

Testo informacija ir jos taikymas

Beatričė ANDZIULIENĖ, Arūnas ANDZIULIS, Sonata BAKANAUSKAITĖ (KU)

el. paštas: beata@ik.ku.lt

1. Įvadas

Testavimą kaip žinių tikrinimo ir vertinimo metodo pasirinkimą veikia bendros švietimo kaitos tendencijos ir vyraujančios ugdymo sampratos, taip pat akivaizdus informacinių technologijų vaidmuo ugdymo proceso kaitai. Kompiuterinių mokymo technologijų iššauktas poreikis taikyti naujus žinių bei gebėjimų vertinimo metodus neatmeta poreikio gauti įvairiapusę informaciją apie testuojamąjį, geriau jį pažinti, diagnozuoti mokymo-mokymosi proceso eigą, išryškinti ugdymo spragas [1, 2]. Kuriant testą ir analizuojant testavimo rezultatus svarbu ne tik įvertinti kiekvieno testuojamojo bei visos grupės balus, jų pasiskirstymą bei įverčių paklaidą, bet ir išanalizuoti testo informacijos kokybę. Šiuolaikinės, testo kokybę bei jo parametrus nagrinėjančios, teorijos leidžia atlikti išsamią tiek testuojamųjų tiek ir testo užduočių analizę. Žinių tikrinimui reikia sukurti tokį testą, kuris tiksliai išaiškintų kiekvieno testuojamojo žinių spragas, įvertintų individualius gebėjimus. Todėl, kuriant testą, reikia siekti tolygaus užduočių sunkumo pasiskirstymo kuo tiksliau suderinto su testuojamųjų grupės žinių lygiu, o tai leidžia padaryti užduočių sunkumo bei jų įnešamos informacijos į testą nagrinėjimas.

Tik sistemingas vertinimas bei greitas ir informatyvus vertinimo procedūrų grįžtamas ryšys leidžia efektyviai organizuoti ugdymo procesą, todėl aktualu naudotis šiuolaikinės testų teorijos pasiekimais taikant testavimą ugdymo procese bei ruošiant kokybiškus kompiuterinius testus.

Šio darbo **tyrimo objektas** – testo informacija. **Darbo tikslas** – testo grįžtamosios informacijos analizė įvertinant testo kokybę bei jos rezultatų panaudojimas realizuojant diagnostinę bei pažintinę tikrinimo funkcijas, siekiant pakelti ugdymo proceso efektyvumą. Darbe remiamasi **užduoties sprendimo teorijos metodologija**.

2. Užduoties sprendimo teorijos metodologija

Klasikinė testų teorija intensyviai vystėsi ir buvo taikoma sudarant testus bei atliekant statistinę testo duomenų analizę iki 1970-ųjų metų. Tačiau ji susidūrė su keletu apribojimų, kurių pati negalėjo išspręsti. Duomenys, gauti klasikinės testų teorijos pagalba, leidžia nustatyti santykinę kiekvieno testuojamo padėtį normaliniame pasiskirstyme, tačiau jų pagalba negalima objektyviai įvertinti parametrų, charakterizuojančių testuojamąjį (žinių lygis, gebėjimai) ar testo užduočių sunkumą [3].

Jau 1950 metais kartu su kompiuterinės technikos atsiradimu išryškėjo klasikinės testų teorijos problemos ir buvo kuriami nauji testų analizės modeliai, juose įvedami tokie parametrai kaip testuojamojo žinių lygis ir testo užduočių sunkumas. Ilgainiui šie modeliai vystėsi ir sudarė vieningą šiuolaikinės testų teorijos kryptį vadinamą testo užduoties sprendimo teorija (Item Response Theory – IRT).

Užduoties sprendimo teorijoje (UST) daroma prielaida, kad tarp gautų testo rezultatų ir testuojamojo gebėjimų bei žinių lygio egzistuoja ryšys – i -ojo testuojamojo žinių lygio θ_i ir j -osios užduoties sunkumo β_j sąveika yra testavimo rezultatai. Tai leidžia išspręsti atvirkštinį praktinį klausimą – pagal testuojamųjų atsakymus nustatyti θ_i ir β_j parametrus. Bendras i -tojo testuojamojo žinių lygis apibūdinamas santykiu:

$$\theta_i = \ln(p_i/q_i), \quad (1)$$

o užduoties sunkumo koeficientas

$$\beta_j = \ln(q_j/p_j). \quad (2)$$

Čia p_i ir q_i – i -tojo testuojamojo teisingų ir klaidingų atsakymų dalis, atlikus visas testo užduotis, kur $p_i = 1 - q_i$; p_j ir q_j – visų grupės testuojamųjų teigiamų ir neigiamų atsakymų j -ąją užduotį skaičius, $p_j = 1 - q_j$.

