

TIKIMYBINĖ TESTŲ TEORIJA SKAIČIAVIMŲ GEBĖJIMŲ DIAGNOSTIKOJE VIRTUALIŲ MOKYMO OBJEKTŲ TAIKYMO EFEKTYVUMUI TIRTI

Orinta Šalkuvienė, Danutė Kiseliova, Arkadijus Kiseliovas,
Šiaulių universitetas, Edukologijos fakultetas

Įvadas

1950 metais klasikinės testų teorijos (CTT) pagrindu pasirodė pirmieji moksleivių gebėjimų tyrimo testai. Auganti psichometrinių matavimų reikmė skatino vis tobulesnių analitinių įrankių, palyginus su tradicine testų teorija (CTT) paklausą. Paskutinį dešimtmetį žymiai suaktyvėjo IRT (*Item Response Theory*) testų teorijos taikymas diagnostinių bei kitų tyrimų matavimams dėl patogių skalių ir užduočių analizei. IRT yra modeliais pagrįstas matavimas, kuriuo įvertinamas respondento (matematinio gebėjimų, kritinio mąstymo, fizinio funkcionavimo ir pan. lygis (Embretson and Reise, 2000).

Pažymėtina, kad Lietuvoje originalūs IRT modeliai nekuriami, o esamus modelius taiko tik pavieniai tyrinėtojai (Kiseliova, 2002; Kiseliova, Kiseliovas, 2004). Šiame kontekste paminėtinas prof. B. Bitinas, kuris IRT modelius pradėjo taikyti vos jiems atsiradus, apie 1960–1970 m., o vėliau kartu su programuotoju R. Paulavičiumi sukūrė programinę įrangą „PAULA“ edukacinės diagnostikos uždaviniams spręsti (Битинас, Паулавичюс, 1987): pavienės programinės funkcijos atitinka IRT nuostatas.

Apskritai daugiau nei 95% testų, skelbiamų komerciniuose testų leidyklų kataloguose, yra parengti CTT bazėje (Brickenkamp, 1997; Testzentrale, 1998). Mokslo organizaciniu požiūriu, ryškus IRT trūkumas – reikalinga tarpdisciplininė socialinių tyrinėtojų, matematikų ir programuotojų santalka. Pastarąją įgyvendinti įprasta specialistų telkimo tvarka, grindžiama tradicine mokslo klasifikacija, yra sunku. Testų konstravimas pagal IRT yra žymiai sudėtingesnis ir brangesnis. Norint pasiekti reikiamą modelio grynumą, būtinos didelės imtys ir t. t. Savo ruožtu praktinis IRT testų taikymas reikalauja gilesnio psichometrinio ir matematinio pasiruošimo. Vis dėlto akivaizdu, jog, palyginus su CTT, kur operuojama „žaliais“ testo balais ir iš esmės dirbama „apgrai-bomis“, IRT leidžia pakilti į analitinį matematinio modeliavimo lygį, konstruoti metrologiškai pilnavertes ir psichometrijos požiūriu labai skaidriai interpretuojamas skales. Anksčiau CTT ir IRT dažniausiai buvo supriešinamos, o dabar socialinių tyrinėtojų stovykloje šiuo požiūriu išivyrėja konvergencinės nuostatos: dirbama pagal CTT, tačiau IRT modeliai pasitelkiami optimaliai testo žingsnių atrankai. Tokia pozicija pagrįsta. Iš tyrimų praktikos žinoma, jeigu logistinis modelis galioja, tai koreliacija tarp „žalių“

testo balų ir asmens parametrų, nuskaitytų iš funkcijos, svyruoja nuo 0,90 iki 0,95 (Rost, 1995). Tokiu atveju testo balų metrologinė kokybė garantuota.

Straipsnio objektas. Homogeninių klasių moksleivių, dalyvavusių pedagoginiame eksperimente virtualių mokymo objektų efektyvumui tirti mokant aritmetikos veiksmų, skaičiavimo gebėjimų prieš eksperimentą diagnostika.

Tikslas. Diagnostiniais aritmetinių skaičiavimų testais prieš pedagoginį eksperimentą apibūdinti tyrimui atrinktų homogeninių klasių moksleivių skaičiavimų gebėjimus, taikant tikimybinę testų teoriją.

Uždaviniai: Apžvelgti *tikimybinės testų teorijos* (IRT) sampratą; remiantis konkrečiais pavyzdžiais, paaiškinti testo kokybės statistinius požymius, taikant tikimybinę testų teoriją.

Metodai: dalykinės literatūros bendrųjų reikalavimų diagnostiniams matematikos pasiekimų testams nagrinėti, taikant *tikimybinę testų teoriją* analizė; diagnostinio testo duomenų sisteminimas, analizė ir interpretacija.

Tyrimo organizavimas

Tyrimas vykdytas 2008 m. sausio mėnesį. Pradiniame etape testuoti 839 ketvirtokai. Remiantis testo rezultatais, išskirta 16 homogeniškų klasių. Tolesniam tyrimui pasirinkta homogeniška grupė, sudaryta iš 14 klasių. Iš šios grupės į eksperimentinę grupę atrinkti 52 mokiniai iš dviejų miesto ir vienos rajono mokyklos, o į kontrolinę grupę – 46 mokiniai iš dviejų miesto ir vienos rajono mokyklos.

Testų administravimas ir surinktų duomenų tvarkymas. Mokiniai testavo mokytojai bei Šiaulių universiteto studentai, vadovaudamiesi standartine testavimo instrukcija. Kiekvieno testo duomenys surinkti per ne ilgesnį, kaip dviejų savaitių laikotarpį.

