

# PSÜHHOMEETRIA JA RIIGIEKSAMID

*Einar Rull*

Riiklik Eksami- ja Kvalifikatsioonikeskus

## 1. Millega tegeleb inimene psühholoogi seisukohast?

Inimene on *laisk tunnetaja* ([4], lk 36). Pilt sündmustest, milles ta osaleb, on varasema kogemuse põhjal valmis konstrueeritud ning meeleorganitega ta vaid üldjoontes kontrollib tegelikkuse vastavust mudelitele. Isegi kui me lihtsalt ringi vaatame, kasutame oma ajast mahukamalt informatsiooni kui seda välismaailmast lisandub. Sellisel asjade korraldusel on evolutsiooniline eelis, sest mõtlemine võtab väga palju energiat, mida on teinekord vaja kasutada ka mujal, et elada üle raskeid aegu.

Kui meie igapäevaelu kulgeb pingevabamas keskkonnas, on meil vähem vajadust kontrollida, mis meie ümber toimub, ning hüppevalmis olla. Me ei pea otsima iga sõna või žesti tagant varjatud signaale või tähendusi ning meil on vähem stressi. See võimaldab meil enam kontsentreeruda konstruktiivsele igapäevatööle. Pingevabamates kultuurides lähebki enam energiat mõistlikele asjadele ning me peaksime seega hindama seda, kui teised pakuvad meile võimalust laisa tunnetajana elades koondada oma tähelepanu Maslow' püramiidi ülemisele otsale. Kultuurides, kus see puudub, elatakse hoopis teistsugust elu.

Inimesed tunnetavad seda maailma, mida füüsik esitab laine tihedusfunktsiooni kaudu, meelte loodud audiovisuaalse paradiisi-na. Me oleme oma paradiisis elamisega harjunud ning peame seda igapäevaseks. See aitab meil suurepäraselt maailmas orienteeruda, ehkki eksisteerib vaid meie peas.

---

Artikkel põhineb 2010. aastal XII Eesti matemaatika päevadel 28. juunil Türi-salus peetud ettekandel „Psühhomeetria ja riigieksamid“.

Hilisemad psühholoogid on leidnud, et väga laisk tunnetaja inimene alati ei olegi. Eriti, kui tema enda huvid on mängus. Teatud informatsiooni on inimene valmis ülimal põhjalikkusega töötlemata ning isegi omal algatusel uusi stiimuleid otsima. See on siis nõnda nimetatud *motiveeritud strateegi* ([4], lk 37) fenomen. Matemaatikas kirjeldab elu seda tahku hästi mänguteooria, kus mängijad käituvad strateegiliselt läbimõeldult ning ennustavad teiste mängijate tegutsemist mitu sammu ette. *Laisk tunnetaja* avaldub aga jälle selles, et me ei soovi alati saada tingimata maksimaalset kasumit, vaid rahuldume sellisega, mis kindlustab enam-vähem meie elu jätkusuutlikkuse.

Kõike seda informatsiooni, mis on seotud inimese enese identiteediga ja eneseesitlusega, töötleb inimene motiveeritud strateegina läbi erilise põhjalikkusega, sest tahab jätta enesest „viimase peal“ head muljet. Sedasi toimub rolliressursi kogumine. Me hindame ka ise teisi inimesi selle põhjal, kui muljetavaldav on nende rolliressurss. Rolliressursi rikastamiseks vahetavad inimesed omavahel väärtuslikke suhtlemiskogemusi ja see võimalus innustab meid suhtlema. Selles kommunikatsioonis pannakse samas paika ka võimusuhted vastavalt sellele, kellel on rohkem, kellel on vähem teistele anda.

Üks kõige hilisemaid väljapakutud mudelid on *paralleelne piirangute ületamine* ([1], lk 211). Mudelite maailma ja reaalsel keskkonda vaadeldakse kahe paralleelse maailmana, mida vaheldumisi külastatakse. Inimene elab tihti fundamentalistina pea „õigetel“ põhimõtetes ning keskkonnast tulevaid vasturääkivusi õnnestub üsna suurel määral ignoreerida. Kuni selle hetkeni, mil saab selgeks, et vanaviisi jätkamine võib teda teistega võrreldes viletsamasse seisusse jätta. Selle järel hakkab ta hoolega keskkonda uurima ning käituma enam keskkondlike stiimulite järgi. Ka see käitumine võib osutada strateegiliselt ebaefektiivseks, sest hea süsteemi puudumise tõttu jäävad mõned olulised stiimulid tähele panemata ning nende kiireks märkamiseks on vaja uusi efektiivsemaid mudelid. Eelmiste mudelite tähenduses erandid hakkavad sobituma uutesse mudelitesse ning inimene hakkab edasi fundamentalistina käituma uute põhimõtete järgi.

Psühholoogid leiavad, et teatud inimestel on kalduvus jääda väga jäigalt kinni oma mudelitesse ning nad võivad sellega endale liiga teha ja ka teistele inimestele lausa ohtlikuks kujuneda. Paindlik inimene suudab ka keskkonnaga arvestada. Seega see inimene, kes

ainult väärtuste järgi käitub, ei tarvitse alati käituda kõige mõistlikumalt. Samas tekitab segadust ka teine äärmus, kui selgroog ja identiteet kaovad. Konkreetsetele meile iga päev vastutulevatele inimestele kui pika evolutsiooni läbi teinud strateegiate kandjatele on omane toimetada paralleelselt usinasti nii põhimõtete maailmas kui ka reaalses keskkonnas.

Psühholoogia uute saavutuste valguses oleme kõik *amatöörteadlased*. Ka need, kes päristeadlased ei ole, ehitavad oma maailmapildi üles meetoditele, mis on nende jaoks hoolsalt läbi kaalutud, ning vaidlevad teistega lähenemisviiside pädevuse üle. Uus ja üsna radikaalne võib tunduda ka arvamus, et paljud üsna komplitseeritud amatöörteadlase eeldused on meil olemas evolutsioonilistel põhjustel sünnipäraselt. Näiteks püüd otsida ja kasutada mudeleid, ruumiline ettekujutusvõime, kõne äraõppimise ja kasutamise võime, arvutamisevõime kümne piires, võime panna ennast teiste olukorda, võime õppida ainult jälgides ilma harjutamata.

## **2. Mis toimub inimese peas probleemide lahendamisel?**

Konkreetne mõte käivitab inimese peas neuronite võrgustiku. Erinevad mõtted käivitavad inimese peas erinevaid võrgustikke. Erinevatel inimestel on eri ajupiirkonnad erinevalt arenenud. Nendel väidetest põhineb psühhomeetria, mis annab seletuse, miks inimeste võimed on erinevat laadi probleemide lahendamisel erinevad või miks inimesed liiguvad õppimisel edasi erineva kiirusega.

Ajus võib eristada jämedalt viit piirkonda, mis on seotud ruumilise, verbaalse, loogilise, kinesteetilise (motoorset tegevust puudutava) ja assotsiatiivse mõtlemisega. Peenemal vaatlemisel võib neid leida isegi 69 või mõne allika põhjal rohkemgi.

Konkreetset mõtet mõeldes käivitub viie piirkonna neuronvõrkude kombinatsioon. Kui nende piirkondade suhe on fikseeritud, siis on tegemist ühe konkreetse *kognitiivse profiiliga*, millel saab õpilaste teadmisi võrrelda või järjestada. Teiste sõnadega: kui meil on küsimused, mis on sama kognitiivse profiiliga, siis õpilaste edukuse järjekord nendele küsimustele vastates tuleb sama, ehkki ülesannete raskused võivad olla erinevad. Teistsuguse kognitiivse profiiliga ülesannete puhul võib õpilaste järjekord tulla veidi teistsugune. Kui me järjestame õpilasi geomeetriaülesannete lahendamise oskuse järgi, siis ei lange see kokku algebraülesannete lahendamise järjekorraga, sest ülesannetele vastavad ajus erinevad kogni-

tiivsed profiilid. Vastavaid võimekusi nendel profiilidel nimetatakse latentseteks, sest neid ei saa otse jälgida, vaid tuleb testide alusel kaudselt arvutada.

Üksikküsimusele vastamisel käivitub üks kognitiivne profiil. Keerulisemate, mitmesammuliste küsimuste puhul jagatakse see järjestikusteks üksikküsimusteks, kusjuures erinevatel üksikküsimustel ei tarvitse reeglina olla sama kognitiivne profiil ning siis on need tegelikult erinevatel skaaladel. Kaks kognitiivset profiili on erinevad, kui leidub kaks õpilast, kes on esimesele profiilile vastaval teljel ühes järjestuses ning teisele profiilile vastaval teljel teises järjestuses.

Reaalselt toimub suur hulk mõtlemisest poolautomaatselt ning lisaks järjestikusele infotöötlemisele esineb ka paralleeltöötlust. Selle modelleerimine on väga keeruline ning pedagoogilistes mõõtmistes kasutatav psühhomeetria selliseid ülesandeid ei käsitle. Et kasutada psühhomeetriat, tuleb ülesanne jagada üksikoperatsioonideks ning vastavad tulemused fikseerida.

Kui meil mingil põhjusel ei õnnestu ülesannet üksikküsimusteks lahti harutada, on võimalik ometi faktoranalüüsi kasutades eri kognitiivsetele profiilidele vastavad faktorskoorid kätte saada.

Kui meil on tegemist ülesannete ja testide põhjal tekkinud õpilaste järjestustega, siis saame neid järjestusi võrrelda ning vaadata, kui palju lõpptulemuse järjestust sisaldub üksikute ülesannete loodud järjestustes. Halvemal juhul on üksikülesande poolt genereeritud järjestus vastupidine testi genereeritud järjestusele. Siis on tavaliselt tegemist vigase ülesandega, millele nõrgemad õpilased vastavad paremini kui tugevad. Kui üksikülesande ja testi järjestused pole üldse seotud, siis me ütleme, et üksikküsimus mõõdab mingit muud konstrukti kui test tervikuna. Psühhomeetriliselt hea test peaks koosnema ülesannetest, mis tekitavad kõik sama õpilaste järjestuse ja seega mõõdavad sama konstrukti. Õppekava seisukohalt oleks selline test aga halb, sest mõõdab paremal juhul vaid 10% õppekava materjalist, sest muu materjal tekitab erineva õpilaste järjestuse.

