



**Vilniaus universitetas
Matematikos ir informatikos
institutas
LIETUVA**



INFORMATIKA (09 P)

INFORMATIKOS EDUKACINIŲ TESTŲ GENERAVIMAS IR VALIDUMO TYRIMAS

Lina Vinikienė

2016 m. spalį

Mokslinė ataskaita MII-DS-09P-16-14

VU Matematikos ir informatikos institutas, Akademijos g. 4, Vilnius LT-08663

www.mii.lt

Santrauka

Ataskaitoje pateikiama testų kūrimo problematika, paminimi svarbiausi aspektai, kurie turėtų būti vertinami tiriant uždavinių sudėtingumą. Tolimesniems tyrimams tikimybinė testų teorija bus naudojama kaip metodas įvertinti testo patikimumą ir sudaryti modelį, kuris aprašytų kaip teisingai turi būti parengtas testas su interaktyviais klausimais ir duomenys, leidžiantys mokytojui analizuoti mokinio pasiekimus, įvertinti testo atitikimą ugdymo programai. Tyrimu siekiama išspręsti testų validumo problemą, kai testo klausimai yra interaktyvūs

Reikšminiai žodžiai: interaktyvumas, testas, įvertinimas, patikimumas

Turinys

1	Įvadas	4
2	Testų kūrimo ir jų rezultatų analizės problematika.....	4
2.1	Testo klausimų/uždavinių interaktyvumas.....	4
2.2	Klausimo sudėtingumas	5
3	Tikimybinė testų teorija	6
4	Literatūra.....	8

1 Įvadas

Svarbiausia šių dienų problema – mokinių įvertinimas. Tradicinį vertinimą atsakant klausimus popieriaus lape keičia testavimas kompiuteriu. Tačiau kuriant klausimus įvairiomis testavimo sistemomis vis sudėtingiau pamatuoti, ar testo klausimai atitiks mokinio žinių lygį, kaip pateikti klausimą kompiuterio ekrane, kuris įvertintų mokinio žinias ir gebėjimą atsakyti klausimą teisingai nesuteikiant mokiniui galimybės atspėti teisingą atsakymą. Taip pat norint patraukti mokinių dėmesį, sudominti dėstomu dalyku, kuriami interaktyvūs klausimai. Atliekami tyrimai, kuriuose analizuojami klausimai su pasirenkamais atsakymais, tačiau trūksta informacijos, kaip kurti interaktyvius testo klausimus, uždavinius, o ypač parinkti tinkamą uždavinio sudėtingumą taip, kad šis atitiktų vertinamo mokinio žinių lygį.

2 Testų kūrimo ir jų rezultatų analizės problematika

Manoma, kad mokytojai mokykloje praleidžia 35 % savo laiko vertinant mokinius ir daugiau nei 10 % - rengiant vertinimo nurodymus, kriterijus (Su, Wang, 2010). Mokinio vertinimas priklauso nuo tinkamo klausimo parinkimo ir būdo kaip klausimas yra konstruojamas (Su, Wang, 2010). Kuriant klausimus turi būti atsižvelgiama į tokius kriterijus, kaip: laikas reikalingas atsakyti klausimą, bendras klausimų skaičius, sąvokų, kurias mokinys turi žinoti, paskirstymas, sudėtingumo lygis. Galima naudotis testavimo sistemomis, kuriose yra galimybė parinkti testo klausimus iš klausimų banko, tačiau kyla klausimas, kaip juos parinkti, kad būtų nuosekliai pagal temas vertinamos mokinių žinios. Mokytojai taip pat turi išlaikyti sąvokų naudojamų teste balansą. Kuriant testą svarbu, kad kiekvienas testo klausimas ar uždavinys būtų susijęs su specifine kurso sąvoka taip, kad teisingam atsakymui parinkti būtų naudojamos žinios (Su, Wang, 2010). Tačiau naudojantis testavimo sistemomis pildosi testo klausimų bazė, mokytojui vis sudėtingiau pasirinkti tinkamą klausimą. Kaupiant testo klausimus, dažnai jie koreguojami atsižvelgiant į kategorijas, bet ne į laiką ar mokymo programą. Norint atlikti statistinius skaičiavimus, rekomenduojama klausimų klasifikacija, raktinių žodžių parinkimas, kuris tiksliau apibūdintų pavienius testo klausimus. Klausimai turi būti išsprendžiami per kelias minutes, sąvokos atitikti klausimo sudėtingumą, klausimai nereikalauja kitokių žinių ar gebėjimų, nei kad reikia klausimui/uždaviniui išspręsti, reikalauti pažingsninio sprendimo, skatinti grįžtamąjį ryšį (Palmer, 2005).

