai	11,11	31,48	5,56	50,00
III. Geometrijos pradžmenys	9,26	7,41	1,85	12,50
IV. Matavimai	0,00	18,52	0,00	25,00
V. Statistikos pradžmenys	0,00	0,00	0,00	0,00
Užduočių skaičius	29,63	57,41	12,96	100,00

2 lentelė. Tyrimo dalyvių imties charakteristikos

Imties tūris	Klasių skaičius	Lytis		Mokyklos tipas			Vietovė		
		bernaiukai	mergaitės	vidurinė	pagrindinė	pradinė	miestas	miestelis	kaimas
1141	64	49.8 %	50.2 %	67.5 %	7.5 %	25 %	70.4 %	4.4 %	25.2 %

Šioje darbo dalyje pateikiame užduočių, sudarančių šį matematikos pasiekimų testą, diagnostinės kokybės skaitines charakteristikas: sunkumo indeksą, standartinį nuokrypį, diskriminacijos indeksą, koreliacijos su testu koeficientą, ryšio su hipotetiniais matematikos pasiekimais įvertį r_{it} . Šia informacija galima remtis atrenkant užduotis kitam testui. Alfa faktorinės analizės metodu tirdami testą sudarančių užduočių visumą, išskyrėme labiausiai susijusias jų grupes. Šių grupių vidaus struktūrą papildomai nagrinėjome alfa faktorinės analizės metodu. Užduotis, kurių faktorinis svoris pagrindiniame faktoriuje buvo labai mažas, atmetėme. Pateikiame užduočių interkoreliacijos (dviejų daliinių skalių koreliacijos (B. Bitinas, 1998)) ir koreliacijos su faktoriumi koeficientus, matematinio faktorinės analizės modelio parametrus: Kaiser-Meyer-Olkin /KMO/ koeficientą, įvertinančių interkoreliacinės matricos tinkamumą faktorinei analizei ir faktoriaus aprašomosios galios įvertį, išreikštą faktoriui tenkančios dispersijos procentais. Šios analizės pagrindu, jei tai yra įmanoma, sudarėme įvairios prigimties skales (B. Bitinas, 1974).

S skalė. Tai klasikinėje testų teorijoje plačiausiai naudojama skalė. Klasikinės testų teorijos požiūriu postuluojuama, kad kiekviena testo užduotis teikia lygiavertę informaciją apie tiriamą savybę (moksleivio matematikos pasiekimus), todėl visų užduočių svoriai yra lygūs vienetui. Jei užduočių yra pakankamai daug, o testas validus, išspręstų uždavinių skaičius parodo moksleivio matematikos pasiekimų lygį. Skalės įverčiai yra nuo nulio iki skaičiaus, lygaus uždavinių (užduočių) skaičiui. Atlikę tiesinę transformaciją, šią skalę pertvarkome į standar-

tinę Z skalę, kurios vidurkis yra 0, o standartinis nuokrypis lygus 1. Moksleivio matematikos pasiekimus matuoti S skale yra nesudėtinga, nes testo vartotojui (mokytojui) reikia tik suskaičiuoti, kiek užduočių atlikta teisingai.

B skalė. Ši skalė yra konstruojama remiantis uždavinio išspręstumo įvertinimu polinėse grupėse. Tos grupės sudaromos iš tiriamųjų, kurių vertinimai yra žemesni nei 27 procentiliai ir aukštesni nei 73 procentiliai. Taikoma iteracinė procedūra. Pirmiausia grupės sudaromos pagal pirminius testavimo rezultatus (S skalė), skaičiuojami šių grupių centrus jungiančios tiesės krypties koeficientai. Teigiama, kad šie koeficientai pirminiu etapu išreiškia užduočių pradinę informacinę vertę, sudaroma naujaskalė ir tikslinama polinių grupių sudėtis. Procedūra atliekama tol, kol polių grupių sudėtis nekinta. Galutiniai rezultatai išreiškia uždavinių informacinę vertę (svorį). Skalės **B1** įverčiai normuoti nuo nulio iki dydžio, lygaus kvadratinei šakniai iš uždavinių (užduočių) skaičiaus, skalės **B2** – nuo 0 iki 1.