PISA uurimuse taolised testid koosnevad faktoranalüüsi mõttes ühedimensionaalsetest testidest (iga ülesanne tekitab sama järjekorra) ning siis on lihtne riike täpselt omavahel võrrelda. Ehkki samavõrd õige oleks mõõta ka mingil sellest erineval teljel ning riikide järjekord võiks tulla mingis ulatuses erinev.

Riigieksamite puhul saaks ühedimensionaalsuse nõude puhul kontrollida vaid umbes 10% õppekavast ning seetõttu lepitakse õppekaval põhinevates testides sellega, et kõik ülesanded ei tekita sama õpilaste järjekorda. Küll aga jälgitakse, et erinevate aastate lõikes oleksid kogu testi summaarsed faktorkaalud võrreldavad. Teiste sõnadega, vaadatakse, et algebra- ja geomeetriaülesandeid oleks igal aastal samas vahekorras.

Psühhomeetrias mõõdetakse õpilase võimekust sellise ülesande raskusega, mille ta lahendab ära 50% tõenäosusega.

Ajuvigastusega patsientide puhul on aju teised osad võimelised üle võtma vigastatud osade tegevuse. Seetõttu väidavad mõned eksperimentaalpsühholoogid, et psühhomeetria ei tarvitse inimese mõtlemist empiirilisel eriti adekvaatselt kirjeldada. Teada on näiteks matemaatikas eriandekaid inimesi, kellel on tõsiseid raskusi igapäevaeluga toimetulekul. Samas aga standardiseeritakse ja skooritakse maailmas igal aastal miljoneid teste psühhomeetrialetuginedes.

Kui puudub seos eksimise tõenäosuse ja õpilase võimekuse vahel, siis ei saa psühhomeetria abil tulemusi arvutada. Näiteks võib tuua Suurbritannias mõned aastad tagasi valminud 6. klassi elektroonilise IKT-oskuste testi, kus kontoritarkvara kasutamisel tehtud eksimused ja õpilaste võimekused ei olnud statistiliselt seotud.

### **3. Kolme lähenduse mudel**

Milliseid teadmisi, oskusi, vilumusi või pädevusi me ka ei tötaks mõõta, tegelikult mõõdame vaid latentset võimekusfaktorit vastaval kognitiivsel profiilil ja see allub psühhomeetria seaduspärasustele. Sest oskused ja muu ei võimalda konstruktide vahel statistilist vahetegemist. Antud teema kõige hilisem ja sügavam käsitlus on Carrolli pakutud faktoranalüüsil põhinev kognitiivsete võimete kolmestraatumiline teooria (*three-stratum theory of human cognitive abilities*). Enamik inimese maksimaalse käitumise mõõtmistest on klassifitseeritav selle alusel. Pedagoogilistes mõõtmistes kasutatakse seega faktoranalüüsil põhinevat kolme lähenduse mudelit. Esimeses lähenduses võib eeldada, et kõik ülesanded, sõltumata sellest, millise kaaluga millist aju osa kasutatakse, tekitavad täpselt sama õpilaste järjekorra. Teiste sõnadega tähendab see käibetõde, et kellel on head eeldused ühes aines, on tihti head eeldused ka teistes ainetes.

Teises lähenduses eristuvad viis faktorit ning kolmandas 69 erinevat sõltumatut faktorit, nagu eespool mainitud. Haridusalastes mõõtmistes otsitakse igal kognitiivsel skaalal ülesannet, millele õpilane vastab 50% tõenäosusega õigesti.

Seda ühte kõige jämedamas lähenduses avalduvat faktorit nimetatakse ka generaalfaktoriks ehk g-faktoriks. See omakorda jaguneb kaasasündinud *muutliku intelligentsuse* faktoriks ja kultuuri poolt mõjutatud *kristalliseerunud intelligentsuse* faktoriks. Kui on tegemist inimestega samast kultuurikeskkonnast, siis võib mudelist kultuuri poolt mõjutatud faktori välja jätta. Eelnimetatud g-faktor on inimese informatsiooni töötlemise kiirus ning see avaldub kõige puhtamalt näiteks funktsionaalses lugemisoskuses. Arusaadavalt korreleeruvad kõik 5 järgmises lähenduses nähtavat faktorit omavahel üsna tugevalt.

Kaasasündinud g-faktor, mida õppimine eriti ei mõjuta, vaid mida mõjutavad pigem bioloogilised faktorid nagu toitumus ning vanus, ei jäta eriti ruumi õppimise mõjule. Ometi väidavad PISA uuringu ideoloogid, et kolm aastat keskmist õppimist annab meile ühe standardhälbe võrra kõrgema tulemuse. Me kasutame õppimise tagajärjel keerukamaid tunnetuslikke skeeme, ehkki meie „protssori taksagedus“ jääb samaks.

Kui me vaatame üleriiklikku statistikat, siis iga konkreetse võimekuse puhul on üle ja alla keskmise õppivaid õpilasi enam vähem võrdselt ning seetõttu õppimist (lisaks keskmisele õppimisele) me histogrammides üldse ei näegi. Näeme ainult võimekuse jagunemist, mis on moduleeritud eksamite struktuuri poolt.

Et õppematerjal on üles ehitatud lihtsamalt keerulisemale, siis neid õpilasi, kes aru saavad, jääb lõpupoole vähemaks. Õpilaste arv sõltub pigem võimekusest kui õppimisest, sest üle- ja alaõppijaid on võrdselt. Andekamatele õpilastele piisab ühekordsest kuulmisest ning nad mäletavad materjali kogu elu, sest fakt satub olemasolevasse süsteemi. Vähemandekad võtavad asja omaks pingsa mõtlemise järel ning unustavad kergemini või raskemini. Veel vähem andekad võivad materjali pähe õppida, kuid see ununeb kähku ning nad ei tarvitse osata oma teadmisi testis õieti kasutadagi.

Punktide jaotus eksamil ei väljenda mitte sedavõrd õppekava läbimist protsentides, vaid õpilaste latentse võimekusparameetri jaotust kontingendis. Seega ei ole eriti mõtet ka seisukohal, et õpilane, kes on õppekava läbinud, peaks saama eksamil 100 punkti. Eksamil võib ju kasutada teatud piirides õppekaval põhinevaid lihtsa-

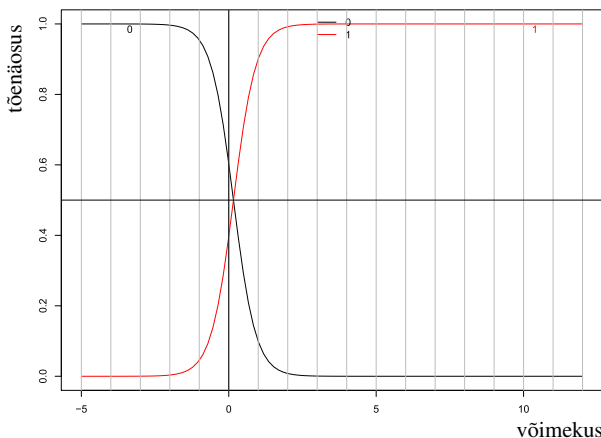
maid ja keerulisemaid ülesandeid. Küll aga saab näiteks hinnata, kui palju on õpilase puhul vaja veel keskmist õppimist, et ta võiks mingi konkreetse tõenäosusega saada eksamil näiteks 100 punkti.

Maksimumtulemuse saanud õpilase puhul teame vaid, et ta on üle teatud nivoo, kuid mitte seda, kui palju. Saja punkti puhul nagu ka 0 punkti puhul on õpilane tegelikult mõõtmisvahemikust väljas. Kuna meie konkreetsetel eksamitel on skaala tihti ülevalt kokku surutud, siis võib 100 punktile vastata intervallskaalal meie arvutuste järgi isegi kuni 120 punkti.

#### **4. Tõenäosuslik testiteooria (*Item Response Theory, IRT*)**

Olgu meil tegemist ühe konkreetse kognitiivse profiiliga ehk ühte konstrukti mõõtvat skaalaga, kus õpilased on üheselt kindlas järjekorras. Vaatleme skaalal ühte juhusliku raskusega ülesannet, näiteks joonisel 1 esitatud 2007. aasta keemia riigieksami variandi A ülesannet 30, mille eest võis saada kas 1 või 0 punkti sõltuvalt sellest, kas vastus oli õige või vale.

Kõigepealt konstrueerime  $x$ -telje. Selle nullkohaks on tinglikult kogu keemia eksami variandi A keskmine tulemus 62 punkti (vastab sellise õpilase võimekusele, kes sai eksamil 62 punkti ning me hindame õpilase tõenäosust vastata sellele küsimusele õigesti). Ühikuks on kogutulemuse standardhälve 21 punkti. Võime-



Joonis 1. 2007. aasta keemia riigieksami variandi A ülesande 30 IRT-kõverad: tõenäosus vastata küsimusele õigesti (tõusev joon) ja valesti (langev joon) kogu eksami  $z$ -skaalal. Arvutused sooritatud Anti Teepeere poolt statistikapaketiga R, kasutades psühhomeetriaalidest.

kusega 1 õpilane sai keemia eksamil 83 punkti ning me hindame selles punktis tema tõenäosust vastata väljavalitud küsimusele õigesti. Kui keemia eksami histogramm oleks täpselt Gaussi joon, saaksime lihtsasti arvutada vastavuse riigieksami punktide ja meie  $x$ -telje vahel, mille järgi näiteks  $-2$  vastab  $x$ -teljel  $62 - 2 \cdot 21 = 20$  punktile.

Kui me aga oletame, et ülesannete raskuste ebahütlase valiku tõttu eksamitöös on histogramm Gaussi joonega võrreldes moonunud (küll aga on säilinud õpilaste järjekord), on meil ometi võimalik õpilaste tulemused Gaussi jaotusele vastavaks ümber arvutada. Nimetame seda üleminekuks 0–100-punktiliselt eksamiskaalalt Gaussi jaotuse keskväärtuse ja standardhälbe määratud  $-\infty$  kuni  $+\infty$  vahemikuga intervallskaalale.