Testo duomenų kodavimas. Testas sudarytas iš 97 invencinių uždavinių, t. y. mokinys turėjo konstruoti atsakymą. Koduojant testavimo rezultatus, remtasi vertinimo normomis. Atsakymai ir sprendimai analizuoti ir, remiantis standartinėmis instrukcijomis, vertinti 3 gradacijų rangine skale: 0 – užduotis nespręsta, 1 – pateiktas teisingas sprendimas, 2 – sprendimas iš esmės klaidingas. Duomenų kodavimo ir perkėlimo procedūrų patikimumui įvertinti pakartotinai koduota ir suvesta 10 % atsitiktinai atrinktų testo lapų. Klaidų skaičius neviršijo 0,30 %.

IRT samprata

CTT statistikos esmė – užduočių sunkumas, užduočių diskriminacija bei patikimumas priklauso nuo respondentų, kuriems buvo užduoti klausimai, imties. IRT užduočių parametrai nepriklauso nuo imties ir laikomi nekintančiais skirtingose tyrinėjamos populiacijos grupėse. Be to, CTT duoda tik vieną patikimumo įvertinimą ir atitinkamą standartinę matavimo paklaidą, tuo tarpu IRT modeliai matuoja skalės tikslumą pagrindiniu nematomu (latentini) kintamuoju θ (Cooke & Michie, 1997; Hays, Morales & Reise, 2000). IRT respondento tikrąjį savybių (pvz., matematinių gebėjimų) lygį kontinume įvertina geriau negu susumuotas skalės balas CTT (Santor & Ramsay, 1998).

IRT teorijos „startą“ įprasta sieti su danų matematiko Georg'o Rasch'o vardu. Plėtodamas IRT modelių teoriją, G. Rasch (1960) aptarė būtinybę kurti statistinius modelius, turinčius specifinio objektyvumo savybę, pagrįstus dominuojančia idėja, kad respondento ir užduoties, kurią jis sprendžia, parametrai būtų apskaičiuojami atskirai, bet lyginami panašiais matais.

Nors nemažai tyrėjų mano, kad IRT yra moderni šiudienė prichometrinė teorija, jos koncepcija ir metodologija plėtojama jau daugiau nei tris ketvirčius amžiaus. Dar 1925 metais L. L. Thurstone (Thurstone, 1925) parengė konceptualius IRT pamatus.

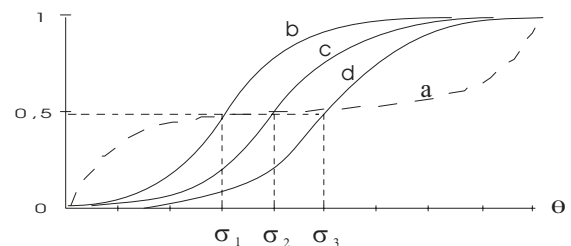
Iš pradžių IRT modeliai buvo kuriami tik dichotominiam atsakymo formatui (išspręsta – neišspręsta) ir plėtojami edukacinės bei kognityvinės diagnostikos rėmuose. Vėliau modeliai buvo pritaikyti nagrinėti ir tokiems testo žingsniams, kurie numato daugiapakopį atsakymo formatą. Tai leido išplėsti modelių taikymą ir nekognityvinės diagnostikos sferoje. Vis dėlto ir šiandien giliau išnagrinėtos tik nesudėtingų IRT bazėje (tai yra monotoniškai didėjančių) funkcijų savybės. Akivaizdu, kad minėto tipo funkcijos be išlygų tinka tik kognityvinei ir mokymosi pasiekimų diagnostikai. Asmenybės savybių, socialinių nuostatų ir vertybių diagnostikoje, kur paprastai remiamasi daugiapakopiu atsakymo formatu, o monotoniškai kylantis ICC modelis nepagrįstas empiriškai arba nepriimtinas teoriškai, ieškoma kitų tyrimo priemonių.

IRT testų teorijoje centrinė kategorija – testo užduoties išsprendimo tikimybė. Todėl ši teorija neretai dar vadinama *tikimybine testų teorija*, kurios esmė – santykiui tarp atsakymo elgsenos (užduoties išsprendimo tikimybės p) ir psichometrinio įverčio, dar vadinamo asmens parametru arba gebėjimo parametru (teta θ), apibūdinti postuluojamą funkcinį ryšį. Grafinę šio ryšio išraišką atspindi testo žingsnio charakteringoji kreivė (*Item Characteristic Cur-*

ve (ICC), vienokia ar kitokia konfigūracija paprastai nulemianti duoto modelio psichometrinę prasmę.

Konstruojant empirinę charakteringąją funkciją pagal testo rezultatą, visi tiriamieji suskirstomi į grupes. Specializuotoje edukacinių tyrimų rezultatų kompiuterinio apdorojimo programoje PAULA (Битинас, Паулавичюс, 1987) tokių grupių yra septynios. Užduoties sunkumo koeficientas kiekvienoje grupėje traktuojamas kaip empirinis tikimybės šios grupės moksleiviams išspręsti duotą užduotį įvertis.

