

Tartu Ülikool
Sotsiaal- ja haridusteaduskond
Psühholoogia instituut

Merle Kõrgesaar

**FLYNNI EFEKTI ESINEMINE EESTI ABITURIENTIDE SEAS
RAVENI TESTI PÕHJAL**

Seminaritöö

Juhendaja: Helle Pullmann, *Ph.D*

Läbiv pealkiri: Flynn'i efekt abiturientidel

Tartu 2013

Kokkuvõte

Flynni efektiks (FE) nimetatakse intelligentsustestide skooride (IQ) keskmiste tõusu ajas. Viimase sajandi vältel kogu maailmas kogutud andmed on näidanud, et keskmiselt on sooritustase tõusnud 3 IQ punkti dekaadi kohta. Samas on avaldatud töid, mille kohaselt on osades riikides FE peatunud või isegi pöördunud. Käesolevas töös uuriti FE võimalikku esinemist Eesti abiturientide näitel mittesõnalise üldvõimekuse Raveni SPM testiga ajavahemiku 2001-2012 jooksul. Kui varasemad tööd Eestis on pikema ajaperioodi ja teise testiga näidanud FE esinemist (Must et al., 2009), siis käesolevas töös ilmnes hoopis testisoorituses langustrend, kus tosina aastaga langes abiturientide keskmine testiskoor 0.56 standardhälvet (8.4 IQ punkti). Seejuures leiti eriti tugev langusefekt tütarlaste testiskoorides. Kuigi võimalikke tulemust mõjutavaid tegureid püüti kontrollida, ei ole siiski võimalik ammendavat vastust langustrendi põhjuste kohta käesoleva töö põhjal anda.

Abstract

Presence of Flynn Effect Among Estonian School-Leavers in Raven Matrices.

The average increase of intelligence test scores (IQ) over time is called Flynn effect (FE). Gathered data from several different countries all over the world during the last century have shown that the average performance level has risen 3 IQ points per decade. However, the published studies from some countries have reported of stopped or even reversed FE. The main goal of this study was to examine the possible presence of FE among Estonian school-leavers with non-verbal Raven SPM test to measure their general intelligence during the period 2001-2012. Although the previous two studies in Estonian cohort in the longer period showed the presence of FE (Must et al., 2009), the present work found a decreasing trend in IQ-gains - over a twelve years Raven SPM test scores fell 0.56 standard deviation units (8.4 IQ points) of the cohort. However, decline was particularly strong in girls' scores. Although several possible confounding factors were controlled for, the current study is not able to offer any solid explanations for the reported decline.

Sissejuhatus

Intelligentsuse all mõistetakse inimese üldise vaimset võimekust, mis muude asjade kõrval eeldab võimet arutleda, planeerida, lahendada ülesandeid, mõelda abstraktselt, saada aru keerulistest ideedest, õppida kiiresti ja õppida kogemustest. See ei ole lihtsalt raamatutarkus, kitsalt akadeemiline või testidele vastamise oskus (Gottfredson, 1997; eestikeelne tõlge Mõttus, Allik, & Realo, 2011).

Kaasaegne intelligentsuse hindamine sai alguse 20. sajandi algusaastatel, kui 1904. aastal avaldasid Prantsuse psühholoogid Alfred Binet ja Theophile Simon laste intelligentsuse mõõtmiseks mõeldud testikogumiku (Binet & Simon, 1916), mille toel hinnati õpilaste vaimset arengutaset eakaaslastega võrreldes. Selle möödunud sajandi vältel on välja töötatud suur hulk erinevaid intelligentsuse teste ja nende rakendamisest kogunenud hulgaliselt andmeid läbi erinevate aastakümnete. Neid andmed süstemaatiliselt analüüsid on selgunud, et testide keskmised tulemused on kogu sajandi vältel märkimisväärselt kasvanud. Millega on tegemist ja millest selline tendents tuleneb?