## 5. Üleminek intervallskaalale

Et reaalselt on eksamite histogramm üsna lähedane Gaussi jaotusele, siis võime esimeses lähenduses vaadelda nullpunkti kui keskmise tulemusega õpilase tulemust ja kohta  $-2$  kui sellise õpilase tulemust, kes on saanud 20 toorpunkti.

Kui me soovime seda joont Gaussi joonega lähendada, siis me peaksime keskmisele punktile, mis vastab mediaanile, omistama keskväärtuse. Mediaan on keemia eksamil 63 punkti. Edasi mõtleme Gaussi joone aluse pindala kohta käivatele seaduspärasustele. Kolmandik õpilastest jääb keskmisest tulemusest standardhälbe võrra ülespoole ja kolmandik standardhälbe võrra allapoole. Vastavatele piiri peal olevatele õpilastele omistamegi intervallskaala tulemused  $62+21$  ja  $62-21$  punkti ning kõigile ülejäänud punktidele vahepeal ja kaugemal aitab meil leida intervallskaalalised väärtused van der Waerdeni teisendus. IRT-programmid teevad seda skaala teisendamist tavaliselt automaatselt. Sellega oleme läinud nullist sajani ulatuvalt riigieksamite skaalalt sisuliselt üle Gaussi jaotusega seotud teljele, mis ulatub  $-\infty$  kuni  $\infty$ . Lisaks arvestame, et meie  $x$ -teljel oli tegemist lisaks intervallskaalale ka sisuliselt  $z$ -skaalaga, kus keskväärtus on 0 ja standardhälve 1.

## 6. IRT-jooned

Järgnevalt vaatleme joonisel 1 sellel võimekusteljel õpilase tõenäosust vastata 2007. aasta keemia riigieksami A variandi 30. küsimu-



sele kas valesti (langev joon) või õigesti (tõusev joon). Kokku on 100% ning vahepeal on punkt, kus jooned lõikuvad ning on 50% tõenäosus vastata nii õigesti kui valesti. Vastavat võimekust nimeatakse ülesande raskuseks. Tegemist on veidi raskemat sorti ülesandega, sest keskmine (võimekusega 0) õpilane vastab sellele õigesti vaid umbes 40% tõenäosusega. Ülesandele õigesti vastamise tõenäosus  $P(\Theta)$  avaldub lihtsa valemiga

$$P(\Theta) = \frac{1}{1 + e^{-a(b-\Theta)}},$$

kus  $\Theta$  on õpilase võimekus skaalal, mida kutsutakse ka logitite skaalaks, sest tegu on eksponentfunktsiooniga. Suurus  $b$  on ülesande raskus mõõdetuna õpilase võimekuse ühikutes ja  $a$  positiivne eristusindeks, mis väljendab, kui järsult ülesanne eristab. Ühele skaalale kuuluvate ülesannete puhul on eristusindeks konstantne.

Urime eelpooltoodud võrrandit. Väga madala võimekuse korral on  $\Theta$  absoluutväärtuselt suur negatiivne arv, eksponentfunktsiooni väärtuseks tuleb suur positiivne arv ja kogu murru väärtus on seega ligikaudu 0. Väga kõrge võimekuse korral on eksponentfunktsiooni väärtus peaaegu võrdne nulliga ning kogu avaldis ühega. Kui õpilase võimekus  $\Theta$  on võrdne ülesande raskusega, siis on eksponentfunktsiooni väärtus võrdne ühega ning tõenäosus ülesandele õigesti vastata on 0,5.

Kui meil on tegemist sama kognitiivse profiiliga, siis kordaja  $a$  on konstantne ning mudelit kutsutakse üheparameetriliseks IRT-mudeliks ehk 1LP-mudeliks ning vastavat IRT-analüüsi haru Raschi analüüsiks Taani matemaatiku Georg Raschi järgi. Kui meil ei ole ühedimensionaalsuse nõue täpselt täidetud, saame ometi teatud täpsusega arvutada, kuid me peame eristusindeksil muutuda laskma. Kui olulised on kaks parameetrit, siis sel juhul nimetatakse mudelit kaheparameetriliseks IRT-mudeliks ehk 2LP-mudeliks. Reaalses elus ei tarvitse eksperimentaalne IRT-kõver täpselt mudelile vastata; mudelile vastavust hinnatakse  $\chi^2$ -meetodiga. Mittevastavus viitab sisulistele probleemidele ülesandega nagu mitmetähenduslikkus või segasus ülesande tekstis.

Kui me eeldame, näiteks valikvastustega küsimuste puhul, äraarvamist, peame sisse tooma äraarvamise indeksi  $c$  (võrdne valikute arvu pöördväärtusega) ning meil on tegemist kolmeparameetrilise IRT-mudeliga ehk 3LP-mudeliga, mis avaldub eelmise mudeli

kaudu järgnevalt:

$$P'(\Theta) = c + (1 - c)P(\Theta).$$

Kui meil on näiteks neli valikvastust, siis tõenäosus õigesti vastata alla 1/4 ei langegi. Õpilasel peab nullskoori saamiseks lisaks teadmiste puudumisele olema ka ebaõnne.

Valikvastustega ülesandeid saab arvutada ka nii, et me joonistame igale peibutusvastusele oma IRT-kõvera, mis on siis negatiivse eristusindeksiga, nagu joonise 1 tumedam joon, sest paremad õpilased valivad seda vastust harvemini kui nõrgad õpilased. Erinevad peibutusvastused erinevad selle poolest, kui võimekas peab olema õpilane, et seda vastust enam mitte valida. Igal peibutusvastusel peaks olema ideaaljuhul erinev võimekusvahemik, kus ta suudab peibutada.

IRT-kõverate parameetrite muutmise mõju saab uurida Eksamikeskuse e-eksamite wikikeskkonnas aadressiga <http://www.ekk.edu.ee/~e-eksam/irt/>. GeoGebra-rakenduse autoriks on Anti Teepere.

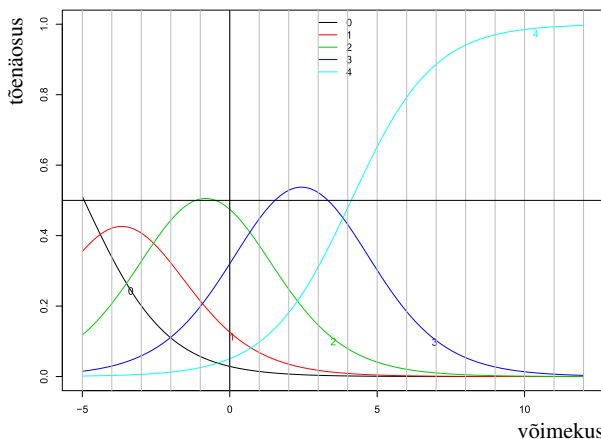
## 7. Polütoomid

Eelpool vaatlesime õige ja vale vastusega dihhotoomsetele ülesannetele vastavat IRT-analüüsi. Mõnikord aga antakse testides osaliselt õigete vastuste eest osalisi punkte ning vastavat mudelit kutsutakse siis polütoomseks mudeliks. IRT-analüüsis tuleb siis välja joonistada näiteks kolm joont – tõenäosus saada 0 punkti, pool punkti ja terve punkt.

Üsna sarnase pildiga on ka täisarvulise hindamisskaalaga ülesanded ehk hinnatava vastusega ülesanded, nagu näiteks võõrkeele suulise osa või kirjandi puhul. Matemaatikaülesanne kuuluks siia siis, kui kogu ülesande ulatuses mõõdetaks sama võimekust. Siis on meil tegemist tõenäosusjoontega iga üksiku skoori saamiseks ehk polütoomidega.

Vaatame joonisel 2 kujutatud 2007. aasta keemia riigieksami variandi A esimest ülesannet, kus oli võimalik saada kas 0, 1, 2, 3 või 4 punkti.

Kõikide tõenäosuste summa on iga võimekuse puhul võrdne ühega, sest mingi arvu punkte on iga võimekusega õpilane saanud. Nullpunkt  $x$ -teljel vastab kogu eksami keskmisele tulemusele



Joonis 2. 2007. aasta keemia riigieksami variandi A ülesande 1 polütoomid. Polütoomid vastavad tõenäosusele saada ülesande eest mingi kindel arv punkte kogu eksami  $z$ -skaalal. Arvutused sooritatud Anti Teepere poolt statistikapaketiga R, kasutades psühhomeetriaalidest.

ja ühik standardhälbele. Võimekusskaala on intervallskaala. Keskmise õpilane sai selle ülesande eest kõige tõenäolisemalt 2 punkti.

Ülesande *raskus* on polütoomide korral nullinda ja viimase polütoomi lõikepunktis. Sellest punktist mõõdetakse kahe kõrvutise polütoomi lõikepunkti kaugusi; neid kaugusi nimetatakse astakuteks. Teine paralleelselt kasutatav lähenemine polütoomidega juhtumit analüüsida on defineerida Thurstone'i lävendid. IRT-programmid pakuvad tavaliselt automaatselt mõlemat statistikut. Näiteks 2 punkti saamise Thurstone'i lävend on võimekus, mille puhul õpilane saab 50% tõenäosusega 2 punkti või rohkem. Dihhotoomse ülesande puhul langevad ülesande raskus ja ainuke lävend kokku.

## 8. Suurima marginaalse tõepära (*Maximum Marginal Likelihood, MML*) meetod

See meetod võimaldab meil arvutada histogrammi, kui on teada üksikküsimuste parameetrid ning õpilaste võimekuse jaotus, mis võib põhimõtteliselt olla suvaline, ka Gaussi jaotusest erinev.

IRT-meetod võimaldas iga õpilase teadaoleva võimekuse korral arvutada tõenäosuse vastata õigesti ülesandele parameetritega  $a$  ja  $b$  ( $\Theta \rightarrow p(a, b)$ ). Üldisemal juhul saame tõenäosused iga erineva skoori jaoks, kui on mitmepunktiline ülesanne.

Õpilaste võimekuste jaotuskõver iseloomustab eri tüüpi võimekuste sagedusjaotusi  $P(\Theta)$  meid huvitavas hulgas. Integreerides esimest vastavust üle kõigi eksamitegijate võimekuste jaotuse, saame ülesande võimalike skooride sagedused. Liites edasi kokku kõik ülesanded, saame kogu testi võimalike skooride sagedused, mis on histogramm. Kui meil on antud ülesande histogramm, uuritakse selle esimese tuletise ekstreemumkohtasid, et leida ülesande raskust. Sellest on saanud see meetod oma nime.