ICC kreivės turi S formos kreivės pavidalą (1 pav.). G. Rascho modelio ICC apima skirtingo sunkumo užduotis (b – lengva, c – vidutinio sunkumo, d – sunki, a – modelio sąlygos netenkinanti užduotis) su horizontaliomis asimptotėmis ($P_j(\theta) = c_j$ ir $P_j(\theta) = 1$). Labai mažoms θ reikšmėms $P_j(\theta)$ didėja lėtai. Kai θ priartėja prie b_j , tai $P_j(\theta)$ paprastai didėja daug greičiau. Didėjimo greitis pasiekia maksimumą, kai $\theta = b_j$. Kai θ tampa žymiai didesnis nei b_j , tada $P_j(\theta)$ augimo greitis sulėtėja ir pasiekia 0 su labai didelėmis θ reikšmėmis. Tai alternatyvus užduoties sudėtingumo b_j , kaip kreivės išlinkimo taško (threshold), susijusio su teisingo atsakymo konstravimo tikimybe, apibūdinimas.



1 pav. Rascho modelio ICC. Skirtingo sunkumo užduotys: b – lengva, c – vidutinio sunkumo, d – sunki, a – modelio sąlygos netenkinanti užduotis

Užduoties *charakteringąja funkcija* (užduoties charakteringąja kreive, ICC) vadinama monotoniškai didėjanti funkcija, išreiškianti ryšį tarp tikimybės, kad respondentas teisingai atsakys į testo klausimą, ir užduoties parametru. Šiuolaikinėje testų teorijoje postuluojuama, kad šį ryšį geriausiai išreiškia logistinė funkcija. Kiekviena užduotis apie testuojamojo gebėjimus (konkrečiu atveju apie matematinius pasiekimus) teikia tam tikrą informacijos kiekį, kurį G. Rasch modelyje, žinant užduoties parametrus, galima apskaičiuoti. Jeigu tikimybė išspręsti užduotį artima 0 arba 1, tai informatyvumas artimas 0. Didžiausią informacijos kiekį užduotis teikia apie tuos tiriamuosius, kurių tikimybė išspręsti užduotį lygi 0,5. Informacinės funkcijos reikšmės priklauso

nuo užduoties diskriminacijos parametro: kuo didesnė diskriminacija, tuo didesnę informacijos kiekį užduotis teikia apie aukščiau minėtus tiriamuosius. Viso testo teikiama informacija aprašoma testo informacijos funkcija:

$$I(\theta) = \sum \frac{(P'_i(\theta))^2}{P_i(\theta)(1-P_i(\theta))}$$

$P_i(\theta)$ – i -tosios užduoties charakteringoji funkcija,

$P'_i(\theta)$ – jos išvestinė.

Aišku, kad testo informacija priklauso nuo užduočių skaičiaus k . Didėjant k , gerėja skaičiuojamos tikimybės tikslumas. Tačiau praktiškai tenka atsakyti tik į iš anksto nustatytam tikslumo lygiui pasiekti reikalingų klausimų skaičių.

Tyrimo rezultatai ir interpretavimas

Testo rezultatų analizei taikyta tikimybinė testų teorija. Respondentų skaičius – 98 ketvirtų klasių moksleiviai.

Skaičiavimų testo užduotys atitinka moksleivių amžių, žinias ir bendruosius bei specialiuosius gebėjimus, apibrėžtus Bendrosiose programose ir išsilavinimo standartuose (Bendrosios programos ir išsilavinimo standartai, 2008).

Tyrimo dalyvavusių homogeniškų klasių moksleivių aritmetinių skaičiavimų gebėjimus iki eksperimento pagal užduočių sunkumą sąlyginai galima suskirstyti į penkis lygius:

1. *Labai lengvos užduotys*, kurių sunkumo koeficientas nuo 0,85 iki 1,00.
2. *Lengvos užduotys*. Sunkumo koeficientas kinta nuo 0,75 iki 0,85.
3. *Optimalaus sunkumo užduotys*, galinčios užtikrinti gerą diskriminaciją. Sunkumo koeficientas nuo 0,35 iki 0,75.
4. *Sunkios užduotys*. Sunkumo koeficientas kinta nuo 0,20 iki 0,35.
5. *Labai sunkios užduotys*. Sunkumo koeficientas nuo 0,00 iki 0,20.

Iki eksperimentiniai moksleivių, dalyvavusių virtualių mokymo objektų efektyvumo, mokant aritmetikos veiksmų, tyrime, skaičiavimo testo užduočių kokybės statistiniai parametrai pateikti 1–4 lentelėse:

1 lentelė. *Labai lengvų skaičiavimo užduočių sklaida*

Užduoties nr.	Nesprendė	Gerai	Klaidingai	Vidurkis	Standartinis nuokrypis	Alpha koeficientas	Koreliacija su testu	Sunkumas
C42	0,00	86,73	13,27	1,13	0,34	0,71	0,07	0,87
C90	1,02	85,71	13,27	1,12	0,36	0,71	-0,04	0,86