Käesolev seminaritöö keskendub intelligentsustestide keskmiste skooride kasvamise uurimisele ajas, andes esmalt ülevaate maailmas viimase sajandi vältel leitud tulemustest ning seejärel uurides, kas ja millisel määral on see kasvutrend täheldatav viimasel dekaadil Eesti abiturientide seas.

Flynni efekt

Otago Ülikooli emeriitprofessor James R Flynn võrdles 1984. aastal Stanford-Binet ja Wechsleri intelligentsustestidega mõõdetud ameeriklaste arhiveeritud rahvuslike andmeid ajavahemikus 1932-1987 ja avastas tulemusi analüüsid fakti, mille kohaselt intelligentsustestide standardskoorid (IQ) tõusid sellel perioodil keskmiselt 0,30 punkti aastas ehk 3 IQ punkti dekaadi jooksul (Flynn, 1999). Selle fenomeni nimetas sama valdkonda uuriv teadlane Charles Murray 'Flynni efektiks' (Herrnstein & Murray, 1994).

Flynni efekt (edaspidises tekstis lühendatult *FE*) seisneb selles, et kui võrrelda teatud vanuses olevate inimeste IQ skooride kümme, kakskümmend, kolmkümmend jne aastat tagasi mõõdetud samas vanuses inimestel, siis täna saavad samaealised keskmiselt kõrgemaid skooride, kui minevikus on nende ealistelt inimestelt mõõtmise tulemusena saadud (J. Flynn, Shaughnessy, & Fulgham, 2012).

Seejuures ei ole tegemist ainult ameeriklaste andmetelt leitud fenomeniga, vaid 1987. aastaks olid Flynnil sarnased tulemused 14–st tööstuslikult arenenud riigi kohta (Trimble, 2011). Sarnased uuringud on jätkunud ka väljaspool arenenud riike ning 29-s riigis on tuvastatud IQ skooride tõus ajas (Kanaya & Ceci, 2011).

Flynni efekt ja erinevad mõõtevahendid

Erinevates kultuurides leitud tendents viitab FE üldisemale universaalsusele, kuigi skooride tõus ajas ei ole ühene ning sõltub ka kasutatavast mõõtevahendist. Kõige sagedamini on FE uuritud traditsiooniliste intelligentsustestide (edaspidi ka IQ-testide) tulemuse vahendusel, nagu näiteks 1916. aastal Stanfordini ülikooli psühholoog Lewis Termani poolt väljatöötatud Stanford-Binet test (Terman, 1921).

David Wechsleri poolt väljatöötatud nii täiskasvanute (WAIS – *Wechsler Adult Intelligence Scale*, viide) kui ka laste (WISC – *Wechsler Intelligence Scale for Children*, viide) intelligentsuse skaalad. Nimetatud testid sisaldavad lisaks mittesõnalistele ülesannetele sõnalisi küsimusi ja võivad seeläbi olla mõjutatud ka kultuurilisest ja hariduslikust taustast. Samas on laialt FE uurimiseks kasutatud 1936. aastal John C. Raveni poolt koostatud mitteverbaalset Raveni Standardsete Progressiivsete Maatriksite (SPM – *Raven Standard Progressive Matrices*; Raven, 2000) testi, mis koosneb loogilist mõtlemist ja arutlusoskust hindavastest pildimaatriksitest ning mida peetakse kõige puhtamaks üldintelligentsuse mõõdikuks sõltumata rassist, rahvusest või kultuuride erinevustest (Jensen, 1998).

Erinevate intelligentsustestidega kogutud tulemusi analüüsides on leitud, et sõnalistes alatestides (nt Wechsleri ja Stanford-Binet' testides) on tendents skooride tõusuks ajas kõige väiksem, samas kui Raveni maatriksite ja Wechsleri testide mittesõnaliste alatestide puhul on saadud ajas oluliselt kõrgemad skooride muutused (J. Flynn et al., 2012). See tendents on leidnud kinnitust erinevates kultuurides, näiteks Lõuna-Aafrika Vabariigis läbiviidud FE uuringust selgus, et Wechsler testide mitteverbaalses osas avaldus FE kõrgema skoorina kui verbaalse testi osa tulemustes (Jan te Nijenhuis, Murphy, & van Eeden, 2011).