Klausimų atsakymus gali įtakoti, ne tik tai, kaip mokytojas parenka klausimus, bet ir tai, ar mokinys yra suinteresuotas atsakyti klausimą teisingai, ar mokinys atsakydamas, pademonstruos savo žinias, ar išmoks naujų dalykų. White (1959) nurodo, kad mokiniai turi vidinę motyvaciją, kai jie išmoksta kažko naujo sprendžiant uždavinius. Čia svarbu, kad mokiniai būtų motyvuojami mąstyti, lyginti realaus gyvenimo situacijas su pateiktu uždaviniu, testo klausimu. Kai kurie tyrėjai pabrėžia, kad mokinio sėkmė sprendžiant uždavinius taip pat gali priklausyti ir nuo aplinkos, kolegų įtakos, noro pasiekti daugiau nei kiti (Palmer, 2005).

Sudol L. & Stunder C. (2010) įvardija neteisingų sprendimų priežastis tokias, kaip: mokiniai nesupranta mokomosios medžiagos, mokinys nesuprato klausimo, klausimas buvo per sudėtingas tikslinei grupei arba buvo vertinamos žinios, kurios nebuvo apibrėžtos mokomojoje medžiagoje.

2.1 Testo klausimų/uždavinių interaktyvumas

Atliekant testus ar sprendžiant uždavinius interaktyvumas apibrėžiamas kaip vartotojo veiksmo vykdymo uždavinys, pavyzdžiui įvedimas klaviatūra, spragtelėjimas pele ir

pan. Sprendžiant tokio tipo uždavinius mokiniui suteikiama galimybė ne tik veikti, bet ir bendrauti ir jis gali pasiekti aukštesnius mokymosi rezultatus. Be to šių dienų švietimo problema: raštingumas. Kuomet mokiniams suteikiama galimybė spręsti uždavinius, kuriuose pateikiama informacija paveikslėliais, animacija, mokinys nėra ribojamas rašytinės kalbos, gali laisviau gilintis į uždavinio prasmę. Gilindamiesi vaikai patys valdo informacijos srautą ir išvengia informacijos perkrovos. Interaktyvūs uždaviniai remiasi tuo, kaip mokinys supras klausimo/uždavinio turinį, kaip mokinys konstruoja sprendinį. Interaktyvių uždavinių privalumais gali būti įvardijama: didesnis patikimumas, augantis mokinių susidomėjimas ir motyvacija, didesnė galimybė pamatuoti mokinio supratimą, įvertinti mąstymą, skatinama mokinių refleksija sprendžiant uždavinius, geriau įvertinti kognityvinius ir problemų sprendimo gebėjimus (Strain-Seymour et al., 2009). Pavyzdžiui, net vaikai nuo gimimo sutelkia dėmesį į judančius objektus. Viena iš priežasčių, kodėl jų dėmesys krypsta į judančius objektus, tai kad, judėjimas padeda identifikuoti objektus. Vieni tyrėjai teigia, kas dėmesys sutelkiamas į objektą, kiti į erdvės padėtį, o dar kiti – į objekto savybes, tokias kaip spalvą, formą (Koong, Wu, 2010). Specifinių uždavinių atlikimas įtakoja mokinio supratimą, sąvokų plėtojimą.

Taip pat, jei mokytojas stebi kaip mokinys atlieka interaktyvius uždavinius, jis gali vertinti mokinio gebėjimus, kaip jis mąsto, analizuoja uždavinį. Proske et al. (2007) nurodo šias interaktyvumo funkcijas: prisijungimas, navigacija, aktyvumo skatinimas ir konstravimas, perengimas, stebėseną, dėmesio sutelkimas.

Interaktyvumas vienas iš elementų, kuris uždavinį daro suprantamesniu ir aiškesniu. Tačiau svarbu žinoti, kaip sukonstruoti uždavinį kompiuterio ekrane, kad interaktyvus uždavinys vertintų mokinio žinias ir gebėjimus.

2.2 Klausimo sudėtingumas

Klausimų sudėtingumas susijęs su mokinio patirtimi. Kuo vyresnis mokinys, tuo sudėtingesnius uždavinius jis turi mokėti spręsti (Forišek, 2010). Uždavinių sudėtingumas susijęs su reikalaujamomis kompetencijomis ir kriterijais, kaip uždavinys pateikiamas mokiniui, vartotojo sąsaja, uždavinio vaizdavimu kompiuterio ekrane ir t.t. (Dagienė, Stupuriene, 2016).

Klausimo sudėtingumo skaičiavimas apima mokinio gebėjimą atsakyti klausimą ar išspręsti uždavinį. Paprasčiausias klausimo sudėtingumo skaičiavimas, tai santykis teisingų atsakymų ir bandymų atsakyti klausimą. Kuo mažesnė ši reikšmė, tuo sudėtingesnis klausimas, kuo didesnė – tuo lengvesnis. Reikšmė 1 identifikuoja labai lengvus klausimus, o 0 – sudėtingus. Tačiau ši reikšmė glaudžiai susijusi su mokiniu ir klausimu, pavyzdžiui, gali būti ribojama bandymų skaičiaus (uždavinio sprendimo kelis kartus), vaizdavimo ekrane ir pan.

