

Tallinna Ülikool
Matemaatika-loodusteaduskond
Informaatika osakond

Kätlin Holm

**TESTITULEMUSTE
STANDARDISEERIMINE EESTI VABARIIGI
RIIGIEKSAMITE NÄITEL**

Magistritöö

Juhendaja: Katrin Niglas, PhD

Autor: _____ ” ” _____ 2006.a.

Juhendaja: _____ ” ” _____ 2006.a.

Osakonnajuhataja: _____ ” ” _____ 2006.a.

Tallinn 2006

SISUKORD

SISSEJUHATUS	3
1. RIIGIEKSAMITE TAUST	5
1.1. RIIGIEKSAMI MÕISTE.....	5
1.2. RIIGIEKSAMITE AJALUGU	6
1.3. RIIKLIK EKSAMI- JA KVALIFIKATSIOONIKESKUS	7
2. TESTI KVALITEET JA TULEMUSTE STANDARDISEERIMINE	9
2.1. TESTI MÕISTE JA LIIGID	9
2.2. TESTIKÜSIMUSTE LIIGID.....	10
2.3. TESTÜLESANNETE RASKUS	11
2.4. TESTI KVALITEEDINÄITAJAD.....	12
2.4.1. TESTI RELIAABLUS	12
2.4.2. TESTI VALIIDSUS	14
2.5. TESTITULEMUSTE STANDARDISEERIMINE.....	15
3. RIIGIEKSAMITE VÕRRELDAVUSE KAJASTAMINE MEEDIAS	27
3.1. RIIGIEKSAMITE ERINEV RASKUSASTE	27
3.2. KESKMISE TULEMUSE VÕRRELDAVUS KOOLITI	27
3.3. LÄVENIPÕHINE VASTUVÕTT ÜLIKOOLIDESSE.....	28
3.4. RIIGIEKSAMITE TULEMUSTE HINDAMISSÜSTEEMI MUUTMINE.....	29
4. PROBLEEMI LAHENDUSI TEISTES RIIKIDES.....	30
4.1. LÄTI VABARIIK.....	30
4.2. AMEERIKA ÜHENDRIIGID	30
4.3. ROOTSI KUNINGRIIK	31
5. UURIMUSE METODOLOOGIA	33
5.1. UURIMISSTRATEEGIA	33
5.2. VALIMI KIRJELDUS.....	33
5.3. ANDMETÖÖTLUS.....	35
6. ÜLEVAADE KOLME RIIGIEKSAMI TULEMUSTEST AASTATE LÕIKES	36
6.1. EESTI KEELE KIRJAND.....	36
6.2. INGLISE KEEL.....	39
6.3. MATEMAATIKA	43
7. RIIGIEKSAMITE TULEMUSTE STANDARDISEERIMINE	47
7.1. STANDARDISEERIMISE VAJALIKKUS	47
7.2. VÕIMALUSED JA OHUD	48

7.2.1. STANDARDISEERIMINE Z-SKAALALE	49
7.2.2. STANDARDISEERIMINE FIKSEERITUD KESKMISE SKAALALE	57
7.2.3. STANDARDISEERIMINE PROTSENTJÄRJESTUSSKAALALE.....	61
7.3. JÄRELDUSED	63
KOKKUVÕTE	66
KASUTATUD KIRJANDUS	68
SUMMARY	71
MÕISTETE JA LÜHENDITE SELGITUSI	72
LISA 1. Z-SKAALA STANDARDISEERITUD TULEMUSED.....	75
LISA 2. FIKSEERITUD KESKMISE SKAALA STANDARDISEERITUD TULEMUSED.....	77

SISSEJUHATUS

Sissejuhatuses tutvustan töö teema aktuaalsust ja valiku põhjendust, töö struktuuri, probleemi sõnastust, töö eesmärgi ning tulemust ja eesmärgi saavutamiseks kasutatavaid meetodeid.

Eesti kogemused suurte üleriigiliste testide koostamisel ja tulemuste hindamisel ei ole kuigi pikad, riigieksameid on tehtud alates 1997. aastast ning tänavu toimuvad nad alles kümnendat korda. Antud valdkonnas on veel mitmeid põhjalikult analüüsimata tahke ning üheks nendest on ka riigieksamite tulemuste standardiseerimine.

Riigieksamite tulemused varieeruvad nii ainete kui aastate lõikes ning nende omavahelise võrreldavuse saavutamiseks on võimalik kasutada erinevaid standardiseerimismeetodeid. Teema on hetkel Eesti Vabariigis aktuaalne ja uudne, varasemate aastate praktilised kogemused üleriigiliste testide standardiseerimisel puuduvad. Lisaks sellele ei ole antud teema kohta väljastatud ühtset ja ülevaatlikku materjali, mis oleks abiks kõigile, kel antud valdkonna vastu sügavam huvi, kuid ei ole eelnevalt testitulemuste standardiseerimisega kokkupuutunud.

Teema valikule aitas kaasa tema uudsus ning sügavama analüüsi vajalikkus ja lisaks sellele ka minu isiklik huvi antud valdkonna vastu. Magistritöö käigus tehtav analüüs aitab kindlasti paremini mõista vastava sammu astumise tagamaid ning analüüsida erinevaid võimalikke stsenaariume.

Vastavalt väljatoodud probleemidele olen oma magistritöös püstitanud järgmised eesmärgid:

- selgitada riigieksami kui testi olemust;
- leida erinevaid võimalusi riigieksamite tulemuste võrreldavuse parandamiseks standardiseerimismeetodite abil;
- analüüsida väljatoodud standardiseerimismeetodite ohte ja võimalusi ning nende mõju testi tulemustele ja vastavalt sellele teha järeldusi;
- koostada ühtne ülevaatlik materjal, mis aitaks valdkonda mõista ka neil, kellel puuduvad sellekohased eelnevad põhjalikud teoreetilise teadmised.

Eesmärgi saavutamiseks kasutan Eesti Vabariigi kolme riigieksami tulemuste statistilisi näitajaid, mille põhjal teen meetodite kasutatavuse kohta järeldusi.

Töö tulemusena selgitan vastavalt väljatoodud meetoditele testitulemuste standardiseerimise võimalusi ning ohte ja teen vastavalt saadud tulemustele järeldusi meetodi kasutamise efektiivsuse kohta.

Magistritöö koosneb sissejuhatusest, seitsmest peatükist, kokkuvõttest, kasutatud kirjanduse loetelust, inglise keelsest sisu ülevaatest ning kasutatud mõistete ja lühendite selgitustest.

Esimeses peatükis, „Riigieksamite taust”, selgitan pikemalt riigieksamite olemust ning ajaloolist külge, mille abil on paremini võimalik mõista, milline on riigieksamite hetkeseis. Lisaks sellel kirjeldan tervikpildi saamiseks riigieksamite haldamisega tegelevat institutsiooni.

Teises peatükis, „Testi kvaliteet ja tulemuste standardiseerimine, selgitan põhjalikult testi mõistet, erinevaid liike, testülesannete raskuse määramise olulisust ning testi kvaliteedinäitajaid. Lisaks sellele selgitan testitulemuste standardiseerimise olemust ja võimalusi.

Kolmandas peatükis, „Riigieksamite võrreldavuse kajastamine meedias“, toon välja ja analüüsin seoses riigieksamite tulemuste võrreldavusega meedias kõlapinda leidnud erinevaid teemasid.

Neljandas peatükis, „Probleemi lahendusi teistes riikides”, toon välja, kuidas on näidetena väljatoodud riikides lahendatud keskkooli lõpetamise ja ülikooli sisseastumiseksamite ühildamine.

Viiendas peatükis, „Uurimuse metodoloogia”, selgitan töö uurimisstrateegiat, valimi kirjelduse ja andmetöötluse meetodeid.

Kuuendas peatükis, „Ülevaade kolme riigieksami tulemustest aastate lõikes“, analüüsin põhjalikult kolme valitud riigieksamite statistilisi näitajaid.

Seitsmenda peatükis, „Riigieksamite tulemuste standardiseerimine”, toon välja vastavalt eelnevatele peatükkidele riigieksamite tulemuste standardiseerimise võimalusi koos nende tugevuste ja nõrkustega ning tulemustest sõltuvalt teen vastavasisulisi järeldusi.

1. RIIGIEKSAMITE TAUST

Temaatika selgitamiseks alustan kõigepealt riigieksamite olemuse, ajaloo ning neid korraldava asutuse kirjeldamisest.

1.1. RIIGIEKSAMI MÕISTE

Lähenedes eksami mõistele üldises plaanis, on ta üks teadmiskontrolli vormidest, mille põhjal hinnatakse õppija teadmiste, oskuste ja vilumuste omandamist. Riigieksam kitsendab pisut antud mõistet. Gümnaasiumi riigieksam on osa õpitulemuste välishindamise süsteemist, mille alla kuuluvad lisaks ka üleriigilised tasemetööd kooliastmete lõpus ja põhikooli lõpueksamid (*REKK. Üldteave, 2006*).

Riigieksamid toimuvad üks kord aastas, aprillist-juunini, ning riigieksamite ained ja toimumiskuupäevad kinnitab haridusminister vastava määrusega. Gümnaasiumi riigieksam on reeglina testivormis (välja arvatud eesti keele ja vene keele riigieksam, mille vorm on kirjand) ning võib olla kas kirjalik või kirjalik ja suuline kaheosalise eksami puhul. Gümnaasiumi lõpetamiseks tuleb sooritada viis eksamit, milledest kolm peavad olema riigieksamid (lõpetaja võib valida ka kõik viis riigieksamit). Sealhulgas on kõikidele lõpetajatele kohustuslik eesti keele (eesti õppekeelega koolides) ja vene keele (vene õppekeelega koolides) riigieksam. Gümnaasiumi lõputunnistus kehtib koos riigieksamitunnistusega. Kolme kohustusliku riigieksami hulgas võib olla üks võõrkeeleksam. Vene õppekeelega koolides õppijatel on kohustuslik sooritada ka riigikeele eksam. Riigieksamitöö ülesanded töötab välja ministri määrusega kinnitatud spetsiaalne eksamit ettevalmistav komisjon, kuhu kuuluvad vastava ala tunnustatud spetsialistid. Gümnaasiumide, kutseõppeasutuste lõpetajatel ja eksternidel on kohustus teatada oma eksamivalikud koolile hiljemalt 15. jaanuariks. Varem lõpetanutel, varasemat riigieksami tulemust parandada soovijatel või välisriigis Eesti Vabariigi keskkaridusele vastava haridustaseme omandanutel on Riiklikule Eksami- ja Kvalifikatsioonikeskusele teatamise tähtjaks 20.jaanuar (*Riigieksamid, 1998, 3:4; ERT, 2005*).

Riigieksamite tulemusi hinnatakse 100-palli süsteemis ning gümnaasiumi lõpetamiseks on positiivne tulemus alates 20 pallist, alla mille on riigieksam sooritatud mitterahuldavalt. 20-palline soorituspiir kehtib alates 2002.aastast, varem oli tulemus positiivne alates ühest pallist.

Oma eksamitulemus on võimalik ühel korral vaidlustada apelleerimiskomisjonis ning selle tulemusena võib, aga ei pruugi, komisjon muuta riigieksami tulemust ja seda mõlemas suunas – nii punktide suurendamise kui vähendamise näol.

Riigieksam kontrollib riiklikus õppekavas määratletud õpitulemuste omandamist ning tema eesmärkideks on:

1. kontrollida riikliku õppekavaga määratletud õpitulemuste saavutatust eksamiaineis;
2. tagada gümnaasiumilõpetajate eksamihinnete võrreldavus;
3. võimaldada õpilasel saada objektiivsem pilt oma õpitulemustest;
4. saada ülevaade õpetamise/õppimise tulemuslikkusest koolis;
5. anda Haridusministeeriumile, haridusosakonnale, koolidele ja lapsevanematele ülevaade õppeprotsessi tulemuslikkusest;
6. ühildada gümnaasiumide lõpueksamid kõrgkoolide sissastumiseksamitega;

(Riigieksamid, 1998, 3).

1.2. RIIGIEKSAMITE AJALUGU

Kuigi riigieksamid viidi Eestis esimest korda katseliselt läbi 1996. aastal ning ametlikult 1997. aastal, siis riigieksamiteks hakati ettevalmistusi tegema juba 1993. aastal.

01.oktoobril.1996 toimus Haridusministeeriumi ja ülikoolide ühisnõupidamine, kus arutati riigieksamite ühitamist ülikoolide sisseastumiseksamitega. 1. novembriks otsustasid ülikoolid ühitamise kasuks, millega riigieksamid muutusid ühtlasi sisseastumiseksamiteks koos kõige sellest tulenevaga. Haridusministri 09.10.1996 käskkirjaga nr. 182 asutati 1. jaanuarist 1997 Haridusministeeriumi valitsemisalas alaliselt tegutsev riigiasutus Riiklik Eksami- ja Kvalifikatsioonikeskus *(Riigieksamid, 1997, 4:5).*

22.märtsil 1999. aastal kinnitati haridusministri määrusega nr 20. “Gümnaasiumi lõpueksamite korraldamise tingimused ja kord”, mille kohaselt tuli gümnaasiumi lõpetajal viiest lõpueksamist vähemalt kolm sooritada riigieksamitena ning emakeeleksam oli kõigile kohustuslik. Lisaks oli kohustuslikuks muu õppekeele koolides ka riigikeeleksam.

2005. aasta seisuga toimuvad riigieksamid järgmistes ainetes: ajalugu, bioloogia, eesti keel (kirjand), füüsika, geograafia, inglise keel, keemia, matemaatika, prantsuse keel, riigikeel,

saksa keel, vene keel (kirjand), vene keel, ühiskonnaõpetus (*Riigieksamid 1997, 5; REKK. Statistika, 2006*).

Aastatel 1997–2005 kõigis aineis kokku sooritatud 460441 riigeksamit, mis teeb aasta keskmiseks eksamisoorituste arvuks 51160 (*REKK. Statistika, 2006*).

1.3. RIIKLIK EKSAMI- JA KVALIFIKATSIOONIKESKUS

Riikliku Eksami- ja Kvalifikatsioonikeskuse eesmärk, nagu nimigi reedab, on aga laiem kui vaid üldhariduskoolide riigieksamite korraldamine. Lisaks sellele koordineeritakse eesti keele tasemeeksamite ettevalmistamist ja läbiviimist, kutsekvalifikatsioonide ja riiklike õppekavade väljatöötamist ning töötatakse välja ja korraldatakse üleriigilisi tasemetöid ja põhikooli ühtlustatud küsimustega lõpueksameid.

Toon, vastavalt Riikliku Eksami- ja Kvalifikatsioonikeskuse (edaspidi REKK) kodulehel märgitule, välja asutuse visiooni ning peamised ülesanded.

Asutuse visioon tulevikust on järgmine: REKK on arendus- ja korraldusorganisatsioon, mis hõlmates oma ala parimaid spetsialiste, teeb laiapõhjalist koostööd avalikkuse ja haridusjuhtidega Eesti hariduse arengu ja kvaliteedi säilitamiseks.

REKK'i peamisteks ülesanneteks on:

- tagada kutse-, eri- ja ametialade riiklike õppekavade väljatöötamine ja arendamine;
- tagada tugisüsteemid riiklike õppekavade rakendamiseks üldhariduskoolides ning kutseõppeasutustes ;
- tagada õpitulemuste välishindamise süsteemi areng ja kõrge tase;
- tagada eesti keele oskuse taseme määratlemise pädev ja efektiivne süsteem;
- tagada tugisüsteem pedagoogide koolitussüsteemi sihipäraseks ja efektiivseks arendamiseks;
- kirjastada metoodilisi ja juhendmaterjale õppekavaarendusest, eksami- ja tasemetööde analüüsist ja tulemustest, õpetaja tööst, ettevalmistusest, enesetäiendamisest ning atesteerimisest.

(*REKK. Üldteave, 2006; Riigieksamid, 1997, 5*).

Riigieksamite korraldamise, koordineerimise ning tulemuste väljatoomise kohapealt on REKK Eesti Vabariigis ainuke asutus, mis vastava valdkonnaga nii süvitsi tegeleb.

Aastatel 1997-1999 avaldas REKK koostöös Eesti Vabariigi Haridusministeeriumiga iga vastava aasta ning aine riigieksami tulemuste kohta eraldi brošüüri, kus kirjeldati ja analüüsiti vastavate riigieksamite töid ning tulemuste statistilisi näitajaid. Ajavahemikus 2000-2004 lisandus vastava aasta aine riigieksami töö kirjeldusele ja statistilisele analüüsile ka tuleva aasta riigieksami töö kirjeldus.

Lisaks sellele, on alates 1998. aastast väljastatud riigieksamite tulemuste statistiliste näitajate põhjalik analüüs ka REKK'i koduleheküljel, kus lisaks tavakriteeriumitele (osalejate arv, keskmine, miinimum ja maksimum tulemused, standardhälve, tühistatud ja apelleerimisel muudetud tööde arv) analüüsitakse andmeid ka: õppkeele, soo, koolilõpetamise ajalise külje pealt, kooli asukoha, koolitüübi, haridusosakonna, maakonna, kooli omanike, aastate, aastahinnete järgi. Lisaks võrreldakse ka variantide ja eksamiülesannete korreleeruvust.

2. TESTI KVALITEET JA TULEMUSTE STANDARDISEERIMINE

Kuna riigieksamid on peaaegu eranditult (välja arvatud eesti keele ja vene keele kirjand) koostatud testivormis, siis selleks, et mõista riigieksami väljundi näol tehtavaid järeldusi, tuleb kõigepealt tutvuda testi olemuse, koostamise, kvaliteedinäitajate ning tulemuste tõlgendamise võimalustega.

2.1. TESTI MÕISTE JA LIIGID

Mis on test? Sõna *test* tuleb inglise keelest ning tähendab katse, kontroll, proov.

Test on vahend mingisuguste omaduste mõõtmiseks. Testi eripäraks on see, et tema ülesannetele on võimalik kiiresti fikseerida vastust. Vastamiseks on sageli tarvis ära märkida ülesande üks etteantud vastuse variantidest, kirjutada lünka üks sõna või arv. Esineb ka lühivastustega testiküsimusi (*Mikk, 2002, 9*).

Pedagoogilises psühholoogias jaotatakse testid põhiliselt kolme gruppi: ainetestid, võimetetestid ja isiksusetestid. Sellise jaotuse aluseks on inimese omaduse liik, mida tahetakse mõõta. Ainetestidega mõõdetakse valitud õppeaine teadmiste ja oskuste taset, võimetetestiga erinevate võimete (näiteks: verbaalsed võimed, kujutlusvõime) taset, isiksuse testid mõõdavad erinevaid isiksuse omadusi (näiteks: temperamenditüüp, seltsivus). Kuna riigieksami tulemus mõõdab valitud aine õppekava sisu omandatust, siis keskendun siinkohal ainetestide koostamise lahtiseletamisega. Ainetest on küsimuste ja ülesannete kogum, millega mõõdetakse inimese õpitulemusi ehk õppeainepädevusi ning kontrollida saab lisaks teadmiste ka arutlus- ning analüüsioskust. Ainetestid võib veel omakorda jagada eristustestideks ning programmitestideks. Eristustest on ainetest, mille abil püütakse õpilasi eristada nende ainealaste teadmiste või oskuste järgi. Taoline test on vajalik näiteks kõrgkooli sisseastumistestina, riigieksamitena. Eristustest peab sisaldama nii kergemaid kui raskemaid ülesanded, et testitulemuste põhjal oleks võimalik testisooritajad tulemuste järgi reastada. Välismaal kasutatavad ainetestid on põhiliselt eristustestid (*Mikk, 2002, 15*).

Programmitest on test, mis vastab oma sisu poolest ainekavale. Ta kontrollib, kuidas ainekavas seatud eesmärgid on täidetud. Ehk kõik õigesti vastanute hindeks on „5“ ning alla

poole õigesti vastanute hindeks „2“. Programmitestist ei jäeta välja ei liiga kergeid ega ka liiga raskeid ülesandeid. Testi koostamise ainsaks kriteeriumiks on järgida vastavust ainekavale. Selline süsteem on meil koolides hindamise aluseks (*Mikk, 2002, 37:40*).

Riigieksami seatud eesmärkidena kontrollitakse nii riikliku õppekavaga määratletud õpitulemuste saavutatust eksamiaineis kui soovitakse ühildada gümnaasiumide lõpueksamid kõrgkoolide sissastumiseksamitega. Esimene osa eeldab, et riigieksam täidaks programmitesti rolli, kuna sellisel moel on võimalik õpilaste õppekava täitmise osakaalu mõõta ning samas teine punkt eeldab jällegi, et riigieksamid oleksid eristustestid. Selge see, et mõlemat eesmärki võrdsel tasemel täita on keeruline ning seega peab üks neist olema jõupositsioonil. Praeguseks on kujunenud olukord, kus ülekaalus on riigieksamite eristustesti roll. Väite aluseks on riigieksami tulemuste mõõtmine skaalal, vahemikus 0-100 punkti. Kui tegu oleks programmitestiga, siis tuleks õpilasele panna riigieksami eest hinne ning sooritus, mis jääks alla 50% oleks puudulik.

Kuna riigieksamite puhul on tegemist eristustestidega, siis lähtuvalt sellest kirjeldan järgnevalt eristustesti iseärasusi.

2.2. TESTIKÜSIMUSTE LIIGID

Enim levinud on valikvastustega ülesanded, kus küsimusele on lisatud võimalikud vastuse variandid ja vastajal tuleb nende seast leida õige. Valikvastustega küsimustest lihtsamad on alternatiivvastuselised küsimused – õpilasele esitatakse väide ning ta peab märkima, kas see on õige või väär. Võimalike valikute arv on tavaliselt siiski piiratud nelja-viiega ning antud viis on ainetestides kasutusel ka välismaal. Siinkohal on vastuse hindamine lihtne ning praktiliselt veatu. Samas on olemas ka variant, kus vastusevariantidest on õigeid rohkem kui üks, mis on aga vähem levinud. Sellisel juhul tuleb kindlasti ka instruksioonis ära märkida, et õigeid vastuseid võib olla rohkem kui üks. Ohuks võib siinkohal pidada seda, et valikvastustega ülesannetele saab märkida vastuseid ka juhuslikult, ilma, et õpilane ülesande enda jaoks lahti mõtestaks ning ära lahendaks ning nii mõnigi juhuslik vastus võib osutada õigeks (*Mikk, 2002, 27*).

Testides on võimalik kasutada ka vabavastuselisi küsimusi ja ülesandeid. Vabavastuselise küsimuse plussiks on see, et ta võimaldab õpilasel end ise väljendada, arendades seeläbi väljendusoskust. Samas puudusena saab kohe esile tuua selle, et neid on raskem hinnata (õpilase käekirja lugemisest tulenevad raskused, vastus võib sageli olla osaliselt õige).

Lisaks on testides kasutusel ka lüünkülesanded, kus tekstis on mõne sõna või arvu asemel lünk, mis tuleb õpilasel täita. Lüünkülesannete miinuseks on see, et nad suunavad õpikuteksti pigem pähe õppima, kui iseseisvalt mõtlema.

Valikvastustega ülesannetest on edasiarendatud kõrvutamisülesanded. Kõrvutamisülesannetes on tavaliselt kaks tulpa objekte ja nende vahel tuleb leida vastavus teatavate reeglite järgi. Ühe lihtsa näitena võiks esile tuua järgmise kõrvutamisülesande:

Leia inglise keelne vaste:

<i>Must</i>	<i>beautiful</i>
<i>Ilus</i>	<i>small</i>
<i>Suur</i>	<i>pig</i>
	<i>dirty</i>
	<i>big.</i>

Kõrvutamisülesandes on oluline ühes tulbas anda rohkem valikuid kui teises, nii välditakse seda, et viimase ülesande vastus automaatselt välja tuleks.