Flynni efekt ja kultuuridevahelised erinevused

Raveni maatriksite tulemuste põhjal 4 riigis – Hollandis, Norras, Iisraelis ja Belgias – oli 3 riigi, v.a Norra, keskmine tõus 30 aasta vältel (vahemikus 1952-1982) 18-20 IQ punkti,

ehk keskmiselt 6 IQ punkti dekaadi kohta. Norras oli sarnane tõus esimesel mõõdetud dekaadil, kuid peale 1968. aastat oli IQ skooride tõus dekaadi jooksul 7.5 punkti (Flynn, 1999). Peale seda perioodi tehtud sarnane uuring Norras aga näitas, et 21.saj alguse testimise andmete põhjal paralleelselt Raveni ja teiste intelligentsustestidega on Norras IQ skooride tõus peale 1990.-ndaid aastaid peatunud (Sundet, Barlaug, & Torjussen, 2004).

Flynn analüüsis Raveni maatriksite skooride põhjal FE Argentiinas (vahemikus 1964-1998) ning sai 34 aasta keskmiseks tõusuks 21.35 IQ punkti, mis on Raveni maatriksitele tüüpiline. See uuring näitas, et kuigi uuring oli tehtud vaid ühes Argentiina linnas ja selle põhjal ei saa teha väga põhjalikke järeldusi kogu Argentiina kohordi kohta, siis võib nende tulemuste põhjal siiski ära märkida, et La Plata linna IQ skoorid olid Norra seisakule vastupidiselt kõrged, mis Flynn'i arvates võivad tuleneda Norra ja Argentiina ühiskonde erinevatest arengustaadiumitest ja stabiilsuse tasemest (Flynn & Rossi-Casé, 2012).

Sama mõtet on edasi arendatud ka uuringutes, mis on läbi viidud Lõuna-Aafrika Vabariigis (LAV) 20. saj andmete põhjal (Jan te Nijenhuis et al., 2011) ja Koreas (vahemikus 1970-1990), et mõõta FE arengumaades ning tõmmata paralleele stabiilsemate sama regiooni riikidega, kus IQ skoorid on saavutanud teatud platoo (Te Nijenhuis, Cho, Murphy, & Lee, 2012). Neis mõlemas uuringus on kasutatud WAIS teste, ent tulemusi mõõdeti verbaalse ja mitteverbaalse testiosade kontekstis eraldi. Tulemusena saadi keskmiseks FE skooriks LAV puhul 79,31 IQ punkti sajandi ehk 7,93 punkti dekaadi kohta (Jan te Nijenhuis et al., 2011) ja 7,7 IQ punkti dekaadi kohta Korea uuringus (Jan te Nijenhuis, Cho, Murphy, & Lee, 2012).

Seega viitavad kultuuridevahelised uuringud FE sõltuvusele nii kultuurist kui mõõtevahendist ja analüüsiperioodist, mis hoiab teemat uurijate jaoks endiselt aktuaalsena ning jätkuvalt otsitakse vastuseid FE esinemise arengumustritele ja põhjustele.

Flynni efekti uurimine Eestis

Ka Eestis on dotsent Olev Musta eestvedamisel juba aastaid FE uuritud. Need uuringud on olnud tänuväärt materjal käesoleva töö sünnil andes ühelt poolt aimu, mida on uuritud ja avastatud ning teiselt poolt impulsi püüda FE teemat Eestis veelgi põhjalikumalt analüüsida, sest autorile teadaolevalt ei ole tänapäeval ühte kindlasse vanusegruppi kuuluvatelt katseisikutelt Raveni SPM testiga andmeid kogutud ega teatud ajaperioodi lõikes varasemate tulemustega võrreldud.