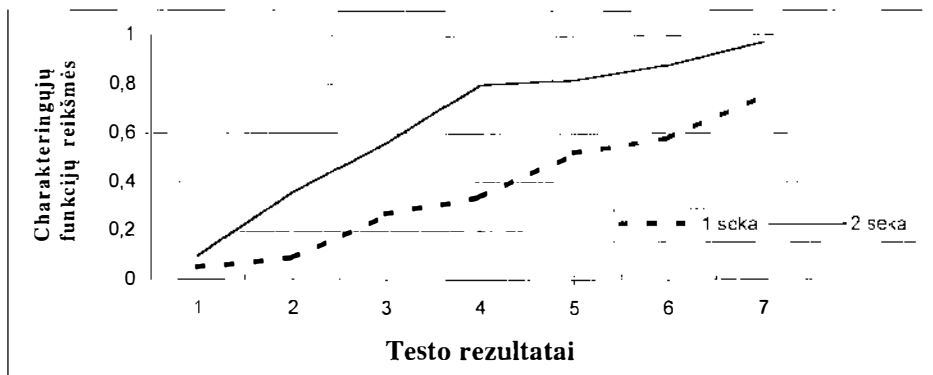
R skalė. Ši skalė (Rašo–Birnbaumo skalė) konstruojama remiantis logistinės funkcijos matematinio modeliu. Skalė normuota taip, kad imties vidurkis būtų lygus nuliui, o standartinis nuokrypis – vienetui.

Matematikos pradinio mokymo sistemoje mokėjimas skaičiuoti yra matematinio išsilavinimo pagrindas, todėl ugdymo turinyje ir testo specifikacijoje šiems operacijoms skiriamas itin didelis dėmesys. Paanalizuosime šią testo dalį detaliau (3–5 lentelės, 1 pav.).

Analizuojamos sekos yra matematinio tyrimo užduotys. Pagal išspręstumą pirmoji priskirtina minimaliam, antroji – pagrindiniam

3 lentelė. Skaičių sekos. Užduočių parametrai

I Prateškite skaičių sekas							
400, 360, 320, 280, (1 seka)							
1000, 100, 900, 200, 800, (1 seka)							
Užduotis	Imties tūris	Užduoties sunkumas	Standartinis nuokrypis	Diskriminacija	Koreliacija su testu	Koreliacijos koeficiento su kitomis testo užduotimis vidurkis	r_n
1 seka	1147	0,82	0,38	0,40	0,45	0,19	0,24
2 seka	1147	0,54	0,50	0,50	0,39	0,16	0,40



1 pav. Sekų sudarymo užduočių empirinių charakteringųjų funkcijų grafikai

4 lentelė. Skaičių sekos užduočių vidinis suderinamumas

		1 seka	2 seka	1 faktorius
Koreliacijos koeficientai	1 seka	1,000	0,292	0,540
	2 seka	0,292	1,000	0,540

KMO koeficientas 0,50. 1 faktorius apima 29 proc. sklaidos. Koreliacijos koeficientų statistinio reikšmingumo lygmuo $p < 0,001$.

pasiekimų lygmeniui. Skiriamosios gebos rodikliai nėra vienareikšmiai: nors antrosios užduoties diskriminacija yra didesnė, tačiau koreliacija su testu mažesnė negu pirmosios. Empirinės charakteringųjų funkcijų kreivės savo forma atitinka IRT (moderniosios testų teorijos) modelį (1 pav.).

Pagal išspręstumą pirmieji trys uždaviniai priskirtini minimaliam, likusieji – baziniam pasiekimų lygmeniui. Antro reiškinio antro skaičiaus rašymo užduoties diskriminatyvumas labai geras, kitų – patenkinamas arba geras. Šios dvi užduočių grupės gali būti pavyzdys, kaip

5 lentelė. Skaičių sekos užduočių svoriai įvairiose skalėse ir skalių patikimumo įverčiai

	1 seka		2 seka	
	Užduotis neišspręsta	Užduotis išspręsta	Užduotis neišspręsta	Užduotis išspręsta
S	1,000	1,000	1,000	1,000
B1	0,1640	0,7650	0,1640	0,6008
B2	0,0968	0,4517	0,0968	0,3547
R	0,2250	0,2180	0,2998	0,2572

G1 & G2 interkoreliacija = 0,692.
Spearmano–Brauno pataisa = 0,818.
Gutmano patikimumo koeficientas = 0,424.

sprendimo rezultatas priklauso nuo uždavinio formos: panaudojus nelygybės ženklą ir tu labiau formalizavus sąlygą, uždavinio išspręstumas mažėja. Pirmame ir šeštame reiškinuose yra tie patys skaičiai, tačiau pastarųjų sunkumo indeksas 0,1 mažesnis. Alfa faktorinės analizės būdu išskirti du faktoriai. Pirmasis faktorius gali būti interpretuotas kaip išreiškiantis gebė