2 lentelė. *Lengvų skaičiavimo užduočių sklaida*

Užduoties nr.	Nesprendė	Gerai	Klaidingai	Vidurkis	Standartinis nuokrypis	Alpha koeficientas	Koreliacija su testu	Sunkumas
C1	0,00	83,67	16,33	1,16	0,37	0,70	0,13	0,84
C13	0,00	83,67	16,33	1,16	0,37	0,71	0,08	0,84
C77	4,08	83,67	12,24	1,08	0,4	0,70	0,33	0,84
C79	4,08	82,65	13,27	1,09	0,41	0,71	0,03	0,83
C43	0,00	81,63	18,37	1,18	0,39	0,71	0,00	0,82
C82	5,10	81,63	13,27	1,08	0,42	0,70	0,15	0,82
C3	0,00	80,61	19,39	1,19	0,40	0,71	0,04	0,81
C7	0,00	80,61	19,39	1,19	0,40	0,7	0,12	0,81
C17	0,00	80,61	19,39	1,19	0,40	0,71	0,03	0,81
C38	2,04	80,61	17,35	1,15	0,41	0,71	-0,01	0,81
C80	1,02	80,61	18,37	1,17	0,41	0,71	0,07	0,81
C10	0,00	79,59	20,41	1,20	0,41	0,71	-0,06	0,80
C40	2,04	79,59	18,37	1,16	0,42	0,70	0,22	0,80
C75	2,04	79,59	18,37	1,16	0,42	0,7	0,19	0,80
C76	3,06	79,59	17,35	1,14	0,43	0,70	0,17	0,80
C2	0,00	78,57	21,43	1,21	0,41	0,70	0,14	0,79
C37	0,00	78,57	21,43	1,21	0,41	0,70	0,21	0,79
C78	5,10	78,57	16,33	1,11	0,45	0,70	0,17	0,79
C84	3,06	78,57	18,37	1,15	0,44	0,70	0,22	0,79
C89	3,06	78,57	18,37	1,15	0,44	0,70	0,15	0,79
C9	0,00	77,55	22,45	1,22	0,42	0,71	0,06	0,78
C11	0,00	77,55	22,45	1,22	0,42	0,71	-0,02	0,78
C12	0,00	77,55	22,45	1,22	0,42	0,71	0,05	0,78
C16	0,00	77,55	22,45	1,22	0,42	0,70	0,13	0,78
C23	0,00	77,55	22,45	1,22	0,42	0,71	-0,01	0,78
C45	1,02	77,55	21,43	1,20	0,43	0,7	0,12	0,78
C66	1,02	77,55	21,43	1,20	0,43	0,71	-0,24	0,78
C85	1,02	77,55	21,43	1,20	0,43	0,71	0,06	0,78
C8	0,00	76,53	23,47	1,23	0,43	0,71	0,01	0,77
C31	0,00	76,53	23,47	1,23	0,43	0,70	0,27	0,77
C81	3,06	76,53	20,41	1,17	0,45	0,71	0,07	0,77
C19	0,00	75,51	24,49	1,24	0,43	0,71	-0,03	0,76
C51	0,00	75,51	24,49	1,24	0,43	0,70	0,15	0,76

3 lentelė. *Optimalaus sunkumo skaičiavimo užduočių sklaida*

Užduoties nr.	Nesprendė	Gerai	Klaidingai	Vidurkis	Standartinis nuokrypis	Alpha koeficientas	Koreliacija su testu	Sunkumas
C4	1,02	74,49	24,49	1,23	0,45	0,70	0,18	0,74
C6	1,02	74,49	24,49	1,23	0,45	0,71	-0,02	0,74
C30	1,02	74,49	24,49	1,23	0,45	0,70	0,27	0,74
C83	3,06	74,49	22,45	1,19	0,47	0,71	0,05	0,74
C32	1,02	73,47	25,51	1,24	0,46	0,71	0,07	0,73
C34	0,00	73,47	26,53	1,27	0,44	0,70	0,14	0,73
C29	1,02	72,45	26,53	1,26	0,46	0,71	0,08	0,72
C36	0,00	72,45	27,55	1,28	0,45	0,71	0,09	0,72
C86	2,04	72,45	25,51	1,23	0,47	0,71	0,10	0,72
C20	0,00	71,43	28,57	1,29	0,45	0,7	0,14	0,71
C28	2,04	71,43	26,53	1,24	0,48	0,71	0,08	0,71
C46	0,00	71,43	28,57	1,29	0,45	0,70	0,12	0,71
C53	0,00	71,43	28,57	1,29	0,45	0,71	0,02	0,71
C74	2,04	71,43	26,53	1,24	0,48	0,71	-0,08	0,71
C26	0,00	70,41	29,59	1,3	0,46	0,71	0,00	0,70
C25	0,00	69,39	30,61	1,31	0,46	0,70	0,12	0,69
C24	0,00	68,37	31,63	1,32	0,47	0,71	0,03	0,68
C33	3,06	68,37	28,57	1,26	0,5	0,70	0,22	0,68
C52	0,00	69,39	30,61	1,31	0,46	0,70	0,14	0,68
C15	0,00	67,35	32,65	1,33	0,47	0,70	0,12	0,67
C27	1,02	67,35	31,63	1,31	0,48	0,70	0,23	0,67
C62	3,06	67,35	29,59	1,27	0,51	0,70	0,24	0,67
C88	4,08	67,35	28,57	1,24	0,52	0,70	0,30	0,67
C22	2,04	66,33	31,63	1,30	0,50	0,70	0,16	0,66
C56	1,02	66,33	32,65	1,32	0,49	0,71	0,05	0,66
C59	2,04	66,33	31,63	1,30	0,50	0,70	0,25	0,66
C41	2,04	65,31	32,65	1,31	0,51	0,71	0,10	0,65
C57	2,04	65,31	32,65	1,31	0,51	0,70	0,24	0,65
C39	1,02	64,29	34,69	1,34	0,50	0,70	0,25	0,64
C14	1,02	63,27	35,71	1,35	0,50	0,71	-0,03	0,63
C18	3,06	63,27	33,67	1,31	0,53	0,70	0,27	0,63
C63	3,06	63,27	33,67	1,31	0,53	0,70	0,35	0,63
C64	3,06	63,27	33,67	1,31	0,53	0,70	0,21	0,63
C35	0,00	62,24	37,76	1,38	0,49	0,71	0,06	0,62
C21	0,00	61,22	38,78	1,39	0,49	0,70	0,20	0,61
C58	2,04	61,22	36,73	1,35	0,52	0,70	0,19	0,61
C65	3,06	61,22	35,71	1,33	0,53	0,71	-0,02	0,61
C44	0,00	60,2	39,80	1,40	0,49	0,71	-0,08	0,60
C54	3,06	60,2	36,73	1,34	0,54	0,71	0,06	0,60
C87	3,06	60,2	36,73	1,34	0,54	0,70	0,12	0,60