Nors atskiri testavimo rezultatai priklauso nuo eilės atsitiktinių ir neatsitiktinių faktorių galima kalbėti apie tam tikrą tikimybę testuojamajam, kurio žinių lygis θ_i teisingai atlikti skirtingo sudėtingumo užduotį. Skiriami vienparametrinis, dviparametrinis ir triparametrinis logistiniai UST modeliai [4–6].

Vienparametrinis G. Rasch logistinis modelis turi tik vieną parametą – užduoties sunkumo koeficientą [6]. Jame tikimybė testuojamajam, kurio žinių lygis θ , atlikti skirtingo sunkumo β_j užduotį:

$$P_j(\theta) = \frac{e^{1,7(\theta - \beta_j)}}{1 + e^{1,7(\theta - \beta_j)}}, \quad (3)$$

čia $j = 1, \dots, k$, j – užduočių skaičius teste.

Tuo tarpu tikimybė, kad atsitiktinai pasirinktas testuojamasis, kurio žinių lygis yra θ , teisingai atliks β sunkumo užduotį aprašoma:

$$P_i(\beta) = \frac{e^{1,7(\theta_i - \beta)}}{1 + e^{1,7(\theta_i - \beta)}}, \quad (4)$$

$i = 1, \dots, N$, N – testuojamųjų skaičius. Reikšmę θ_i galima vertinti kaip i -tojo testuojamojo padėtį, o β_j – kaip j -osios užduoties padėtį toje pat kintamųjų θ, β ašyje. Tokiu atveju $|\theta_i - \beta_j|$ – tai atstumas tarp testuojamojo žinių lygio θ_i ir užduoties sunkumo β_j . Šis skirtumas rodo kiek užduotis per sunki, ar per lengva pagal testuojamojo žinių lygį.

Dviparametrinis Birnbaum logistinis modelis įvertina du parametrus – užduoties sunkumo koeficientą β_j ir užduoties skiriamąją gebą a_j , parametą rodanti, kiek gerai už-

duotis atskiria geresnius nuo blogesnių tikrinamo dalyko atžvilgiu. Šiame modelyje tikimybė atlikti užduotį:

$$P_j(\theta) = \frac{e^{a_j(\theta - \beta_j)}}{1 + e^{a_j(\theta - \beta_j)}}, \quad (5)$$

$$P_i(\beta) = \frac{e^{a_i(\theta_i - \beta)}}{1 + e^{a_i(\theta_i - \beta)}}. \quad (6)$$

Modelis su trim parametrais yra labiausiai realistinis, jame šalia dviejų pirmųjų parametru įvedamas trečias parametras nusakantis tikimybę atspėti užduoties atsakymą – užduoties nuspėjimo matas c_j . Tuomet triparametriniame logistiniame Birnbaum modelyje santykinė tikimybė skirtingiems θ_i gebėjimų lygio testuojamiesiems teisingai išspręsti j -ąją sunkumo β_j užduotį transformuojasi į tokią formulę:

$$P_j(\theta) = c_j + \frac{1 - c_j}{1 + e^{a_j(\theta - \beta_j)}}. \quad (7)$$

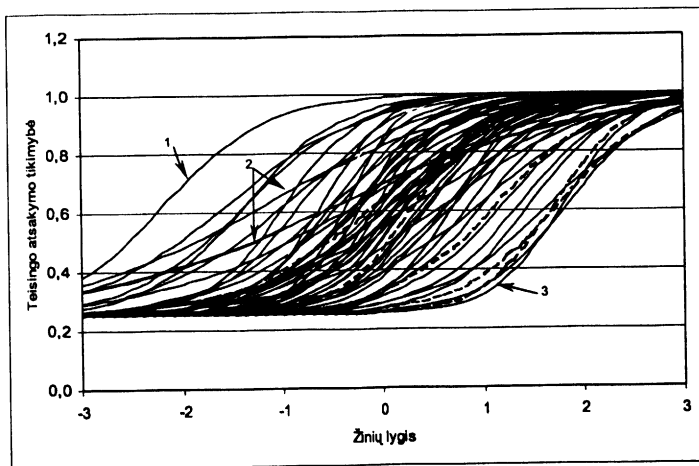
3. Užduoties charakteringoji kreivė

Užduoties sprendimų teorijoje kiekvienai užduočiai priskiriama funkcija $P_j(\theta)$, nusakanti sąryšį tarp testuojamųjų individualių gebėjimų ir tikimybės teisingai išspręsti užduotį. Ji vadinama charakteringąja užduoties funkcija, o jos grafikas – charakteringąja užduoties kreive (ICC – Item Characteristic Curve). Tai S formos kreivė turinti horizontalias asimptotes intervaluose, kai $P = c_j$ ir $P(\theta) = 1$. Esant labai žemoms θ reikšmėms, $P(\theta)$ auga lėtai, o artėjant prie β_j , teisingo atsakymo tikimybė $P(\theta)$ didėja sparčiau ir pasiekia maksimalų greitį, kai $\theta = \beta_j$, toliau greitis mažėja ir artėja prie nulio, esant labai didelėms θ reikšmėms (1 pav.). Užduoties charakteringoji kreivė pateikia alternatyvų užduoties sunkumo vertinimo kriterijų – kreivės lūžio tašką, surišančią teisingo atsakymo tikimybę ir žinių lygį [3].

Užduoties skiriamoji geba yra proporcinga kreivės polinkiui taške $\theta = \beta_j$, kai $P(\theta) = 0,5$ – tai kriterijus, kuris atspindi užduoties informatyvumą, pernešamos informacijos kiekį, testuojamojo gebėjimų ir žinių lygio atžvilgiu [4].

Nuspėjamumo parametras c_j , kuris visuomet $c_j < 1$, yra įvestas triparametriniame modelyje, siekiant įvertinti galimybę atspėti teisingą atsakymą į testo klausimą, neturint jokių žinių ir supratimo. Klausimai, kurių spėjimo tikimybė didesnė, bus mažiau informatyvūs, nei tie, kur spėjimo tikimybė nulinė, – jos nėra. Šis parametras svarbus naudojant uždaro tipo testus [7].

Testo užduočių charakteringosios kreivės, pavaizduotos 1 pav., gautos taikant triparametrinį Birnbaum modelį. Kreivės padėtis žinių lygio skalėje charakterizuoja i -tosios užduoties sunkumą. Ruošiant testą siekiama, kad užduočių charakteringosios kreivės tolygiai pasiskirstytų per visą žinių intervalą, t.y., kad skirtingą žinių lygį turintiems testuojamiesiems būtų apibrėžta tikimybė atlikti tam tikrą užduočių skaičių. Todėl vienodo sunkumo užduotys iš testo yra šalinamos. Šalinant iš testo užduotis, reikia atsižvelgti į



1 pav. Testo užduočių charakteringųjų kreivių (ICC) eiga ir pasiskirstymas:
1 – per lengvos, 2 – per sunkios, 3 – mažai diferencijuojančios užduotys

užduoties skiriamąją gebą testuojamųjų žinių atžvilgiu, t.y., atrinkti didesnes a_j reikšmes, turinčias ir šalinti mažiau diferencijuojančias (gulstesnės ICC) užduotis. Taip pat gerinant testo kokybę, keliant jo efektyvumą atliekamas testo „valymas“ pašalinant per lengvas (1), per sunkias (2) bei mažai diferencijuojančias užduotis (3) (1 pav.).

4. Charakteringoji testuojamojo kreivė

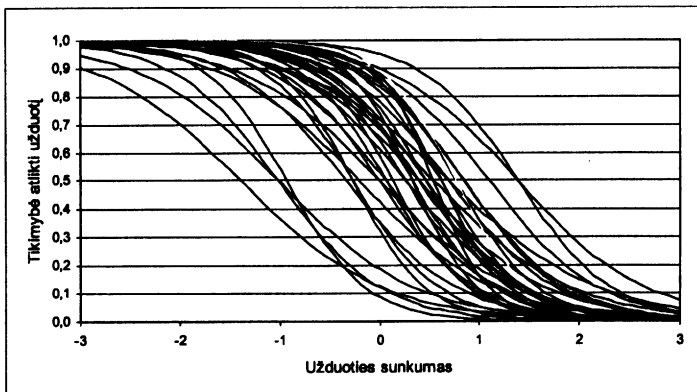
Aprašant tikimybę tam tikrų gebėjimų testuojamajam išspręsti skirtingo sunkumo užduotį UST modelis pateikia analogišką logistinių kreivių $P_i(\beta)$ šeimą.