## 9. Gruppide võrdlemine

Kui üks grupp teeb kahte erinevat testi, siis on võimalik tulemuste põhjal võrrelda kahe testi raskust. Kui kaks gruppi teevad ühte testi, siis on võimalik võrrelda gruppide tugevust.

Riigieksamite puhul teeb igal aastal erinev grupp erinevat ülesannetekomplekti ning keskmise tulemuse varieeruvust põhjustavad potentsiaalselt mõlemad efektid.

Kui kahes testis kasutatakse ühte kokkulangevat ülesannet (ankurülesannet), siis on mõlema testi puhul pärast toimumist võimalik arvutada selle ülesande raskus ja eristusindeks. Seejärel võime nihutada ja venitada lineaarselt teise eksami skaalat nii, et selle ülesande parameetrid langeksid kokku tema parameetritega esimeses testis. Selle teisendusega satuvad ka ülejäänud teise testi ülesannete parameetrid ja õpilaste võimekusparameetrid esimese testiga samale skaalale. Sellist teisendust saab aga teha vaid siis, kui mõlemas testis on erinevat liiki ülesannete vahekord sama ehk teiste sõnadega: hinnatavate konstruktide faktorprofiilid on samad. Vastasel juhul ainult nihutamisesest ja venitamisest ei piisa, sest muutub ka õpilaste järjekord. Õppekaval põhinevate testide puhul ei ole tavaliselt kognitiivsete profiilide samasuse tingimus täidetud ning sellist teisendust ei kasutata – erinevalt näiteks PISA testist, kus on alati tegemist sama kognitiivse skaalaga.

Kui võrrelda õpilase mingi aine eksamisooritust tema kõigi ainete keskmise eksamisooritusega, siis see võib osutada rikkalikumaks indikaatoriks kui üksikud eksamisooritused eraldi vaadelduna. Kõigi eksamite keskmine varieerub üldjuhul läbi aastate vähem kui ühe eksami keskmine. Kõigi ainete keskmine on seega parema puudumisel etalon, millega erinevate aastate eksameid võrrelda. Kui kõigi ainete keskmine on mingil aastal matemaatikat valinud kontingendil kõrgem, siis võib suurema tõenäosusega oletada,

et on tegemist umbes samavõrd võimekama kontingendiga ning täpselt samasuguse eksami puhul peaksid nad saama ka samavõrra parema tulemuse.

See lähenemine eeldab, et erinevaid eksameid valinud õpilaste suhtarvud aastast aastasse kõiguvad üsna vähe. Üks õpilane ei tee küll kõiki eksameid, küll aga teevad kõik õpilased kokku kõiki eksameid üsna samas vahekorras.

Teiseks tuleb arvestada, et keeleeksamitel saavutatakse üldjuhul kõrgemaid tulemusi kui teistes ainetes. Küll aga on keele komponendi osakaal aastast aastasse kõikide eksamite puhul keskmise eksamite arvutamisel enam-vähem sama. Riigikeel on kõigile vene õppekeelega õpilastele kohustuslik, inglise keele eksami valivad peaaegu kõik õpilased ning ülejäänud keeleeksamite sooritajate arvude summa on väiksem kui kõige väiksema sooritajate arvuga üldharidusaine, füüsika sooritajate arv.

Eksamiained on erinevad ning see erinevus tuleb ka selgelt ilmsiks võrdluses keskmise eksamiga. Meid aga huvitab edaspidi, kuidas see erinevus käitub eri aastate lõikes, kus keeleainete mõju on enam-vähem sama.

Kui me üritame võrrelda eri ainete reaalseid eksamiskaalasisid vastava intervallskaalaga, mille saaksime Gaussi jaotusele üleminekuga, siis näeksime, et vahemikus 20–90 punkti on meil tegemist ligikaudu intervallskaalaga. Samas aga vastab sajale punktile pigem 120 punkti ning seega keskmise tulemuse arvutamine skaala ülemises otsas on alahinnatud nagu alumises otsas ülehinnatud. Üle 90 punkti ja alla 20 punkti saab aga siiski küllaltki väike hulk õpilasi ning ka viga keskmise arvutamisel ei ole liiga suur selleks, et seda võrdlust siiski mitte üritada teha.

See lähenemine arvestab kontingendi võimekuse varieeruvust aastast aastasse paremini kui vaid keskmiste tulemuste kasutamine, mis kuulutab varjatud eeldusena kontingendi tugevuse varieeruvuse lausa nulliks, põhjustades sellega suuremat ebatäpsust kui käesolev lähenemine.

## **10. Keskmine eksam**

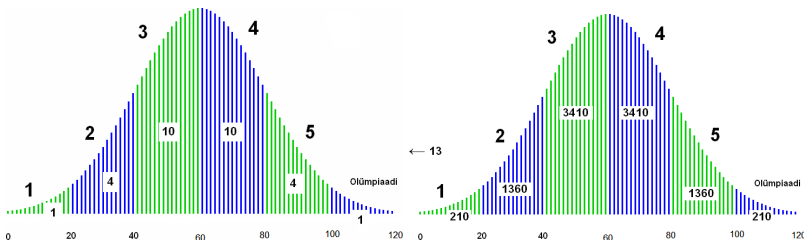
Kui me vaatame kõiki eksameid üle kõikide aastate, siis keskmine eksamitulemus on umbes 60 punkti ja standardhälve 20 punkti.

Joonisel 3 vasakul on kujutatud Gaussi joon parameetritega 60 ja 20 riigieksamite skaalal. Skaala on välja joonistatud 120 punkti-

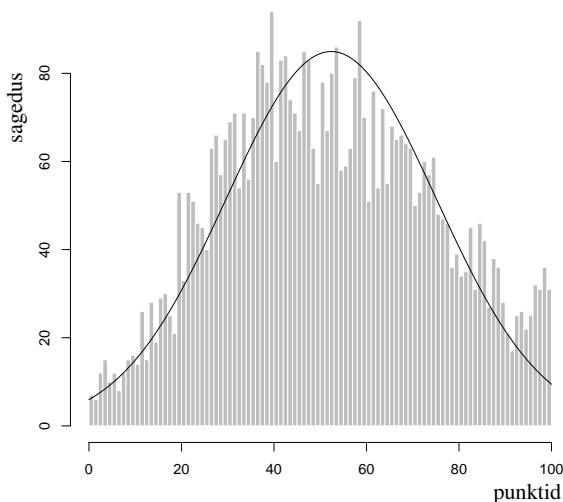
ni, sest Gaussi jaotuse järgi leiduks ka sellise võimekusega õpilasi. Küll aga ei saa selle raskusega ülesandeid testidesse panna, sest neid peetakse tihti mitte ainult huvitavateks ülesanneteks, vaid ka olümpiaadiülesanneteks. Võrdleme seda jaotust õpilaste võimekuse jaotusega klassis, kus on 30 õpilast.

Ühe kolmandiku õpilaste tulemus on vahemikus 60–80 punkti ning tunnis saaksid need õpilased hindeks „nelja“, neid oleks kolmekümnest umbes kümme. Kolmandiku õpilaste tulemus on vahemikus 40–60 punkti ning tunnis saavad need õpilased tavaliselt hindeks „kolme“, neid on samuti kolmekümnest kümme. Ülemine kuuendik õpilastest saab hindeks „viie“ (viis kolmekümnest õpilasest) ja alumine kuuendik „kahe“ (samuti viis kolmekümnest õpilasest). Üks õpilane on olümpiaaditasemel ning üks langeb isegi alla hinde „kaks“ alampiiri. Et „kaheste“ õpilaste arvu vähendada, on koolis leiutatud „kolm miinus“, mida kasutatakse tulemuste vahemikus 30st 40 punktini ning umbes kaks õpilast kolmekümnest saavad „kolm miinuse“. Ülejäänud kolm nõrgimat õpilast jäävad suvetööle ning kaks nendest jõuavad järgmisse klassi; üks õpilane, kelle tulemus on vahemikus 0–20 punkti, jääb klassikursust kordama. Seega on riigieksamite skaala ligilähedaselt ühilduv koolis kasutatava skaalaga. Kui üks õpilane 30st ehk 2,5% õpilastest jääb klassikursust kordama, siis ollakse sellega harjunud. Kui aga 10000 inglise keele eksamit teinud õpilasest 2,5% kukub läbi, tähendab see üle 200 õpilase läbikukkumist ning see tekitab ajakirjanduses rasvaseid pealkirju.

Kui me üritame täpsemalt välja tuua õpilaste arvud standardhälvete kaupa 10000 õpilase puhul (vastab umbes inglise keele eksa-



Joonis 3. Gaussi jaotus keskväärtusega 60 ja standardhלבega 20, mis vastab kõikide aastate ja kõikide eksamite keskmisele eksamisooritusele. Vasakul: koolis igapäevaselt kasutatavad hinded on lähedase jaotusega. Paremal: õpilaste arvud standardhלבete kaupa, kui õpilasi oleks 10000. Reaalsed skaalad on ülalt 100 punkti peale lükatud.



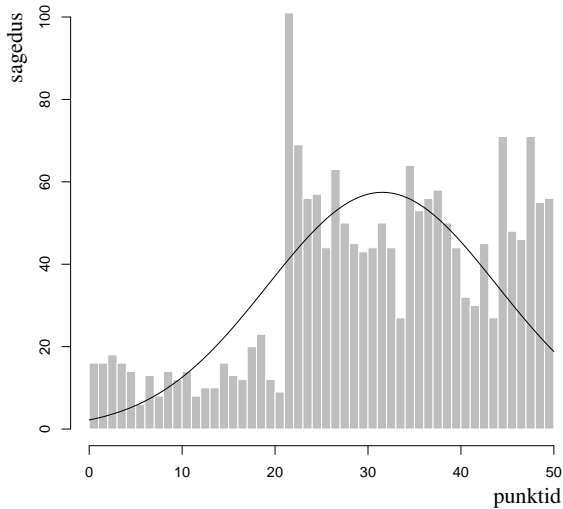
$$\bar{x} = 52,4, \sigma = 22,7, N = 4838$$

Joonis 4. Matemaatika 2009. aasta riigieksami tulemuste histogramm.

mile), siis saame pildi nagu joonisel 3 paremal. Nagu näeme jooniselt 3, peaks Gaussi lähenduse järgi saama 13 õpilast tulemuse alla nulli. Üle 100 punkti saaks 210 õpilast ning ebatäpse mõõtmise tagajärjel on nad surutud allapoole. 13 õpilast saaks intervallskaalal üle 120 punkti ning paar õpilast isegi üle 140 punkti, kui eksamisse nende mõõtmiseks vastavad ülesanded leida.