6 lentelė. Gretimi skaičiai. Užduočių parametrai

2. Užrašykite duotiems skaičiams gretimus kaimynus:

1) 999 2) 800 3) 101

4) 540 5) 111 6) 999

Užduotis	Imties tūris	Sunkumas	Standartinis nuokrypis	Diskriminacija	Koreliacija su testu	Koreliacijos koeficiento su kitomis testo užduotimis vidurkis	r_{rr}
1 reiškinys	1147	0.90	0.30	0.24	0.41	0.21	0.46
2 reiškinys	1147	0.88	0.32	0.31	0.49	0.25	0.50
3 reiškinys	1147	0.92	0.28	0.22	0.41	0.21	0.46
4 reiškinys	1147	0.72	0.46	0.46	0.44	0.22	0.46
5 reiškinys	1147	0.78	0.42	0.38	0.42	0.21	0.45
6 reiškinys	1147	0.81	0.40	0.35	0.41	0.20	0.45

7 lentelė. Gretimų skaičių rašymo užduočių vidinis suderinamumas

		1 reiškinys	2 reiškinys	3 reiškinys	4 reiškinys	5 reiškinys	6 reiškinys	1 faktorius	2 faktorius
Koreliacijos koeficientai	1 reiškinys	1,000	0.699	0.657	0.332	0.306	0.390	0.695	-0.411
	2 reiškinys	0.699	1,000	0.710	0.409	0.379	0.396	0.764	-0.410
	3 reiškinys	0.657	0.710	1,000	0.371	0.353	0.401	0.722	-0.383
	4 reiškinys	0.332	0.409	0.371	1,000	0.741	0.664	0.710	0.363
	5 reiškinys	0.306	0.379	0.353	0.741	1,000	0.774	0.771	0.535
	6 reiškinys	0.390	0.396	0.401	0.664	0.774	1,000	0.748	0.366

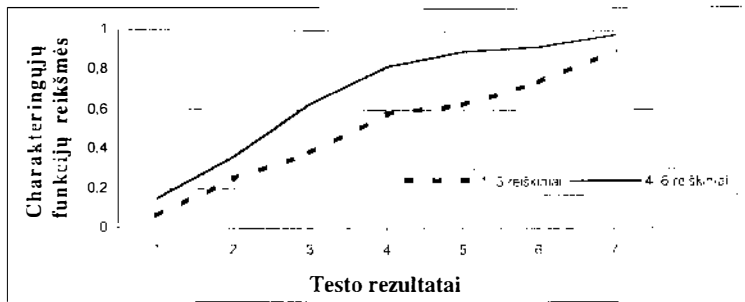
Koreliacijos koeficientų statistinio reikšmingumo lygmuo $p < 0,001$. KMO koeficientas = 0,80.

1 faktorius apima 54 proc. sklaidos, 2 faktorius apima 17 proc. sklaidos. Faktorinio modelio aprašomoji galia 71 proc.

8 lentelė. Natūraliųjų skaičių numeracija. Gretimi skaičiai. Užduočių svoriai įvairiose skalėse ir skalių patikimumo įverčiai

	1 reiškinys		2 reiškinys		3 reiškinys	
	Užduotis neišspręsta	Užduotis išspręsta	Užduotis neišspręsta	Užduotis išspręsta	Užduotis neišspręsta	Užduotis išspręsta
S	1.000	1.000	1.000	1.000	1.000	1.000
B1	1.1052	0.2713	0.2427	0.3380	0.1419	0.2914
B2	0.0318	0.0819	0.0733	0.1020	0.0428	0.0880
R	0.0740	0.0930	0.0766	0.0901	0.0686	0.0827
	4 reiškinys		5 reiškinys		6 reiškinys	
	Užduotis neišspręsta	Užduotis išspręsta	Užduotis neišspręsta	Užduotis išspręsta	Užduotis neišspręsta	Užduotis išspręsta
S	1.000	1.000	1.000	1.000	1.000	1.000
B1	0.3820	0.3663	0.2474	0.3384	0.2289	0.3589
B2	0.1153	0.1106	0.0747	0.1022	0.0691	0.1084
R	0.0859	0.0897	0.0845	0.0864	0.0843	0.0841

G1 & G2 interkoreliacija = 0,908. Spearmano–Brauno pataisa = 0,952. Gutmano patikimumo koeficientas = 0,892. Kronbacho alfa koeficientas = 0,847.



2 pav. Gretimų skaičių rašymo užduočių empirinių charakteringųjų funkcijų grafikai

imą rašyti gretimus skaičius, antrasis – kaip nelygybės sąryšio samprata. Remdamiesi tuo, kad pirmasis faktorius paaiškina didžiąją dalį sklaidos, o visų užduočių faktoriniai svoriai beveik vienodi, visas užduotis sujungiame viena skale. Šios skalės parametrai yra geri. 2 paveiksle yra pavaizduotos pirmos ir antros užduočių eilučių išspręstumą išreiškiančios empirinės charakteringųjų funkcijų kreivės. Jos atitinka IRT modeliui keliamus reikalavimus.