Siiani Eestis tehtud FE uuringud baseeruvad ühelt poolt Juhan Torki poolt kogutud andmetel (Tork, 1940), mis pärinevad aastatest 1933/36 Eesti kooliõpilastelt (keskmine vanus 13 aastat) kogutud *NIT*-testiga, mis on USA-s kasutatud *National Intelligence Tests* Eesti jaoks adapteeritud versioon. Torki toonast uuringut on just Eestis FE uurimise eesmärgil korranud sama mõõtvahendit kasutades 1997/98. aastatel Tartu Ülikooli psühholoogia instituudi töörühm Olev Musta juhtimisel. Hoolikalt arvestades nii toonaseid kui tänaseid erinevusi ühiskonnas, kasvukeskkonnas ja hariduse kättesaadavuse osas, lisaks veel mõõtmisvea ja juhuslikkuse arvestus, andis 381 Torki ja 449 Musta uuringus osalenud õpilase tulemuste võrdluse põhjal Flynni efektiks üle kõigi alltestide mõõdetud üldintelligentsuse tulemuseks 0.44 standardhälbe (SD) ühikut, mis IQ punktides teeb ligikaudu 6 IQ punkti dekaadi kohta (Must, Must, & Raudik, 2003).

Olev Musta uurimisrühma poolt on Eesti NIT testi kasutades uuringut korratud ka 2006. aastal ja tulemusi võrreldi varasemate tulemustega. Leiti taas, et IQ skoorid Eestis on analüüsivate perioodide lõikes tõusuteel - viimase kahe uuringu andmeid kokku võttes leiti FE tulemuseks Eestis mõõdetud tulemuste põhjal 1997/98 kuni 2006 SD ühikuks 1/6, mis on ca 5 IQ punkti kaheksa aasta kohta (Must, te Nijenhuis, Must, & van Vianen, 2009).

Esile väärrib siinkohal toomist seni tehtud Eesti uuringutes tulemuse esitamine standardhälbe ehk SD-ühikutes, mis muudab eri ajastutel Eesti NIT testiga kogutud andmed omavahel enam tähenduslikult võrreldavaks. Arvestades asjaolu, et mõõtmisi aastatest 1933/36, 1997/98 ja 2006 ei saa võtta invariantsetena ja hinnatud kohordid ei ole omavahel otseselt võrreldavad, siis järeldusi on saadud teha eelkõige kahe teguri põhjal, milleks on testimiste ajaline vahe aastates ja kohortide vahelised ühiskondlikud/hariduslikud muutused (Must et al., 2003; Must et al., 2009).

Eesti uurimuste kokkuvõtteks selgub, et seni mõõdetud Eesti tulemused ei ole näidanud sarnast seisakut IQ skooride tõusus, nagu seda leiti Skandinaaviamaades tehtud FE (Sundet et al., 2004) ja Eesti tänapäeva (2006. aastal hinnatuna) oludes on skoorid liikunud tõusujoones. Samas puudub info viimasel kümnendil toimunud arengute hindamiseks ja käesolev töö püüab seda lünka täita.

Flynni efekti põhjused

Põhjused, miks FE toimib ja miks IQ skoorid ajas on tõusnud, pole veel siiani selged ega üheselt mõistetavad. Teooriaid on pakutud FE uurijate ning kriitikute poolt erinevaid -

psühhomeetriline artefakt, muutused keskkonnas, parem toitumine, meditsiini edusammud, hariduse olulisus ning kättesaadavus, geneetiline heteroos. Neist arvestatavamateks peetakse nii geneetilisi põhjuseid kui keskkondlikke muutusi, sealhulgas hariduslik sekkumine, millest saab kokkuvõttes pärilikkuse-keskkonna koostoimeteooria, mida praegusel hetkel peetakse ka kõige adekvaatsemaks põhjenduseks, kuid kaugeltki mitte lõplikuks seletuseks FE tekkel (Trimble, 2011).