Järgnevalt vaatleme jooniselt 4 võrdluseks 2009. aasta matemaatika tulemuste histogrammi koos keskmise tulemuse ja standardhälbe põhjal konstrueeritud Gaussi joonega – on sarnasusi ja erinevusi.

Õpetlik võib olla jooniselt 5 vaadata 2010. aasta matemaatika põhikooli lõpueksami histogrammi, mis kahjuks ei erine oluliselt teiste ainete ja teiste aastate põhikooli lõpueksami tulemuste histogrammidest. Eksamit hinnatakse koolis. Kõigepealt torkavad silma hindapiiril olevate õpilaste venitamisest tekkinud „korstnad“. Õpilaste võimekus peaks ka põhikoolis Gaussi järgi jagunema, kuid 10–20 punkti vahemikus on histogrammis teooriaga raskesti seletatav auk, mis on põhjustatud veidi kõrgest lävendist. Sama murega hakkab ilmselt maadlema ka tulevane matemaatika kohustuslik riigieksam, kus lävend on 50% maksimumist. Et test eristaks hästi, peaks Gaussi joon mõistlikus ulatuses skaalale ära mahtuma.



$$\bar{x} = 31,5, \sigma = 12,4, N = 1779$$

Joonis 5. Matemaatika 2010. aasta põhikooli lõpueksami tulemuste histogramm.

## 11. ECTS-skaala

Euroopa kõrgkoolide vahel punktide ülekandmise ECTS-skaalat kasutavad Norra ja Taani kõrgkoolid ka vastuvõtul, kui on tegu teiste riikide gümnaasiumi lõputulemuste tõlgendamisega. ECTS-skaala tulemusi arvutatakse keskmise riigieksami tulemuse kohta, kusjuures arvesse lähevad ainult need õpilased, kellel on kõigi riigieksamite tulemus vähemalt 20 punkti ning ka muud gümnaasiumi lõpetamiseks esitatavad koolipoolsed tingimused täidetud.

Tabelis 1 on esitatud ECTS-skaala hindepiirid aastatel 2006–

Tabel 1. Õpilaste keskmiste riigieksamitulemuste jaotuse ECTS-skaala lävendid aastatel 2006–2010. Arvesse on võetud ainult õpilased, kelle kõigi riigieksamite tulemus on vähemalt 20 punkti.

|      | A<br>90%–100% | B<br>65%–90% | C<br>35%–65% | D<br>10%–35% | E<br>0%–10% |
|------|---------------|--------------|--------------|--------------|-------------|
| 2006 | 80            | 67           | 55           | 42           | 20          |
| 2007 | 80            | 67           | 56           | 44           | 20          |
| 2008 | 80            | 68           | 56           | 43           | 20          |
| 2009 | 81            | 68           | 56           | 43           | 20          |
| 2010 | 82            | 69           | 57           | 44           | 20          |



Tabel 2. Kontingentide tugevused aastatel 2003–2010 kohustuslikes ainetes (eesti- ja venekeelsed kirjandid ning riigikeel) gümnaasiumiõpilaste puhul; \* = valikeksam.

|      | Kirjand | Vene kirjand | Riigikeel |
|------|---------|--------------|-----------|
| 2003 | 57,82   | 59,59        | 59,61     |
| 2004 | 59,77   | 60,99        | 60,91     |
| 2005 | 61,62   | 59,47        | 59,54     |
| 2006 | 60,64   | 59,58        | 59,49     |
| 2007 | 60,78   | 60,53        | 60,44     |
| 2008 | 61,48   | 58,22        | 58,07     |
| 2009 | 61,46   | 57,19        | 57,29     |
| 2010 | 63,24   | 60,24*       | 59,05     |

2010. Näeme muuhulgas, et need on püsinud aastast aastasse enam-vähem paigal, kõikudes nelja aasta jooksul vaid paari punkti võrra üles-alla. Kui keskmine eksamitulemus käitub aastast aastasse üsna stabiilselt, siis peaks see sobima heaks ankruks erinevate aastate üksikute eksamite võrdlemisel.

ECTS-skaala hindepiirid keskmise eksami jaoks on esitatud eksamikeskuse koduleheküljel aadressil <http://www.ekk.edu.ee/exam-results-on-ects-scale>.

Keskmise eksami käitumist ajas võiks alternatiivselt analüüsida ka aritmeetilise keskmise muutumise alusel aastate lõikes. Vaatleme ainult gümnaasiumiõpilasi, sest riigikeele kohustuslikkus ka vene kutsekoolides moonutaks veidi pilti. Samuti on välja jäetud varemlõpetanud ning eksternid. Kogu õpilaskontingendi saaksime kätte, kui vaataksime kirjandit, venekeelset kirjandit ja riigikeelt, sest need on ainukesed kohustuslikud eksamid.

Keskmisses tulemuses näeme tabelis 2 veidi suuremaid kui ühepunktiseid fluktuatsioone, kuid need on väiksemad kui fluktuatsioonid keskmiste tulemuste vahel eri aastate lõikes.

Algaastatel oli vene õpilaste kontingent vaevumärgatavalt tugevam kui eesti õpilaste kontingent, kuid viimastel aastatel on vahe pigem kaldunud eesti õpilaste kasuks. Põhimõtteliselt oleks niisugust efekti võimalik seletada ka sellega, et eesti õpilased valivad eksameid, kus nad on tugevamad, ning vene õpilased eksameid, kus nad on nõrgemad. Kuid et riigikeele eksamil saavutatakse peaaegu 10 punkti võrra kõrgemaid tulemusi kui kõigil teistel eksamitel ning eesti õppekeelega õpilastel ei ole sellega võrreldavat eksamit, siis võib selle hüpoteesi pigem kõrvale jätta. Pigem on asi

selles, et tugevamad vene koolide õpilased on asunud eestikeelsesesse õppesse või on kuidagi teisiti vene õppekeelega koolide tase langenud (näiteks õpetajate puudulikuma enesetäiendamise tõttu). Tabelist 2 näeme ka, et eesti ja vene õppekeelega keskmised ei liigu sünkroonis. Järelikult ei ole see efekt tingitud eksamite raskusest, vaid kontingendi tugevuse kõikumisest ning eksamite võrdluses tuleks kontingendi tugevuse efekt välja arvata. Kui eesti õppekeele puhul võiksime täheldada eksamitulemuste inflatsiooni, siis seda ei tähelda me vene õppekeele puhul ja järelikult peaksid vastavate kontingentide võimekused liikuma eri suundades.

## 12. Milliseid eksameid teevad tugevamad õpilased?

Kui tahame teada, kui tugevad õpilased valivad ühte või teist eksamit, siis võtame õpilast iseloomustavaks tugevuse parameetriks tema kõigi eksamitulemuste keskmise ning vaatleme, kui suur on selle parameetri keskmine mingit ainet valinud õpilaste puhul. Nime-tame seda kontingendi tugevuseks. Vaatleme tabelis 3 neid keskmisi kõigis ainetes kahe viimase aasta jooksul (siinsetes rehkendustes on kutsekoolid ja varem lõpetanud lisatud). Konkreetse õpilase puhul on tema keskmise eksamitulemuse arvutamise juures võetud arvesse ka eelmistel aastatel sooritatud eksamid, sest geograafia

Tabel 3. 2009. ja 2010. aasta riigieksamite kontingentide tugevused (koos kutsekoolide õpilaste ja varem lõpetanutega).

| 2009 kontingent |        | 2010 kontingent |       |
|-----------------|--------|-----------------|-------|
| Prantsuse keel  | 77,39  | Prantsuse keel  | 74,50 |
| Saksa keel      | 72,19  | Saksa keel      | 71,03 |
| Keemia          | 70,00* | Füüsika         | 69,68 |
| Ajalugu         | 66,95  | Ajalugu         | 67,56 |
| Füüsika         | 68,57  | Inglise keel    | 65,57 |
| Inglise keel    | 65,14  | Matemaatika     | 65,02 |
| Matemaatika     | 63,34  | Keemia          | 63,85 |
| Vene keel       | 62,00  | Bioloogia       | 62,01 |
| Kirjand         | 61,23  | <i>Keskmine</i> | 61,82 |
| Bioloogia       | 62,21  | Kirjand         | 61,50 |
| Geograafia      | 61,06* | Vene keel       | 60,54 |
| <i>Keskmine</i> | 59,71  | Ühiskonnaõpetus | 60,08 |
| Ühiskonnaõpetus | 59,57  | Vene kirjand    | 59,61 |
| Vene kirjand    | 57,37  | Geograafia      | 59,10 |
| Riigikeel       | 50,33  | Riigikeel       | 51,42 |

ja keemia eksamit tehaksegi tihti 10. või 11. klassis. Seega näiteks sellel aastal sooritatud geograafia eksami kontingendi tugevus muutub veel ka järgmisel ja ülejärgmisel aastal, sest kõik geograafia eksamit valinud õpilased ei ole veel gümnaasiumi lõpetanud. Eksami ümbertegemine läheb arvesse kui veel üks täiendav eksam.

Kui eelpool nimetatud põhjusel keemia ja geograafia eksamid vaatluse alt kõrvale jätta, siis näeme ülejäänud ainete kontingendi tugevustes kuni paaripunktiseid kõikumisi, mille suund on tihti erinev võrreldes keskmise tulemuse kõikumise suunaga.