Lisaks sellele on tegelikult võimalik ainetesti osaks lugeda ka esseed. Essee teema ning kirjutamise aeg on sel puhul täpselt ettemääratud, kasutusel näiteks Ameerika Ühendriikides. Essee puhul hinnatakse üldist meeldivust, sõnavalikut, lause ehitust, struktuuri, näidete sobivust ja ideede loogikat (*Mikk, 2002, 30:34*).

2.3. TESTÜLESANNETE RASKUS

Selleks, et kõikide õpilaste teadmisi teineteisest eristada, tuleb testi kaasata võimalikult palju erineva raskusega ülesandeid. Ei ole ju mõtet testil, mille küsimused on nii lihtsad, et kõik neid lahendada suudavad või vastupidi nii rasked, et mitte keegi neile vastata ei oska. Oluline on, et küsimuste raskusaste oleks diferentseeritav, nii on võimalik õpilasi teineteisest

teadmiste taseme poolest eristada. Selgitamaks välja sihtgrupile määratud ülesannete raskusastet koostatakse vastava rühma esindusgrupp, mis esindab proportsionaalselt reaaleluga kõiki antud gruppi kuuluvaid liikmeid (poisid, tüdrukud, erinevate õpitulemustega õpilased, linna- ja maakonna ning maakoolide õpilased) ning ülesannete raskusastet hinnatakse antud ülesannetele vastamata jätnute või valesti vastanute hulga protsendi põhjal (*Kärner, 1976, 19*).

2.4. TESTI KVALITEEDINÄITAJAD

2.4.1. TESTI RELIAABLUS

Testi reliaablus (inglise keeles: *reliability*) ehk usaldatavus näitab seda, kuivõrd täpselt test mõõdab seda, mida ta mõõdab, ilma et küsitaks mis see „see“ on. Reliaablus on testi mõõtmisel olulisem omadus, mis väljendub mõõtmistulemuste stabiilsuses. On oluline, et korduval mõõtmisel saadaks ikka sama tulemus. Testiülesannete koostamisel tuleb seega jälgida, et need nõuaksid õpilaselt ühtesid ja samu võimeid, millest johtuvalt peaksid analoogsete testide tulemused olema ka omavahel võrreldavad. Seega, kui testi tulemused korduvatel mõõtmistel kõiguvad ning ei korreleeru teiste samasugustega, siis ei saa testi pidada usaldusväärseks (*Toim, 1981, 19*).

Testitulemused võivad kõikuda nii õpilase teadmistest kui mõõtmisvigadest tulenevatel põhjustel. Kui mõõtmisvead on suured võrreldes õpilase teadmiste erinevustega, siis see test ei sobi mõõtmiseks, sest selle reliaablus on madal. Reliaabluse määramiseks kasutatakse põhiliselt kolme meetodit: kordustest, paralleeltest ja poolitamise meetodit .

Kordustesti meetod määrab kindlaks testi stabiilsuse – kuivõrd püsivalt mõõdab test mingisugust omadust. Testi esmakordsel sooritusel mõõdetakse selle tulemused ning teatud aja möödudes sooritavad samad isikud sama testi uuesti ning seejärel võrreldakse testi tulemusi – tulemuste korreleeruvuse järgi mõõdetaksegi antud juhul testi reliaablust. Siin on aga mitu ohtu, esiteks võisid testi sooritajad testi tegemisest õppida ning nende teadmiste tase ei ole siis enam täpselt sama ja teiseks ei tohiks kahe testi sooritamise vahele jääda liiga pikk ajavahemik, kuna väliskeskkonnast tingituna võivad tekkida mitmesugused muutused testi tegijates endis. Kahe testi vaheline optimaalne aeg võiks olla mitte üle kahe nädala. (*Toim, 1981, 19; Metsämuuronen, 2000, 36*).

Paralleeltesti ja poolitamise meetodi puhul mõõdetakse testi seesmist püsivust – kas testi kõik elemendid mõõdavad ühte ja sama omadust. Paralleeltesti puhul tehakse test samal ajal, samade gruppide poolt, mitmes variandis. Testi reliaablust mõõdetakse paralleeltesti tulemuste korreleeruvusega ning mida suurem on korreleeruvus, seda suurem on ka reliaablus.

Poolitamise meetodi puhul poolitatakse test kaheks ja erinevaid pooli lahendavad erinevad grupid. Reliaablust mõõdetakse kahe grupi tulemuste vahelise korrelatsiooniga. Poolitamise meetodit saab kasutada vaid juhul, kui mõlema poole dispersioonid oluliselt ei erine. (*Toim, 1981, 19:20; Metsämuuronen, 2000, 37*).

Kuna erinevad reliaabluse määramise meetodid on rakendatavad eri tingimustes, siis saab üldjuhul korraka kasutada vaid ühte reliaabluse mõõtmise meetodit. Kui aga võimalik, siis tuleb kasutada ühe meetodina kordustesti ning teise meetodina kas paralleeltesti või poolitamise meetodit, kuna kordustest mõõdab testi stabiilsust ning teised meetodid testi sisemist konstantsust.

Testi reliaablus sõltub enamasti järgmistest aspektidest:

- testi küsimuste arvust – mida rohkem on testis küsimusi, seda kõrgem on testi reliaablus;
- väliskeskonna mõjutustest – kas testi tegemise ajal esines mingisuguseid takistavaid tegureid;
- oskuse hajuvusest katseisikute grupis – mida enam erinevad grupi siseselt katseisikute oskused, seda väiksem on testi mõõtmisviga võrreldes oskuse hajuvusega ja seda suurem on testi reliaablus. Tulemus on ootamatu, kuid see on nii seetõttu, et testi reliaablus on defineeritud testi mõõtmisvea ja mõõdetava suuruse hajuvuse suhte kaudu. Seetõttu on oluline koos reliaabluse koefitsiendiga näidata ka testitulemuste hajuvus vastaval katsel. Seega, suure oskuste hajuvusega grupis võime saada soovitud tulemusi ka madalama reliaabluse korral, kuid kui hajuvus on väike ning oskused rühmasiseselt suhteliselt võrdsed, peab testi reliaablus olema väga kõrge, et üksiktulemusi teineteisest eristada;
- testi koostamise kvaliteedist – kas kõik testi küsimused on mõistetavad ühtemoodi ja juhend on koostatud korrektselt ning arusaadavalt;
- kas testi hindajad on hindamissüsteemist ühemoodi arusaanud ning hindavad õigete kriteeriumite järgi.

(*Mikk, 2002, 71:72*).

2.4.2. TESTI VALIIDSUS

Testi valiidsus (inglise keeles: *validity*) ehk kehtivus näitab seda, kuivõrd test mõõdab seda omadust, mida me mõõta tahame. Näiteks, koostame ajaloos kirjaliku testi II Maailmasõja faktiteadmiste peale. Mõõtmistulemustest selgub, et test on usaldusväärne, mis näitab seda, et ta mõõdab midagi üsna täpselt, kuid kas see midagi, on õpilase teoreetilised teadmised, või loogiline mõtlemis- ja analüüsimisvõime? Valiidsus on testi kvaliteedi põhinäitaja, reliaablus on suunatud pigem sellele, et valiidsust saavutada. Mõlemad, nii reliaablus kui valiidsus, näitavad testi mõõtmistäpsust, kuid nende vahe seisneb selles, et reliaablus näitab seda, kui hästi test mõõdab seda, mida ta mõõdab, olenemata sellest mida ta mõõdab. Valiidsus näitab aga just seda, et test mõõdaks täpselt seda, mida on tarvis mõõta. Seega test võib olla reliaabne, kui ei pruugi valiidsus olla, vastupidi aga olla ei saa (*Toim, 1981, 20*).

Valiidsust on nelja liiki – sisuline ehk loogiline, võrdlev ehk konkureeriv valiidsus, faktorvaliidsus ja ennustav valiidsus.

Testi sisuline valiidsus sõltub sellest, kas suudetakse valida testi ülesanded selliselt, et nende lahendamine näitaks mõõdetava omaduse olemasolu testi sooritajal. Lõpliku sisulise valiku võimalikest ülesannetest teeb (3-5 spetsialistist koosnev) ekspertgrupp, kes on uuritava omaduse asjatundjad (*Mikk, 2002, 74*).

Peale seda, kui testi ülesanded on valitud ning testi reliaablus on kinnitust leidnud, tuleb testitulemusi korreleerida mõne väliskriteeriumiga. Selleks võib kasutada võrdlevat ehk konkureerivat valiidsust. Kõigepealt peab olema olema kindlat omadust mõõtev test ning seejärel tuleb valida kriteerium, mille järgi võrrelda. Võimalus on kriteeriumina kasutada näiteks teist ning varem läbiproovitud valiidsust testi ning seejärel võrrelda kahe testi tulemuste korreleeruvust.

Kolmas valiidsuse liik on faktorvaliidsus. Antud meetodi puhul arvestatakse seda, et inimeste psüühilised omadused avalduvad mitmesugustes tingimustes, tegudes, ülesannete lahendamise oskustes. Faktoranalüüsi läbiviimise käigus jaotatakse ülesanded vastavalt faktoritele (võime või omadus) ning ülesanded, millel on suured faktorlaadungid (kaalud), mõõdavad antud võimet kõige paremini. Valiidseks tuleb lugeda selliseid teste mille varieeruvus on suurel määral faktorite poolt kirjeldatud (*Toim, 1981, 76*).

Neljas valiidsuse liik on ennustav valiidsus. Ennustav valiidsus näitab kui hästi võimaldab test prognoosida edukust mingisuguses ainevaldkonnas. Ennustava valiidsuse määramiseks jagatakse kõigepealt grupp testitulemuste põhjal tugevamateks ja nõrgemateks ning seejärel jälgitakse nende praktilisi sooritusi teatud aja jooksul. Näiteks samamoodi võib ennustada õpilase riigieksamite tulemusi vastavalt sellele, milline oli ta õppeedukus valitud ainetes gümnaasiumi jooksul (*Toim, 1981, 77*).

2.5. TESTITULEMUSTE STANDARDISEERIMINE

Vajadus testitulemuste standardiseerimisest tuleneb testi mõõtmistulemuste kaudsest iseloomust. Õpilane, sooritades riigieksami saab oma tulemuse toorpunktides. Näiteks kui võõrkeele testi maksimum punktide arvuks oli 100 punkti ning õpilane sai lõpptulemuseks 70 punkti, siis need märgivadki toorpunktide arvu. See annab meile vastuse ühe õpilase tulemustest 100-punktilisel skaalal, kuid ei ütle väga palju üldise saavutuse kohta. Kas 70 punkti vastaval eksamil kuulub väga heade tulemuste hulka või mitte? Me saame võrrelda omavahel erinevaid tulemusi ka toorpunktide abil, et üks õpilane sai 70 toorpunkti ja teine sai 50 toorpunkti ning selle abil võime öelda, et teine õpilane sai 20 punkti nõrgema tulemuse, kuid üldisi järeldusi saavutuse taseme kohta teha ei saa. Vastuse küsimusele annab võrdlus ülejäänud grupiga – milline on grupi tulemuste keskmine tase ja hajuvus.

Kuna riigieksami näol on tegu eristustestiga, kus soovitakse määrata õpilaste paremusjärjestust, siis on tulemuste omavaheline võrdlemine väga vajalik.

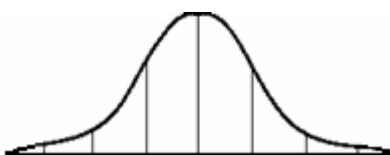
Käsitledes riigieksameid, siis ühe aasta tulemuste aritmeetiline keskmine iseloomustab vaid selle aasta lõpetajate lendu, kuid õpilaste õpitulemused ja/ või testi raskusaste varieeruvad teataval määral aastate lõikes. Näiteks: inglise keele riigieksami keskmine skoor 1998. aastal oli 58,84 ning 2005. aastal 71,93. Seega, kui õpilane sai inglise keele riigieksamil toorpunktideks 67,00 punkti, siis tema positsioon võrreldes 1998. aastal ja 2005. aastal lõpetanutega on erinev. Seega on oluline arvesse võtta, et testitavate grupis võivad aastate jooksul toimuda muudatused. Vastavad vaatlused on näidanud, et muutused populatsioonis on siiski nii väikesed, et kord väljatöötatud normid kehtivad ligikaudu viis aastat. Pärast seda on aga normide parandamine ehk täpsustamine hädavajalik. Sellest johtuvalt tuleks ka riigieksamite keskmiste tulemuste võrdlemisel arvestada viieaastaste intervallidega (*REKK*).

Statistika, 2006; Toim, 1981, 33). Antud teemat käsitlen pikemalt peatükis „Riigieksamite tulemuste standardiseerimine”.

Lisaks eelpool toodule on toorpunktidel veel teinegi miinus. Erinevatel testidel võib maksimaalne punktide arv olla erinev. Näiteks väljastati kuni 2001. aastani eesti keele (kirjand) ja vene keele (kirjand) riigieksami tulemus 10-punktsel skaalal, ülejäänud ainetel aga 100-punktsel skaalal. Seega, kui ei teata keskmist tulemust ning punktide varieeruvust, ei ole toorpunktide arv väga informatiivne.

Kolmandaks pole toorpunktide abil võimalik võrrelda erinevate ainete tulemusi omavahel. Kui sooritaja sai näiteks 65 punkti inglise keele riigieksamil ja 55 punkti matemaatika riigieksamil, pruugi see tähendada, et ta oli inglise keele riigieksamil teiste sooritajatega võrreldes tugevam. Inglise keele riigieksami keskmiseks tulemuseks võis olla näiteks 70 punkti ja matemaatikas 45, mis näitab, et sooritaja tulemus oli inglise keele riigieksamil alla keskmise ning matemaatikas üle keskmise.

Testitulemuste standardiseerimiseks on matemaatilises statistikas rida erinevaid võimalusi. Kõige levinum neist on meetod, mis tugineb eeldusele, et testitulemused jagunevad normaaljaotuse lähedaselt¹, kuigi vastavat meetodit on võimalik kasutada ka asümmeetriliste jaotuste korral.



Joonis 1. normaaljaotus

Nagu väljatoodud joonise pealt näha, on normaaljaotuse puhul tegu ideaalselt sümmeetrilise jaotusega – kummalegi poole keskväärtust ühepalju tulemusi (ehk 50% tulemustest). Tänu normaaljaotuse omadusele saab määrata ka, mitu protsenti tulemustest jääb suvalise kahe väärtuse vahemikku. Selleks on tarvis teada tulemuste keskväärtust ning standardhälvet ning nende näitajate kaudu saab iga üksikväärtuse puhul välja tuua kui mitme standardhälbe ühiku

¹ Normaaljaotuskõver on ideaalselt sümmeetriline ulatuse keskpunkti suhtes ning nii keskväärtus, mediaan kui ka mood asuvad kõik täpselt ulatuse keskpunktis.

kaugusel asub ta keskväärtusest. Sellest johtuvalt saab standardhälvet kasutada kui mõõtühikut (Niglas, 1997, 25:27) ning kasutada standardiseerimiseks järgmist valemit:

$$z = \frac{x_i - \bar{x}}{s} \quad (\text{valem 1}),$$

kus z – tulemused standardhälbe ühikutes

x_i – testi toorpunktid

\bar{x} – toorpunktide aritmeetiline keskmine katsegrupis

s – standardhälve.

(Toim 1981, 34).

Ehk selleks, et testi toorpunkte (x_i) ülekanda standardhälbe ühikute skaalale (z), tuleb katsegrupis arvutada aritmeetiline keskmine (\bar{x}) ja standardhälve (s). Saadud skaalat nimetatakse z -skaalaks. Seega, kui testi sooritaja sai toorpunktide arvuks 40 ($x_i = 40$) ning grupi keskmine oli 64 punkti ($\bar{x} = 64$) ja standardhälve oli 13 ($s = 13$), oleks tulemus standardskaalal järgmine:

$$z = 40 - \frac{64}{13} = -1,85 \quad (\text{vt. valem 1})^2.$$

Negatiivne tulemus tähendab seda, et vaadeldava õpilase tulemus oli 1,85 standardhälbe võrra madalam grupi keskmisest. Oletame aga, et õpilase toorpunktide arv oli 76 punkti ning grupi keskmine tulemus oli 70 punkti ning standardhälve oli 11, siis oleks tulemus standardskaalal järgmine:

$$z = 76 - \frac{70}{11} = 0,55 \quad (\text{vt. valem 1}).$$

Positiivne tulemus tähendab aga seda, et vaadeldava õpilase tulemus oli 0,55 standardhälbe võrra kõrgem grupi keskmisest. Kui tulemus oleks olnud null, siis oleks õpilase tulemus võrdunud grupi keskmisega.

Peale standardiseerimist antud kujule saab grupi keskväärtuseks null. Standardiseerimine võrdsustab siinkohal testi(de) lõikes keskmise tulemuse ning seega eeldab, et erinevused tulenevad testist, kuid ei võta arvesse, et erinevused võivad tuleneda ka grupist, kes testi sooritas. Peale testi standardiseerimist on grupi standardhälve alati 1, mis tähendab, et ka

² Tulemuste väljatoomisel ümardasin lõppresultaadi, tuues välja kaks kohta peale koma.

hajuvused erinevates gruppides võrdsustuvad. Tulemuste asümmeetria siiski säilib ning ka erinevuste proportsioonid grupi sees – ehk kui kellegi tulemused grupis olid toorpunktadena väljendatult teistest väga palju paremad või halvemad, siis kajastub see ka standardiseeritud tulemustes. Samas on antud skaala miinusteks see, et keskmisest madalamad tulemused on negatiivsed ning standardiseeritud tulemus on murdarvuline. Murdarvulise tulemuse võiks põhimõtteliselt küll ümardada täisarvuks, kuid kuna z-skaalal on standardhäbeühikuks üks, siis sellise ümardamise korral kaotaks tulemus oluliselt oma täpsuses. Näiteks, on oluline vahe, kas tulemus z-skaalal on 1,22 või 0,90, kuid täisarvuks ümardatud tulemus oleks mõlema puhul 1.

Selleks, et vältida tulemuste väljatoomisel murd- ja negatiivseid arve ning tulemuse täpsuses mitte nii oluliselt kaotada minnakse z-skaalalt üle teistele skaaladele: z-skaalal saadud tulemused, korrutatakse läbi mingisuguste konstantidega (täpsuse säilitamiseks) ja liidetakse juurde teine konstant (vältimaks negatiivseid arve) ning kui seejärel lõpptulemus ümardada täisarvuks, ei kaota tulemus oma täpsuses enam nii oluliselt.

Käsitlen võimalikest skaaladest kolme: kolledžiskaalat, T-skaalat ning staniinide-skaalat.

Kolledžiskaalat kasutatakse näiteks Ameerika Ühendriikides kolledži sisseastumistestides, keskmiseks tulemuseks on 500 ja standardhälve on 100. Üleminekuvalem z-skaalalt on järgmine:

$$k = 100z + 500 \quad (\text{valem 2}).$$

(Mikk, 2001, 63).

Ehk, kui z-skaalal oleks testi sooritaja saanud tulemuseks 0,80 (tulemus üle keskmise), siis sama tulemus kolledži skaalal annaks järgmise tulemuse:

$$k = 100 \cdot 0,80 + 500 = 580 \quad (\text{vt. valem 2}),$$

mis näitab, et testisooritaja tulemus sama suuruse võrra üle keskmise (z-skaalal standardhälbeks 1 ning kolledžiskaalal 100, sellest ka tulemuste üle keskmiste erinevuse 100-kordne vahe).

Kui aga testi sooritaja oleks z-skaalal saanud alla keskmise tulemuse, näiteks -0,80, siis kolledžiskaalale ümberviidult saaks järgmise tulemi:

$$k = (100 \cdot -0,80) + 500 = 420 \quad (\text{vt. valem 2}),$$

mis näitab, et testisooritaja tulemus on sama ühiku võrra alla keskmise. Nagu näha, kolledžiskaalale üleminekul kaovad ära negatiivsed arvud ning tulemus täisarvu kujul on annab täpsema resultaadi kui sama tulemus z-skaalal.

Teine võimalus z-skaala asemel on kasutada T-skaalat, mille keskmine on 50 ja standardhälve 10. Üleminekuvalem z-skaalalt on järgmine:

$$T = 10z + 50 \quad (\text{valem 3}).$$

(Mikk, 2001, 64).

Ehk, kui z-skaalal oleks testi sooritaja saanud tulemuseks 0,80 (tulemus üle keskmise), siis sama tulemus T-skaalal annaks järgmise tulemuse:

$$T = 10 \cdot 0,80 + 50 = 58^3 \quad (\text{vt. valem 3}),$$

mis näitab, et testisooritaja tulemus sama suuruse võrra üle keskmise (z-skaalal standardhälveks 1 ning T-skaalal 10, sellest ka tulemuste üle keskmiste erinevuse 10-kordne vahe).

Kui aga testi sooritaja oleks z-skaalal saanud alla keskmise tulemuse, näiteks -0,80, siis T-skaalale ümberviidult saaks järgmise tulemi:

$$T = 10 \cdot -0,80 + 50 = 42 \quad (\text{vt. valem 3}),$$

mis näitab, et testisooritaja tulemus on sama ühiku võrra alla keskmise.

Kolmanda võimalusena võib z-skaala asemel kasutada staniinide-skaalat. Vastava skaala keskväärtuseks on 5 ja standardhälve on 2. Üleminekuvalem z-skaalalt on järgmine:

$$C = 2z + 5 \quad (\text{valem 4}).$$

(Mikk, 2001, 63).

Ehk, kui z-skaalal oleks testi sooritaja saanud tulemuseks 0,80 (tulemus üle keskmise), siis sama tulemus staniinide-skaalal annaks järgmise tulemuse:

$$C = 2 \cdot 0,80 + 5 = 7^3 \quad (\text{vt. valem 4}),$$

³ Tulemuste väljatoomisel ümardasin lõppresultaadi täisarvuks.

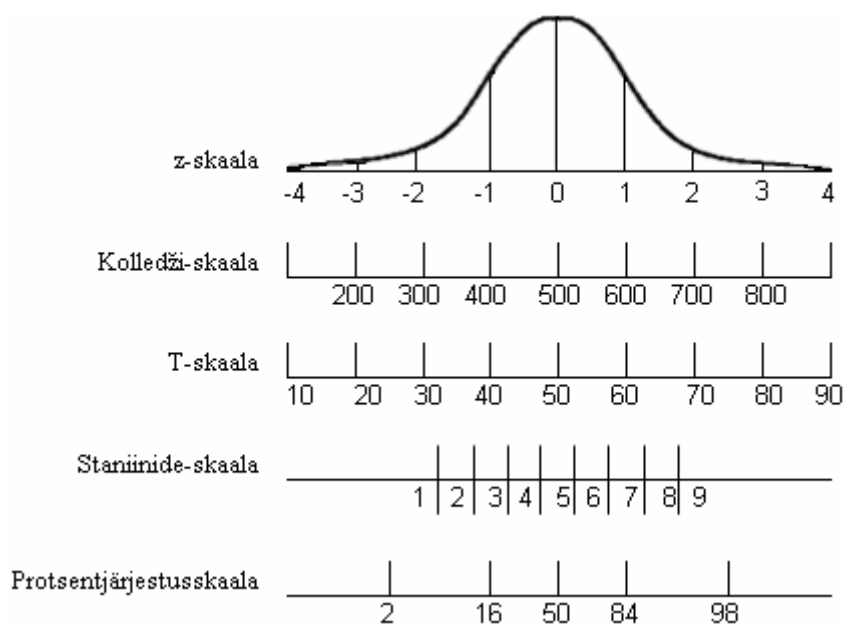
mis näitab, et testisooritaja tulemus on üle keskmise. Kui aga testi sooritaja oleks z-skaalal saanud alla keskmise tulemuse, näiteks -0,80, siis staniinide-skaalale ümberviidult saaksime järgmise tulemi:

$$C = 2 \cdot -0,80 + 5 = 3 \quad (\text{vt. valem 4, lk 19}),$$

mis näitab, et testisooritaja tulemus on sama alla keskmise. Staniinide-skaala puhul ei pääse tulemuse täpsuse säilitamiseks paraku murdarvudest, samuti väga madalate või kõrgete tulemuste korral ei jää tulemused väljatoodud skaala vahemiku algus- ja otspunktide (0-9) piiridesse. Näiteks, kui sooritaja oleks z-skaalal saanud tulemuseks -3,1, siis sama tulemus staniinide-skaalal annaks negatiivse tulemuse:

$$C = (2 \cdot -3,1) + 5 = -1 \quad (\text{vt. valem 4, lk 19}),$$

Skaalade võrdlemiseks esitan nad järgneval joonisel graafilisel kujul:



Joonis 2. erinevate skaalade võrdlus

(Lindgren & Suter, 1994:379; Mikk, 2002, 63)

Kui soovitakse aga testitulemustele kehtestada mingisugust soorituspiiri, nagu näiteks riigieksamite puhul on selleks (toorpunktidena) 20 punkti, siis väljatoodud standardiseerimismeetodid (z-skaala, kolledžiskaala, T-skaala, staniinide-skaala) ei võimalda peale standardiseerimist otsustada tulemuse absoluutse tasemel üle. Võimalik oleks vaid

määrata suhtelise kauguse järgi keskmisest. Näiteks need, kelle tulemus on -2, ei ole testi sooritanud. Teine võimalus oleks viia ka soorituspiir standardskaalale üle, mis tooks kaasa soorituspiiri kõikumise eksamite lõikes.