Esiteks, geneetilise võimaliku põhjusena nimetatakse heteroosi. Heteroos ilmneb tavaliselt sama liigi inbriidsete liinide ristamisel ning väljendub iga põlvkonna kui hübriidide (ristandite) suurenenud kasvus, kasvukiiruses, produktiivsuses, vastupidavuses jms. Nii on inimkonnaga juhtunud, kuna väikesed suletud kogukonnad on lagunened ning elanike linnastumine ja mobiilsus võimaldavad paaritumist geneetiliselt erinevate rühmade vahel (Mingroni, 2007). Kuigi heteroosi kui võimalikku põhjust FE ilmnemisel arvestatakse ühe faktorina, ei peeta seda siiski peamiseks põhjustajaks, kuna FE-d seostatakse ennekõike inimese arengulise staatusega (Woodley, 2011).

Teiseks, keskkondlike muutuste teooria hüpotees väidab, et võrreldes sotsiaalsete oludega minevikus nõuab tänapäeva ühiskond inimeselt rohkem arenenud kognitiivseid oskusi keerulisemas ning enam arenenud keskkonnas hakkama saamiseks ning neid kognitiivseid oskusi kajastab voolav intelligentsus, mille tõus FE ka sisuliselt kõige enam on (Trimble, 2011).

Hariduslik sekkumine on areneva ühiskonna üks tunnuseid. See kannab endas ideed muuta haridus kättesaadavamaks ning arendada inimestes analüüsi ja kriitikasokusi, mis omakorda on seotud kognitiivsete protsessidega ning kahtlemata on seotud IQ-ga (Flynn, 1998). Kuna haridus on laste ja teismeliste sotsiaalses ja intellektuaalses arengus järjest päevakohasem, on kasvanud ka individuaalne panus haridusse. Seda fenomeni peegeldavad kõik riigid, kus Flynn on oma uurimust läbi viies andmeid kogunud - inimesed nendes riikides pühendavad järjest rohkem ressursse haridusse ja koolellu (Flynn, 1998). Selgemini väljendudes: suureneb nende inimeste hulk, kelle koolitee kestab järjest enam aastaid.

Paljud FE põhjustesse süvenenud uurijaid (näiteks Jensen, 1998; Williams, 1998; Ang, Rodgers, & Wanstrom, 2010; Flynn, 1998; Trimble, 2011) toetavad mitme erineva põhjuse koosmõju FE tekkel. Põhjuse paljususe koosmõju teooria FE põhjuse seletamises on atraktiivne seepärast, et see annab piisava selgituse FE järjepidevusele läbi aja ning ka kultuuride erinevusest lähtuvalt põhjuse FE-d kui nähtust ja selle võimalikke

tekkepõhjusi edasi uurimaks (Ang et al., 2010). Erinevate põhjuste koosmõju teooria tugineb peamiselt järgmistel väidetest:

- 1) Haridustaseme paranemine on toonud kaasa hariduse kättesaadavuse ning akadeemilise haritustee pikema kestvuse. Kuna hariduse omandamine on tihedalt seotud intelligentsusega, siis eelduste kohaselt need asjaolud põhjustavad IQ skooride tõusu (Williams, 1998);
- 2) Toitumise ning tervishoiuteenuste paranemine arenenud riikides on parandanud elukvaliteeti esmavajaduste tasandil ning see annab ressursse tegelemaks vaimsete võimete arendamisega. Lisaks sellele võimaldab paremini arenenud ühiskond oma parema elukvaliteedi tasemega parandada oluliselt mittesõnalise intelligentsuse taset (Trimble, 2011);
- 3) Paranenud sünnitusabi ning vaksineerimised arenenud riikides on kõrvaldanud palju takistusi, mis on seni olnud hariduse ja intellektuaalsuse arenemisel olulised (Trimble, 2011);
- 4) Kohanemine testimisega ja IQ testide laialdane kasutamine tänapäeva elus (e.g. töövestlustel, koolidesse astumisel jm) ning üldised kasvavad nõudmised vaimsele võimekusele, mis motiveerivad testitavaid enam pingutama, põhjustavad IQ skooride tõusu ajas (Jensen, 1998; Trimble, 2011; J. Flynn et al., 2012).