3 lentelės tęsinys

C68	9,18	59,18	31,63	1,22	0,60	0,70	0,13	0,59
C67	2,04	57,14	40,82	1,39	0,53	0,71	-0,02	0,57
C5	1,02	56,12	42,86	1,42	0,52	0,70	0,17	0,56
C49	0,00	56,12	43,88	1,44	0,50	0,71	0,07	0,56
C69	5,10	55,10	39,8	1,35	0,58	0,71	0,06	0,55
C70	6,12	55,10	38,78	1,33	0,59	0,70	0,12	0,55
C50	0,00	54,08	45,92	1,46	0,50	0,70	0,34	0,54
C55	0,00	54,08	45,92	1,46	0,50	0,71	0,02	0,54
C97	11,22	54,08	34,69	1,23	0,64	0,70	0,13	0,54
C60	8,16	52,04	39,8	1,32	0,62	0,70	0,33	0,52
C47	0,00	51,02	48,98	1,49	0,50	0,70	0,12	0,51
C91	5,10	51,02	43,88	1,39	0,59	0,70	0,13	0,51
C73	23,47	50,00	26,53	1,03	0,71	0,70	0,23	0,5
C92	4,08	48,98	46,94	1,43	0,57	0,71	-0,11	0,49
C94	17,35	47,96	34,69	1,17	0,70	0,69	0,42	0,48
C72	24,49	45,92	29,59	1,05	0,74	0,70	0,25	0,46
C61	5,10	44,90	50,00	1,45	0,59	0,70	0,27	0,45
C71	24,49	44,90	30,61	1,06	0,74	0,70	0,22	0,45
C96	12,24	42,86	44,9	1,33	0,69	0,69	0,40	0,43
C95	13,27	38,78	47,96	1,35	0,70	0,69	0,51	0,39

4 lentelė. *Sunkios skaičiavimo užduoties sklaida*

Užduoties nr.	Nesprendė	Gerai	Klaidingai	Vidurkis	Standartinis nuokrypis	Alpha koeficientas	Koreliacija su testu	Sunkumas
C93	3,06	22,45	74,49	1,71	0,52	0,71	0,08	0,22

Taikytame teste labai sunkių uždavinių nėra.

Iš 1–4 lentelių IRT metodu išanalizuosime 4 optimalaus sunkumo užduotis 71–74 užduotys su gera skiriamąja geba (4 pav.).

71–73 užduotys. Sudėję skaičius A ir B, gauname 700. Sudėję skaičius B ir C, gauname 550, o sudėję skaičius A, B ir C, gauname 1000. Kokie tai skaičiai? Parodykite, kaip nustatėte.

71) $A = \dots\dots\dots$; 72) $B = \dots\dots\dots$; 73) $C = \dots\dots\dots$;

74 užduotis. Skaičių 870 ir 430 skirtumą sumažinkite 4 kartus.

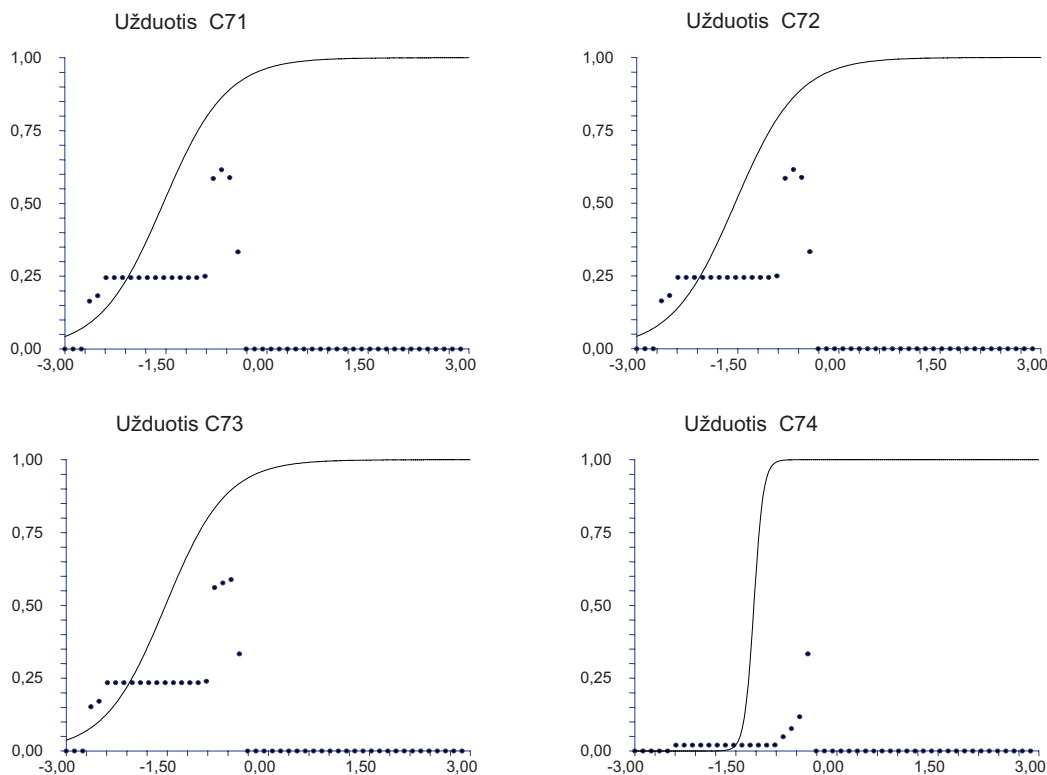
Analizei parinktų užduočių išspręstumo lygmenys bei diagramos pateikti 5 lentelėje.