Kuriant testo klausimus svarbu, pateikti trumpus sakinius, aiškius žodžius ar frazes, aiškius apibrėžimus, ryšius tarp teksto ir paveikslėlių, jei tokie yra, vengti dviprasmybių (Pohl, Hein, 2015).

Pavyzdžiui nagrinėjant „Bebro“ konkurso uždavinius, pateikiama, kad mokiniai greičiau ir lengviau išsprendžia klausimus su pasirenkamais atsakymais, daugiau dėmesio skiria interaktyviems klausimams.

Literatūros šaltiniai, kuriuose tiriami informatikos ir informatinio mąstymo konkurso duomenys taip pat nagrinėjimas uždavinių sudėtingumas. Sudėtingumas suprantamas, kaip mokinio sėkmė, kai jis sprendžia uždavinį ir dažnai skaičiuojamas kaip teisingų atsakymų santykis. Vaniček J. (2016) išskiria šiuos aspektus, kurie įtakoja klausimų sudėtingumą: klausimo ilgis, susidomėjimas skaitomu tekstu, uždavinio formulavimas,

paveikslėlių naudojimas, iliustruojantys pavyzdžiai. Specifiniai aspektai taikomi informatikoje, kur klausimo sudėtingumas įtakojamas: srities (pavyzdžiui optimizavimo uždaviniai yra sudėtingesni nei algoritmų), sprendimo būdo (kompetencijų reikalingų uždavinio sprendimui), situacijos (kaip uždavinys susijęs su testuojamojo patirtimi), aprašymo (kodas, programa, formulės), įgūdžių (nepakankamos žodyno žinios, nežinomi terminai) (Vaniček, 2016). Autorius apibrėžia 5 indikatorius, aprašančius ir leidžiančius nuspėti informatikos konkurso uždavinių sudėtingumą: procentas mokinių, kurie atsakė klausimą teisingai; procentas mokinių, kurie neatsakė klausimo, kaip uždavinio autorius apibrėžė klausimo sudėtingumą, uždavinio sprendimo laikas, testuojamojo nuomonė (kaip mokinys įvertino klausimo sudėtingumą, kuris sudėtingiausias). Šiuo atveju rezultatus gali įtakoti: teisingo atsakymo spėjimas, nepakanka laiko atsakyti klausimą, uždavinys skirtas kitai amžiaus grupei, nevisi mokiniai pildė klausimyną apie uždavinių sudėtingumą.

Informatikos testo klausimai yra sudėtingesni, jei juose naudojami kodai, formulės, programa; sudėtinga klausimo struktūra, mokiniams sunku pasirinkti teisingą atsakymą, kai jis turi pasirinkti vieną iš galimų variantų, reikalauja daugiau skaitymo gebėjimų. Sudėtingesni tie uždaviniai, kuriuose sprendimas reikalauja prisiminimo ir naudojamas tekstas. Taip pat klausimo sudėtingumą gali įtakoti iliustracijos, diagramos, grafikai, žemėlapiai, schemos, teksto įvedimas, neigiami klausimai (Vaniček, 2016).

Nagrinėjant interaktyvių uždavinių/klausimų sprendimo ir sudėtingumo ryšį svarbu nagrinėti, kokius veiksmus mokiniai atlieka norėdami išspręsti interaktyvius uždavinius, kokio tipo klausimai gali būti taikomi įvairaus amžiaus grupėms ir kaip jie turėtų būti pateikiami.

3 Tikimybinė testų teorija

Vertinimas formuoja požiūrį apie mokinio žinias ir rezultatus, taip pat suteikia informaciją apie tai, kaip mokymo programa atitinka mokymo tikslus. Todėl būtina tirti problemas, kurios įtakoja mokinio gebėjimą atsakyti klausimą teisingai. Tikimybinė testų teorija (angl. *Item Response theory*, IRT) gali būti naudojama, kaip būdas apibrėžti situaciją, kuri įtakoja mokinių sprendimus.

IRT naudojama įvertinti testo patikimumą ir analizuoti kaip susijęs klausimo sudėtingumas su mokinio gebėjimais.