Väljatoodud standardiseerimismeetodid eeldavad, et tulemused on normaaljaotusega, kuid vastavaid meetodeid võib põhimõtteliselt kasutada ka asümmeetriliste jaotuste puhul. Sellisel juhul tuleb aga arvestada sellega, et asümmeetria säilib ka peale standardiseerimist ning standardiseeritud tulemuste tõlgendamisel ei tohi lähtuda reeglitest, miks kehtivad normaaljaotuse puhul (näiteks, et keskväärtusest kummalegi poole jääb 50% tulemustest – asümmeetrilise jaotuse puhul sellised tingimused ei kehti).

Sellisel juhul võib normaalskaala asemel kasutada protsentjärjestusskaalat ehk protsentiile. Testis saadud tulemus protsentskaalal näitab, mitu protsenti õpilastest sai temast nõrgema tulemuse. Protsentskaala keskpunktiks on 50. Kui keegi saab sellise tulemuse, siis tähendab see seda, et 50% testi sooritajatest on tulemused madalamad või samad ning ülejäänutel kõrgemad. Kui näiteks õpilane riigieksamil 62 punkti ning 3500-st sooritajast oli temast nõrgem tulemus 1600-l õpilasel, siis leides mitu protsenti moodustab 1600 3500-st, saame kätte õpilaste protsendi, kes said temast nõrgema tulemuse ning selleks on 45,71%, seega 54,29% õpilastest said temast parema tulemuse (*Toim, 1981, 37*).

Protsentskaala kasutamine võrdsustab testi(de) lõikes keskmise tulemuse ning seega eeldab, et erinevused tulenevad testist, kuid ei võta arvesse, et erinevused võivad tuleneda ka grupist, kes testi sooritas. Protsentskaala väärtused asuvad vahemikus 0-100, negatiivseid arve ei esine. Skaala on kasutajasõbralik ning selgesti mõistetav. Protsentskaala kasutamise korral on hajuvus grupisiseselt alati ligikaudu sama, seega toorpunktide viimine protsentskaalale kaotab hajuvuste erinevused nii grupi sees kui gruppide vahel.

Esmapilgul võivad T-skaalal ja protsentskaalal saadud tulemused tunduda sarnased, skaala keskväärtuseks on mõlemal 50. Siiski, T-skaala puhul on tegemist intervallskaalaga, protsentskaala puhul aga järjestusskaalaga, mistõttu tulemuste omavaheliste erinevuste proportsioonid muutuvad. Näiteks tulemusele 84 protsentskaalal võib T-skaalal vastata tulemus 60.

Protsentjärjestusskaalal jagunevad tulemused üldiselt ühtlaselt, kuid kui esineb korduvaid väärtusi, siis nende juures võib esineda ka kuhjumisi. Toorpunktide viimine protsentskaalale ühtlustab erinevuste proportsioone („venitab välja“ jaotuse kuhjumiskohad ja „surub kokku“ hõredad kohad. Kellel oli toorpunktides kõrgeim või madalaim tulemus, saab vastava tulemuse ka protsentjärjestusskaalal. Kui aga sooviks kehtestada testi tulemustele soorituse alampiiri (nagu riigieksamitel: 20 palli), siis kuna tulemus sõltub sooritajate arvust ning nõrgemate tulemuste osahulgast, siis pärast protsentskaalale viimist ei saa enam määrata ühest soorituspiiri. Võimalik oleks standardiseerida iga aine ja aasta tulemuste soorituspiir eraldi. (Niglas, 2005).

Lisaks z-skaala, kolledžiskaala, T-skaala, staniinide-skaala ning protsentskaalale on võimalus testitulemusi standardiseerida veel fikseeritud keskmisega skaala abil. Selleks tuleks testi toorpunktide keskmine viia skaalal kokkulepitud punkti ning „venitada“ ühe poole tulemusi ja „kokku suruda“ teise poole tulemusi säilitades esialgse skaala otspunktid. Standardiseeritud skaala algus- ja lõpp-punkt on kokkuleppelised. Toorpunktidele standardiseeritud punktidele minnakse üle kahes etapis – eraldi standardiseeritakse tulemused alla keskmise ja tulemused üle keskmise ning mõlema standardiseerimise puhul kasutatakse sama loogikat. Kõigepealt arvutatakse koefitsient, mille võrra erinetakse uuest keskmisest, arvutuseks kasutatakse järgmisi valemeid:

$$k_{\max} = \frac{(Sk_{\max} - \bar{x}_{fix})}{(Sk_{\max} - \bar{x})} \quad (\text{valem 5}),$$

$$k_{\min} = \frac{(Sk_{\min} - \bar{x}_{fix})}{(Sk_{\min} - \bar{x})} \quad (\text{valem 6}),$$

kus: k_{\max} – üle keskmise tulemuse koefitsient

k_{\min} – alla keskmise tulemuse koefitsient

Sk_{\max} – uue skaala maksimumväärtus

Sk_{\min} – uue skaala miinimumväärtus

\bar{x}_{fix} – uus fikseeritud keskmine

\bar{x} – toorpunktide aritmeetiline keskmine;

(Niglas, 2005).

Vastavalt sellele, kas tegu on üle või alla toorpunktide keskmise tulemusega lahutatakse skaala maksimum (Sk_{\max}) / alla keskmise tulemuse puhul skaala miinimum (Sk_{\min}), tulemusest uus fikseeritud keskmine (\bar{x}_{fix}) ning saadud vahe jagatakse skaala maksimum (Sk_{\max}) / alla keskmise tulemuse puhul skaala miinimum tulemuse (Sk_{\min}) ja toorpunktide aritmeetilise keskmise (\bar{x}) vahega. Kui testi tulemus võrdub aritmeetilise keskmisega, siis võib koefitsiendi arvutamisel kasutada vabal valikul ühte ülaltoodud valemitest.

Testitulemuste standardiseerimisel standardiseeritakse eraldi (kuid sama loogika järgi) üle ja alla toorpunktide aritmeetilise keskmise tulemused:

$$St_{\max} = (T - \bar{x}) \cdot k_{\max} + \bar{x}_{fix} \quad (\text{valem 7}),$$

$$St_{\min} = (T - \bar{x}) \cdot k_{\min} + \bar{x}_{fix} \quad (\text{valem 8}),$$

kus: St_{\max} – standardiseeritud tulemus, kui üksiktulemus oli üle toorpunktide aritmeetilise keskmise

St_{\min} – standardiseeritud tulemus, kui üksiktulemus oli alla toorpunktide aritmeetilise keskmise

k_{\max} – üle keskmise tulemuse koefitsient

k_{\min} – alla keskmise tulemuse koefitsient

T – üksiktulemus toorpunktide kujul

\bar{x}_{fix} – uus fikseeritud keskmine

\bar{x} – toorpunktide aritmeetiline keskmine;

(Niglas, 2005).

Vastavalt sellele, kas tegu on üle või alla toorpunktide keskmise tulemusega lahutatakse testi üksiktulemusest (T) toorpunktide aritmeetiline keskmine (\bar{x}) ning saadud vahe korrutatakse vastavalt tulemuse koefitsiendiga (üle keskmise tulemuse puhul üle keskmise tulemuse koefitsiendiga (k_{\max}), alla keskmise tulemuse puhul alla keskmise tulemuse koefitsiendiga (k_{\min}) ning liidetakse uus fikseeritud keskmine (\bar{x}_{fix}).

Kui testi tulemus võrdub aritmeetilise keskmisega, siis võib standardiseerimisel kasutada vabal valikul ühte ülaltoodud algoritmidest.

Selgitan antud standardiseerimismeetodit kolme näite puhul, kus ühel juhul on testi üksiktulemus toorpunktide kujul üle testi toorpunktide aritmeetilise keskmise, teisel juhul alla ning kolmandal juhul võrdne testi toorpunktide aritmeetilise keskmisega.

Näiteks: olgu testi toorpunktide aritmeetiline keskmine (\bar{x}) 70 punkti, testi uus fikseeritud keskmine (\bar{x}_{fix}) aga 65 punkti. Testisooritaja A tulemus toorpunktidena (T_1) on 95 punkti, testisooritaja B tulemus toorpunktidena (T_2) on 45 punkti ning testisooritaja C tulemus toorpunktidena (T_3) on 70 punkti. Testi punktiskaala asub vahemikus 0-100 (ehk 0 on miinimumväärtus ja 100 on maksimumväärtus). Kuna testisooritaja A tulemus toorpunktide kujul on üle toorpunktide aritmeetilise keskmise, siis kasutan vastavalt üle keskmise tulemuste koefitsiendi ning standardiseerimise algoritmi:

$$k_{\max} = \frac{(100 - 65)}{(100 - 70)} = \frac{35}{30} \quad (\text{vt. valem 5, 22}).$$

$$St_{\max} = (95 - 70) \cdot \frac{35}{30} + 65 = 94,2 \quad (\text{vt. valem 7, 23}),$$

saan tulemuseks, et testisooritaja A tulemus standardiseeritud kujul on 94,2, mis on samuti kõrgem, kui uus fikseeritud keskmine (65), kuid muutus on toimunud tulemuse ja keskmise vahelises kauguses. Kui standardiseerimata skaalal oli aritmeetilise keskmise (70) ja toorpunktide (95) vahe 25 punkti, siis standardiseeritud kujul on uue fikseeritud keskmise (65) ja saadud punktide (94,2) vahe suurem: 29,2. Kuna uus fikseeritud keskmine on toorpunktide aritmeetilisest keskmisest väiksem, siis skaala otspunktide säilitamiseks „venitatakse” välja üle keskmise tulemuste vahemik.

Testisooritaja B tulemus toorpunktide kujul on aga alla aritmeetilise keskmise ning seetõttu kasutan antud tulemuse standardiseerimiseks alla keskmise tulemuse koefitsiendi arvutamise ning standardiseerimise algoritmi:

$$k_{\min} = \frac{(0 - 65)}{(0 - 70)} = \frac{-65}{-70} = \frac{65}{70} \quad (\text{vt. valem 6, 22}).$$

$$St_{\min} = (45 - 70) \cdot \frac{65}{70} + 65 = 41,8 \quad (\text{vt. valem 8, 23}),$$

saan tulemuseks, et testisooritaja B tulemus standardiseeritud kujul on 41,8, mis on samuti madalam, kui uus fikseeritud keskmine (65), muutunud on aga taas tulemuse ja keskmise vaheline kaugus. Kui standardiseerimata skaalal oli aritmeetilise keskmise (70) ja toorpunktide (45) vahe 25 punkti, siis standardiseeritud kujul on uue fikseeritud keskmise (65) ja saadud punktide (41,8) vahe väiksem: 23,2. Kuna uus fikseeritud keskmine on toorpunktide aritmeetilisest keskmisest väiksem, siis skaala otspunktide säilitamiseks „surutakse” kokku alla keskmise tulemuste vahemik.

Kuna testisooritaja C tulemus toorpunktide kujul on võrdne toorpunktide aritmeetilise keskmisega (70), siis võib koefitsiendina kasutada vabal valikul kas üle keskmise tulemuse koefitsienti või alla keskmise tulemuse koefitsienti, mõlemad annavad standardiseerimisel sama tulemuse. Mõlemad koefitsiendid on juba välja arvatatud, jääb üle veel tulemus standardiseerida. Veendumaks selles, et mõlema koefitsiendi kasutamine annab sama tulemuse, proovin läbi mõlemad variandid:

$$St_{\max} = (70 - 70) \cdot \frac{35}{30} + 65 = 65,0 \quad (\text{vt. valem 7, 23}),$$

$$St_{\min} = (70 - 70) \cdot \frac{65}{70} + 65 = 65,0 \quad (\text{vt. valem 8, 23}),$$

saan mõlema variandi korral tulemuseks, et testisooritaja C tulemus standardiseeritud kujul on 65, mis on võrdne uue fikseeritud keskmisega (65). Kui tulemus standardiseerimata skaalal on võrdne keskväärtusega, siis standardiseeritud kujul jääb tulemus samuti võrdseks keskväärtusega.

Kuna fikseeritud keskmise skaalal on algus- ja lõpp-punkt kokkuleppelised, siis on võimalik siin seada ka testitulemustele nõutav soorituspiir, nagu näiteks riigieksamitel on 20 punkti. Antud meetodi puhul saab standardiseerida ka nii, et skaala otspunkti asemel jäetakse paigale minimaalne fikseeritud soorituspiir, mis välistaks olukorra, kus kellegi toorpunktides üle minimaalse soorituspiiri tulemus satuks peale standardiseerimist alla soorituspiiri. Väljatoodud

standardiseerimisviis vähendab testist tulenevaid erinevusi, säilitades testitavate tasemete erinevused. Peale standardiseerimist muutub pisut üldine hajuvus. Fikseeritust oluliselt erineva keskmisega jaotuses muutub hajuvus erinevas suunas ühel ja teisel pool fikseeritud keskmist. Standardiseerimise tulemusena väheneb asümmeetria ning muutuvad erinevuste proportsioonid grupi sees: asümmeetrilises muutuvad erinevuste proportsioonid ühtlasemaks, kuid fikseeritud keskmisest oluliselt erineva keskmisega sümmeetrilise jaotuse korral suurenevad erinevuste proportsioonid ühel ning teisel pool keskmist. Siiski, kelle tulemus oli toorpunktide skaalal kõrgeim/madalaim, siis nii on see ka standardiseeritud skaalal (*Niglas, 2005*).

3. RIIGIEKSAMITE VÕRRELDAVUSE KAJASTAMINE MEEDIAS

Järgnevalt käsitlen riigieksamite võrreldavuse kajastamist meedias – toon välja ning analüüsin riigieksamite ajaloo jooksul poleemikat tekitanud küsimusi koos omapoolse aruteluga.

3.1. RIIGIEKSAMITE ERINEV RASKUSASTE

Läbi riigieksamite ajaloo on aktuaalsena püsinud küsimus, et riigieksamite raskusaste on aastate lõikes erinev, mis on mõjutanud omakorda eksamitulemusi (*Rebane & Järve, 1999; Ots, 2000; Lukason, 2000*).

Kui vaadelda riigieksamite keskmisi tulemusi aastate lõikes, siis esineb seal teatavat kõikumist küll. Leitakse, et aastal, kus vastavas aines on madalam keskmine hinne, siis tähendab see kohe ka seda, et riigieksami ülesanded oli raskemad. Koolis õpitu sisu on jäänud ju siiski samaks (vastavalt Riiklikus Õppekavas sätestatule) ning ka õpetajad ei muutu kõigis koolides igal aastal. Seega võiks põhjuste otsimisele läheneda laiema nurga alt, mis ei välista loomulikult, et tulemuste erinevus võis tekkida ka eksami erinevast raskusastmest. Näiteks võib üks võimalus olla, et vastava aasta lõpetajate lend oligi tugevam või nõrgem oma üldtasemelt ning erinevused tulemustes tekkisid sellest. Igal aastal järjest ei ole võimalik teha täpselt ühesuguseid eksamitöid ning võimalus, et riigieksamid erinevad teineteisest pisut (kuid mitte oluliselt) raskusastme poolest, on reaalsed. Riigieksamite erinevast tasemest tuleneva probleemi üheks lahenduseks olekski riigieksamite tulemused standardiseerimismeetodite abil võrreldavaks muuta

3.2. KESKMISE TULEMUSE VÕRRELDAVUS KOOLITI

Aja jooksul on küsimusi tekitanud riigieksamite keskmiste tulemuste võrreldavus koolide kaupa, mille sõlmkohti järgnevalt lahkan (*Tamm & Tamm, 2000*).

Leitakse, et ainuüksi (pallide) keskväärtuse väljatoomine ei ole piisav näitaja andmehulga iseloomustamiseks. Oluline on teada ka eksami sooritajate arvu ning tulemuste jaotust. Tuuakse näitena välja kahe kooli tulemuste võrdlus, kus ühes saavutati muidu kõrgeid tulemusi, kuid mõnel õpilasel on mitterahuldav tulemus ning teises koolis oleksid kõigi

õpilaste tulemused keskpärased. Statistilise analüüsi kaudu tuleksid aga mõlema kooli keskmised tulemused võrdsed ning leitakse, et selline võrdlus ei anna objektiivset ülevaadet kooli tasemest. Leitakse, et lisaks tuleks analüüsida ka sooritajate arvu – kui eksamit sooritab koolis viis väga heade tulemustega viis või 50 õpilast, kelle seas on nii paremaid kui kehvemaid tulemusi.

Leian, et viie õpilase tulemused ei anna kindlasti kooli kohta objektiivset ülevaadet. Sellises olukorras tuleks võrdluseks tuua pigem mitmete aastate tulemused ning seejärel leida nendevahelisi seoseid. Palju oleneb ka koolitüübist – omavahel ei saa ju üheselt võrrelda ka vastava aine kallakuga koole ja tavakoole, kuna ainemahud on erinevad. Koolide omavahelisel võrdlemisel tuleks kõik väljatoodud kriteeriume kindlasti arvesse võtta.

3.3. LÄVENDIPÕHINE VASTUVÕTT ÜLIKOOLIDESSE

Seoses Tallinna Ülikooli ja Tartu Ülikooli üleminekuga lävendipõhisele vastuvõtule, on antud teema sattunud arutelu alla meedias, mida järgnevalt ka kajastan (*Lehtsaar, 2004; Laev, 2005; Sootak 2005; EÜL, 2004*).

Algul tervitati lävendipõhist vastuvõttu ülikoolidesse väga soojalt, leiti, et õpilasele tagatakse suurepärase võimalus koheselt teada saada, kuhu ning milliselt erialale ta sisse pääseb, et ei ole tarvis närviliselt jälgida vastuvõetavate pingeridu. 2005.aastal toimus lävendipõhine vastuvõtt osadele erialadele nii Tallinna kui Tartu Ülikoolis ning tekitas sellega parajat segadust, kuna vastu tuli võtta kõik soovijad, kes lävendi ületasid. Mõned erialad olid soovijatest mitmekordselt ülepaistatud ning loengute pidamiseks tuli moodustada mitmeid poolrühmi või kui seda ei õnnestunud, siis tuli osadel üliõpilastel, kellele lauda/ tooli ei jagunud, istuda kas aknalaual või suisa põrandal. Ülikoolides ei osatud kõigis ainetes paika panna õiget lävendipiiri, mis tekitaski ebaproportsionaalse jagunemise ning osadele erialadele tuli riigieelarvelistele kohtadele vastu võtta mitmed kordi suurem tudengitearv, mis tekitas ülikoolidele lisaks ka materiaalseid kulutusi.

2005. aastal korraldati lävendipõhist vastuvõttu alles esimest korda ning on loomulik, et iga uue süsteemi juurutamise juures kerkivad algusfaasis üles ka mitmed probleemid, kuid eks

igast kogemusest õpitakse. Näiteks võetakse 2006. aastal Tallinna Ülikoolis üliõpilasi riigieelarvelistele kohtadele vastu vastavalt riiklikule koolitustellimusele (*TLÜ, 2006*).

3.4. RIIGIEKSAMITE TULEMUSTE HINDAMISSÜSTEEMI MUUTMINE

Viimase punktina kajastan hiljuti meedias kõlapinda leidnud teemat, kus väidetakse, et REKK'il on plaanis riigieksamite tulemuste hindamissüsteemi muuta (*Päärt, 2006*).

REKK valmistuvat riigieksamite hindamissüsteemi muutmiseks ning täpsemalt plaanitakse muuta riigieksamite hinnete esitamise korda. Kui varem esitati tulemused toorpunktidenä, siis tulevikus (2006. aasta kevadel katseliselt) plaanitakse esitada riigieksamite tulemused keskmise tulemusega läbikaalutuna. Edaspidi võidakse riigieksamite tulemusi esitada kujul, kus näidatakse ära sooritaja protsent või pallide arv keskmisega võrreldes (näiteks: 15% üle keskmise või 16 palli alla keskmise). Antud suunas tegutsema pani eelkõige 2005.aastal Tallinna Ülikoolis ja Tartu Ülikoolis kehtestatud lävendipõhine vastuvõtt, kus osade erialadele vastuvõetute arv ületas mitmekordselt riikliku tellimuse ning koolidel tuli ülejäänud õppekohad oma eelarvest finantseerida.

Vastavalt meie vestlustele (22. aprillil 2006) REKK'i üldhariduse õppekavade ja eksamite osakonna juhataja asetäitja Aimi Püüaga, on siinkohal saadud infot pisut valesti tõlgendatud ja ajas etterutatud, 2006. aastal riigieksamite hindamissüsteemis sellist üleminekut veel ettenähtud ei ole, küll aga on sellised variandid võimalikud tulevikus.

Tegelikult on riigieksamite võrreldavuse üle juba varasematelgi aastatel arutletud ning tegu on ilmselt asjade loogilise käiguga. Standardiseeritud tulemusi kasutatakse ka mujal maailmas, näiteks Ameerika Ühendriikides, kõrgkoolidesse sisseastumistel. Riigieksamite standardiseerimine peaks välistama iga-aastastest eksamitöödest sissetulevad erinevused ning muutma nad omavahel võrreldavaks. See oleks jälle samm edasi riigieksamite täiustamisel. Nüüd jääb veel üle välja töötada õiged ja läbikaalutletud standardiseerimise alused, millele pakun omapoolseid lahendusi ka antud magistritöös.

4. PROBLEEMI LAHENDUSI TEISTES RIIKIDES

Kui Eesti Vabariigis on alles tulevikus plaanis hakata riigieksamite tulemusi standardiseerima, siis mujal maailmas, on testitulemuste standardiseerimist kasutatud juba ammu. Järgnevalt toon välja mõned näited erinevates riikides toimivatest eksamitulemuste standardiseerimismeetoditest. Üldpildi laiendamiseks valisin välja kolm riiki, kus on kasutusel teineteisest erinevad meetodid.

4.1. LÄTI VABARIIGI

Läti gümnaasiumi lõpetamiseks tuleb lõpetajal, nagu ka Eestis, valida viis riigieksamit, millest kaks määratakse Haridus- ja Teadusministeeriumi poolt, üks eksam kooli poolt ning kaks õpilase omal vabal valikul, kuid riigieksamite toimumisaeg on abiturientidel peale õppetöö lõppu, juunikuus.