Panašias diagnostinės kokybės charakteristikas turi ir kitos testo užduotys.

Išvados

Pateiktų užduočių analizė parodo gerą mokinių gebėjimą skaityti ir rašyti skaičius iki mili-

jono, įvardyti ir pervardyti kiekius, panaudoti skaičių dešimtainę išraišką jų palyginimui, nuskaityti skaičiavimo, grupavimo ir dešimtainės struktūros ryšius. Užduočių sunkumo ir diskriminacijos įverčiai klasikinės testų teorijos požiūriu yra geri.

Užduočių empirinės charakteringosios kreivės atitinka IRT modelio reikalavimus. Matematinio faktoriinės analizės modelio parametrai ir skalės patikimumo įverčiai patenkinami.

Analizuojamos užduotys atitinka Lietuvos bendrojo lavinimo mokyklos bendrųjų programų bei bendrojo išsilavinimo standartų reikalavimus ir gali būti panaudotos baigiamajam ketvirtos klasės matematikos testo konstravimui.

LITERATŪRA

1. Balčiūnas S., Kiseliovas A., Kiseliova D. Ketvirtos klasės moksleivių matematinių pasiekimų preliminari diagnostika. Lietuvos matematikų draugijos mokslo darbai. Specialus „Lietuvos matematikos rinkinio“ priedas. II tomas. Vilnius, 1998.
2. Bendrojo išsilavinimo standartai, I–X klasės. Projektas. 2 dalis. Vilnius: Lietuvos Respublikos švietimo ir mokslo ministerijos Leidybos centras, 1998.
3. Bitinas B. Statistiniai metodai pedagogikoje ir psichologijoje. Kaunas: Šviesa, 1974.
4. Bitinas B. Ugdymo tyrimų metodologija. Vilnius: Jošara, 1998.
5. Cole N. Conception of educational achievement // Educational researcher. 1990, No 3.
6. Depaepe M. Zum Whol des Kindes? Pädologie, Pädagogische Psychologie und Experimentall Pädagogik in Europa und USA. 1890–1940. Weinheim, 1993.
7. Kiseliova D. IV klasės mokinių matematikos pasiekimų kitimo/raidos testai // Žvirblių takas. 1999, Nr. 1, p. 39–47.
8. Kiseliova D., Kiseliovas A., Traškevičius K. TIMSS programos Respublikos pradinėse klasėse preliminariai duomenys, 35 LMD konferencijos tezės. Vilnius, 1994.

9. Lietuvos bendrojo lavinimo mokyklos bendrosios programos. Vilnius: Lietuvos Respublikos švietimo ir mokslo ministerijos Leidybos centras, 1997.

10. Merkys G. Dalykų didaktikos ir edukacinis autoritarizmas // Mokykla. 1997, Nr. 12.

11. Trečioji tarptautinė matematikos ir gamtos mokslų studija. 7–8 kl. moksleivių tyrimo statistinė ataskai-

ta / Parengė: V. Čekanavičius, G. Trakas, A. Zabulionis. Vilnius, 1997.

12. Киселёва Д., Киселёв А. Классификация и анализ тестов и заданий международной программы ТИМСС по математике III–IV классов. Республиканская конференция „Проблемы профессиональной ориентации молодежи“. Минск, 1994.

COMPOSING A TEST FOR THE FOURTH FORM PUPILS. FEATURES OF THE TEST DIAGNOSTIC QUALITY

Arkadijus Kiseliovas, Danutė Kiseliova

Summary

The investigation is a first attempt to define the concept of the primary school achievements in mathematics, to formulate a methodological frame, to compose and ground the technology of the task selection, to foresee the regulations and conditions of the test administration and to choose statistical methods for the test results analysis.

The following diagnostic quality statistical features of the tasks composing the test for the achievements in mathematics meant for the fourth form pupils have been presented: the index of complexity, the standard deviations, the discrimination index, the correlation coefficient with the test, and the estimation of the relationship with hypothetical achieve-

ments in mathematics. The most interrelated test task groups have been distinguished by Alfa factor analysis method. In addition the inner structure of these groups has been analyzed by Alfa factor analysis method. The tables present the coefficients of the tasks with the factor, the parameters of the mathematical factor analysis model: Kaiser–Meyer–Olkin /KMO/ coefficient, estimating the usability of the intercorrelation matrix in factor (al) analysis and the estimation of the factor descriptive power, expressed by the dispersion per cent received by the factor.

The information obtained may serve as basis in selecting tasks for another test.

Gauta 2001 12 06

Priimta 2002 03 12