Flynni efekti praktiline väljund

Intelligentsustestid, nagu kõik kognitiivsed testid, vajavad ajapikku süstemaatilist kontrolli, sest testi normid võivad olla ajas muutuvad. Nende devalveerumine avaldub testitavate kohanemises testidega ja testide põhimõtetega või populatsiooni kognitiivse arengutaseme muutumises. Muud põhjused testide normide uuendamiseks e. renormimiseks võivad olla testide ülesehituse muutmine vastavalt teadussuundade arengule või testide valideerimine (Resing & Tunteler, 2007). Uurijate arvates põhjustab selline testide normide uuendamine intelligentsustestide vananemise umbes ühe põlvkonna vältel (Kanaya, Ceci, & Scullin, 2003). Vastuseks sellele püüavad IQ-testide koostajad koostada oma uued versioonid testidest selliselt, et vastajatel on vaja kas vastata keerulisematele küsimustele või koguda suuremad toorpunktiskoorid kui see kohort, kes tegid vanema versiooni teste. Seega eelmisel sajandil tehtud testimiste puhul hoiti renormimisega IQ keskmine 100 punkti peal teste perioodiliselt raskemaks muutes (Kanaya, Ceci, et al., 2003).

FE on ilmselt kõige paremini dokumenteeritud näide testi normide ajalisest ebastabiilsusest (Fischer & Milfont, 2010) ning kasutades vananenud norme võib väga lihtsalt teha valesid järeldusi testitute kohta. FE IQ skoorides on USA-s osutunud määravaks iseäranis finants-, sotsiaal- või kohtusfääri puudutavates otsustes, mille langetamiseks kasutatakse alusena IQ skoores: õpiraskuste diagnoosimisel, sõjaväeteenistuseks sobivuse hindamisel, surmanuhtluse määramisel jm (Kanaya, Ceci, et al., 2003).

Kuna Flynn (1998) avastas, et alati kui testiti paralleelselt äsja normide uuenduse e renormimise läbinud testidega ning vananenud normidele vastava testiga (Wechsler ja Stanford-Binet' puhul), saadi madalamad skoorid alati hiljuti renormitud testidelt (Flynn, 1998). Ta avastas seega lisaks IQ skooride tõusule ajas ka IQ skooride languse pärast iga testide renormimise tsükli. Eelnevast järeldatult tuleks erilise ettevaatusega tulemuste tõlgendamisel suhtuda ennekõike juhul, kui testi normid on äsja uuendatud, sest siis saadud skoorid on oluliselt madalamad kui eelmises tsükli kogutud skoorid. Või juhtudel, kui testi normeerimistsükli periood on peatselt lõppemas, sest mõõdetud skoorid osutuvad tingituna normide devalveerumisest kõrgemateks (Kanaya, Scullin, & Ceci, 2003).