IRT tai metodas, kuriuo galima įvertinti tikimybę, kaip teisingai mokinys atsakys testo klausimą. Yra žinomi trys metodai: vieno, dviejų ir trijų parametrų metodai. Vieno parametro metodas yra gerai žinomas Rasch modelis. Naudojamas neparimetrinių testų patikimumui tirti (Berges, Hubwieser, 2015). Šis modelis tiria dichotominių klausimų kategoriją (požymio buvimą arba nebuvimą), pavyzdžiui *tiesa/netiesa* klausimus. Jame yra vienintelis parametras – uždavinio sudėtingumo koeficientas. Pagrindinė šio modelio idėja yra ta, kad tikimybė atsakyti klausimą teisingai apibrėžiama latentiniu konstruktu (mokinio gebėjimu atsakyti testo klausimą) ir uždavinio sudėtingumu, t.y. parodo, kaip testuojamasis nutolęs nuo sudėtingumo, kokio sudėtingumo uždavinys yra pagal mokinio turimą žinių lygį (Berges, Hubwieser, 2015). Latentinis konstruktas gali būti vienmatis arba daugiamatis. Jo negalima stebėti ar pamatuoti tiesiogiai, todėl reikia nustatyti jo rodiklius. Priklausomai nuo psichometrinio konstrukto gali būti naudojami skirtingi IRT matematiniai modeliai. Yra du klausimo analizės metodai: atrenkama aibė vienodų klausimų atitinkanti vieną latentinį psichometrinių konstrukto ir klausimų bankas turi atitikti bandymus ir modelį.

Klausimų atsakymai gali būti vaizduojami atsakymų matrica (Klausimas x Individas). IRT vienas iš būdų analizuoti testo patikimumą ir validumą, tačiau tai nėra vertinimo metodas (Sudol, Studer, 2010). Be to prastai suformuluotas klausimas gali įtakoti mokinio įvertinimą ir gebėjimą atsakyti klausimą teisingai. Taip pat įtaką gali daryti mokinių atsakančių testo klausimus skaičius. Sudol L. & Stunder C. (2010) nurodo, kad IRT parodo tikslesnius matavimus, kai yra kuo didesnis testuojamųjų skaičius priklausomai nuo klausimų.

Priešingai nei klasikinė testų teorija, IRT metodas naudojamas atskiriems klausimams tirti, o ne apibendrinti testo rezultatus (Sudol, Studer, 2010). Rasch modelis dažnai naudojamas įvertinti sąlygas intervalinėje skalėje (Schmettow, Vietze, 2008).

4 Literatūra

1. Berges, M., Hubwieser, P. (2015). Evaluation of Source Code with Item Response Theory. In: ITiCSE'15 Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education. 51-56
2. Dagienė, V., Stupurienė, G.: Bebras- a Sustainable Community Building Model for the Concept Based Learning of Informatics and Computational Thinking. In Informatics in Education, Vol. 15, No.1, 25-44 (2016)
3. Forišek, M.: The Difficulty of Programming Contests Increases. In Teaching Fundamentals Concepts of Informatics. 4th International Conference on Informatics in Secondary Schools - Evolution and Perspectives, ISSEP 2010, Zurich, Switzerland, pp. 72-85 (2010)
4. Koong, Ch., Wu, CH (2010). The applicability of interactive item templates in varied knowledge types. In: Computers & Education, 781-801
5. Palmer, D.: Research report: A Motivation view of Constructivist-informed Teaching. In International Journal of Science Education, Vol.27, No.15,16 December, 1853-1881 (2005)
6. Pohl, W., Hein, H.W., Aspects of quality in the presentation of informatics challenge tasks, LNCS, vol. 9378, 21–32, 2015.
7. Proske, A., Narciss, S., & Körndler, H. (2007). Interactivity and learners' achievement in web-based learning. Journal of Interactive Learning Research, 18(4), 511–531.
8. Schmettow, M., Vietze, W. (2008). Introducing Item response theory for measuring usability inspection processes. In: CHI '08 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 893-902
9. Su, C.Y., Wang, T.I. (2010). Construction and analysis of educational assessments using knowledge maps with weight appraisal of concepts. In: Computer & Education, 1300-1311
10. Sudol, A., L. Studer, C. (2010). Analyzing test items: using item response theory to validate assessments. In: SIGCSE'10 Proceedings of 41st ACM technical symposium on Computer science education, ACM new York, 436-440
11. Strain-Seymour, E., Way, W., Dolan, R. P.: Strategies and processes for developing innovative Items in Large-Scale Assessments. Pearson Education, Inc. Via Internet:
<http://images.pearsonassessments.com/images/tmrs/StrategiesandProcessesforDevelopingInnovativeItems.pdf> (2009)
12. Vaniček, J. (2016). What makes situational informatics task difficult? In: International Conference on Informatics in Schools: Situation, Evolution, and Perspectives (pp. 90-101). LNCS, Vol. 9973, Springer International Publishing
13. White, R. W.: Motivation reconsidered: The concept of competence. Psychological Review, Vol. 66, pp. 297–333 (1959)