Riigieksamite haldamine on tsentraliseeritud ning nende koordineerimine toimub, Õppekava Arendus- ja Eksamikeskuses (inglise keeles: *The Centre for Curriculum Development and Examinations*). Tegu on riikliku asutusega, mida haldab Läti Vabariigi Haridus- ja Teadusministeerium ning mis on finantseeritud riigieelarvest (*CCDE, 2006*).

Riigieksamite süsteem loodi 1995.aastal ning 2004.aastast alates on gümnaasiumi lõpetajatel võimalik sooritada riigieksamid kokku viieteistkümnes õppeaines, milleks on: inglise keel, saksa keel, prantsuse keel, vene keel, riigikeel, matemaatika, keemia, füüsika, ajalugu, bioloogia, majanduse ja äri alused, kultuuriajalugu, läti keel, loodusteadused. Riigieksamite sooritamisel saadakse tulemused toorpunktidena ning vastavalt etteantud juhiste viiakse üle kuuepallisele skaalale tähistustega: A, B, C, D, E, F, kus A tähistab parimat tulemust ning F halvimat. Riigieksami tulemus tuuakse välja nii toorpunktide kujul kui ka kuuepallisel skaalal vastava tähisega. Vastuvõtul ülikoolidesse arvestatakse riigieksamite tulemusi ning kehtestatakse vastavalt erialale sisseastumislävend (*Institute, 2006; CCDE, 2006*).

4.2 AMEERIKA ÜHENDRIIGID

Ameerika Ühendriikides sooritavad õpilased keskkooli lõpetamisel Koolijõudluse Testi SAT (inglise keeles: *Scholastic Assessment Test*), mille tulemused, ületades kõrgkooli poolt seatud

lävendi, tagavad pääsu ülikooli. Vastavate testide koostamise, arendamise ja hindamisega tegeleb suurim ja tuntuim testimise keskus maailmas Pedagoogilise Testimise Teenistus (inglise keeles: *Educational Testing Service (ETS)*), mille keskus asub Ameerika Ühendriikides Princetonis ning lisaks omab testimiskeskusi rohkem kui 180 erinevas riigis (näiteks: Belgias, Prantsusmaal, Saksamaal, Hollandis, Poolas ja Inglismaal) ja kokku üle 9000 üksuse ning sooritajaid aastas on üle 2 miljoni. Hetkel on võimalik valida üle 30 erineva testi vahel, milledest populaarseimateks ja ka keskuse tähtsaimateks on siiani olnud Koolijõudluse Testid (SAT I ja SAT II), mõõdavad keskkoolilõpetajate lugemise, matemaatilise mõtlemise ja kirjutamise oskust. Testid on kõigile tasuta ning sooritamiseks tuleb eelnevalt registreeruda. Ainetestide valikul lähtub koolilõpetaja vastavalt erialale, kuhu ta soovib astuda. Peale testi sooritamist testimise keskus teatatakse tulemustest nii õpilasele kui ka kõrgkoolile, mille õpilane märkis registreerumisel. Kuna aastast-aastasse on testiküsimused erinevad, siis erinevate aastate tulemuste võrdlemiseks kasutatakse tulemuste hindamisel standardiseeritud skaalat ning sisse on viidud ühisosa, mis on ligikaudu 20% ulatuses erinevatel aastatel sama. Selleks võrreldakse testi sooritaja tulemusi ülejäänud grupiga, võttes arvesse nii grupi aritmeetilist keskmist kui ka standardhälvet. Tulemused avaldatakse täisarvudena ning standardiseerimismeetodina kasutatakse kolledži-skaalat. Iga osa eest on võimalik saada 200-800 punkti. Lisaks sellele tuuakse juurde ka võrdlus protsentides – kui suur hulk eksamisooritajaid oli protsentuaalselt alla vastavat (skaleeritud) tulemust (*ETS, 2006*).

4.3. ROOTSI KUNINGRIIK

Sarnane asutus Ameerika Ühendriikide Pedagoogilise Testimise Teenistusega asub ka Rootsis, Umeå Ülikooli juures, üksuse nimeks on Hariduse Juhtimise Osakond. Pedagoogiliste mõõtmistega tegeleva üksuse eesmärkideks on arendada meetodeid, teooriaid ja süsteeme, mille abil on indiviidide ja gruppide teadmisi, oskusi ning vilumusi võimalik mõõta. Näiteks on välja töötatud üleriigilised matemaatika testid, liikluseeskirjade tundmise testid ning lisaks ka Rootsi Koolijõudluse Test SweSAT, mis on olnud kasutusel juba 1977. aastast alates. Vastava testi tulemusi kasutatakse kõrgkooli sisseastujate valikul. Test mõõdab keskkoolilõpetajate taset viies erinevas aines ning koosneb viiest alamtestist, mis mõõdavad: sõnavara laiahaardelisust, matemaatilist mõtlemist, rootsi keele loetu mõistmist, diagrammide,

tabelite ja kaartidetõlgendamist ja inglisekeelse teksti mõistmist. Testid viiakse läbi kõrgharidusasutustes ning peale seda kogutakse kõigi testisooritajate andmed kokku ja neid hindavad Umeå Ülikoolis spetsiaalsed arvutiprogrammid. Selleks, et erinevate aastate tulemused oleksid omavahel võrreldavad, standardiseeritakse testi eest saadud toorpunktid. Standardiseerimiseks kasutatakse skaleerimist, kus tulemused viiakse skaalale, mis jääb vahemikku 0,0-2,0 (*SweSAT, 2006*).

5. UURIMUSE METODOLOOGIA

Uurimuse metodoloogia all toon koos põhjendustega välja töös kasutatud: uurimisstrateegia, valimite kirjelduse ja andmetöötluse viisi.

Uurimuse eesmärkideks on saada põhjalik valimite statistilistest näitajatest ning seda selleks, et teha järeldusi üldkogumi kohta ja vastavalt tulemustele valida sobilik(ud) standardiseerimismeetod(id).

5.1. UURIMISSTRATEEGIA

Tegu on empiirilise uurimisega, kus uurimisstrateegiana kasutan kvantitatiivset läbilõikeuuringut – uurimisobjektiks on vaadeldava perioodi eksaminandide riigieksamite tulemused valitud ainetes.

Üldkogumiks on kolme valitud aine vaadeldava ajavahemiku kõikide eksaminandide riigieksamite tulemused. Vaadeldavaks ajavahemikuks on aastad: 2000-2005 ning valitud aineteks: eesti keele kirjand, inglise keel ja matemaatika.

Valimitesse kuuluvad:

- kõik eesti keele kirjandi riigieksamite tulemused ajavahemikus 2001-2005 (2000. aasta on välja jäetud tänu teistsugusele hindamissüsteemile – tulemusi ei arvatud ümber 100-punktilisele skaalale);
- kõik inglise keele ja matemaatika riigieksamite tulemused ajavahemikus 2000-2005, välja arvatud 2003. aastal sooritatud 50-punktilised tööd (seetõttu, et kõik andmed oleksid ühtsetel alustel võrreldavad).

5.2. VALIMI KIRJELDUS

Valimites sisalduvad andmed on saadud REKK'i koduleheküljelt (*REKK, Statistika, 2006*).

Kuigi riigieksamid toimuvad ametlikult juba 1997. aastast alates, võtan uuritavasse valimisse kolme riigieksami (eesti keele kirjand, inglise keel, matemaatika) andmed alates 2000. aastast. 1997–2000. aastate andmed jätsin välja seetõttu, et ei saanud kõigile vajalikele andmetele

juurdepääsu. Lisaks vajab iga uus süsteem sisseelamisperioodi, mille jooksul saavutatakse riigieksamite läbiviimises ja tulemuste analüüsis teatav stabiilsus.

Valitud kolme (eesti keele kirjand, inglise keel, matemaatika) riigieksami kasuks otsustasin seetõttu, et kõigis kolmes aines sooritati riigieksam esimest korda juba 1997. aastal, mistõttu võib öelda, et 2000. aastaks oli olemas nendes ainetes kolmeaastane kogemus ning teiseks põhjuseks oli eksamisooritajate arv – vastavad kolm ainet on sooritajate arvu poolest kõige suuremad. Mida suurem on aga valim (siinkohal eksamisooritajate arv), seda täpsemalt on võimalik anda üldkogumi kohta hinnanguid.

Valimisse kuuluvad kõik eesti keele kirjandi, inglise keele ja matemaatika riigieksamite tulemused, mis on sooritatud ajavahemikus 2000-2005, välja arvatud 2003. aastal sooritatud inglise keele ja matemaatika riigieksamite 50-punktilised tööd.

2003. aastal oli võimalik katselisena sooritada inglise keele ja matemaatika (lisaks ka keemia) riigieksam kahes erinevas mahus – kas, nagu eelmistelgi aastatel, 100-punktilisel skaalal või uue võimalusena 50-punktilisel skaalal. Inglise keeles sooritas 50-punktilist tööd 250 ja matemaatikas 338 eksaminandi. Selleks, et kõigi eksamisooritajate tulemused oleksid ühtsetel alustel võrreldavad, kuuluvad erandina valimisse 2003. aastal sooritatud inglise keele ja matemaatika riigieksamite puhul vaid 100-punktiste eksamitööde sooritajate tulemused. 50-punktise eksami vahe võrreldes 100-punktise eksamiga seisnes eksami tasemes, kus 50-punktiline eksam eeldas sooritajalt vaid ainealaseid põhiteadmisi, 100-punktise eksami sisu oli aga oluliselt nõudlikum. Vastavat katselist riigieksamit ei ole peale 2003. aastat enam sooritatud.

Valim koosneb kolme aine tulemuste alamhulkadest, mille suurused on järgnevad:

- eesti keel (kirjand) – kokku (2000.a. 10622, 2001.a. 9356, 2002.a. 9874, 2003.a. 9585, 2004.a. 9472, 2005.a. 9757);
- inglise keel – kokku (2000.a. 9461, 2001.a. 8488, 2002.a. 9299, 2003.a. 9134, 2004.a. 9099, 2005.a. 9415);
- matemaatika – kokku (2000.a. 8023, 2001.a. 7248, 2002.a. 7267, 2003.a. 6904, 2004.a. 6894, 2005.a. 6934).

Valimi suuruseks on 156832 eksaminandi tulemused, mille hulk kuuluvad nii abiturientide kui ka varem lõpetanud (*REKK. Statistika, 2006*).

5.3. ANDMETÖÖTLUS

Andmed sain REKK'i koduleheküljelt, kus nad on juba eelnevalt aastate, koolide, õppeainete kaupa töödeldud. Andmete analüüsiks ja töötlemiseks kasutasin statistilist andmetöötlusprogrammi Ms Excel.

Analüüsi käigus püüdsin leida iseloomulikke jooni valitud riigieksamite sooritajate arvulise ja keskmiste toorpunkti-tulemuste kõikumise kohta kuue aasta jooksul. Lisaks sellele võrdlesin erinevate ainete ja aastate standardhälbeid ning tulemuste jaotust.

Antud meetodi valiku kasuks otsustasin lähtudes uurimisprobleemist, kus tõin esile vajaduse riigieksamite tulemuste standardiseerimiseks. Sellest johtuvalt on vajalik põhjalikuma ülevaate saamiseks eelnevalt analüüsida riigieksamite tulemusi.

6. ÜLEVAADE KOLME RIIGIEKSAMI TULEMUSTEST AASTATE LÕIKES

Järgnevalt analüüsin ajavahemikus 2000–2005 tehtud eesti keele kirjandi, inglise keele ja matemaatika riigieksamite tulemusi: eksaminandide arvulise jaotuse, tulemuste jaotuse, keskmiste tulemuste ning nende standardhälvete põhjal.

6.1. EESTI KEELE KIRJAND

Eesti keele kirjand on olnud kohustuslik riigieksam alates 1997. aastast eesti õppekeelega koolidele. Seoses sellega on selgitatav, miks on ta läbi riigieksamite ajaloo olnud enim sooritatud riigieksam - ajavahemikus 2000–2005 sooritatud kokku 58666 korral (*REKK. Statistika, 2006*).

Kirjandit hinnatakse kümnepallisüsteemis ning sajaballisüsteemi teisendamiseks korrutatakse alates 2001. aastast saadud tulemus kümnega. Kirjandit hindab kaks hindajat ja tulemuseks on kahe hindaja keskmine hinnang. Kui tulemuste vahe on suurem kui kaks punkti (10-punktilisel skaalal), siis saadetakse kirjand kolmandale lugemisele ning lõplik hinne kujuneb kolmanda hindaja ja tema hindete kõige lähemal seisva hindaja tulemuste aritmeetilisest keskmisest. Hindajad on kõrgelt kvalifitseeritud ja spetsiaalse väljaõppe saanud ning juhust, kus ka kolmanda hindaja tulemus erineks kahest esimesest hindajast üle kolme punkti ei ole juhtunud (*REKK, Eesti keel 2003, 28; REKK, Eesti keel 2002, 4*).

Eesti keele kirjand algab kell 10:00 ning kestab 6 astronoomilist tundi. Eeldatakse, et eksamikirjand on arutlev, analüüsiv. Eeldatavaks pikkuseks on 3–4 lehekülge (A4 formaat).

Kirjandi hindamisel võetakse arvesse järgmisi aspekte:

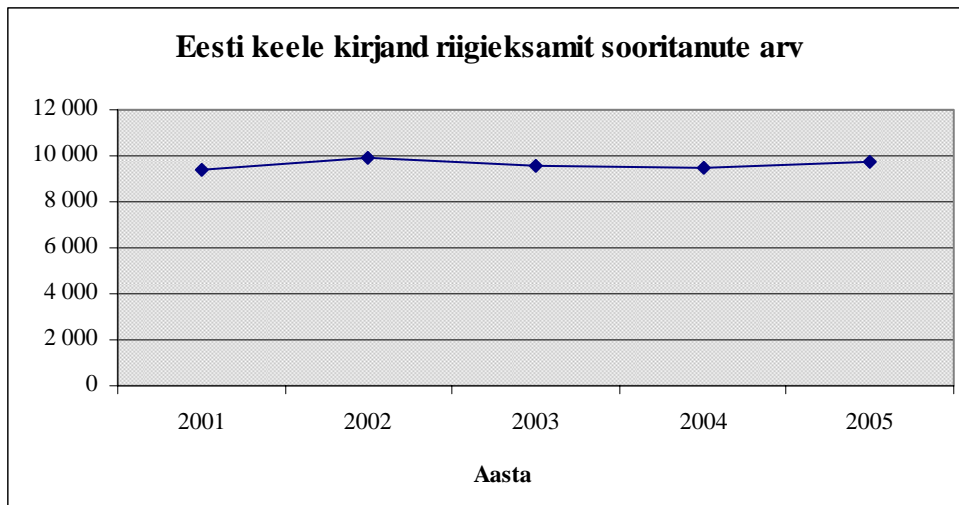
1. sisukus;
2. teemale vastavus, teema avatus;
3. põhjendamis- ja järeldamisoskus;
4. ülesehitus;
5. teksti sidusus, lausestus;
6. sõnavara ja stiil;
7. õigekeelsus ja interpunktsioon;
8. isikupära;

9. töö välimus;

(REKK, Eesti keel 2003, 28; REKK, Eesti keel 2002, 4).

Eesti keele kirjandiga ei mõõdeta konkreetseid teoreetilisi ainealaseid teadmisi, - kirjand mõõdab: eksaminandi küpsust, õigekirja- ja eneseväljendusoskust ning ka üldist lugemust (eeldatakse, et kirjandis kasutatakse kirjanduslikke allikaid). Eksamisooritajal on igal aastal valida kümne teema vahel, millest igaüks esindab üht kindlat rubriiki. Lõpliku sisulise valiku võimalikest teemadest teeb oma ala spetsialistidest koosnev ekspertgrupp, kes on uuritava aine asjatundjad ning sellega tagatakse ka riigieksami valiidsus.

Toon välja eesti keele kirjandi riigieksami eksamindandide arvulise jaotuse uuritavate aastate lõikes – soovimaks välja tuua aastatevahelist võrdlust, esitan tulemi joondiagrammi kujul.

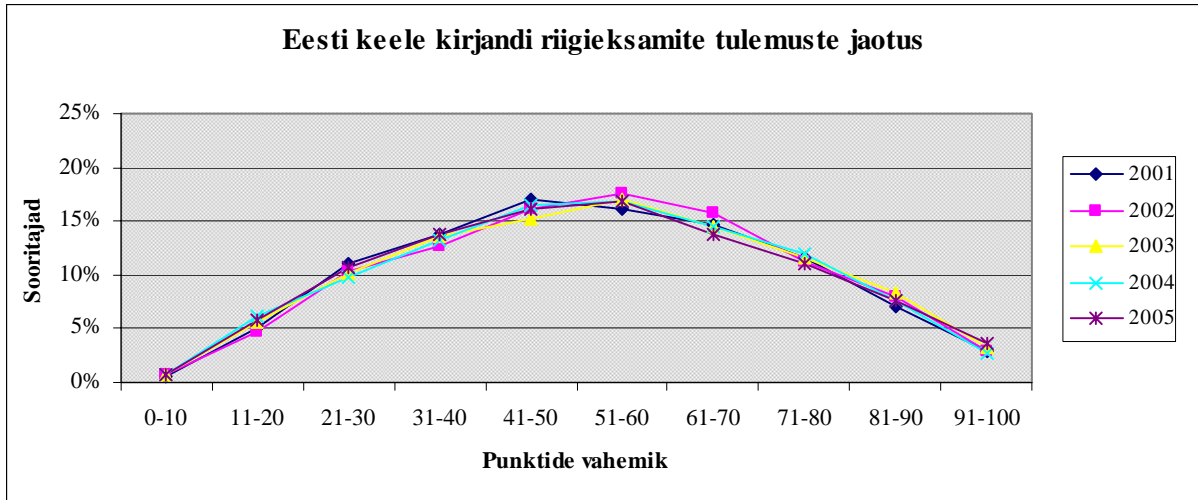


Joonis 3: eesti keele kirjandi riigieksamit sooritanute arv

(REKK. Statistika, 2006).

Jooniselt on näha, et eesti keele kirjandi riigieksami eksaminandide arv on uuritavas ajavahemikus püsinud ühtlasena. Nähtust võib selgitada sellega, et eesti keele kirjandi sooritamine on kohustuslik kõigile eesti õppekeelele gümnaasiumide lõpetajatele.

Järgnevalt esitan eesti keele kirjandi riigieksamite tulemuste jaotuse ajavahemikus 2001-2005 joondiagrammi kujul. Tulemused on esitatud 10-punktiste vahemikena.



Joonis 4: eesti keele kirjandi riigieksamite tulemuste jaotus

(REKK. Statistika, 2006).

Joonise põhjal on näha, et eesti keele kirjandi riigieksamite tulemused on väljatoodud ajavahemiku kõigil aastatel jagunenud ühtlaselt, normaaljaotusele sarnaselt, kus enim tulemusi on koondunud keskmiste punktide arvu juurde (41-50 ja 51-60 punkti) ning kõige vähem on saavutatud väga madalaid (0-10 punkti) ja väga kõrgeid (91-100 punkti) tulemusi. Eesti keele kirjand on selles suhtes erandlik, et kirjandi tulemuste järgi ei saa otsustada eksami raskusastme üle, kirjand peegeldab sooritajate küpsust ning kuna ta on kõigile eesti õppekeelega koolide lõpetajatele kohustuslik, siis on normaaljaotusele sarnane kõver igati ootuspärane (Niglas, 1997, 25).

Järgmiseks toon tabeli kujul välja uuritavate aastate eksamitulemuste keskmised ja standardhälbed⁴.

Tabel 1
Eesti keele kirjandi riigieksamite keskmiste tulemuste võrdlus

Aasta	Keskmine	Standardhälve
2000	56,54	21,20
2001	54,48	20,62
2002	55,35	20,54
2003	55,20	21,09
2004	54,54	20,86
2005	54,51	21,20

(REKK. Statistika, 2006).

⁴ Standardhälve – hajuvuse näitaja, näitab kui palju üksikud tulemused aritmeetilisest keskmisest (keskmiselt) erinevad.

Tabeli põhjal on näha, et eesti keele riigieksami keskmine punktide arv on uuritavas ajavahemikus (2000–2005) püsinud, väikeste kõikumistega, üsna stabiilsena. Madalaim keskmine punktide arv oli 2001.aastal (54,48 punkti) ning kõrgeim 2000. aastal (56,54 punkti), mis teeb suurima ja väikseima keskmise punktide arvu vaheks vaid 2,06. Kõrgeim tulemus on saavutatud uuritava ajavahemiku esimesel aastal (2000. aastal 56,54 punkti), siiski on viimastel aastatel (2004.aastal 54,54, 2005.aastal 54,51) saavutatud mõnevõrra madalamaid tulemusi, mis peegeldab seda, et eksamisooritajate küpsuse tase on mingisugusel põhjusel pisut langenud.

Standardhälbed on uuritavas ajavahemikus olnud suhteliselt sarnased, kogu uuritava ajavahemiku vältel ligilähedased 21 punktile. Suurim tulemuste kõikumine oli 2000. ja 2005. aastal kui standardhälve oli 21,2 ning väikseim kõikumine 2002. aastal kui standardhälve oli 20,5. Standardhälvete stabiilsus näitab ühtlasi ka riigieksami stabiilsust läbi aastate ning keskmiselt 21-punktine kõikumine keskväärtuse suhtes 100-punktise riigieksami piires ei ole liiga suur.

Keskmise tulemuse analüüsimisel tuleb arvestada ka seda, et emakeele kirjandi puhul ei saa õpilane valida, kas seda sooritada või mitte – kohustuslikus korras peavad selle sooritama kõik, ka need, kes antud valdkonnas end kindlana ei tunne. Samas kui valikeksamid valib õpilane enamasti vastavalt oma tugevatele külgedele.

6.2. INGLISE KEEL

Inglise keele riigieksam on olnud kavas alates 1997. aastast ning ajavahemikus 2000–2005 on teda sooritatud kokku 54896 korral, mis on kõikide riigieksamite arvestuses arvukuselt teine tulemus.

Inglise keele riigieksamil mõõdetakse võõrkeele oskust viiest aspektist lähtudes:

1. suulisest kõnest arusaamine (kuulamisoskus);
2. suulise suhtlemise oskus (kõnelemisoskus);
3. kirjalik tekstist arusaamine (lugemisoskus);
4. kirjalik suhtlemine (kirjutamisoskus);
5. keelestruktuuri tundmine;

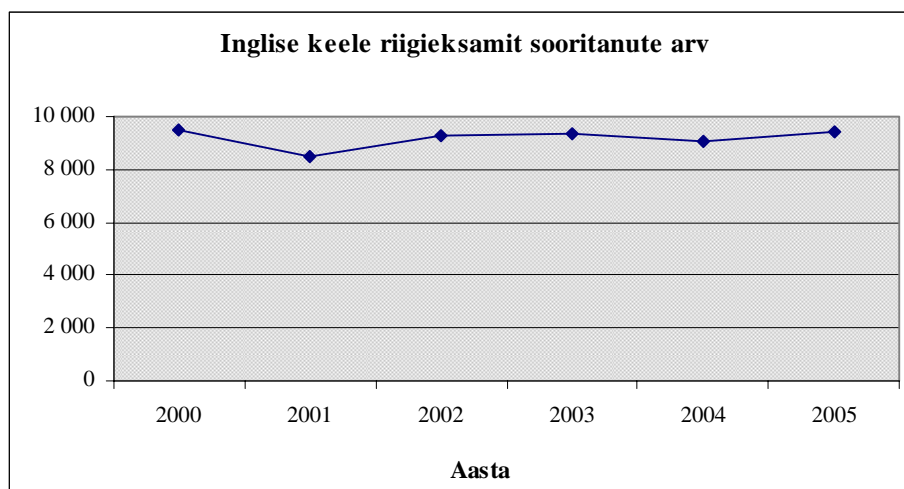
(REKK, *Inglise keel 2003*, lk 4).