Funkcijos $P_i(\beta)$ grafikas – tai charakteringoji i -tojo testuojamojo kreivė (PCC – Person characteristic curve). Tikimybė i -tajam testuojamajam teisingai išspręsti skirtingo sunkumo užduotį yra mažėjanti kintamojo β funkcija. Kreivės lūžio taške, kur $\beta = \theta_j$, funkcijos P_j reikšmė lygi 0,5. Testuojamojo charakteringosios kreivės poslinkis į dešinę rodo didesnę testuojamojo tikimybę atlikti tam tikro sudėtingumo užduotį. Mokymosi eigoje, įgyjant naujas žinias bei gebėjimus charakteringoji testuojamojo kreivė slenka į dešinę.

Kreivių išsidėstymas ir eiga 2 pav. rodo individualius testuojamojo sugebėjimus atlikti skirtingo sudėtingumo užduotis, taip pat rodo jo žinių lygį kitų testuojamųjų atžvilgiu. Atitinkamai grupuojant duomenis galime gauti informaciją apie norimą respondentų grupę. Palyginus parengiamojo ir baigiamojo testavimo rezultatus galima vizualiai pamatyti žinių lygio dinamiką.

5. Testo informacijos funkcija

Kiek teisingai įvertinti testuojamojo žinios bei gebėjimai tai esminis kiekvieno testavimo klausimas. Bet kokio parametro paklaidą nusako dispersija. Kuo didesnė θ dispersija, tuo



2 pav. Testuojamųjų individualių kreivių pasiskirstymas.

didesnė paklaida, kuo mažesnė θ dispersija, tuo didesnis tikslumas įvertinant gebėjimus ir žinias. Pagal A. Birnbaumą gebėjimų paklaida išreiškiama kaip dydis atvirkščiai proporcingas testo informacijos funkcijai (Test Information Function) [4]:

$$I(\theta) = \sum_{j=1}^k \frac{(P'_j(\theta))^2}{P_j(\theta) \cdot (1 - P_j(\theta))}, \quad (8)$$

čia $P'_j(\theta)$ – pirmoji $P_j(\theta)$ išvestinė pagal θ .

Šitaip apibrėžta testo informacija nusako, kaip tiksliai testas matuoja tikrinamo dalyko atžvilgiu skirtingų gebėjimų testuojamuosius. Kuo informacija atitinkamame gebėjimų lygyje yra didesnė, tuo mažesnis yra to gebėjimų lygio įvertinimo patikimumo intervalas ir paklaida

$$\Delta(\theta) = \frac{1}{\sqrt{I(\theta)}}. \quad (9)$$

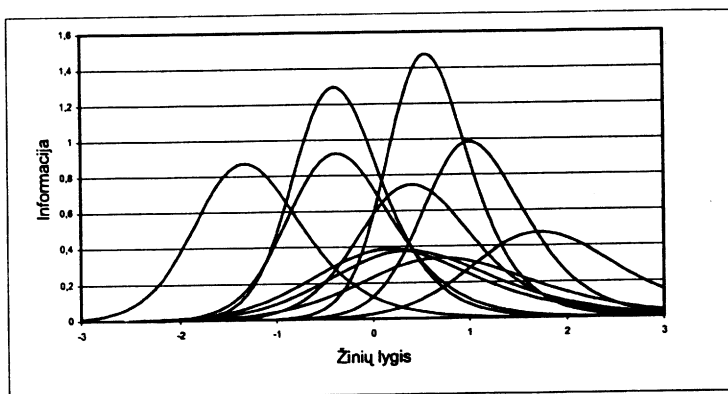
Kiekvienas uždavinys į testo informacijos kiekį „įneša“ savo informacijos dalį, nepriklausomą nuo kitų užduočių. Tuomet j -osios užduoties informacinė funkcija (Item Information Function) skirtingiems θ aprašoma taip:

$$I_j(\theta) = \frac{(P'_j(\theta))^2}{P_j(\theta) \cdot (1 - P_j(\theta))}. \quad (10)$$