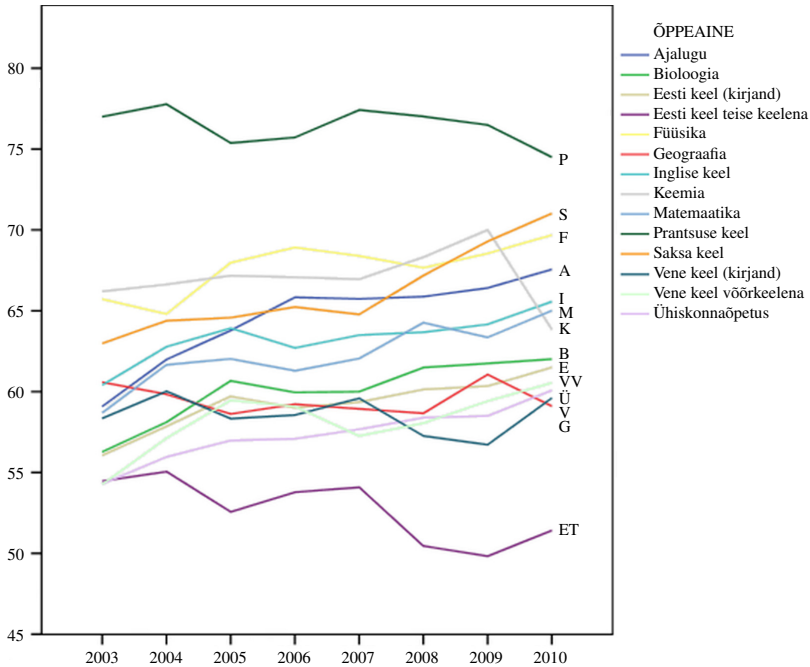
Prantsuse keele eksamit sooritanud kontingent on olnud igal aastal kõige tugevam. Varasematel aastatel on vahe võrreldes teisel kohal oleva ainega olnud isegi üle 10 punkti. Nüüd on see vahe kahanenud alla 5 punkti. Samavõrra on langenud ka prantsuse keele keskmine tulemus. See ei tähenda, et prantsuse keele õpetamisel oleks midagi muutunud, pigem on eksamit hakanud valima nõrgemad õpilased. Samas on prantsuse keele eksamit valinud õpilased teinud tihti matemaatika eksamit üle 5 punkti nõrgemini kui füüsikat valinud õpilased, kes on kontingentide edetabelis veidi tagapool. Matemaatika ongi aga olnud ainus erand, kõigis ülejäänud aintes garanteerib kontingendi tugevus prantsuse keele eksami tegijatele esikoha – kõiki ülejäänud eksameid teeb prantsuse keelt valinud kontingent kõige tugevamini.

Nagu näeme edaspidi, on suur osa eksami keskmise tulemuse kõikumisest seletatav kontingendi tugevuse kõikumisega ning mis sellest üle jääb, alles selle kohta võiksime rakendada neid seletamisviise, mida tavaliselt kasutame puhta keskmise tulemuse puhul.

Et vaadelda kontingentide tugevuse kõikumist pikemas ajavahemikus, on need tugevused kujutatud graafikutena joonisel 6. Õppeainete esitähed on kirjutatud õppeainete graafikute juurde, et hõlbustada nende leidmist.

Prantsuse keele eksamit valinud kontingent on olnud alati kõige tugevam. Keemia, füüsika, saksa keel ja ajalugu on võistelnud oma kontingentidega järgmiste kohtade eest. Riigikeele kontingent on olnud kõige nõrgem selletõttu, et see on kohustuslik kutsekoolis. Traditsiooniliselt valivad nõrgad õpilased ühiskonnaõpetuse ja geograafia eksamit. Vene keeles kirjutatud kirjandi kontingendi tugevuses toimus viimasel aastal selge hüpe ülespoole, sest see eksam on muutunud vabatahtlikuks.

Ülejäänud ainete kontingentide tugevused muutuvad vahemikus 55–65 punkti ning erinevad üksteisest üsna vähe.



Joonis 6. Ainete kontingentide tugevuse muutumine aastatel 2003–2010. Aine nime esitähthälg näitab aine kontingendi tugevust 2010. aastal.

### 13. Milliseid eksameid sooritatakse nõrgemini või tugevamini kui kõiki eksameid keskmiselt?

Me saame võrrelda mingi aine kontingendi tugevust ja kõigi ainete keskmist tulemust, kui see puudutab samu õpilasi. Vahe näitab, kui võrd on konkreetset eksamit tehtud paremini või nõrgemini kui kõiki eksameid keskmiselt (seda ainet valinud õpilaste puhul). Näiteks matemaatika ja füüsika eksamite keskmisi tulemusi ei saaks me selles mõttes otse võrrelda seetõttu, et aineid on teinud erineva tugevusega kontingendid. Matemaatika ja füüsika eksamite tegijate ühisosa teeb pikemas perspektiivis vaadelduna mõlemat eksamit samale keskmisele tulemusele – umbes 70 punktile.

Vaatleme 2010. aasta riigieksamite sooritamist tabelis 4. Ained on tabelis järjestatud selle põhjal, kui suur on vahe selle eksami tulemuse ja kõigi ainete keskmise tulemuse vahel.

Mõned kommentaarid tabeli kohta. Asjaolu, et vene keele kui võõrkeele eksamil saavutatakse 15,34 punkti võrra tugevam tulemus kui kõigil eksamitel keskmiselt, on seletatav sellega, et ainet

Tabel 4. 2010. aasta riigieksamate ainete kontingendi tugevus, keskmine tulemus ja nende vahe (koos kutsekoolide õpilaste ja varem lõpetanutega).

| Aine            | Kontingendi tugevus | Aine keskmine | Vahe  |
|-----------------|---------------------|---------------|-------|
| Vene keel       | 60,54               | 75,88         | 15,34 |
| Saksa keel      | 71,03               | 77,98         | 6,95  |
| Riigikeel       | 51,42               | 57,60         | 6,18  |
| Ühiskonnaõpetus | 60,08               | 63,96         | 3,87  |
| Inglise keel    | 65,57               | 69,28         | 3,71  |
| Geograafia      | 59,10               | 59,31         | 0,20  |
| Vene kirjand    | 59,61               | 59,69         | 0,08  |
| Ajalugu         | 67,56               | 65,58         | -1,99 |
| Prantsuse keel  | 74,50               | 72,41         | -2,09 |
| Bioloogia       | 62,01               | 59,18         | -2,83 |
| Kirjand         | 61,50               | 58,61         | -2,89 |
| Keemia          | 63,85               | 59,45         | -4,40 |
| Füüsika         | 69,68               | 63,95         | -5,74 |
| Matemaatika     | 65,02               | 58,43         | -6,58 |

valivad ulatuslikult õpilased, kelle kodus räägitakse ka vene keelt. Saksa keele eksami kontingent on sellega võrreldes üle 10 punkti tugevam, kuid kõigi ainete keskmine tulemus on vaid kahe punkti võrra kõrgem. Eksamitulemuse kõrgus võrreldes kontingendi tugevusega on saksa keele puhul seletatav sellega, et õpilased on olnud motiveeritud saksa keelt õppima enam kui teisi aineid. Samas kõige tugevama kontingendiga prantsuse keele eksami tegijad sooritasid kõiki aineid keskmiselt isegi 2 punkti nõrgemini kui prantsuse keelt ennast. Prantsuse keele eksamil pole keegi kunagi saanud maksimumpunkte, vaatamata kontingendi tugevusele, tingituna eksamitegijate vähesusest.

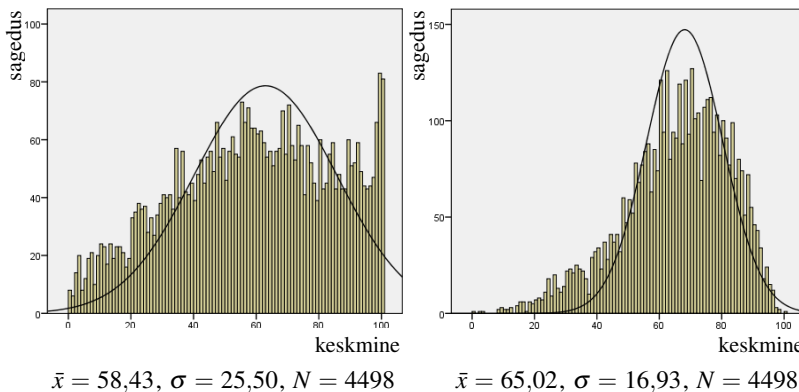
Riigikeele suurt vahet saab seletada sellega, et igapäevase suhtlemise käigus omandatakse eesti keelt umbes aasta keskmise õppimise jagu rohkem kui koolis ametlikult ette nähtud.

Ühiskonnaõpetuse puhul on õigem vaadelda eesti ja vene õppekeelega õpilasi eraldi. Keskmiste tulemuste vahe on õppekeeleli 13,46 punkti ning see on igal aastal jäänud 10 punkti kanti. Eesti õpilastele oli eksam kontingendi tugevusega võrreldes 4,31 punkti liiga lihtne, vene õpilastele 1,19 punkti liiga raske (aastal 2010 vene õppekeelega eksamit teinute arv vähenes ning aasta varem oli eksam vene õppekeelega õpilastele isegi 4,66 punkti liiga raske). Eesti ja vene õppekeelega õpilaste kontingentide vahe oli

7,98 punkti (vene õpilastest valisid eksamit nõrgemad õpilased) ning seega jääb tulemuste erinevuseks 5,5 punkti, kui oleksid olnud võrdsed kontingendid. Keskmise tulemuse erinevus 13,46 punkti on seletatav 7,98 punkti ulatuses kontingentide tugevuste erinevusega ning 5,5 punkti ulatuses õppimise erinevusega.

Geograafia eksam oli lõpuklassi õpilastele vaid 0,86 punkti raskem kui kõik eksamid keskmiselt, varasemate klasside õpilastele aga 4,83 punkti raskem. Ka geograafia eksamile on olnud traditsiooniliselt omane vene ja eesti õppekeelega õpilaste tulemuste suur erinevus. Eesti õppekeelega õpilastele oli eksam 4,04 punkti liiga raske ning vene õppekeelega õpilastele 13,17 punkti liiga raske. Õppekeelte kontingentide vahe oli 7,6 punkti. Tabelist me aga näeme geograafias vahet ainult 12. klassi õpilastele, kes on teinud ka muid eksameid; teiste õpilaste puhul võrdub kontingendi keskmine eksami enda keskmisega, sest nad pole veel teisi eksameid valinud.