Inglise keele riigieksam koosneb viiest osast ning iga osa kohta on testis mitu erinevat ülesannet, millede temaatika on valitud vastavalt õppekavale. Lõpliku sisulise valiku võimalikest ülesannetest teeb oma ala spetsialistidest koosnev ekspertgrupp, kes on uuritava ala asjatundjad ning sellega tagatakse ka testi valiidsus. Iga ülesande kohta on eksamisooritajal teada täpselt võimalike punktide arv. Riigieksamitöö eest on maksimaalselt võimalik saada 100-punkti (minimaalselt null punkti), iga osa eest 20 (REKK, *Inglise keel 2002*, lk 7).

Korraldusliku poole pealt toimub eksam kahe päeval – esimesel päeval on eksami kirjalik osa ning teisel suuline. Kirjalik eksam jaguneb kaheks, alates kirjutamisosaga, seejärel, peale lühikest pausi, sooritatakse ülejäänud osad: lugemis-, kuulamis- ning struktuuriosa.

Eksamil kasutatakse kahte tüüpi hindamissüsteemi – lugemine, kuulamine ja keelestruktuuri tundmine, on kindlate võtmete abil üheselt hinnatavad ning suulist ja kirjutamisosa hindavad aga etteantud hindamisskaala abil vastava väljaõppe saanud hindajad.

Toon välja inglise keele riigieksami eksamindandide arvulise jaotuse uuritavate aastate lõikes – soovimaks välja tuua aastatevahelist võrdlust, esitan tulemi joondiagrammi kujul.



Joonis 5: inglise keele riigieksamit sooritanute arv

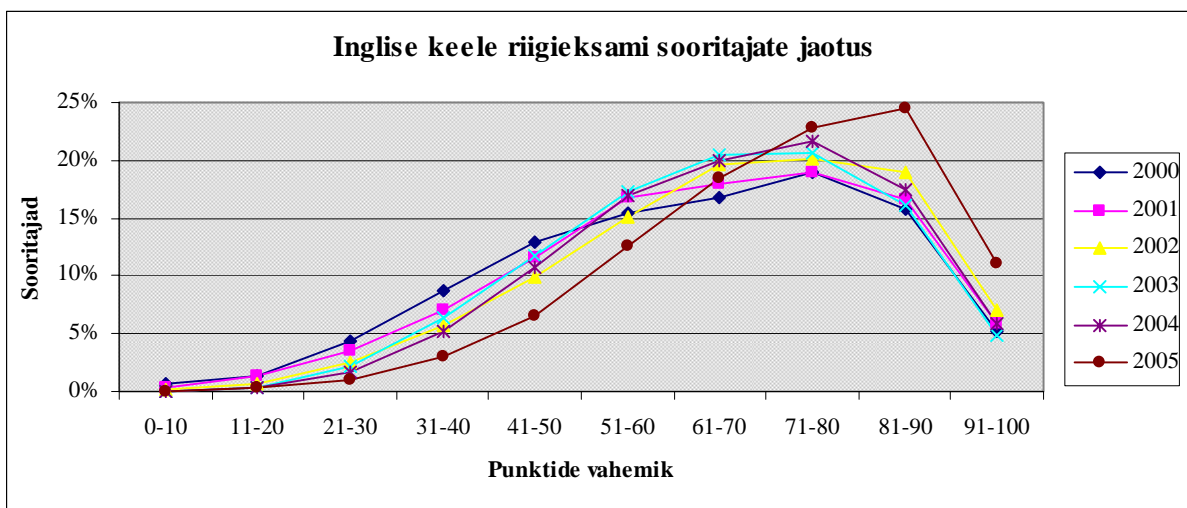
(REKK. *Statistika, 2006*).

Jooniselt on näha, on eksaminandide arv olnud suurim uuritava ajavahemiku esimesel aastal (9461 eksaminandi) ning 2001. aastal elanud üle suure languse, kasvades järgmisel

(2002.aastal) taas ligikaudu oma endisele tasemele ning püsinud seal kuni uuritava perioodi (2005.aasta) lõpuni. Suurima sooritajate arvu (2000.aasta, sooritajaid 9461) ning väikseima sooritajate arvu (2001. aasta, sooritajaid 8488) vahelise kõikumise ulatus oli vaid 973, mis näitab samuti, et eksaminandide arvu jaotumine ei ole väga hajuv (*REKK. Statistika, 2006*).

Inglise keele riigieksami jätkuvat populaarsust võib selgitada õpingute jätkamisel nõutavate riigieksamitega, pea kõigil erialadel on selleks ka võõrkeele riigieksami tulemus ning inglise keel on Eesti Vabariigis A-keelena enim õpitud võõrkeel.

Järgnevalt esitan inglise keele riigieksamite tulemuste jaotuse ajavahemikus 2000-2005 joondiagrammi kujul. Tulemused on esitatud 10-punktiste vahemikena.



Joonis 6: inglise keele riigieksamite tulemuste jaotus

(*REKK. Statistika, 2006*).

Joonise põhjal on näha, et inglise keele riigieksamite tulemused on väljatoodud ajavahemiku erinevatel aastatel on olnud küllaltki kõikumised ning tulemuste jaotus erineb normaaljaotuse kujust, olles tugeva negatiivse asümmeetriaga. Väiksemad tulemuste kõikumised esinevad ka vahemikusiseselt, kuid eriti tõuseb esile just 2005. aasta tulemuste erinevus teistest. Näiteks tulemusi, vahemikus 31-40, 41-50 ja 51-60 punkti, saadi sellel aastal teiste aastatega võrreldes vähem ning samas kõrgeid tulemusi, vahemikus 81-90 ning 91-100, jällegi oluliselt rohkem. Eriti torkavad silma 2005. aasta puhul just väga kõrgete tulemuste ülekaal – 24,5% tulemusi

jäid vahemikku 81-90 punkti ning 22,8% tulemustest vahemikku 71-80 punkti, mis näitab selgelt, et eksami raskusaste oli sooritajate jaoks pigem madal.

Inglise keele riigieksam on valikeksam, siis seetõttu ei saa tulemuste asümmeetria põhjal teha päris üheseid järeldusi. Kuna tulemused on negatiivse asümmeetriaga⁵, siis võiks järeldada, et riigieksami raskusaste on vaadeldava perioodi jooksul olnud madalam, aga samas võib selline nähtus olla tingitud ka sellest, et riigieksamit sooritasid vaid need eksaminandid, kelle tase ongi kõrge. Riigieksam mõõdab sooritaja pädevusi vastavalt riiklikule õppekavale ning on loomulik, et inglise keele süvaklasside tase on võrreldes tavaklassiga tunduvalt kõrgem ning nende taseme objektiivseks võrdlemiseks jäävad riigieksami piirid natuke kitsaks.

Samas, vaadetes inglise keele riigieksami sooritajate suur hulka, võib järeldada, et ilmselt sooritasid eksamit ka kõrgest erineva tasemega eksaminandid ning eksami üldine raskusaste on olnud ikkagi pigem madalam.

Järgmiseks toon tabeli kujul välja uuritavate aastate keskmised tulemused ja standardhälbed.

Tabel 2
Inglise keele riigieksamite tulemuste võrdlus

<i>Aasta</i>	<i>Keskmine</i>	<i>Standardhälve</i>
2000	62,53	19,64
2001	64,09	18,78
2002	66,69	17,84
2003	64,15	16,90
2004	66,60	16,73
2005	71,93	16,02

(REKK. Statistika, 2006).

Tabelist ilmneb, kõigub keskmine punktide arv mõõdukalt. Aastatel 2000–2002 on keskmine hinne järk-järgult tõusnud, 2003. aastal 2,54 punkti võrra langenud ning alates 2004.aastast taas tõusnud. Kõige madalam oligi keskmine hinne 2000.aastal (62,53 punkti) ning kõige kõrgem 2005.aastal (71,93 punkti), vahe kõrgeima ja madalaima keskmise tulemuse vahel on 9,4 punkti. 2005. aastal on keskmine tulemus võrreldes 2004. aastaga päris intensiivselt, 5,33 punkti võrra, tõusnud.

⁵ Negatiivne asümmeetria – tulemuste jaotus on välja venitatud nii, et enamik tulemusi on koondunud skaalal paremale poole.

Standardhälbed on uuritava ajavahemiku jooksul järjest vähenenud, olles uuritava ajavahemiku esimesel aastal 19,64 ning kahanenud viimaseks aastaks 16,02'ni. Vastavad standardhälbed on 100-punktise riigieksami puhul täiesti normaalsed ning kahanemine 16,02 punktini näitab tulemuste jaotumise ühtlustumise tendentsi võrreldes keskväärtusega.

6.3. MATEMAATIKA

Matemaatika riigieksam on olnud kavas alates 1997. aastast ning ajavahemikus 2000–2005 on teda sooritatud kokku 43270 korral, mis on kõikide riigieksamite arvestuses arvukuselt kolmas tulemus.

Matemaatika riigieksami eesmärkideks on:

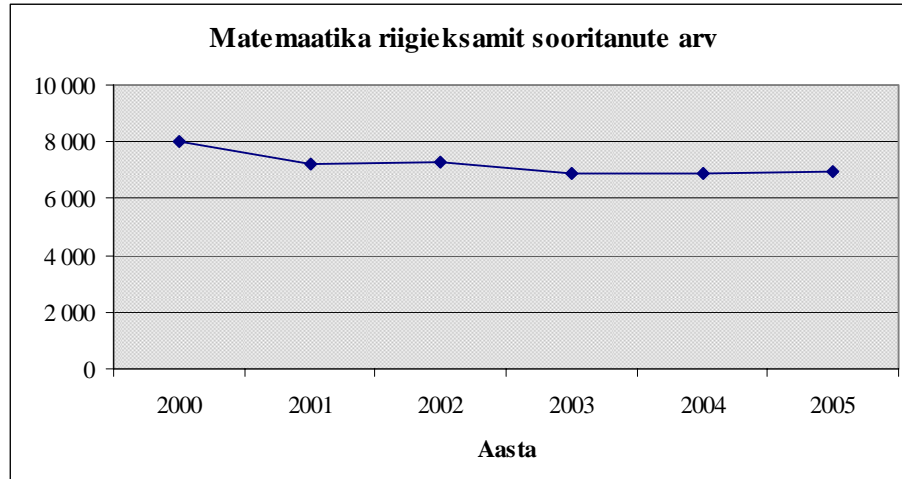
1. välja selgitada kui hästi saab gümnaasiumilõpetaja aru matemaatika mõistetest, faktidest, printsiipidest ja protseduuridest;
2. kuivõrd struktureeritud ja korrastatud on tema teadmised;
3. kui hästi ja suudab õpitud rakendada, lahendada mitterutiinseid ülesandeid;
4. milline on gümnaasiumilõpetaja matemaatikaalane ettevalmistus õpingute jätkamiseks järgmisel astmel;
5. anda gümnaasiumilõpetajale võimalus hinnata oma valmidust ja suutlikkust langetada iseseisvalt otsustusi.

(REKK, Matemaatika 2004, 4).

Eksami kestvus on 5 astronoomilist tundi. Eksamitööd on koostatud testi kujul kahes variandis. Töid hinnatakse ülesannete kaupa ning iga ülesande kohta on eksamisooritajal teada võimalike punktide arv. Lõpliku sisulise valiku võimalikest ülesannetest teeb oma ala spetsialistidest koosnev ekspertgrupp, kes on uuritava omaduse asjatundjad ning sellega tagatakse ka riigieksami valiidsus *(REKK, Matemaatika 2000, 5)*.

Riigieksamitöö eest on maksimaalselt võimalik saada 100-punkti (minimaalselt null punkti).

Toon välja matemaatika riigieksami eksamindandide arvulise jaotuse uuritavate aastate lõikes – soovimaks välja tuua aastatevahelist võrdlust, esitan tulemi joondiagrammi kujul.

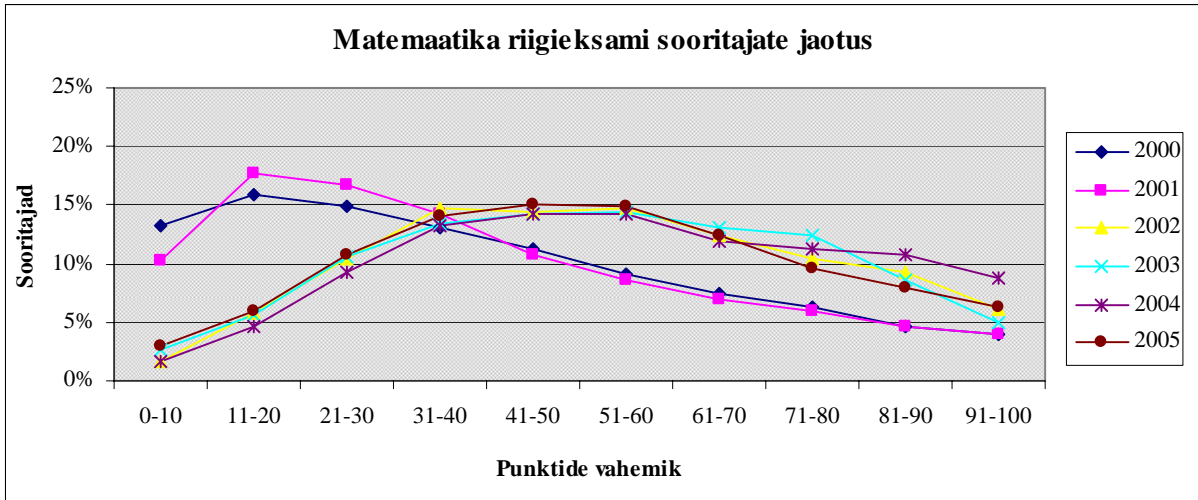


Joonis 7: matemaatika riigeksamit sooritanute arv

(REKK. Statistika, 2006).

Jooniselt on näha, et matemaatika riigeksami sooritajate arv uurita ajavahemiku jooksul näidanud üles kahanevat tendentsi. Suurim oliksamindandide arv uuritava ajavahemiku esimesel aastal (8023 eksaminandi), seejärel langes hüppeliselt 775 võrra (2001. aastal) ning püsis sarnasel tasemel ka 2002.aastal, seejärel langes 363 võrra ning on 6900 piirimail püsinud kuni 2005. aastani. Vastavat nähtust saab taas selgitada 2000. aasta ning 2005. aasta riigeksamite koosseisulise erinevusega. Vastavas ajavahemikus on riigeksamite nimistusse juurde lisandunud kolm sooritajate arvukuselt suurt riigeksamit (geograafia, riigikeel, ühiskonnaõpetus). Teine aspekt oleks kõrgkoolidesse õppima minemisel nõutavad riigeksamid ning erialade populaarsus – enamikel erialadest nõutakse emakeele kirjandi ning võõrkeele riigeksamit. Arvesse tuleks võtta ka, et vene õppekeele koolides on kolmest kohustuslikust riigeksamist kaks ettemääratud (emakeele kirjand ja riigikeele eksam) ning lõpetajal on minimaalselt võimalik valida vaid üks riigeksam omal valikul, juhul kui ei sooritata enam kui kolme riigeksamit (maksimum viis) (REKK. Statistika, 2006).

Järgnevalt esitan matemaatika riigeksamite tulemuste jaotuse ajavahemikus 2000-2005 joondiagrammi kujul. Tulemused on esitatud 10-punktiste vahemikena.



Joonis 8: matemaatika riigieksamite tulemuste jaotus

(REKK. Statistika, 2006).

Joonise põhjal on näha, et riigieksamite tulemused on väljatoodud ajavahemiku erinevatel erinevatel aastatel on olnud küllaltki kõikuvad just punktivahemike siseselt. 2000. ja 2001. aastal on madalate tulemuste (0-10, 11-20 ja 21-30 punkti) osakaal teiste aastatega võrreldes oluliselt suurem ning kõrgemate tulemuste osakaal jällegi madalam. Selline nähtus, kus sooritajate koguarvust saavad keskmiselt 12% eksaminandidest 0-10 punkti ja keskmiselt 17% tulemuseks 11-20 ning vaid keskmiselt 4% 81-90 või 91-100 punkti, näitab selgelt seda, et sooritajate jaoks on eksami raskusaste olnud pigem kõrge. Antud nähtust saab põhjendada 2002. aastal kehtestatud 20-punktilise soorituspiiriga, peale mida sooritasid matemaatika riigieksamit pigem eksaminandid, kes tundsid end aines tugevamalt. Lisaks sellele on teistest eristatav ka 2004. aastal riigieksamit sooritanute suurem osahulk väga kõrgete tulemuste hulgas (81-90 ja 91-100 punkti), mis viitab võimalusele, et suure hulga sooritajate jaoks oli eksami raskusaste võrreldes teiste aastatega madalam. Kui arvesse võtta vaid 2002-2005. aasta tulemused, siis sarnaneb tulemuste jaotuskõver normaaljaotusega.

Järgmiseks toon tabeli kujul välja uuritavate aastate keskmised tulemused ja standardhälbed.

Tabel 3
Matemaatika riigieksamite tulemuste võrdlus

<i>Aasta</i>	<i>Keskmine</i>	<i>Standardhälve</i>
2000	39,19	25,79
2001	39,22	25,15
2002	53,41	22,91
2003	53,00	23,09
2004	56,20	23,57
2005	52,04	23,39

Tabeli põhjal on näha, et matemaatika riigieksami keskmine punktide arv on uuritavas ajavahemikus (2000–2005) olnud küllaltki kõikuv. Madalaim keskmine punktide arv 2000.aastal (39,19 punkti) ning kõrgeim 2004.aastal (56,20 punkti), mis teeb suurima ja väikseima keskmise punktide arvu vaheks juba 17,01. Võrreldes uuritava ajavahemiku algusaastatega (2000.aastal 39,19, 2001.aastal 39,22), on keskmine punktide arv alates 2002. aastast (keskmine 53,41) hüppeliselt tõusnud, saavutades 2004. aastal maksimumi keskmise tulemusega 56,20 ning seejärel 4,16 punkti võrra langenud. Keskmise hinde hüppelist tõusu võib seletada juba eelpool mainitud 2002. aastal riigieksamitele kehtestatud 20-punktise soorituspiiriga.

Võrreldes teiste uuritavate riigieksamitega (eesti keele kirjand, inglise keel) on matemaatika riigieksamite standardhälve suurim, olles uuritava ajavahemiku kahel esimesel aastal (2000 ja 2001) üle 25 punkti, mis näitab eksamitulemuste suhteliselt suurt kõikumist võrreldes keskmise tulemusega. Viimaste aastate jooksul on standardhälve püsinud siiski 23 punkti kandis, mis näitab tulemuste mõningast stabiliseerumist võrreldes keskväärtusega. 23-24-punktine standardhälve ei ole võrreldes 100-punktise riigieksamiga nii suur, et keskmiste tulemuste põhjal järeldusi teha ei saaks.

7. RIIGIEKSAMITE TULEMUSTE STANDARDISEERIMINE

Standardiseerimiseks kasutan analüüsitud riigieksamite tulemusi ning meetoditena peatükis, “2.5. Testitulemuste standardiseerimine“, väljatoodud kuute erinevat standardiseerimismeetodit. Lõppeesmärgiks on saada ülevaade erinevate standardiseerimismeetodite mõjust tulemuste jaotusele ning vastavalt sellele teha järeldusi meetodi sobivuse üle.

7.1. STANDARDISEERIMISE VAJALIKKUS

Seoses sellega, et riigieksamite tulemused on erinevate ainete ning aastate lõikes erinevad, ei ole need omavahel üheselt võrreldavad. Teravalt kerkib problemaatika üles just kõrgkooli sisseastumisel, mis seab erinevate aastate ning eksamiainete sooritajad ebavõrdsesse olukorda. Näiteks matemaatika riigieksami keskmine tulemus 2000. aastal oli 39,19 punkti ning 2004. aastal 56,20 punkti (vt. Tabel 3, lk 46). Vaadates tulemuste üldist jaotuvust on selgelt näha, et kõrgkooli sisseastumisel on 2004. aastal lõpetajatel tuntav eelis 2000. aastal lõpetanute ees. Ligi 17-punktilisel erinevusel on mitmeid põhjuseid: 2001. aastal kehtestatud 20-punktiline soorituspiir paneb eksaminandi valima aine kasuks, milles ta end kindlamini tunneb, lisaks sellele oli 2000. aastal riigieksamianete valik väiksem. Tulemuste erinevused võivad tekkida ka õpilaste paremast ettevalmistusest või tasemest, riigieksami testi erinevustest. Teise näitena võib kõrgkooli sisseastumisel tuua võõrkeelte riigieksamite tulemuste erinevused. Näiteks REKK'i koduleheküljel avaldatud riigieksamite statistika põhjal on näha, et 2005. aastal sooritanute inglise keele võõrkeele riigieksami keskmine tulemus oli 71,93 punkti (sooritajate arv 9415), vene keele võõrkeele riigieksami keskmine tulemus oli aga 78,22 punkti (sooritajate arv 485) (REKK. Statistika, 2006). Kõrgkooli sisseastumisel aga, kus erinevate võõrkeelte osakaalud on võrdsed, on siinkohal eelis neil, kes on 2005. aastal sooritanud võõrkeele riigieksami vene keeles. Tegu on küll ühel ja samal aastal sooritatud võõrkeele riigieksamitega, kuid keskmiste punktide vahe on 6,29. Lisaks sellele tuleb tähelepanu pöörata ka eksamisooritajate arvule, vaadates riigieksamite statistikat, kaldub tendents olema sarnaspoole, et (samalaadsete ainete puhul) väiksema sooritajate arvuga riigieksamite keskmised tulemused kipuvad olema kõrgemad, kuna eksamit sooritavad vaid kõrgema tasemega õpilased.

Selleks, et tekkinud olukorda muuta, tulekski seni väljastatud toorpunktide asemel kasutada standardiseeritud punkte, mis muudaksid tulemused omavahel (ainete ja aastate lõikes) võrreldavaks.

7.2. VÕIMALUSED JA OHUD

Diferentseerimaks eksamisooritaja tegelikku tulemust, oleks tarvis võrrelda grupisiseseelt (ainete kaupa) tema tulemust teistega. Võrrelda võib korraga omavahel mitme aasta tulemusi või iga aasta tulemusi eraldi. Mõlemal variandil on omad plussid ja miinused. Kui me arvestame grupiks ühes aines samal aastal sooritatud riigieksamite tulemused, siis võtame arvesse riigieksamist tulenevaid võimalikke erinevusi, küll aga ei võta me arvesse seda, et antud aastal võiski eksamisooritajate üldine tase olla teine. Võtan näiteks inglise keele riigieksami. 2004. aasta keskmine tulemus oli 66,60 punkti ning 2005. aasta keskmine tulemus oli 71,93 punkti (*vt. Tabel 2, lk 42*). Tulemuste jaotused erinevad teineteisest mõnevõrra, kuid seda, kas tulemuste erinevused tulenevad riigieksami testi erinevast raskusastmest või 2004. aasta ja 2005. aasta eksamisooritajate tasemelisest erinevusest, ei saa vastavate andmete põhjal välja lugeda. Kui arvestada, et erinevused tulenesid testi raskusastmest, siis on tulemuste võrdlemiseks õige kasutada ühe aasta tulemusi, kui aga erinevused tulenesid eksamisooritajate tasemest, siis oleks õige võrrelda mõlema aasta tulemusi koos.