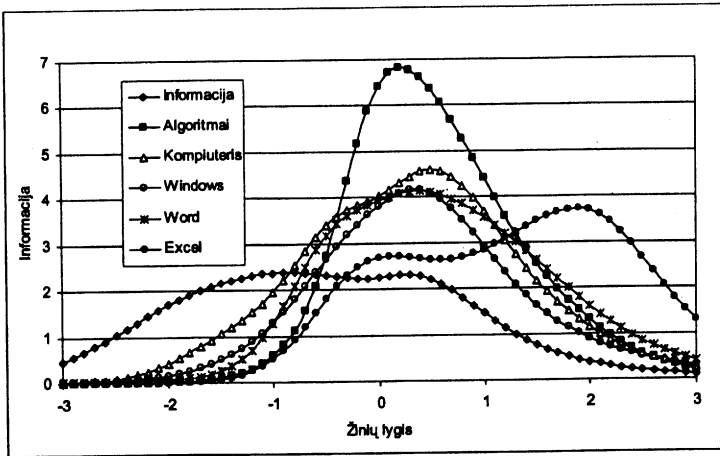
Augant skaitiklio reikšmei auga $I_j(\theta)$ ir pasiekia maksimumą charakteringosios kreivės lūžio taške, esant šiai θ_j reikšmei užduotys, kurių sunkumo parametras $\beta \sim \theta_j$ yra labiausiai informatyvios (3 pav.). Reikia pažymėti, kad ši funkcija yra simetriška maksimumo atžvilgiu, kurį pasiekia taške $\theta = \beta_j$ vienparametriniame ir dviparametriniame modelyje ir $\theta = \beta_j + g(a_j, c_j)$ – triparametriniame modelyje. Tai leidžia nustatyti užduotis, kurios yra per sunkios ar per lengvos duotam testuojamojo žinių lygiui bei tolygiai paskirstyti užduočių kiekį sutinkamai su testuojamųjų gebėjimais.

Tiek dviparametriniams ir triparametriniams modeliams testo informacijos maksimumas yra užduoties skiriamosios gebos matas. Visiems UST modeliams užduotys, turinčios didesnes a_j reikšmes, yra labiau informatyvios nei užduotys turinčios mažesnes a_j reikšmes. Užduoties skiriamoji geba atspindi užduoties pernešamos informacijos kiekį gebėjimų skalės atžvilgiu. Tačiau triparametriniame modelyje užduoties informacijos funkcijos maksimumas priklauso nuo dviejų kintamųjų – a_j ir c_j – nuspėjamumo. Augant c_j užduoties informatyvumas mažėja ir funkcijos maksimumas slenka žemyn – kreivė lėkštėja.

Grįžtant prie testo informacijos funkcijos, jos adityvumas atskirų testo užduočių atžvilgiu leidžia parinkti tinkamas užduotis siekiant optimizuoti bendrą testo informaciją. Paskaičiavus atskirų užduočių informacijos funkcijas ir nubraižius grafikus atrenkamos skirtingo sunkumo, informatyvumo ir nuspėjamumo užduotys, tolygiai pasiskirsčiusios pagal tiriamos respondentų grupės gebėjimus (3 pav.). Analizuojant paruošto testo informaciją, tikslinga išskirti pagrindines testo temas, bei patikrinti jų informacijos kreivių eigą. Jei temos informacinė kreivė žymiai nukrypsta nuo normalinio pasiskirstymo, ar turi išryškintus du maksimumus (4 pav.). Tokį testą reikia papildyti užduotimis, įnešančiomis informacijos atitinkamoje gebėjimų skalės srityje, arba padalinti testą į dvi dalis, du testus, tikrinančius skirtingo žinių bei gebėjimų lygio testuojamuosius. Geros testo užduočių atrankos rezultatas – testo informacijos funkcija turinti normalinį ar jam artimą informacijos pasiskirstymą testuojamos respondentų grupės gebėjimų intervale. Tolygus užduočių pasiskirstymas, užtikrina pakankamą testavimo rezultatų tikslumą pasirinktame gebėjimų lygyje. Testo patikimumas, priklausantis nuo testo užduočių skaičiaus, didėjant testo užduočių skaičiui, auga [3]. Iš gautų testo atlikimo duomenų vienu metu yra įvertinami uždavinių parametrai ir testuojamųjų gebėjimai. Paskaičiavus viso testo informaciją $I(\theta)$ ir nubraižius informacinės funkcijos grafiką galima koreguoti testą įvedant užduotis, kurios geriausiai įvertins norimą respondentų grupę.