Huvitav on võrrelda eesti- ja venekeelset kirjandit. Kohustuslike eksameid sooritatakse alati nõrgemalt kui vabatahtlikke. Selle tõttu on eesti kirjandi keskmine tulemus alla kontingendi tugevuse 2,89 punkti ning vene keele kirjandi puhul on keskmise ja kontingendi tugevuse näitajad praktiliselt võrdsed. Varasematel aastatel on kontingendi ja keskmise vahe olnud mõlemal kirjandil samas kandis.



Joonis 7. Vasakul: aasta 2010 matemaatika riigieksami histogramm. Paremal: aastal 2010 matemaatika riigieksamil osalenud õpilaste kõigi eksamite keskmise tulemuse (kontingendi tugevuse) histogramm.

Keemia ja füüsika on olnud teiste eksamitega võrreldes traditsiooniliselt rasked eksamid. Vahe kontingendi tugevusega on keemia ja füüsika eksami puhul olnud varasematel aastatel umbes 2 punkti kandis, sellel aastal aga isegi 5–6 punkti kandis. Matemaatika jaoks on aga selle aasta vahe 6,58 punkti isegi väike, sest varasematel aastatel on vahe ületanud ka 10 punkti piiri.

Kui võrrelda matemaatika 2010 tulemuse ja kontingendi histogramme (joonis 7), siis näeme, et suurele osale õpilastest on matemaatika eksam isegi liiga lihtne. Umbes 10% õpilaste jaoks oleks õige teha hoopis raskem eksam. Keskmiste õpilaste jaoks on aga eksam juba umbes 10 punkti raske ning kutsekoolide õpilaste matemaatika sooritus on üle 20 punkti allpool kõigi eksamite sooritust. Kanada uurijad on leidnud, et kõige nõrgema viiendiku õpilaste matemaatika tulemused pärast 6. klassi isegi halvenevad järjest, sest nad ei õpi matemaatikat oma võimete tasemel.

Selleks, et detailsemalt vaadelda keskmise tulemuse ja kontingendi vahe kõikumist, on need vahed erinevatel aastatel toodud tabelis 5.

Matemaatikat valinud õpilased on saanud aastate lõikes kõigi eksamite keskmisena 61 kuni 65 punkti. Ootuspärane oleks näha matemaatikatumust kõikumis samas vahemikus ning vahesid

Tabel 5. Millisel määral tehti erinevatel aastatel konkreetse aine riigieksamit paremini kui kõiki eksameid keskmiselt? Tärniga märgitud arvude puhul on õpilaste keskmise eksamitulemuse arvutamisel võetud arvesse vaid samal aastal toimunud eksamid.

|                 | 2010  | 2009   | 2008   | 2007   | 2006    |
|-----------------|-------|--------|--------|--------|---------|
| Vene keel       | 15,34 | 17,37  | 14,14  | 16,83  | 21,65*  |
| Saksa keel      | 6,95  | 4,41   | 6,24   | 5,54   | 5,17*   |
| Riigikeel       | 6,18  | 5,95   | 5,09   | 8,79   | 9,71*   |
| Ühiskonnaõpetus | 3,87  | 1,98   | 4,48   | 0,63   | 0,74*   |
| Inglise keel    | 3,71  | 4,75   | 3,24   | 5,17   | 1,71*   |
| Geograafia      | 0,20  | 1,39*  | -4,83* | -2,58* | 0,68*   |
| Vene kirjand    | -0,08 | -1,05  | -0,90  | -1,79  | -1,70*  |
| Ajalugu         | -1,99 | -4,33  | -5,35  | -1,13  | 0,65*   |
| Prantsuse keel  | -2,09 | -1,62  | -0,70  | -1,74  | -0,54*  |
| Bioloogia       | -2,83 | 0,39   | -1,01  | -0,96  | 2,23*   |
| Kirjand         | -2,89 | -2,59  | -2,93  | -2,66  | -3,09*  |
| Keemia          | -4,40 | -1,93* | -2,82* | -4,99* | -1,57*  |
| Füüsika         | -5,74 | -2,31  | -4,19  | 1,41   | 0,31*   |
| Matemaatika     | -6,58 | -11,10 | -5,75  | -11,74 | -10,57* |

nulli ligiduses. Võib tekkida soov standardiseerida eksamid viisil, et see vahe oleks igal aastal null. Muidugi kaob siis praegune selgus. Õppekava seab piirid ka kõvasti kergema eksami tegemisele.

Kogu nihutamise protseduur näeks siis välja selline, et kõigepealt läheksime eksamiskaalaga seotud vahemikust  $[0, 100]$  punkti Gaussi joone keskmise ja standardhälbega seotud intervallskaalale vahemikus  $(-\infty, \infty)$ , siis teostaksime arvutatud nihke ja teisendaksime nihutatud tulemused Gaussi joonega seotud joonelt tagasi  $[0, 100]$  skaalale eelmise teisenduse pöördteisendusega ning vajadusel ümardaksime tulemusi.

Me võime kasutada ka mõnevõrra keerukamat teisendust. Kujutame eksamitulemuse sõltuvuse eksamite keskmisest tulemusest hajuvusdiagrammil, leiame regressioonivõrrandi tõusu ja algordinaadi ning teisendame tulemusi nii, et teisendatud tulemustel oleks tõus 1 ja algordinaat 0. Keerulisemal juhul võiksime lineaarse regressiooni asemel kasutada täiendavalt ka astmefunktsioone. Eksaminandide vahesus võib põhjustada mõnede eksamite puhul selle teisenduse ebatäpsust.

Neid teisendusi saaks teha eeldusel, et meil on erinevatel aastatel tegemist ainult erineva raskusega eksamitöödega ning muid efekte võime ignoreerida. See ei tarvitse aga nii olla. Näiteks ühiskonnaõpetuse ja geograafia eksamil käituvad vene ja eesti õppekeelega grupid erinevalt – ühele on eksam umbes 5 punkti liiga raske, teistele 5 punkti liiga kerge. Kui eesti ja vene õppekeelega õpilaste arvude suhe muutub, siis muutub ka vahe.

#### **14. Mis aine eksamit valinud õpilased teevad matemaatika eksamit kõige kõrgemale tulemusele?**

Tabelis 6 vasakul on ained järjestatud kontingendi tugevuse alusel. Vaatluse all on kõik õpilased aastatel 2003–2009. Paremal on eksamikontingendid järjestatud selle põhjal, kuidas nendesse kuuluvad õpilased on sooritanud matemaatika eksamit. Kõige tublimad on füüsika eksamit valinud õpilased peaaegu 70-punktise tulemusega. Tabeli parema osa esimeses veerus on matemaatika tulemus ja teises veerus matemaatika tulemuse ja vastava aine eksamitulemuse vahe seda ainet valinud õpilaste puhul. Näiteks füüsikat valinud õpilased said keskmise matemaatikatumuse 69,91 ning see oli 0,59 punkti vähem kui füüsika eksami keskmine tulemus ajavahe-  
mikul 2003–2009.



Tabel 6. Vasakul: kontingendi tugevused, kui arvesse on võetud kõik eksamid aastatel 2003–2009. Paremal: eelnev tabel ümber järjestatud selle põhjal, kui tugevasti on õpilased nimetatud kontingentidest sooritanud matemaatika eksamit (esimene veerg); teises veerus on matemaatika ja reanimele vastava aine eksamitulemuste vahe.

|                 |              |                 |              |              |
|-----------------|--------------|-----------------|--------------|--------------|
| Prantsuse keel  | 76,11        | Füüsika         | 69,91        | −0,59        |
| Keemia          | 67,08        | Keemia          | 64,62        | −4,80        |
| Füüsika         | 67,24        | Prantsuse keel  | 63,59        | −10,75       |
| Saksa keel      | 65,03        | Ajalugu         | 59,09        | −8,84        |
| Ajalugu         | 64,31        | Saksa keel      | 58,37        | −8,78        |
| Inglise keel    | 63,09        | Bioloogia       | 57,35        | −8,32        |
| Matemaatika     | 61,93        | Inglise keel    | 56,35        | −8,73        |
| <i>Keskmine</i> | <i>60,08</i> | <i>Keskmine</i> | <i>55,07</i> | <i>−8,65</i> |
| Bioloogia       | 59,84        | Kirjand         | 53,88        | −8,50        |
| Kirjand         | 59,05        | Vene kirjand    | 53,60        | −9,11        |
| Geograafia      | 59,04        | Geograafia      | 53,30        | −9,91        |
| Vene kirjand    | 58,17        | Riigikeel       | 53,14        | −9,24        |
| Vene keel       | 58,12        | Matemaatika     | 52,91        | −9,01        |
| Ühiskonnaõpetus | 57,28        | Ühiskonnaõpetus | 50,42        | −10,77       |
| Riigikeel       | 52,52        | Vene keel       | 48,88        | −13,03       |

## 15. Koolide võrdlemine

Eksami kontingendi tugevuse, keskmise tulemuse ja nende vahe saame arvutada ka iga kooli korral ainult selle kooli õpilaste jaoks. Koolide tabelis näeme siis koolide keskmisi tulemusi ning nende õpilaste kõigi eksamite keskmisi tulemusi, kes valisid seda ainet. Koolide jaoks arvatud vahed omandavad mõtte võrdluses kogu Eesti eksaminandide tabeliga (tabel 7), sest erinevad eksamid on veidi ebavõrdses olukorras tulenevalt õppekavast ning eksamile esitatavatest nõudmistest. Kui näiteks 2010. aastal tehti matemaatikat keskmiselt umbes 6,59 punkti nõrgemini kui muid aineid, siis samas leidub isegi suhteliselt nõrkade õpilastega koole, kus matemaatika eest teeniti keskmiselt rohkem punkte kui teiste ainete eksamite eest matemaatikat valinud õpilaste puhul.