Eksaminandide üldist taset oleks võimalik tuvastada, võttes aluseks eksamisooritajate gümnaasiumi keskmised hinded ning vastavalt sellele on näha tasemete erinevust aastate lõikes. Samas kerkib siit üles aga koolide hinnete võrreldavuse küsimus – kas on omavahel üheselt võimalik võrrelda inglise keele hinnet süvaklassi ja tavaklassi õpilaste vahel? Kui igal aastal oleks süva- ja tavaklasside inglise keele eksamisooritajate arvud proportsionaalselt ligilähedased, siis oleks selline võrdlus üldistel alustel teostatav, kui aga mitte, siis ilmselt tuleks lisaks veel eraldi võrrelda erinevate aastate süva- ja tavaklasside õpilaste tulemusi.

Teise võimalusena võiks riigieksamitesse järjestikustel aastatel sisse viia ühisosa, mis võimaldaks erinevate aastate eksamisooritajate tasemeid omavahel võrrelda ning nende põhjal järeltõlgida. Samas kerkib üles probleem varem tehtud eksamite suhtes, kuna nende tööde puhul ühisosa sisseviimist enam kasutada ei saa. Asjaolu, et riigieksamite tulemuste erinevused võivad tuleneda nii sooritajate tasemest kui ka riigieksami sisust, muudab ühtsete

standardiseerimisaluste väljatöötamise mõnevõrra keerukamaks, kuid näiteks ühisosa sisseviimisel oleks eksamisooritajate tasemed omavahel võrreldavad ilma täiendavate andmete kogumiseta. Paraku ei ole erinevatel aastatel võimalik (ega ka vajalik) teha täpselt ühesuguseid riigieksami teste ning sellest johtuvalt jääb alles võimalus, et riigieksamite tulemused toorpunktide kujul jäävad aastate lõikes kõikumata. Erinevuste, mis tulenevad testi sisust, ühtlustamiseks on võimalik kasutada aga mitmeid erinevaid standardiseerimise meetodeid, mille võimalusi ma järgnevalt riigieksamite andmete põhjal analüüsingi.'

Võtan vaatluse alla kolme analüüsitud aine: eesti keele kirjand, inglise keel ja matemaatika, riigieksamite tulemuste standardiseerimise võimalused vastavalt peatükis, „Testitulemuste standardiseerimine“, väljatoodule.

Riigieksamite analüüsi käigus selgus, et eesti keele kirjandi tulemuste jaotused vastasid üldjoontes normaaljaotusele, matemaatika riigieksami puhul oli üldine tendents, kui välja arvata 2000. ja 2001. aasta tulemused, samuti normaaljaotuse sarnane, kuid inglise keele riigieksamite tulemuste jaotus oli märgatavalt negatiivse asümmeetriaga.

Järgnevalt analüüsin erinevate näidete abil eelpool väljatoodud standardiseerimismeetodite mõju tulemuste jaotusele. Selleks võrdlen omavahel nii ühe aine erinevate aastate kui ka erinevate ainete (sama aasta) tulemusi.

Erinevate aastate ja ainete tulemuste võrdlemiseks standardiseerin kõigepealt valimi kõik väärtused ning toon nad välja ülevaate saamiseks tabelkujul 10-toorpunktiliste intervallidena. Seejärel võrdlen jooniste abil standardiseerimismeetodi mõju tulemuste jaotusele.

Valimite kõik standardiseeritud väärtused z-skaalal on välja toodud magistritöös, „LISA 1.“, all ning fikseeritud keskmise skaalal, „LISA 2.“, all.

7.2.1. STANDARDISEERIMINE Z-SKAALALE

Esimese näitena vaatlen eesti keele kirjandi 2003. ja 2004. aasta riigieksamite tulemuste standardiseerimist z-skaalale ning saadud tulemused esitan tabelkujul (*vt. valem 1, lk 17*).

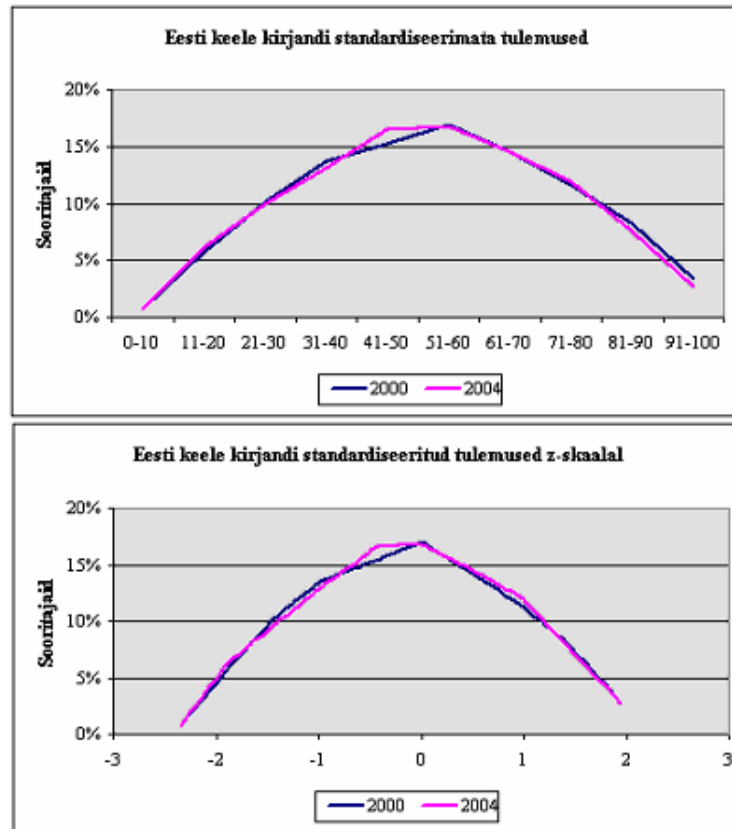
Tabel 4
Eesti keele kirjandi standardiseeritud tulemused z-skaalal

<i>Toorpunktid</i>	<i>z-skaala</i>	
	<i>2003</i>	<i>2004</i>
0	-2,62	-2,61
10	-2,14	-2,14
20	-1,67	-1,66
30	-1,19	-1,18
40	-0,72	-0,70
50	-0,25	-0,22
60	0,23	0,26
70	0,70	0,74
80	1,18	1,22
90	1,65	1,70
100	2,12	2,18

(vt. Tabel 1, lk 39).

Tabeli põhjal on näha, et kui tulemused toorpunktidenä olid võrdsed, siis ka standardiseeritud kujul esineb siin vaid väikseid erinevusi. Kuna eesti keele kirjandi riigieksamite tulemuste jaotus oli ligilähedane normaaljaotusele ning ühtlasi erinevate aastate tulemuste jaotus toorpunktide kujul oli sarnane, siis suuri erinevusi erinevate aastate standardiseeritud tulemuste vahel ei teki. Näiteks eksamisooritaja 60-punktiline tulemus toorpunktide kujul annab standardiseeritud kujul: 2003. aastal tulemuseks 0,23 ja 2004. aastal tulemuseks 0,26. Tulemused, mis olid enne alla/ üle keskmise, on seda ka peale standardiseerimist.

Järgnevalt võrdlen diagrammide abil standardiseerimismeetodi mõju tulemuste jaotusele.



Joonis 8: eesti keele kirjandi standardiseeritud ja standardiseerimata tulemuste võrdlus

Võrreldes joonise abil eesti keele kirjandi standardiseerimata ja standardiseeritud tulemuste jaotust, on näha, et tulemuste jaotus ei muutu oluliselt peale standardiseerimist ning seda seetõttu, et keskväärtused ja standardhälbed ei erine aastate lõikes teineteisest oluliselt. Sarnased, oluliste muutusteta tulemused on saavutatavad ka teiste, antud töös käsitlemist leidvate, standardiseerimismeetoditega. Tulemuste jaotuse ühtluse tõttu ei ole vastava standardiseerimismeetodi kasutamiseks mitte mingisuguseid takistusi või kitsendusi. Kuna tulemustes esineb negatiivseid arve, siis tuleks kasutajasõbralikkuse saavutamiseks andmed teistele skaaladele üle viia, kasutades selleks näiteks peatükis, „Testitulemuste standardiseerimine“, väljatoodud skaalasisid.

Järgnevalt esitan tabelkujul standardiseeritud eesti keele kirjandi 2003. ja 2004. aasta riigieksamite tulemused: staniinide-skaalale (C), T-skaalale (T) ja kolledži-skaalale (k) (vt. valem 2, lk 18; valem 3, 19; valem 4, lk 19).

Tabel 5
Eesti keele kirjandi standardiseeritud tulemused staniinide-, T- ja kolledži-skaalal

Toorpunktid	2003			2004		
	C	T	k	C	T	k
0	0	24	238	0	24	239
10	1	29	286	1	29	286
20	2	33	333	2	33	334
30	3	38	381	3	38	382
40	4	43	428	4	43	430
50	5	48	475	5	48	478
60	5	52	523	6	53	526
70	6	57	570	6	57	574
80	7	62	618	7	62	622
90	8	67	665	8	67	670
100	9	71	712	9	72	718

(vt. Tabel 4, lk 50).

Tabeli põhjal on näha, et peale z-skaalal standardiseeritud tulemuste üleviimist staniinide-, T- ja kolledži-skaalale, ei esine jaotuses enam negatiivseid arve.

Staniinide-skaala tulemused avaldatakse ühekohaliste arvudena ning sellisel kujul ei erista tulemusi teineteisest väga täpselt.

Kuna T-skaala tulemused on sarnased toorpunktidele, siis tuleks arvestada, et antud skaala puhul ei ole otstarbekas tulemusi korruga väljastada nii toorpunktide kui standardiseeritud punktide kujul, vastasel juhul võivad tulemused teineteisega segamini minna.

Kolledži-skaala tulemuste puhul ei ohtu, et neid võiks segamini ajada toorpunktliste tulemustega. Antud skaala puhul on vabalt võimalik avaldada riigieksamite tulemused korruga nii toor- kui standardpunktide kujul.

Järgnevalt võrdlen aga matemaatika kahe 2000. ja 2004. aasta tulemuste standardiseerimist z-skaalale, mis on teineteisest märgatavalt erinevad. Saadud tulemused esitan 10-punktiliste vahemikena tabelkujul.

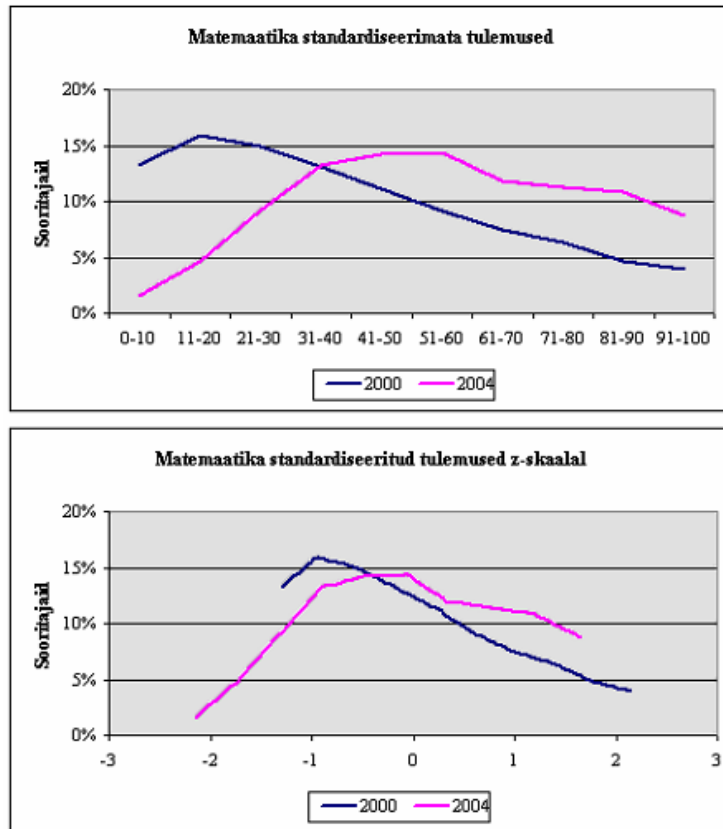
Tabel 6
Matemaatika standardiseeritud tulemused z-skaalal

Toorpunktid	z-skaala	
	2000	2004
0	-1,52	-2,38
10	-1,13	-1,96
20	-0,74	-1,54
30	-0,36	-1,11
40	0,03	-0,69
50	0,42	-0,26
60	0,81	0,16
70	1,19	0,59
80	1,58	1,01
90	1,97	1,43
100	2,36	1,86

(vt. Tabel 3, lk 46).

Tabeli põhjal on näha, et kui tulemused toorpunktidena olid võrdsed, siis standardiseeritud kujul esineb siin juba märgatavaid erinevusi. Näiteks eksamisooritaja 30-punktiline tulemus toorpunktide kujul annab standardiseeritud kujul: 2000. aastal tulemuseks -0,36 ja 2004. aastal tulemuseks -1,11. Tulemused aga, mis olid enne alla/ üle keskmise, on seda ka peale standardiseerimist. Selline tulemus oli etteaimatav, kuna matemaatika riigieksamite jaotuses toorpunktide kujul olid väljatoodud aastate tulemuste jaotused ka toorpunktidena tuntavalt erinevad.

Järgnevalt võrdlen diagrammide abil standardiseerimismeetodi mõju tulemuste jaotusele.



Joonis 9: matemaatika standardiseeritud ja standardiseerimata tulemuste võrdlus

Võrreldes joonise abil matemaatika standardiseerimata ja standardiseeritud tulemuste jaotust, on näha, et tulemuste jaotus on peale standardiseerimist muutunud – võrdsustunud on mõlema aasta keskmised tulemused ning vähenenud on hajuvused. Standardiseerimine muudab võrreldavaks omavahel ühe aine kahe erineva aasta tulemused.

Järgnevalt esitan tabelkujul standardiseeritud matemaatika 2000. ja 2004. aasta riigieksamite tulemused: staniinide-skaalale (C), T-skaalale (T) ja kolledži-skaalale (k) (vt. valem 2, lk 18; valem 3, 19; valem 4, lk 19).

Tabel 7
Matemaatika standardiseeritud tulemused staniinide-, T- ja kolledži-skaalal

Toorpunktid	2000			2004		
	C	T	k	C	T	k
0	2	35	348	0	26	262
10	3	39	387	1	30	304
20	4	43	426	2	35	346
30	4	46	464	3	39	389
40	5	50	503	4	43	431
50	6	54	542	4	47	474
60	7	58	581	5	52	516
70	7	62	619	6	56	559
80	8	66	658	7	60	601
90	9	70	697	8	64	643
100	10	74	736	9	69	686

(Tabel 6, lk 53).

Tabeli põhjal on näha, et peale z-skaalal standardiseeritud tulemuste üleviimist staniinide-, T- ja kolledži-skaalale, ei esine jaotuses enam negatiivseid arve.

Kolmanda näitena vaatlen eesti keele kirjandi, inglise keele ja matemaatika 2001. aasta tulemuste standardiseerimist z-skaalale. Kõigepealt esitan toorpunktide standardiseeritud tulemused tabelkujul.

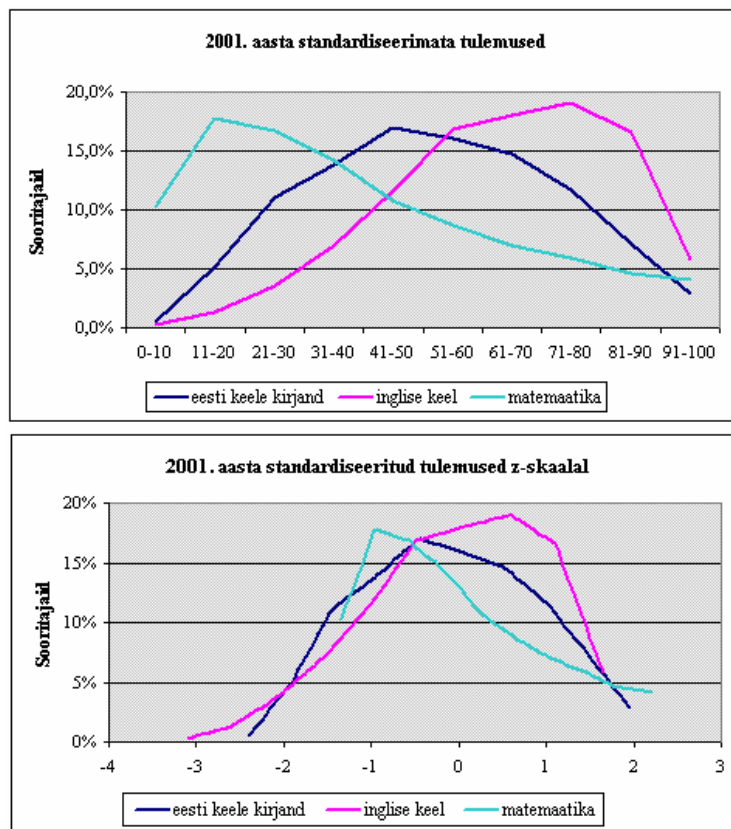
Tabel 8
Eesti keele kirjandi, inglise keele, matemaatika standardiseeritud tulemused z-skaalal

Toorpunktid	2001. aasta tulemused z-skaalal		
	eesti keel	inglise keel	matemaatika
0	-2,64	-3,41	-1,56
10	-2,16	-2,88	-1,16
20	-1,67	-2,35	-0,76
30	-1,19	-1,82	-0,37
40	-0,70	-1,28	0,03
50	-0,22	-0,75	0,43
60	0,27	-0,22	0,83
70	0,75	0,31	1,22
80	1,24	0,85	1,62
90	1,72	1,38	2,02
100	2,21	1,91	2,42

(vt. Tabel 1, lk 38; Tabel 2, lk 42; Tabel 3, lk 46).

Tabeli põhjal on näha kolme erineva aine samade aastate tulemused on teineteisest tuntavalt erinevad. Tulemus oli ette näha tänu sellele, et jaotused toorpunktidenä olid samuti märgatavalt erinevad. Näiteks 50-toorpunktiline tulemus eesti keeles annaks z-skaalal -0,22 punkti, inglise keeles -0,75 ja matemaatikas 0,43. Tulemused, mis olid enne alla/ üle keskmise, on seda ka peale standardiseerimist.

Nüüd võrdlen diagrammide abil standardiseerimismeetodi mõju tulemuste jaotusele.



Joonis 10: kolme aine standardiseeritud ja standardiseerimata tulemuste võrdlus

Tabeli põhjal on näha, et kolme erineva aine jaotuse keskmised on peale standardiseerimist ühtlustunud ning hajuvused on samuti vähenenud. Säilib grupisene asümmeetria ning erinevuste proportsioonid. Standardiseerimine muudab võrreldavaks omavahel kolme erineva aine tulemused.

Järgnevalt esitan tabelkujul standardiseeritud matemaatika 2000. ja 2004. aasta riigieksamite tulemused: staniinide-skaalale (C), T-skaalale (T) ja kolledži-skaalale (k) (vt. valem 2, lk 18; valem 3, 19; valem 4, lk 19).

Tabel 9
Kolme aine standardiseeritud tulemused staniinide-, T- ja kolledži-skaalal

Toorpunktid	eesti keel			inglise keel			matemaatika		
	C	T	k	C	T	k	C	T	k
0	0	24	236	-2	16	159	2	34	344
10	1	28	284	-1	21	212	3	38	384
20	2	33	333	0	27	265	3	42	424
30	3	38	381	1	32	318	4	46	463
40	4	43	430	2	37	372	5	50	503
50	5	48	478	3	42	425	6	54	543
60	6	53	527	5	48	478	7	58	583
70	7	58	575	6	53	531	7	62	622
80	7	62	624	7	58	585	8	66	662
90	8	67	672	8	64	638	9	70	702
100	9	72	721	9	69	691	10	74	742

(vt. Tabel 8, lk 55).

Tabeli põhjal on näha, et z-skaalal standardiseeritud tulemuste üleviimine staniinide-skaalale ei päästa inglise keele tulemuste puhul negatiivsetest tulemustest. Toorpunktide kujul on juba kuni 10-punktised standardiseeritud tulemused negatiivsed ning seetõttu ei ole antud skaala kasutajasõbralikkust arvesse võttes samuti kõige parem. Suurem negatiivsete tulemuste osakaal võrreldes eesti keele kirjandi ja matemaatika tulemustega tähistab seda, et riigieksamite keskmised toorpunktide kujul olid arvestavalt kõrgemad.

T-skaala tulemused on kõik positiivsed ning skaalavahemik on sarnane toorpunktidele – kõrgeimad tulemused standardiseeritud kujul ulatuvad 69-74 punktini.

Tulemused kolledži-skaalal on kõik positiivsed ning ohtu, et neid võiks segamini ajada toorpunktliste tulemustega ei ole. Antud skaala puhul on vabalt võimalik avaldada riigieksamite tulemused korruga nii toor- kui standardpunktide kujul.

7.2.2. STANDARDISEERIMINE FIKSEERITUD KESKMISE SKAALALE

Esmalt vaatlen ühe aine piires erinevate aastate tulemuste standardiseerimist. Võrdlen näitena (tulemuste jaotuste poolest teineteisest erinevaid) matemaatika riigieksami 2000. ja 2004. aasta tulemusi.

Meetodi sisuks on viia grupi toorpunktide keskmine kokkulepitud punkti ja „venitada“ ühe poole tulemusi ning „kokku suruda“ teise poole tulemusi säilitades esialgse skaala otspunktid. (vt. valem 5, lk 22; vt. valem 6, lk 22 ;vt. valem 7 ,lk 23; vt. valem 8, lk 25).

Võtan uueks fikseeritud keskmiseks kahe aasta keskmise tulemuse, milleks on 47,7 ning skaala otspunktideks jäävad 0-100. Kõigepealt esitan saadud standardiseeritud tulemused tabelkujul 10-punktiliste intervallidena (vt. Tabel 3, lk 46).

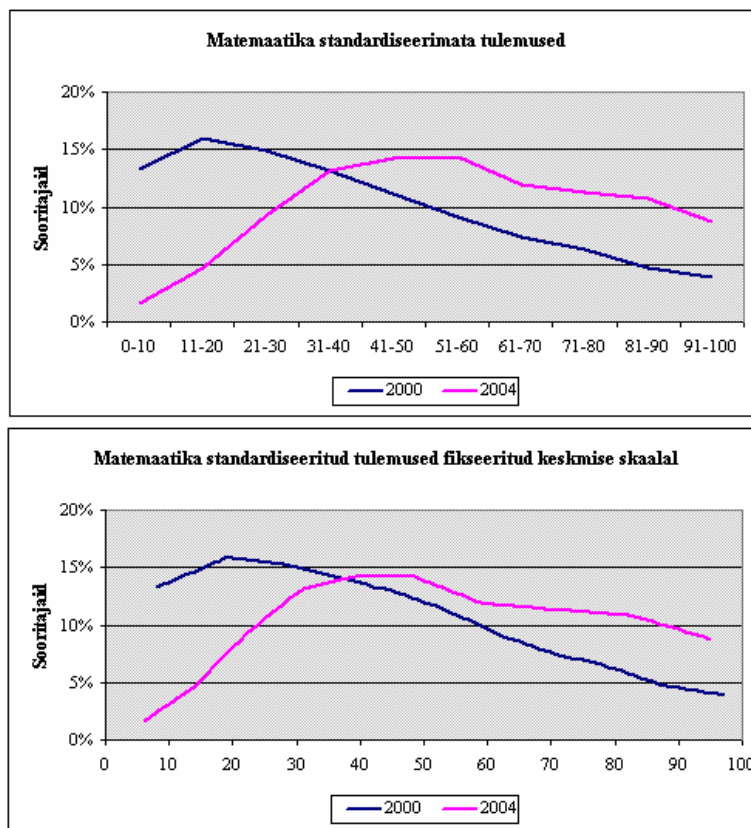
Tabel 10
Matemaatika standardiseeritud tulemused fikseeritud keskmise skaalal

Toorpunktid	Fikseeritud keskmise skaala	
	2000	2004
0	0,0	0,0
10	12,2	8,5
20	24,3	17,0
30	36,5	25,5
40	48,4	33,9
50	57,0	42,4
60	65,6	52,2
70	74,2	64,2
80	82,8	76,1
90	91,4	88,1
100	100,0	100,0

(vt. Tabel 3, lk 46).