5 lentelė. *Skaiciavimo užduočių išspręstumo lygmenys*

	Reikšmės	Dažniai	Procentai	Suminiai procentai	Diagrama
Užduotis C71	0 (nesprendė)	24,00	24,49	24,49	
	1 (gerai)	44,00	44,90	69,39	
	2 (klaidingai)	30,00	30,61	100,00	
Užduotis C72	0 (nesprendė)	24,00	24,49	24,49	
	1 (gerai)	45,00	45,92	70,41	
	2 (klaidingai)	29,00	29,59	100,00	
Užduotis C73	0 (nesprendė)	23,00	23,47	23,47	
	1 (gerai)	49,00	50,00	73,47	
	2 (klaidingai)	26,00	26,53	100,00	
Užduotis C74	0 (nesprendė)	2,00	2,04	2,04	I
	1 (gerai)	70,00	71,43	73,47	
	2 (klaidingai)	26,00	26,53	100,00	

IRT charakteringosios kreivės, parodančios, kuriuos veiksmus (nesprendė, sprendė su klaidomis, išsprendė teisingai) ir su kokia tikimybe pateiktos 2–3 paveiksluose: vertikaliuoje ašyje pavaizduota tikimybė, su kuria moksleiviai atlieka / neatlieka kurį

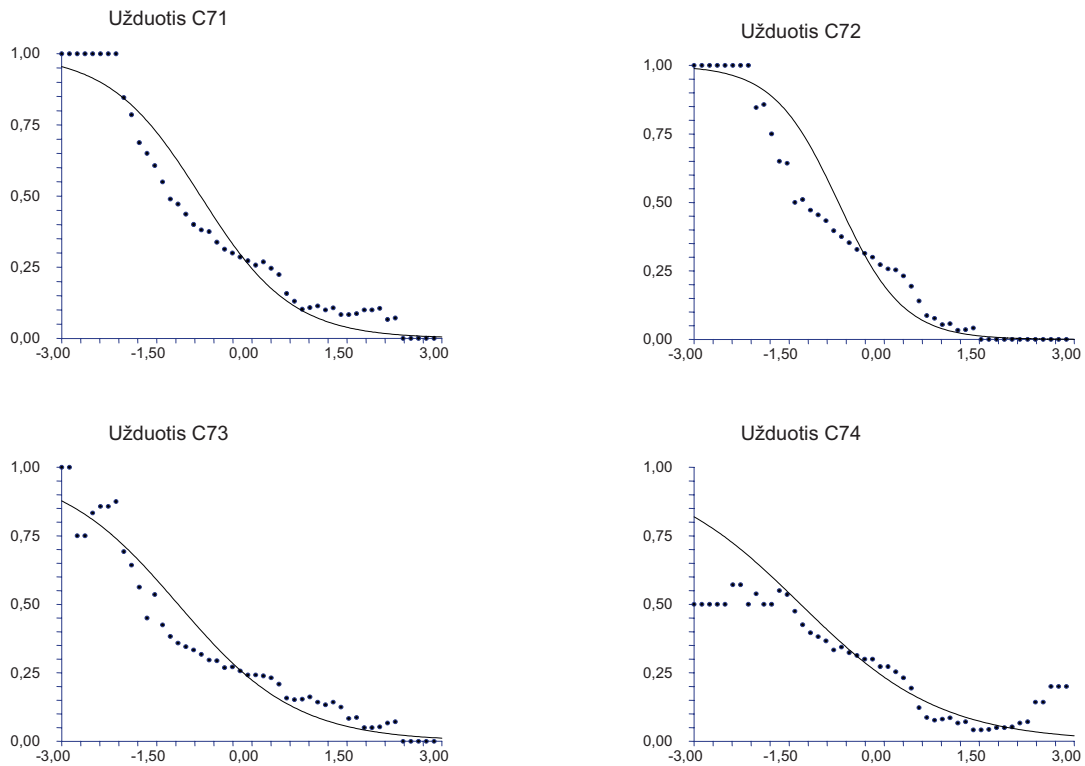
nors veiksmą; horizontaliojoje ašyje atidėti moksleivių skaičiavimų gebėjimai atlikti / neatlikti kurį nors veiksmą su pateikta užduotimi Z skalėje. Taškais pažymėti faktiniai gebėjimai.



2 pav. Moksleivių, nesprendusių skaičiavimo užduočių, gebėjimai Z skalėje

2 paveikslo duomenys rodo, kad užduotis C71–74 su maksimalia tikimybe 0,69 nė nebandė