Tabelis 8 on näitena valitud võrdluseks üks Lõuna-Eesti kool, mis on siiski veidi tugevam kui keskmine Eesti kool. Kui 2010. aastal oli õpilase võimekus ühe eksamisoorituse kohta 61,82 punkti (väärtsu tabeli 7 reas „Kokku“ ja veerus „Kõik“), siis nimetatud koolis said õpilased keskmiseks võimekuseks ühe eksamisoorituse kohta 64,25 punkti (ligi kaks ja pool punkti rohkem). Silma torkab,

Tabel 7. 2010. aasta gümnaasiumiastme õpilaste riigieksamite tulemused ja õpilaste arvud (eksternid ja varem lõpetanud välja arvatud). Veergudes OT1, . . . , OT5 on eksamiosade tulemused, eelviimases veerus kõigi eksamite keskmine (kontingendi tugevus) ning viimases veerus aine keskmise ja kontingendi tugevuse vahe. K = keskmine, N = sooritajate arv.

|                             |   | Tulemus | OT1   | OT2   | OT3   | OT4   | OT5   | Kõik  | Vahe  |
|-----------------------------|---|---------|-------|-------|-------|-------|-------|-------|-------|
| Ajalugu                     | K | 65,58   | 12,95 | 16,99 | 16,72 | 18,70 |       | 67,56 | -1,99 |
|                             | N | 1778    | 1778  | 1778  | 1778  | 1778  |       | 1778  | 1778  |
| Bioloogia                   | K | 59,18   | 58,94 |       |       |       |       | 62,01 | -2,83 |
|                             | N | 3247    | 3247  |       |       |       |       | 3247  | 3247  |
| Eesti keel<br>(kirjand)     | K | 58,61   | 58,36 |       |       |       |       | 61,50 | -2,89 |
|                             | N | 10025   | 10025 |       |       |       |       | 10025 | 10025 |
| Eesti keel<br>teise keelena | K | 57,60   | 9,33  | 11,98 | 12,60 | 10,58 | 13,18 | 51,42 | 6,18  |
|                             | N | 3281    | 3281  | 3273  | 3273  | 3273  | 3273  | 3281  | 3281  |
| Füüsika                     | K | 63,95   | 14,33 | 20,67 | 28,94 |       |       | 69,68 | -5,74 |
|                             | N | 717     | 717   | 717   | 717   |       |       | 717   | 717   |
| Geograafia                  | K | 59,31   | 59,05 |       |       |       |       | 59,10 | 0,20  |
|                             | N | 6737    | 6737  |       |       |       |       | 6737  | 6737  |
| Inglise keel                | K | 69,28   | 14,15 | 13,13 | 13,30 | 11,77 | 16,89 | 65,57 | 3,71  |
|                             | N | 8379    | 8379  | 8377  | 8377  | 8377  | 8377  | 8379  | 8379  |
| Keemia                      | K | 59,45   | 59,19 |       |       |       |       | 63,85 | -4,40 |
|                             | N | 2055    | 2055  |       |       |       |       | 2055  | 2055  |
| Matemaatika                 | K | 58,43   | 32,93 | 25,64 |       |       |       | 65,02 | -6,58 |
|                             | N | 4498    | 4498  | 4473  |       |       |       | 4498  | 4498  |
| Prantsuse<br>keel           | K | 72,41   | 12,28 | 14,30 | 14,87 | 13,07 | 17,88 | 74,50 | -2,09 |
|                             | N | 34      | 34    | 34    | 34    | 34    | 34    | 34    | 34    |
| Saksa keel                  | K | 77,98   | 18,76 | 16,35 | 14,61 | 14,01 | 17,13 | 71,03 | 6,95  |
|                             | N | 381     | 381   | 356   | 356   | 356   | 381   | 381   | 381   |
| Vene keel<br>(kirjand)      | K | 59,69   | 59,46 |       |       |       |       | 59,61 | 0,08  |
|                             | N | 1687    | 1687  |       |       |       |       | 1687  | 1687  |
| Vene keel<br>võrkeelena     | K | 75,88   | 13,25 | 13,56 | 14,53 | 15,06 | 19,44 | 60,54 | 15,34 |
|                             | N | 346     | 346   | 346   | 346   | 346   | 346   | 346   | 346   |
| Ühiskonna-<br>õpetus        | K | 63,96   | 10,27 | 18,54 | 19,01 | 15,94 |       | 60,08 | 3,87  |
|                             | N | 6050    | 6050  | 6043  | 6043  | 6043  |       | 6050  | 6050  |
| Kokku                       | K | 61,82   |       |       |       |       |       | 61,82 | -0,01 |
|                             | N | 49215   |       |       |       |       |       | 49215 | 49215 |

et matemaatika eksamil saadi selles koolis 1,34 punkti rohkem kui teiste eksamite eest kokku nende õpilaste puhul, kes olid valinud matemaatika eksami. Kui arvestada, et Eestis keskmiselt saadi matemaatika eest 6,58 punkti vähem, siis on see kool oma matemaatika tulemusega kogu Eesti tulemusega võrreldes umbes 8 punkti tugevam (eeldusel, et võrdleme võrdseid kontingente), mis vastab umbes aastajagu keskmisele õppimisele.

Tabel 8. Ühe Lõuna-Eesti kooli riigieksamite tulemused ja õpilaste arvud 2010. aastal (eksternid ja varem lõpetanud välja arvatud). Veergudes OT1, ..., OT5 on eksamiosade tulemused, eelviimases veerus kõigi eksamite keskmine (kontingendi tugevus) ning viimases veerus aine keskmise ja kontingendi tugevuse vahe. K = keskmine, N = sooritajate arv.

| Õppeaine                |   | Tulemus | OT1   | OT2   | OT3   | OT4   | OT5   | Kõik  | Vahe  |
|-------------------------|---|---------|-------|-------|-------|-------|-------|-------|-------|
| Ajalugu                 | K | 72,74   | 15,15 | 18,00 | 19,17 | 20,09 |       | 71,74 | 1,00  |
|                         | N | 23      | 23    | 23    | 23    | 23    |       | 23    | 23    |
| Bioloogia               | K | 55,89   | 55,69 |       |       |       |       | 57,39 | -1,50 |
|                         | N | 18      | 18    |       |       |       |       | 18    | 18    |
| Eesti keel<br>(kirjand) | K | 58,52   | 58,24 |       |       |       |       | 62,98 | -4,45 |
|                         | N | 84      | 84    |       |       |       |       | 84    | 84    |
| Füüsika                 | K | 60,80   | 12,80 | 18,40 | 29,60 |       |       | 65,80 | -5,00 |
|                         | N | 5       | 5     | 5     | 5     |       |       | 5     | 5     |
| Geograafia              | K | 67,19   | 66,92 |       |       |       |       | 66,24 | 0,95  |
|                         | N | 42      | 42    |       |       |       |       | 42    | 42    |
| Inglise keel            | K | 67,34   | 14,60 | 12,66 | 13,53 | 12,06 | 14,43 | 66,29 | 1,06  |
|                         | N | 70      | 70    | 70    | 70    | 70    | 70    | 70    | 70    |
| Keemia                  | K | 51,21   | 50,82 |       |       |       |       | 54,36 | -3,14 |
|                         | N | 14      | 14    |       |       |       |       | 14    | 14    |
| Matemaatika             | K | 65,49   | 34,86 | 30,63 |       |       |       | 64,14 | 1,35  |
|                         | N | 43      | 43    | 43    |       |       |       | 43    | 43    |
| Saksa keel              | K | 75,00   | 15,25 | 15,63 | 15,63 | 14,00 | 14,00 | 71,00 | 4,00  |
|                         | N | 1       | 1     | 1     | 1     | 1     | 1     | 1     | 1     |
| Ühiskonna-<br>õpetus    | K | 67,46   | 10,20 | 18,98 | 20,68 | 17,38 |       | 63,52 | 3,94  |
|                         | N | 50      | 50    | 50    | 50    | 50    |       | 50    | 50    |
| Kokku                   | K | 64,05   | 36,79 | 19,13 | 17,34 | 15,20 | 14,42 | 64,26 | -0,21 |
|                         | N | 350     | 350   | 192   | 149   | 144   | 71    | 350   | 350   |

Koolide ja kogu Eesti tulemustes on mitmest osast koosnevate eksamite puhul eraldi välja toodud ka osatulemused OT1, OT2, OT3, OT4 ja OT5. Üks oluline strateegiline näitaja, mis nendest osatulemustest ilmneb, puudutab võõrkeelte suulise osa hindamist koolis. Veerg OT5 on võõrkeelte puhul eksami suulise osa keskmine punktide arv, mis peaks olema üpris lähedane kuulamisosa OT2, sest kui õpilane ei saa keerulisest tekstist kuulamisel aru, ei tohiks ta osata ka sama keerulist teksti rääkida. Eesti keele teise keelena eksami suuline osa salvestatakse ning seoses sellega on suulise osa tulemus üle riigi vaid veidi üle punkti kõrgem kui kuulamisosa tulemus, sest hindajate tegevus on ka hiljem kontrollitav. Inglise keeles on see vahe aga üle Eesti üle kolme punkti ning mõnedes koolides on see olnud isegi üle 10 punkti. Liigne lahus võõrkeelte suulise osa hindamisel seab ju tegelikult ka matemaati-

ka eksamit valinud õpilased ebavõrdsemasse olukorda. Antud väljavalitud koolis on inglise keele kuulamisosa ja suulise osa vahe 1,76 punkti, mis viitab suuremale rangusele kui Eestis keskmiselt.

Eeltoodud lähenemise peamine eelis on see, et ta võimaldab ausamalt hinnata õpilaste panust konkreetsetesse eksamitesse, sest igas koolis ei tarvitse olla keskmist õpilast.

## Kirjandus

1. D. Abrams. Social identity, social cognition, and the self: The flexibility and stability of self-categorization. D. Abrams, M. A. Hogg, editors, *Social Identity and Social Cognition*. Blackwell, 1999, 197–229.
2. M. du Toit. *IRT from SSI: Bilog-MG, Multilog, Parscale and Testfact*. Scientific Software International, 2003.
3. *PISA Data Analysis Manual. SPSS, Second edition*. OECD, 2009.
4. D. Operario, S. T. Fiske. Integrating social identity and social cognition: A framework for bridging diverse perspectives. D. Abrams, M. A. Hogg, editors, *Social Identity and Social Cognition*. Blackwell, 1999, 26–54.
5. E. Rull. Eksamitulemustest ja standardiseerimisest. *LoTe*, **2**, 2008, 25–27.