Tabeli põhjal on näha, et vaadeldud aastate tulemustes esinevad erinevused. Näiteks toorpunktide kujul saadud 30 punkti annab 2000. aastal standardiseeritud kujul tulemuseks 36,5 punkti ja 2004. aastal 25,5. punkti. Skaala otspunktid – miinimum ja maksimumväärtused on 0 ja 100 ka standardiseeritud kujul. Vastavalt sellele, kui uus fikseeritud keskvärtus on aritmeetilisest keskmisest suurem, siis venitatakse alla fikseeritud keskmise tulemusi ning surutakse kokku üle fikseeritud keskmise tulemused. Kui uus fikseeritud keskmine on aritmeetilisest keskmisest väiksem, siis vastupidi.

Järgnevalt võrdlen diagrammide abil standardiseerimismeetodi mõju tulemuste jaotusele.



Joonis 11: matemaatika standardiseeritud ja standardiseerimata tulemuste võrdlus

Joonise põhjal on näha, et matemaatika 2000. ja 2004. aasta tulemuste jaotuste keskmised on peale standardiseerimist lähedasemad, kuid siiski mitte võrdsed (erinevus on säilinud esialgses suunas). Säilivad grupisisesed (aasta) proportsioonid – kellel oli toorpunktidenä kõrgeim tulemus, on seda ka peale standardiseerimist. Antud meetod võrdsustab aastate lõikes riigieksamite erinevused. Tulemuste jaotuse kuju muutub vastavalt fikseeritud keskmisele. Kui uueks keskmiseks oleks valitud aritmeetilisest keskmisest tunduvalt erinev väärtus, siis oleks jaotuse kujul ka vastavalt sellele muutunud. Meetod on kasutajasõbralik ning võimaldab sätestada ka 20-punktilist soorituspiiri. Selleks, et välistada olukorda, kus sooritaja sai toorpunktidenä tulemuse üle 20-punkti, kuid standardiseerimisel satuks tulemus alla 20-punkti, ei standardiseerita alumise vahemikuna alla 20-punktilisi tulemusi – alumiseks vahemikuks oleksid seega tulemused alates 20 punktist kuni keskmiseni.

Järgmisena vaatlen kolme erineva aine ühe aasta tulemusi ning selleks võtan uueks fikseeritud keskmiseks kolme aine keskmise tulemuse, milleks on 52,6 ning skaala otspunktideks jäävad

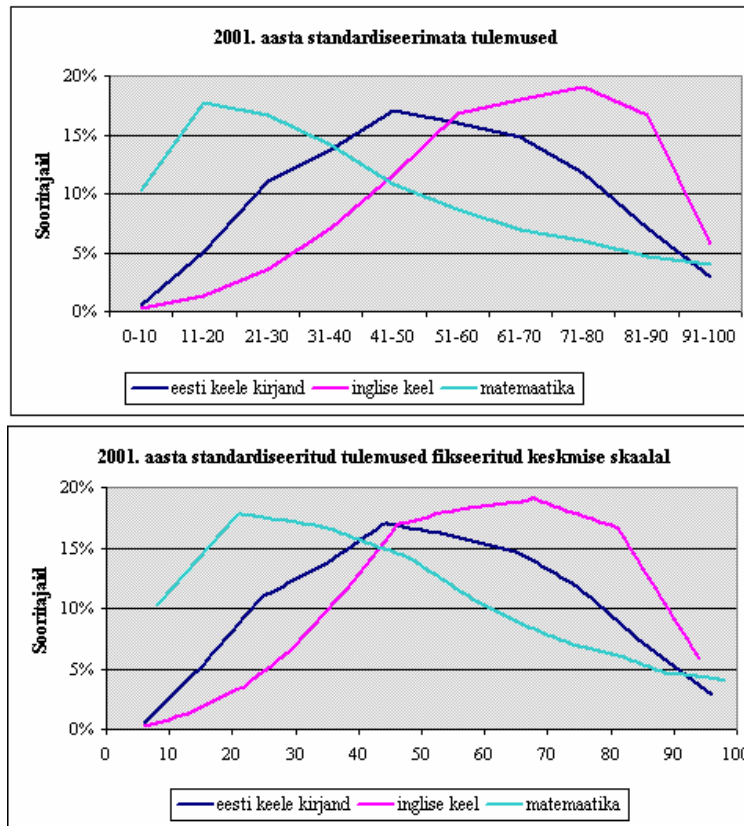
0-100. Kõigepealt esitan saadud standardiseeritud tulemused tabelkujul 10-punktiliste intervallidena.

Tabel 10
Matemaatika standardiseeritud tulemused fikseeritud keskmise skaalal

<i>Toorpunktid</i>	<i>fikseeritud keskmise skaala</i>		
	<i>eesti keele kirjand</i>	<i>inglise keel</i>	<i>matemaatika</i>
0	0,0	0,0	0,0
10	9,7	8,2	13,4
20	19,3	16,4	26,8
30	29,0	24,6	40,2
40	38,6	32,8	53,2
50	48,3	41,0	61,0
60	58,3	49,2	68,8
70	68,8	60,4	76,6
80	79,2	73,6	84,4
90	89,6	86,8	92,2
100	100,0	100,0	100,0

Tabeli põhjal on näha, et on standardiseeritud kujul teineteisest ainete lõikes erinevad. Näiteks toorpunktide kujul saadud 50 punkti annab standardiseeritud kujul eesti keele tulemuseks 48,3 punkti, inglise keeles 41,0 punkti ja matemaatika 61,0 punkti. Fikseeritud skaala otspunktid (0 ja 100) jäävad aga alati paigale.

Järgnevalt võrdlen diagrammide abil standardiseerimismeetodi mõju tulemuste jaotusele.



Joonis 12: kolme aine standardiseeritud ja standardiseerimata tulemuste võrdlus

Joonise põhjal on näha, et peale standardiseerimist on tulemuste jaotuses toimunud muudatused – keskväärtused on teineteisele lähedasemad ning hajuvused on vähenenud. Kuna uus fikseeritud keskmine on määratud vastavalt kolme aine keskväärtustele, siis selline standardiseerimisalgoritm võrdsustab erinevate ainete tulemused omavahel. Peale standardiseerimist ühtlustuvad erinevuste proportsioonid, kuid jäävad siiski eristatavaks – kellel oli toorpunktides suurim tulemus on seda ka peale standardiseerimist.

7.2.3. STANDARDISEERIMINE PROTSENTJÄRJESTUSSKAALALE

Võrdlen omavahel 2005. aasta eesti keele kirjandi, inglise keele ja matemaatika tulemuste standardiseerimise mõju tulemuste jaotusele protsentjärjestusskaalal. Selleks esitan kõigepealt tulemuste jaotuse protsentjärjestusskaalal tabelkujul.

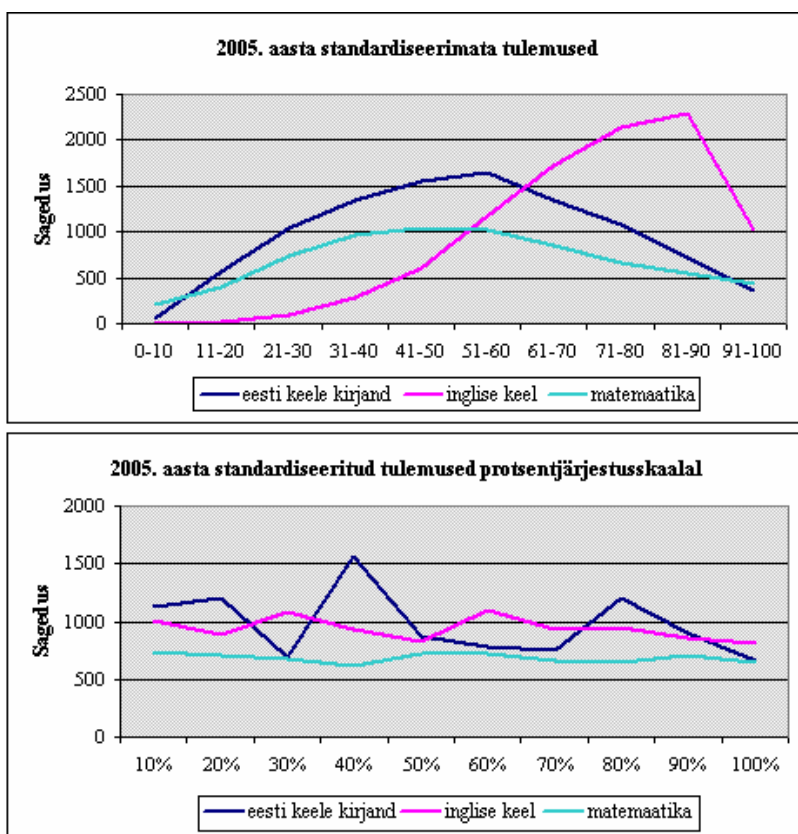
Tabel 11
Kolme aine standardiseeritud tulemused fikseeritud keskmise skaalal

Protsent	eesti keel	inglise keel	matemaatika
10%	1125	1009	731
20%	1202	881	710
30%	693	1087	677
40%	1560	928	623
50%	865	824	728
60%	775	1095	725
70%	745	930	664
80%	1197	953	653
90%	900	854	713
100%	656	814	642

(REKK. Statistika, 2006).

Tabeli põhjal on näha, et eesti keele kirjandi ja inglise keele tulemuste osas esineb teatavaid kõikumisi, kuid matemaatika tulemused on jagunenud suhteliselt ühtlaselt.

Järgnevalt võrdlen diagrammide abil standardiseerimismeetodi mõju tulemuste jaotusele.



Joonis 13: kolme aine standardiseeritud ja standardiseerimata tulemuste võrdlus

Joonise põhjal on näha, et toortulemuste standardiseerimine protsentjärjestusskaalale on jaotuste kaju tunduvalt muutnud. Võrdsustanud ainete lõikes keskmised tulemused, vähenenud on hajuvuste erinevused. Korduvate väärtuste juures (näiteks eesti keele kirjandi 40 protsentil) on toimunud kuhjumised, kuid üldjoontes on jaotus ühtlane. Kellel oli toorpunktides kõrgeim tulemus, on seda ka peale standardiseerimist. Standardiseerimine protsentjärjestusskaalale muudab siinkohal võrreldavaks erinevate ainete tulemused, kuid ei võta arvesse, et erinevused võisid tekkida ka testi sooritajatest.

Protsentjärjestusskaala osas vaatlesin vaid ühe aasta (kolme erineva aine) tulemusi, kuna üldjoontes jääb tulemuste jaotus peale standardiseerimist sarnaseks, erinevuste proportsioonid ühtlustuvad, hajuvus erinevatel gruppidel ligikaudu sama.

7.3. JÄRELDUSED

Standardiseerisin valikuliselt näiteandmeid erinevate standardiseerimismeetodite abil ning tulemused tulid mõnevõrra erinevad.

Vaadates meetodi mõju tulemuste jaotusele selgus, et kui tulemused olid toorpunktidena jagunenud normaaljaotusele sarnaselt (eesti keele kirjandi puhul) ning erinevate aastate vahel ei esinenud suuri kõikumisi, siis olid ühtviisi sobivad kõik standardiseerimismeetodid. Kuna standardiseerimise käigus olid tulemuste muutused väga väikesed, siis on tegelikult võimalik ka toorskoorilisi tulemusi omavahel aastate lõikes objektiivselt võrrelda (*vt. joonis 8, lk 51*).

Kui tulemused olid asümmeetrilised, siis andsid erinevad meetodid ka teistsuguseid tulemusi. Standardiseerimine z-skaalale võrdsustas omavahel gruppide hajuvused, kuid säilitas asümmeetria – kelle tulemused erinesid teistest toorpunktide kujul väga palju, jäi see nii ka peale standardiseerimist. Seega, kui gruppide tulemused on tugevalt asümmeetrilised, ei toimu standardiseerimisel tulemuste ühtlustumist olulisel määral. Kellel oli toorpunktides skaala kõrgeim tulemus saab ka standardiseeritud skaalal kõrgeima tulemuse. Meetod võtab arvesse, et erinevused tulenesid testist, kuid ei arvesta, et erinevused võisid tekkida ka sooritajatest (*vt. joonis. 9, lk 54; vt. joonis. 10, lk 56*).

Z-skaala keskväärtuseks on null ning tulemused alla keskmise väärtused on negatiivsed, seetõttu skaala kasutajasõbralikkuse saavutamiseks tuleks tulemused üle viia mõnele teisele

skaalale. Antud töös tõin näitena välja kolm sellist võimalust: staniinide-skaala, T-skaala ja kolledži-skaala. Staniinide-skaala puhul selgus, et tulemustes võib esineda ka negatiivseid väärtusi ning kuna tulemus väljastatakse ühekohalise arvuna, siis ei ole antud skaala eristusvõime väga suur.

T-skaala tulemused on kahekohalised, positiivsed arvud ning sarnanevad kirjutuspildilt toorpunktidele. Seetõttu, näiteks tulevikus riigieksamite puhul, peab väljastamisel arvesse võtma, et korraga ei väljastataks standardiseeritud ja toorpunktilisi tulemusi, kuna nende sarnasus võib segadust tekitada. Kolledži-skaala tulemused olid samuti positiivsed kolmekohalised arvud ning nende väljastamisel ei olnud ohtu, et standardiseeritud tulemused võiksid toorpunktidega segamini minna (*vt. Tabel 6, lk 52; vt. Tabel 7, lk 55; vt. Tabel 9, lk 57*).

Fikseeritud keskmise skaala puhul olenesid tulemused sellest, milline uus fikseeritud keskmine valiti. Kui uus fikseeritud keskmine erines oluliselt toorpunktide keskväärtusest, siis muutus märgatavalt ka tulemuste jaotus. Standardiseerimine vähendas pisut asümmeetriat ning ühtlustas jaotuste proportsioone. Üldine hajuvus peale standardiseerimist muutub pisut, kuid suuremad erinevused gruppide vahel jäävad siiski alles. Teineteisest sümmeetriliselt väga erinevate tulemuste korral annab antud meetod seega parema tulemuse kui z-skaalale viimisel. Kellel oli toorpunktides skaala kõrgeim tulemus saab ka standardiseeritud skaalal kõrgeima tulemuse. Skaala otspunktid on kokkuleppelised, millest lähtuvalt võib vahemikena kasutada ka teisi võimalusi peale töös võrreldud 0-100 vahemiku (*vt. joonis. 11, lk 59; vt. joonis. 12, lk 61*).

Protsentjärjestusskaala kasutamisel võrdsustub gruppidevaheline asümmeetria. Skaala otspunktid on fikseeritud (0-10) keskväärtused jäävad ligikaudu 50 kanti. Seetõttu on meetodi kasutamine väga sobilik asümmeetriliste jaotuste puhul. Standardiseerimine kaotab hajuvuste erinevused, kuid on palju korduvaid väärtusi, võib esineda vastavas piirkonnas ka kuhjumisi. Kellel oli toorpunktides skaala kõrgeim tulemus saab ka standardiseeritud skaalal kõrgeima tulemuse. Meetod võtab arvesse, et erinevused tulenesid testist, kuid ei arvesta, et erinevused võisid tekkida ka sooritajatest (*vt. joonis. 13, lk 62*).

Standardiseerimise eesmärkideks on võrdsustada ja muuta teineteisega võrreldavaks erinevate ainete ja aastate omavahelised tulemused. Seetõttu tuleb meetodi valikul lähtuda sellest, et standardiseerimise tulemusena tasanduksid aastate lõikes tulemuste erinevusest tekkinud kõikumised ning ühtlustuksid tulemuste jaotused. Teisest küljest peab meetod olema ka kasutajasõbralik ja võtma arvesse nii riigieksami sisust kui ka sooritajatest tulenevaid erinevusi. Täiesti vastuvõetav oleks ka idee Ameerika Ühendriikide Koolijõudluse Testi tulemuste standardiseerimisest, kus avaldati tulemus paralleelselt nii kolledži-skaalal kui ka protsentjärjestusskaalal (*ETS, 2006*).

KOKKUVÕTE

Oma magistritöös kajastasin ja analüüsisin testitulemuste standardiseerimise võimalusi, lähtuvalt Eesti Vabariigi riigieksamitest. Olles selgitanud nii riigieksamate kui testiteooria tausta, analüüsisin standardiseerimiseks vajaliku info kogumiseks kolme valitud riigieksami tulemuste statistilisi näitajaid ning seejärel vastavate näitajate abil erinevaid standardiseerimismeetodeid.

Töö seisukohast oli väga hea, et analüüsiks sai valitud kolm teineteisest oma tulemuste poolest erinevat ainet, sellest johtuvalt oli võimalik jälgida standardiseerimismeetodite mõju tulemuste jaotusele.

Tulemuste standardiseerimiseks valisin välja kolm meetodit, millest kolledži-skaala ja protsentjärjestusskaala on laialt kasutusel ka Ameerika Ühendriikides Koolijõudluse Testide hindamisel ning kolmanda meetodina analüüsisin ka Katrin Niglase poolt väljatoodud fikseeritud keskmise meetodit. Fikseeritud keskmise standardiseerimismeetodi analüüsimise kasuks otsustasin, kuna meetod võttis lisaandmete kogumiseta arvesse nii testist kui testitavatest tulenevaid erinevusi (vähendas testi erinevuste mõju ja säilitas testitavate tasemete erinevused).

Analüüsid ja võrreldes kolme erineva standardiseerimismeetodi mõju tulemuste jaotusele jõudsin töö käigus järgmistele järeldustele.

- Kui tulemuste jaotus toorpunktidenä oli normaaljaotusele sarnane, siis standardiseerimise käigus olulisi muutusi ei tekkinud ning erinevate aastate testitulemused on ka toorpunktide kujul tegelikult omavahel võrreldavad. Standardiseerimise kohapealt sobivad kõik väljatoodud meetodid, kuna tulemuste jaotus ja aastevaheline erinevus jäävad üldjoontes.
- Kui tulemuste jaotus oli tugevalt asümmeetriline, siis z-skaala kasutamise puhul võrdsustasid gruppidevahelised hajuvused, kuid asümmeetria säilis. Seega tugevalt asümmeetriliste jaotuste puhul ei anna vastav skaala kasutamine kõige paremaid tulemusi. Kuna z-skaala tulemustes esineb negatiivseid arve, siis kasutajasõbralikkuse saavutamiseks tuleks tulemused teistele skaaladele teisendada.

- Asümmeetrilise jaotuse korral muutis (keskmisest erineva) fikseeritud keskmise standardiseerimismeetod tulemuste jaotuse proportsioone ühtlasemaks ning vähendas asümmeetriat, seega sobib kasutamiseks ka asümmeetriliste jaotuste puhul. Skaala otspunktide on kokkuleppelised, kasutajasõbralik.
- Protsentjärjestusskaala kasutamine ühtlustas jaotuste asümmeetria ning ühtlustas erinevuste proportsioone. Tulemused olid jaotunud üldiselt ühtlaselt, kuid korduvate väärtuste juures esines kuhjumisi. Skaala otspunktid on 0-100, tulemused on lihtsalt tõlgendatavad.

Töö tegemisel sätestasin järgmised eesmärgid:

- selgitada riigieksami kui testi olemust;
- leida erinevaid võimalusi riigieksamite tulemuste võrreldavuse parandamiseks standardiseerimismeetodite abil;
- analüüsida väljatoodud standardiseerimismeetodite ohte ja võimalusi ning nende mõju testi tulemustele ja vastavalt sellele teha järeldusi;
- koostada ühtne ülevaatlik materjal, mis aitaks valdkonda mõista ka neil, kellel puuduvad sellekohased eelnevad põhjalikud teoreetilise teadmised.

Arvesse võttes püstitatud eesmäärke võib eelpool väljatoodule tuginedes väita, et töö eesmärgid said täidetud.

KASUTATUD KIRJANDUS

(Adamson) Adamson, Andres (1998). *Kümme küsimust Jaan Mikule*. Õpetajate Leht (1998, juuni 19), URL: (2005, juuli 1).

(CCDE) The Centre for Curriculum Development and Examinations (Latvia), URL: <http://isec.gov.lv/en/exam.shtml> (2006, jaanuar 05).

(ERT) Elektrooniline Riigiteataja (RTL, 01.02.2002, 20, 249), *Õpitulemuste välishindamise põhimõtted, riigieksamitööde, põhikooli lõpueksamitööde ja üleriigiliste tasemetööde koostamise, hindamise ja tulemuste analüüsi alused*, URL: <https://www.riigiteataja.ee/ert/act.jsp?id=163425&replstring=33> (2005, juuli 01).

(ETS) Educational Testing Service (ETS), URL: <http://www.ets.org> (2006, jaanuar 10).

(EÜL) Eesti Üliõpilaskondade Liit, Tallinna toimetis (2004). *TÜ lävendipõhine vastuvõtt võib ülikoolile palju maksta*, URL: <http://www.eyl.ee/index.php?lang=est&mid=2,29&nid=,282> (2005, juuli 09).

(Hennoste) Hennoste, Märt, Riiklik Eksami- ja Kvalifikatsioonikeskus (2002), *Eesti keel. Riigieksam 2002/2003*, Tallinn: REKK trükikoda.

(Hennoste) Hennoste, Märt, Riiklik Eksami- ja Kvalifikatsioonikeskus (2003), *Eesti keel. Riigieksam 2003/2004*, Tallinn: REKK trükikoda.

(Institute) The Latvian Institute. *Education and Science in Latvia*, URL: http://www.li.lv/old/education_science.htm (2006, veebruar 15)

(Jõemaa) Jõemaa, Lehte (1998). *Halb unenägu kestab. Riigieksamid*. Õpetajate Leht, URL: <http://greta.cs.ioc.ee/~opleht/Arhiiv/98Jun19/paevateemadel.html#3> (2005, juuni 1).

(Kärner) Kärner, Olev (1976). *Standardiseeritud kontrolltöid matemaatikast. V klass*. Tallinn Eesti NSV Haridusministeerium.

(Kübar) Kübar, Eva (2005). *Tartu Ülikool tahab riigieksameid normeerida*. Postimees Online, URL: <http://www.postimees.ee/050905/esileht/siseuudised/175935.php> (2005, september 03).

(Laev). Laev, Sigrid. *Lävendipõhine vastuvõtt ummistas ülikoole*. Eesti Päevaleht Online (2005), URL: http://www.epl.ee/artikkel_299037.html (2005, august 24).