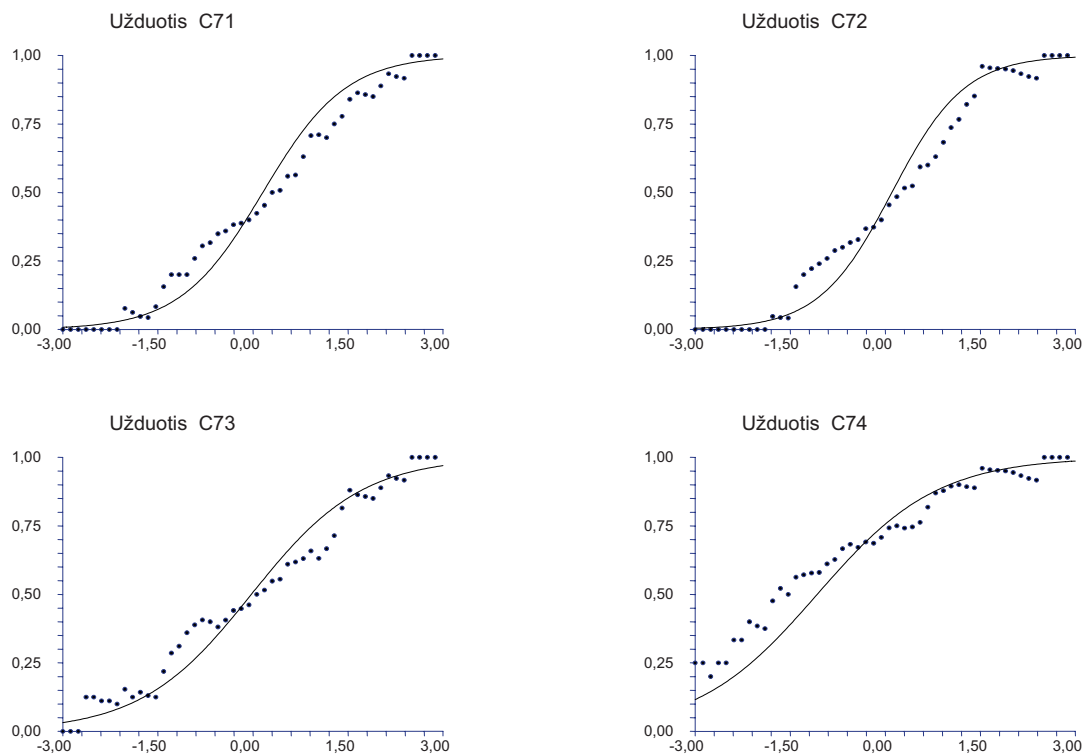
spręsti moksleiviai, turintys tik minimalius aritmetinių skaičiavimų gebėjimus.



3 pav. Tiriamųjų negebėjimas teisingai išspręsti užduotį (Z skalė)

3 paveikslo duomenys rodo, kad, sprenddami užduotis C71–74, klaidų gali daryti visų gebėjimo lygių moksleiviai, tačiau didžiausia tikimybė, kad

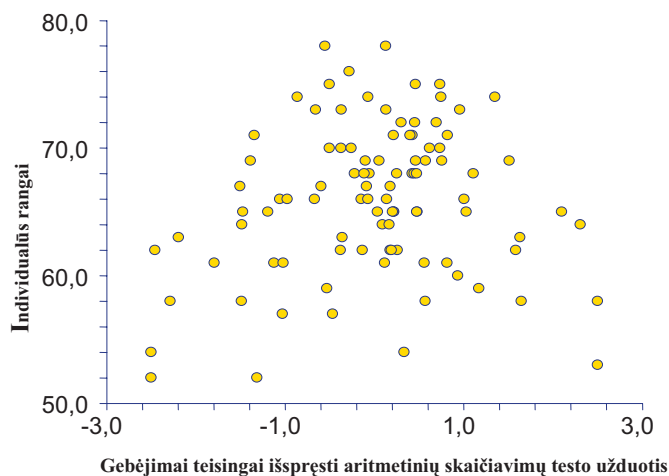
daugiausia klaidų darys minimalių aritmetinių skaičių gebėjimų moksleiviai.



4 pav. Moksleivių gebėjimai teisingai išspręsti aritmetinių skaičių užduotis (Z skalė)

4 paveikslo kreivės rodo gerą skaičiavimo gebėjimų skiriamąją gebą. Tikėtina, kad aukštesnių gebėjimų moksleiviai geriau sprendžia skaičiavimo

uždavinius. C74 užduotis yra lengvesnė. Ją geriau sprendė ir minimalių gebėjimų moksleiviai.



5 pav. Moksleivių gebėjimų teisingai išspręsti aritmetinių skaičiavimų užduotis (Z skalė) ir individualių rangų skirstinys

5 paveikslo duomenys rodo, kad pateiktas skaičiavimo testas iš esmės yra skirtas vidutinių (normalių) aritmetinių skaičiavimų gebėjimų moksleiviams. Tai atitinka klasių pedagoginę realybę – mokymo procesas mokykloje orientuotas į vidutinių gebėjimo moksleivius. Tikėtina, kad tiek aukštesnių, tiek minimalių gebėjimų moksleivių klasėse nėra daug. Iš 5 paveikslo duomenų matyti, kad tiek minimalių, tiek aukštesnių skaičiavimo gebėjimų turinčių, bet silpnai testo užduotis spendusių moksleivių pasiskirstymas yra maždaug vienodas. Dominuoja normalių gebėjimų moksleiviai.

Empirinių tyrimų duomenys, taikant tikimybinę testų teoriją, kaip ir klasikinę testų teoriją, atskleidžia tiriamųjų skaičiavimų gebėjimus bei jų žinių spragas ir teikia galimybių pradinėms klasių mokytojams diferencijuoti moksleivių matematinio ugdymo programas bei tikslingai taikyti virtualius mokymo objektus.