3 pav. Testo klausimų informacinės funkcijos (pagal A. Birbaum modelį).



4 pav. Temų informacijos kreivės.

6. Išvados

Užduoties sprendimo teorijos naudojimas suteikia daugiau galimybių sudaryti efektyvų testą:

- įvertinti ne tik kiekvienos testo užduoties sunkumą, bet ir jos įnešamą informacijos kiekį;
- pasiekti tolygų užduočių sunkumo pasiskirstymą suderintą su testuojamųjų žinių lygiu, individualiais testuojamojo gebėjimais;
- žinant kiek užduotis įneša informacijos, galima paruošti testą kuris geriausiai įvertins norimą respondentų grupę.

Testo rezultatų analizė suteikia daugiau grįžtamosios informacijos ugdymo procesui koreguoti, jo efektyvumui didinti. Leidžia diagnozuoti ugdymo proceso spragas:

- išaiškinti blogiausiai įsisavintas temas;
- išryškinti sunkiausius klausimus.

Grįžtamosios testo informacijos analizė leidžia pažinti testuojamuosius

- įvertinti bendrus testuojamųjų gebėjimus;
- išaiškinti kiekvieno testuojamojo žinių spragas;
- išsiaiškinti testuojamojo gebėjimą atlikti skirtingo tipo užduotis.

UST metodologijos taikymas tikslingas įvairiuose testavimo proceso etapuose. Jis padėtų objektyviai įvertinant naudojamus užsienio ar savo specialistų paruoštus testus bei paruošti kokybiškus naujus testus, didinti ugdymo proceso efektyvumą išanalizavus grįžtamąją informaciją.

Testuojamojo charakteristikų ir testo užduočių parametrų vertinimas toj pačioj skalėje, leidžiantis palyginti testuojamojo žinių lygį su jam duotos užduoties sunkumo matu, sudaro pagrindą kurti šiuolaikines adaptuotas kompiuterizuoto mokymo žinių kontrolės sistemas.

Literatūra

- [1] В.С. Аванесов, *Методологические и теоретические основы тестового педагогического контроля*, Дисс. д-ра под. наук, Санкт-Петербург (1994).
- [2] B. Andziulienė, E. Valionienė, Kompiuterinių technologijų integracijos į ugdymo procesą tyrimas, *Informacijos mokslai*, 2, 18–24 (2001).
- [3] B. Andziulienė, A. Andziulis, S. Bakanauskaitė, Testų kokybinės charakteristikos ir jų vertinimas, *Informacinės technologijos – 2002*, Konferencijos pranešimų medžiaga, Kaunas, pp. 96–100.
- [3] R.K. Hambleton, H. Swaminathan, H. Rorers, *Fundamentals of Item Response Theory*, Sage Publications (1991).
- [4] G. Rasch, *Probabilistic Models for Some Intelligence and Attainment Tests*, MESSA Press (1993).
- [5] D. Wright, Solving measurement problem with the Rasch model, *J. of Educational Measurements*, 97–116 (1977).
- [6] D.G. Myers, *Psichologija. Poligrafija ir informatika*, 364–371 (2000).

Test information and it using

B. Andziulienė, A. Andziulis, S. Bakanauskaitė

The questions of quality of test and test information by using Items Response Theory (IRT) are discussed in this work. Using the IRT gives many chances to make an effective test:

- to measure not only difficult of everyone test's item, but what number of information it brings in test;
- to get an even distribution of items by its difficulty which is harmonize with abilities of testing person;
- if it's known how many item gets information in test, it's possible to make a test which will evaluate the abilities of testing person group best.

Analysė of test's results gives more reversible information for adjustment of education process.

Also it lets to diagnose the vacancy of education process:

- to detect what themes are mastered worst;
- to find the more difficult items.

The reversible information of test lets to recognize the testing persons:

- to measure the total testing person's abilities;
- to detect the vacancies of everyone testing person;
- to know how everyone of testing person can to work the different type of items.

The employing of IRT technologies is correct in different stages of testing process. It helps in objectively evaluating the test quality and increasing efficiency of education process.

The estimation of characteristics of tested persons and parameters of test items in the same scale enables to compare the ability of person with the difficult of item. It enables to create the modern potted systems of paperless learning control.