- (Lehtsaar) Lehtsaar, Tõnu (2004). *Üks avaldus ühele erialale*. Postimees Online, URL: <http://www.koolielu.ee/pages.php/020406,10729> (2004, juuli 29).
- (Lindgren & Suter) Lindgren, Henry Clay ja Suter, W. Newton (1994). *Pedagoogiline psühholoogia koolipraktikas*, Tartu: Tartu Ülikool.
- (Lukason & Karlova) Lukason, Anneli ja Karlova, Taru (2000). *Riigieksami võlu ja valu õpetaja vaatevinklist*. Postimees Online, URL: <http://vana.www.postimees.ee:8080/leht/00/07/20/arvamus.htm#teine> (2000, juuli 20).
- (Metsämuuronen) Metsämuuronen, Jari (2000), *Mittarin rakentaminen ja testiteorian perusteet*, Võru : Jaabes OÜ.
- (Mikk) Mikk, Jaan (2002). *Ainetestid*, URL: <http://raud.ut.ee/~jaanm/ainetestid.pdf> (2006, veebruar 03).
- (Niglas) Niglas, Katrin (1997). *Statistika loengumaterjale*, URL: http://www.cs.tlu.ee/~katrin/sisu/materjalid/opmat/stat_loeng.pdf (2006, märts 10).
- (Niglas) Niglas, Katrin (2005). *Eksamitulemuste standardiseerimine*. REKK'ile esitatud materjal.
- (Ots) Ots, Diana-Dolores (2000). *Riigieksamid – mõte või mõttetatus?*, Eesti Päevaleht, URL: http://www.epl.ee/artikkel_68461.html?PHPSESSID=976247da822cacc0b43b3195a88f1a69 (2005, juuli 1).
- (Päärt) Päärt, Villu (2006). *Riigieksamite hindamine võib muutuda*. Postimees Online, URL: <http://www.postimees.ee/090106/esileht/siseuudised/188197.php> (2006, jaanuar 09).
- (Rebane & Järve) Rebane, Epp ja Järve, Linda (1999). *Riigieksam on osa õpitulemuste välishindamise süsteemist*. Õpetajate Leht (1999), URL: <http://www.cs.ioc.ee/~opleht/Arhiiv/99Nov26/haridusdialoog.html#3> (2005, juuli 1).
- (REKK. Inglise keel) Riiklik Eksami- ja Kvalifikatsioonikeskus (2002), *Inglise keel. Riigieksam 2002/2003*, Tallinn: REKK trükikoda.
- (REKK. Inglise keel) Riiklik Eksami- ja Kvalifikatsioonikeskus (2003), *Inglise keel. Riigieksam 2003/2004*, Tallinn: REKK trükikoda.

(REKK. *Statistika*) Riikliku Eksami- ja Kvalifikatsioonikeskus (REKK). *Statistika*, URL: <http://www.ekk.edu.ee/statistika/index.html> (2005, juuli 01).

(REKK. *Üldteave*) Riikliku Eksami- ja Kvalifikatsioonikeskus (REKK). *Üldteave*, URL: <http://www.ekk.edu.ee/riigieksamid/index.html> (2005, juuli 01).

(*Riigieksamid, 1997*) Adamson, Andres ja Kond, Meelis, Eesti Vabariigi Haridusministeerium, Riiklik Eksami- ja Kvalifikatsioonikeskus (1997), *Riigieksamid 1997*, Tallinn: EKK trükikoda.

(*Riigieksamid, 1998*) Adamson, Andres, Kond, Meelis, Eesti Vabariigi Haridusministeerium, Riiklik Eksami- ja Kvalifikatsioonikeskus (1999), *Riigieksamid 1998*, Tallinn: EKK trükikoda.

(*Sootak*) Sootak, Varje (2005). *Lävend toob paremaid õpilasi*. Tartu Ülikooli ajaleht nr 28, URL: <http://www.ajaleht.ut.ee/120155> (2005, september 16).

(*SweSAT*) Department of Educational Measurement, *The Swedish Scholastic Assessment Test – SweSAT*, URL: http://www.umu.se/edmeas/hprov/index_eng.html (2006, veebruar, 20).

(*Tamm & Tamm*) Tamm, Lembi ja Tamm, Tarmo (2000). *Riigieksamite võlu... ja valu*. Postimees Online, URL: <http://vana.www.postimees.ee:8080/leht/00/07/20/arvamus.htm#teine> (2005, juuli 1).

(*TLÜ*) Tallinna Ülikool. *Vastuvõtu info*, URL: <http://www.tlu.ee/?LangID=1&CatID=656> (2006, jaanuar 10).

(*Toim*) Toim, Kalju (1981), *Psühhomeetria teoreetilised alused*, Tartu: Tartu Riiklik Ülikool.

(*Uudelepp & Lõhmus*) Uudelepp, Helgi ja Lõhmus, Ahto, Riiklik Eksami- ja Kvalifikatsioonikeskus (2000), *Matemaatika. Riigieksam 2000/ 2001*, Tallinn: REKK trükikoda.

(*Uudelepp & Lõhmus*) Uudelepp, Helgi, Lõhmus, Ahto, Riiklik Eksami- ja Kvalifikatsioonikeskus (2003), *Matemaatika. Riigieksam 2003/ 2004*, Tallinn: REKK trükikoda.

SUMMARY

Capture of this Master Thesis is “Standardization of test results. The case on Estonian National Exams”.

It consists of seven chapters, summary (written in Estonian and English), listing of used literature, different terms and shortenings. In this research I presented and analyzed possibilities of standardization test results by example of Estonian National exams. Having explained background of test theories and national exams, I analyzed statistical results on three chosen national exams collecting the information needed for standardization.

For standardization of test results I chose three methods, from which college scale and percentile scale are widely used in evaluating Scholastic Assessment Tests in United States of America and as a third method I analyzed fixed average method presented by Katrin Niglas. The third method was chosen because its possibilities to take into consideration both characteristics, as well as differences becoming from tests and test takers.

According to hypothesizing the problem I set goals of thesis:

- Explain the nature of national exam as test;
- Find different possibilities to improve comparability of results of national exams by standardization methods;
- Analyze dangers and possibilities of these standardization methods according on results make assumptions about their effectiveness;
- Set up an uniform and summarised material, that would help to understand this particular field even for those who have never heard about standardization of test results before.

As a result of this master thesis, possibilities of standardization of national exams in Republic of Estonia and strong and weak points of standardization methods used for that are becoming clear.

MÕISTETE JA LÜHENDITE SELGITUSI

Ainetest – test, millega mõõdetakse valitud (õppe)aine teadmiste ja oskuste taset.

Aritmeetiline keskmine – ehk keskväärtus.

Asümmeetria – sümmeetria antonüüm; antud töös kasutusel tulemuse jaotust iseloomustava karakteristikuna, kus jaotus on kas ühele või teisele poole väljavenitatud, mittesümmeetriline.

Bimodaalne asümmeetria – jaotusel on kaks tippu, mis asetsevad üks ühel ja teine teisel pool ulatuse keskpunkti.

Dispersioon – ruuthälvete aritmeetiline keskmine. Standardhälbe ruut.

Eksam – üks teadmistekontrolli vormidest, mille põhjal hinnatakse õpitu teadmiste, oskuste ja vilumuste omandamist.

Eristustest – ainetest, mille abil püütakse õpilasi eristada nende ainealaste teadmiste või oskuste järgi.

Joondiagramm – diagramm (ehk arvjoonis), mille abil näidatakse andmete muutumist ajas.

Korrelatsioon – kahe muutuva suuruse vahelise seose tugevuse näitaja (kus korrelatsioonikordaja 1 tähendab täielikku seost, 0 seose puudumist).

Lävend – on üliõpilaste vastuvõtuks (kõrgkooli) kehtestatud minimaalne nõutav punktisumma sajast võimalikust. Kõrgkooli kandideerijatel, kellel sisseastumisel arvestatavate riigieksamite põhjal arvatud punktisumma ületab kooli poolt seatud lävendi või on sellega võrdne, võetakse kõrgkooli vastu.

Mediaan – karakteristik, mis jagab tulemused täpselt kahte ossa – täpselt pooled väärtused on mediaanist väiksemad ja täpselt pooled suuremad.

Negatiivne asümmeetria – tulemuste jaotus on välja venitatud nii, et enamik tulemusi on koondunud skaalal paremale poole.

Normaaljaotus – jaotus, mille puhul tulemused on koondunud jaotuse keskosa ümber ning hajuvus ja asümmeetria on väikesed.

Programmitest – test, mis vastab oma sisu poolest ainekavale. Kontrollib, kuidas ainekavas seatud eesmärgid on täidetud.

REKK – Riiklik Eksami- ja Kvalifikatsioonikeskus.

Reliaablus – ehk usaldatavus näitab seda, kuivõrd test mõõdab seda, mida ta mõõdab, ilma et küsitaks mis see „see“ on.

Riigieelarveline koht (RE) – üliõpilase koolitamise kulud tasub riik.

Riigieelarve väline koht (REV) – üliõpilase koolitamise kulud tuleb õpilasel ise tasuda.

SAT – inglise keeles: *Scholastic Assessment Test* – Koolijõudluse test Ameerika Ühendriikides.

Standardhälve – hajuvuse näitaja, näitab kui palju üksikud tulemused aritmeetilisest keskmisest (keskmiselt) erinevad. Ruutjuut dispersioonist.

Standardiseerimine – toorpunktide viimine standardiseeritud skaalale.

Standardpunktid – testitulemused standardskaalal (tulemus võrrelduna teistega).

Standardskoor – standardpunktid.

Standardviga – näitab, kui suur on tõenäosus, et meie poolt moodustatud valimi keskvärtus on üldkogumi keskvärtusest märksa suurem või väiksem.

Statistika – kogutud andmete töötlemine ning nende põhjal järelduste tegemine.

SweSAT – Koolijõudluse test Rootsi Kuningriigis.

Test – lühivastuseliste ülesannete kogum, millede lahendatavuse järgi saab otsustada mõõdetava omaduse taseme üle.

Toorpunktid – õigete vastuste eest testis saadud punktide summa (kokku).

Ulatus – jaotuse maksimaalse ja minimaalse väärtuse vahe.

Valiidsus – ehk kehtivus näitab seda, kuivõrd test mõõdab seda omadust, mida me mõõta tahame.

Valim – on teatava eeskirja järgi moodustatud hulk üldkogumisse kuuluvatest objektidest ehk mõõtmiseks valitud üldkogumi osa.

Vt. – lühend sõnast vaata; kasutatakse viitamisel.

Üldkogum – ehk populatsioon on objektide hulk, mille/ kelle kohta soovitakse uurimuse tulemusena sisulisi järeldusi teha.

LISA 1. Z-SKAALA STANDARDISEERITUD TULEMUSED

Z-skaala							
Toorpunktid	Eesti keele kirjand			Matemaatika			Inglise keel
	2001	2003	2004	2000	2001	2004	2001
0	-2,64	-2,62	-2,61	-1,52	-1,56	-2,38	-3,41
1	-2,59	-2,57	-2,57	-1,48	-1,52	-2,34	-3,36
2	-2,55	-2,52	-2,52	-1,44	-1,48	-2,30	-3,31
3	-2,50	-2,48	-2,47	-1,40	-1,44	-2,26	-3,25
4	-2,45	-2,43	-2,42	-1,36	-1,40	-2,21	-3,20
5	-2,40	-2,38	-2,37	-1,33	-1,36	-2,17	-3,15
6	-2,35	-2,33	-2,33	-1,29	-1,32	-2,13	-3,09
7	-2,30	-2,29	-2,28	-1,25	-1,28	-2,09	-3,04
8	-2,25	-2,24	-2,23	-1,21	-1,24	-2,04	-2,99
9	-2,21	-2,19	-2,18	-1,17	-1,20	-2,00	-2,93
10	-2,16	-2,14	-2,14	-1,13	-1,16	-1,96	-2,88
11	-2,11	-2,10	-2,09	-1,09	-1,12	-1,92	-2,83
12	-2,06	-2,05	-2,04	-1,05	-1,08	-1,88	-2,77
13	-2,01	-2,00	-1,99	-1,02	-1,04	-1,83	-2,72
14	-1,96	-1,95	-1,94	-0,98	-1,00	-1,79	-2,67
15	-1,91	-1,91	-1,90	-0,94	-0,96	-1,75	-2,61
16	-1,87	-1,86	-1,85	-0,90	-0,92	-1,71	-2,56
17	-1,82	-1,81	-1,80	-0,86	-0,88	-1,66	-2,51
18	-1,77	-1,76	-1,75	-0,82	-0,84	-1,62	-2,45
19	-1,72	-1,72	-1,70	-0,78	-0,80	-1,58	-2,40
20	-1,67	-1,67	-1,66	-0,74	-0,76	-1,54	-2,35
21	-1,62	-1,62	-1,61	-0,71	-0,72	-1,49	-2,29
22	-1,58	-1,57	-1,56	-0,67	-0,68	-1,45	-2,24
23	-1,53	-1,53	-1,51	-0,63	-0,64	-1,41	-2,19
24	-1,48	-1,48	-1,46	-0,59	-0,61	-1,37	-2,13
25	-1,43	-1,43	-1,42	-0,55	-0,57	-1,32	-2,08
26	-1,38	-1,38	-1,37	-0,51	-0,53	-1,28	-2,03
27	-1,33	-1,34	-1,32	-0,47	-0,49	-1,24	-1,97
28	-1,28	-1,29	-1,27	-0,43	-0,45	-1,20	-1,92
29	-1,24	-1,24	-1,22	-0,40	-0,41	-1,15	-1,87
30	-1,19	-1,19	-1,18	-0,36	-0,37	-1,11	-1,82
31	-1,14	-1,15	-1,13	-0,32	-0,33	-1,07	-1,76
32	-1,09	-1,10	-1,08	-0,28	-0,29	-1,03	-1,71
33	-1,04	-1,05	-1,03	-0,24	-0,25	-0,98	-1,66
34	-0,99	-1,01	-0,98	-0,20	-0,21	-0,94	-1,60
35	-0,94	-0,96	-0,94	-0,16	-0,17	-0,90	-1,55
36	-0,90	-0,91	-0,89	-0,12	-0,13	-0,86	-1,50
37	-0,85	-0,86	-0,84	-0,08	-0,09	-0,81	-1,44
38	-0,80	-0,82	-0,79	-0,05	-0,05	-0,77	-1,39
39	-0,75	-0,77	-0,74	-0,01	-0,01	-0,73	-1,34
40	-0,70	-0,72	-0,70	0,03	0,03	-0,69	-1,28
41	-0,65	-0,67	-0,65	0,07	0,07	-0,64	-1,23
42	-0,61	-0,63	-0,60	0,11	0,11	-0,60	-1,18
43	-0,56	-0,58	-0,55	0,15	0,15	-0,56	-1,12
44	-0,51	-0,53	-0,51	0,19	0,19	-0,52	-1,07
45	-0,46	-0,48	-0,46	0,23	0,23	-0,48	-1,02
46	-0,41	-0,44	-0,41	0,26	0,27	-0,43	-0,96
47	-0,36	-0,39	-0,36	0,30	0,31	-0,39	-0,91
48	-0,31	-0,34	-0,31	0,34	0,35	-0,35	-0,86
49	-0,27	-0,29	-0,27	0,38	0,39	-0,31	-0,80
50	-0,22	-0,25	-0,22	0,42	0,43	-0,26	-0,75

51	-0,17	-0,20	-0,17	0,46	0,47	-0,22	-0,70
52	-0,12	-0,15	-0,12	0,50	0,51	-0,18	-0,64
53	-0,07	-0,10	-0,07	0,54	0,55	-0,14	-0,59
54	-0,02	-0,06	-0,03	0,57	0,59	-0,09	-0,54
55	0,03	-0,01	0,02	0,61	0,63	-0,05	-0,48
56	0,07	0,04	0,07	0,65	0,67	-0,01	-0,43
57	0,12	0,09	0,12	0,69	0,71	0,03	-0,38
58	0,17	0,13	0,17	0,73	0,75	0,08	-0,32
59	0,22	0,18	0,21	0,77	0,79	0,12	-0,27
60	0,27	0,23	0,26	0,81	0,83	0,16	-0,22
61	0,32	0,28	0,31	0,85	0,87	0,20	-0,16
62	0,36	0,32	0,36	0,88	0,91	0,25	-0,11
63	0,41	0,37	0,41	0,92	0,95	0,29	-0,06
64	0,46	0,42	0,45	0,96	0,99	0,33	0,00
65	0,51	0,46	0,50	1,00	1,03	0,37	0,05
66	0,56	0,51	0,55	1,04	1,06	0,42	0,10
67	0,61	0,56	0,60	1,08	1,10	0,46	0,15
68	0,66	0,61	0,65	1,12	1,14	0,50	0,21
69	0,70	0,65	0,69	1,16	1,18	0,54	0,26
70	0,75	0,70	0,74	1,19	1,22	0,59	0,31
71	0,80	0,75	0,79	1,23	1,26	0,63	0,37
72	0,85	0,80	0,84	1,27	1,30	0,67	0,42
73	0,90	0,84	0,88	1,31	1,34	0,71	0,47
74	0,95	0,89	0,93	1,35	1,38	0,76	0,53
75	1,00	0,94	0,98	1,39	1,42	0,80	0,58
76	1,04	0,99	1,03	1,43	1,46	0,84	0,63
77	1,09	1,03	1,08	1,47	1,50	0,88	0,69
78	1,14	1,08	1,12	1,50	1,54	0,92	0,74
79	1,19	1,13	1,17	1,54	1,58	0,97	0,79
80	1,24	1,18	1,22	1,58	1,62	1,01	0,85
81	1,29	1,22	1,27	1,62	1,66	1,05	0,90
82	1,33	1,27	1,32	1,66	1,70	1,09	0,95
83	1,38	1,32	1,36	1,70	1,74	1,14	1,01
84	1,43	1,37	1,41	1,74	1,78	1,18	1,06
85	1,48	1,41	1,46	1,78	1,82	1,22	1,11
86	1,53	1,46	1,51	1,82	1,86	1,26	1,17
87	1,58	1,51	1,56	1,85	1,90	1,31	1,22
88	1,63	1,56	1,60	1,89	1,94	1,35	1,27
89	1,67	1,60	1,65	1,93	1,98	1,39	1,33
90	1,72	1,65	1,70	1,97	2,02	1,43	1,38
91	1,77	1,70	1,75	2,01	2,06	1,48	1,43
92	1,82	1,74	1,80	2,05	2,10	1,52	1,49
93	1,87	1,79	1,84	2,09	2,14	1,56	1,54
94	1,92	1,84	1,89	2,13	2,18	1,60	1,59
95	1,97	1,89	1,94	2,16	2,22	1,65	1,65
96	2,01	1,93	1,99	2,20	2,26	1,69	1,70
97	2,06	1,98	2,04	2,24	2,30	1,73	1,75
98	2,11	2,03	2,08	2,28	2,34	1,77	1,81
99	2,16	2,08	2,13	2,32	2,38	1,82	1,86
100	2,21	2,12	2,18	2,36	2,42	1,86	1,91

LISA 2. FIKSEERITUD KESKMISE SKAALA STANDARDISEERITUD TULEMUSED

Fikseeritud keskmise skaala					
Toorpunktid	<i>eesti keele kirjand</i>	<i>inglise keel</i>	<i>matemaatika</i>		
	2001	2001	2000	2001	2004
	($\bar{x}_{fix} = 52,6$)	($\bar{x}_{fix} = 52,6$)	($\bar{x}_{fix} = 47,7$)	($\bar{x}_{fix} = 52,6$)	($\bar{x}_{fix} = 47,7$)
0	0,0	0,0	0,0	0,0	0,0
1	1,0	0,8	1,2	1,3	0,8
2	1,9	1,6	2,4	2,7	1,7
3	2,9	2,5	3,7	4,0	2,5
4	3,9	3,3	4,9	5,4	3,4
5	4,8	4,1	6,1	6,7	4,2
6	5,8	4,9	7,3	8,0	5,1
7	6,8	5,7	8,5	9,4	5,9
8	7,7	6,6	9,7	10,7	6,8
9	8,7	7,4	11,0	12,1	7,6
10	9,7	8,2	12,2	13,4	8,5
11	10,6	9,0	13,4	14,8	9,3
12	11,6	9,8	14,6	16,1	10,2
13	12,6	10,7	15,8	17,4	11,0
14	13,5	11,5	17,0	18,8	11,9
15	14,5	12,3	18,3	20,1	12,7
16	15,4	13,1	19,5	21,5	13,6
17	16,4	14,0	20,7	22,8	14,4
18	17,4	14,8	21,9	24,1	15,3
19	18,3	15,6	23,1	25,5	16,1
20	19,3	16,4	24,3	26,8	17,0
21	20,3	17,2	25,6	28,2	17,8
22	21,2	18,1	26,8	29,5	18,7
23	22,2	18,9	28,0	30,8	19,5
24	23,2	19,7	29,2	32,2	20,4
25	24,1	20,5	30,4	33,5	21,2
26	25,1	21,3	31,6	34,9	22,1
27	26,1	22,2	32,9	36,2	22,9
28	27,0	23,0	34,1	37,5	23,8
29	28,0	23,8	35,3	38,9	24,6
30	29,0	24,6	36,5	40,2	25,5
31	29,9	25,4	37,7	41,6	26,3
32	30,9	26,3	38,9	42,9	27,2
33	31,9	27,1	40,2	44,3	28,0
34	32,8	27,9	41,4	45,6	28,9
35	33,8	28,7	42,6	46,9	29,7
36	34,8	29,5	43,8	48,3	30,6
37	35,7	30,4	45,0	49,6	31,4
38	36,7	31,2	46,2	51,0	32,2
39	37,7	32,0	47,5	52,4	33,1
40	38,6	32,8	48,4	53,2	33,9
41	39,6	33,6	49,3	54,0	34,8
42	40,5	34,5	50,1	54,8	35,6
43	41,5	35,3	51,0	55,5	36,5
44	42,5	36,1	51,8	56,3	37,3
45	43,4	36,9	52,7	57,1	38,2
46	44,4	37,8	53,6	57,9	39,0
47	45,4	38,6	54,4	58,7	39,9
48	46,3	39,4	55,3	59,4	40,7
49	47,3	40,2	56,1	60,2	41,6
50	48,3	41,0	57,0	61,0	42,4
51	49,2	41,9	57,9	61,8	43,3
52	50,2	42,7	58,7	62,6	44,1
53	51,2	43,5	59,6	63,3	45,0
54	52,1	44,3	60,4	64,1	45,8
55	53,1	45,1	61,3	64,9	46,7

56	54,2	46,0	62,2	65,7	47,5
57	55,2	46,8	63,0	66,5	48,7
58	56,3	47,6	63,9	67,2	49,8
59	57,3	48,4	64,7	68,0	51,0
60	58,3	49,2	65,6	68,8	52,2
61	59,4	50,1	66,5	69,6	53,4
62	60,4	50,9	67,3	70,4	54,6
63	61,5	51,7	68,2	71,1	55,8
64	62,5	52,5	69,0	71,9	57,0
65	63,6	53,3	69,9	72,7	58,2
66	64,6	55,1	70,8	73,5	59,4
67	65,6	56,4	71,6	74,3	60,6
68	66,7	57,8	72,5	75,0	61,8
69	67,7	59,1	73,3	75,8	63,0
70	68,8	60,4	74,2	76,6	64,2
71	69,8	61,7	75,1	77,4	65,4
72	70,8	63,0	75,9	78,2	66,6
73	71,9	64,4	76,8	78,9	67,8
74	72,9	65,7	77,6	79,7	69,0
75	74,0	67,0	78,5	80,5	70,1
76	75,0	68,3	79,4	81,3	71,3
77	76,0	69,6	80,2	82,1	72,5
78	77,1	71,0	81,1	82,8	73,7
79	78,1	72,3	81,9	83,6	74,9
80	79,2	73,6	82,8	84,4	76,1
81	80,2	74,9	83,7	85,2	77,3
82	81,3	76,2	84,5	86,0	78,5
83	82,3	77,6	85,4	86,7	79,7
84	83,3	78,9	86,2	87,5	80,9
85	84,4	80,2	87,1	88,3	82,1
86	85,4	81,5	88,0	89,1	83,3
87	86,5	82,8	88,8	89,9	84,5
88	87,5	84,2	89,7	90,6	85,7
89	88,5	85,5	90,5	91,4	86,9
90	89,6	86,8	91,4	92,2	88,1
91	90,6	88,1	92,3	93,0	89,3
92	91,7	89,4	93,1	93,8	90,4
93	92,7	90,8	94,0	94,5	91,6
94	93,8	92,1	94,8	95,3	92,8
95	94,8	93,4	95,7	96,1	94,0
96	95,8	94,7	96,6	96,9	95,2
97	96,9	96,0	97,4	97,7	96,4
98	97,9	97,4	98,3	98,4	97,6
99	99,0	98,7	99,1	99,2	98,8
100	100,0	100,0	100,0	100,0	100,0