Išvados

1. Remiantis literatūros analize bei testo kokybės tikimybinėmis charakteristikomis (pateiktos straipsnio lentelėse ir paveiksluose), galima tvirtinti, kad tyrime taikyti pakankamai kokybiški ketvirtos klasės mokinių skaičiavimo gebėjimų diagnostiniai testai.
2. Pateisinti reikalavimai testo kokybei, o tai leidžia tvirtinti, kad eksperimentui išskirti homogeninių klasių moksleiviai parinkti tinkamai, t. y. moksleiviai turi maždaug vienodus matematinis gebėjimus skaičiavimo užduotims atlikti.

3. Tikėtina, kad atlikus eksperimentą, t. y. mokant aritmetikos veiksmų, panaudojant virtualius mokymo objektus, galima bus objektyviai fiksuoti virtualių mokymo objektų poveikius aritmetinių skaičiavimų gebėjimams, tobulinti.

Literatūra

1. *Bendrosios programos ir išsilavinimo standartai*, 2008. Vilnius.
2. Brackenamp R., 1997, *Handbuch Psychologischer und Pädagogischer Tests*. Goettingen, Seattle: Hogrefe.
3. Cooke D. J., Michie C., 1997, An item response theory analysis of the Hare Psychopathy Checklist – Revised. *Psychological Assessment*. Nr. 9. P. 3–14.
4. Embretson S. E., Reise S. P., 2000, *Item Response Theory for Psychologists*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
5. Hays R. D., Morales L. S., Reise S. P., 2000, Item response theory and health outcomes measurement in the 21st century. *Medical Care*. Nr. 38 (9 Supplement). P. 28–42.
6. Kiseliova D., 2002, Ketvirtų klasių mokinių matematiniai gebėjimai kaip didaktinės diagnostikos objektas. *Daktaro disertacija*. Šiauliai.
7. Kiseliova D., Kiseliovas A., 2004, Matematinis gebėjimų diagnostika. *Mokslinė monografija*. Pirmoji knyga. Šiaulių universiteto leidykla.
8. Rasch G., 1960, *Probabilistic models for some intelligence and attainment tests*. Chicago: MESA.
9. Rost J., 1995, *Lehrbuch Testtheorie Testkonstruktion*. Verlag Hans Huber, Bern.
10. Santor D. A., Ramsay J. O., 1998, Progress in the technology of measurement: applications of item response models. *Psychological Assessment*. Nr. 10. P. 345–359.

11. Testzentrale (Hrsg.), 1998. Testkatalog 1998/1999. *European test Publishers Group*. Testzentrale, Goettingen; Bern: Hogrefe.
12. Thurstone L. L., 1925, A method of scaling psychological and educational tests. *Journal of Educational Psychology*. P. 16. P. 433–451.
13. Битинас Б., Паулавичюс Р., 1987, Система анализа педагогических данных на ЭВМ. Вильнюс.

ITEMS RESPONSE THEORY OF COMPUTING SKILLS TESTS IN DIAGNOSIS OF VIRTUAL OBJECTS IN THE APPLICATION OF EDUCATIONAL EFFECTIVENESS STUDY

Orinta Šalkuvienė, Danutė Kiseliuva, Arkadijus Kiseliovas

Summary

One of the most important principles in educological analysis is separation of homogeneous groups, control groups, and experimental groups. When analyzing virtual objects in the application form of computing, performance, it is necessary to select schoolchildren with equal computing skills. The aim of this article is to describe the analysis of necessary counting skills in homogeneous groups of schoolchildren prior to pedagogical experiment, using diagnostic tests in mathematics. Methods used were: literary analysis of the general requirements for diagnostic tests of achievements in mathematics; the design, analysis and interpretation of diagnostic test of the fourth year schoolchildren; identification of homogeneous groups on the basis of results of testing. The article discusses the characteristics of probability of arithmetical calculation and its suitability for diagnostics of calculation abilities. The results obtained confirm ability of homogenous groups to calculate and efficiency of the research using virtual teaching with special objects when dealing with arithmetic task. Schoolchildren of groups IV-V were researched.

Keywords: test, mathematical ability, probability, test theory, characteristic curve.

TIKIMYBINĖ TESTŲ TEORIJA SKAIČIAVIMŲ GEBĖJIMŲ DIAGNOSTIKOJE VIRTUALIŲ MOKYMO OBJEKTŲ TAIKYMO EFEKTYVUMUI TIRTI

Orinta Šalkuvienė, Danutė Kiseliuva, Arkadijus Kiseliovas

Santrauka

Vienas svarbiausių principų edukologiniams tyrimams atlikti – homogeninių grupių sudarymas bei kontrolinių ir eksperimentinių grupių išskyrimas.

Šio straipsnio tikslas – diagnostiniais matematikos testais apibūdinti tyrimui būtinų homogeninių klasių moksleivių skaičiavimų gebėjimus prieš pedagoginį eksperimentą.

Taikyti metodai: dalykinės literatūros analizė bendrųjų reikalavimų diagnostiniams matematikos pasiekimų testams konstruoti tema; diagnostinio testo IV klasės mokiniams konstravimas, analizė ir interpretacija; homogeninių klasių išaiškinimas, remiantis testavimo rezultatais.

Straipsnyje aptariamos aritmetinių skaičiavimų testo tikimybinės charakteristikos ir tinkamumas skaičiavimų gebėjimų diagnostikai. Tyrimo rezultatai patvirtina eksperimentui parinktų klasių skaičiavimo gebėjimų homogeniškumą bei sukonstruoto skaičiavimų testo tinkamumą virtualių mokymo objektų taikymo, mokant aritmetikos veiksmų IV–V klasėse, efektyvumo tyrimui.

Prasminiai žodžiai: testas, matematiniai gebėjimai, tikimybinė testų teorija, charakteringosios kreivės.

[teikta 2011-02-14