FE praktilise väljundi analüüsimise eesmärgiks on olnud tõdemus, et testide normid aeguvad ning vajavad kontrollimist-uuendamist. FE põhjalik uurimine annab võimaluse hinnata kuidas, millal ja kui palju on vaja norme korrigeerida. FE kui nähtuse esinemise uurimine pole enam niivõrd aktuaalne, kuna selle olemasolu on läbi aegade mitmetes uuringutes tõestud. Ülaltoodud eesmärgist lähtuvalt on märksa olulisem uurida täpsemalt FE tekkepõhjust, et neist tulenevalt nähtust paremini kontrollida ning arvesse võtta, iseäranis olukordades, kus kaalul on edasine elu või koguni surmanuhtlus (Trimble, 2011). Samasugune praktiline vajadus FE uurida, õnneks küll teises kontekstis kui surmanuhtluse üle otsustamine, on Eestiski, sest tänapäeva Eestis kasutatakse paljude karjäärinõustajate ning koolipsühholoogide poolt veel 1980. aastatel suhteliselt piiratud arvu õpilastelt kogutud andmete põhjal normeeritud teste (Pullmann et al., 2004). Kuna nende normide järgi tehakse siiani testitavate kohta järeldusi, siis võib tekkida olukord, kus aegunud normide põhjal on Eestis peagi lapsed psühhomeetriliselt hinnatuna andekad. Seega kannab normide uuendamise teema endas ka Eesti jaoks väga olulist ja kriitilist küsimust, mida käesolev töö püüab samuti pisut avada.

Töö eesmärgid

Kuivõrd FE Eestis ei ole varem Raveni SPM testiga uuritud, siis on käesoleva töö unikaalne panus esmalt abiturientide näitel välja selgitada, kas ja millisel määral esineb FE 2012. aastal kogutud andmete võrdlusel 11.5 aastat varem kogutud Raveni testiga kogutud andmetega.

Kuna Raveni SPM rahvuslik normeerimine toimus 2001. aastal (Pullmann et al., 2004) ja seega on möödunud ligikaudu tosin aastat normandmete kogumisest, siis on töö teine eesmärk analüüsida vajadust kehtivate testinormide ülevaatamiseks, et pakkuda vastus praktiseerivatele psühholoogidele, kes kasutavad Raveni testi oma igapäevases töös koolinoorte hindamiseks ja ootavad hinnangut vanusegrupi normide kehtivusele. Kui töö tulemusel leiab kinnitust märkimisväärne skooride muutus selle perioodi vältel, siis võimaldab see anda hinnangu Eestis rakendatavate SPM testi normide ajakohasusele või vajadusele normide korrigeerimiseks. Seega on käesoleval seminaritööl on nii teoreetiline kui praktiline eesmärk.

Tuginedes varasematele uuringutele eeldan töö peamise hüpoteesina, et võrreldes Eesti abiturientide Raveni SPM testisooritust ligikaudu 12 aastase intervalliga, ilmneb Flynn'i efekt. Seejuures eeldan, et skooride tõus ajas võiks olla kuni 3 IQ punkti, kuna viimane dekaad Eestis on olnud sotsiaalsete faktorite poolest suhteliselt stabiilne.

Siiski on tulemused ennustamatud, sest näiteks üldine tehnika (sh nutitehnika) areng Eestis on olnud hüppeline ja see võiks tähendada noortel teatud tänapäevaste oskuste komplekti kiiret arengut, mis võiks olla üldintelligentsusskooride tõusu ajas soodustav.

Hariduse kättesaadavuse ja selle kvaliteedi olukord on piirkonniti Eestis erinev – Ühelt poolt väiksemates maakohtades suletakse koole ning koondatakse koolid suurematesse keskustesse, siis lastel kulub kooli ja koju jõudmiseks füüsiliselt rohkem aega, mis jätab neile vähem aega tegeleda õppimise ning muude arendavate tegevustega nagu trennid, muusikakool või huviringid. Teisalt näitab riigieksamite tulemuste pingerida (Postimees, 2012), et suuremates keskustes asuvates koolides võiks olla hariduse kvaliteet parem ehk kõrgema tasemega koolides on õpilaste õpitulemused kõrgemad (Kikas, Toomela, & Mõttus, 2006).

Seepärast arvestades käesoleva uurimuse valimit üle Eesti nii, kus on esindatud nii väiksemaid maakonna koole kui ka linnade kõrgema tasemega koole, võib eeldada, et FE Eesti abiturientidel avaldub, kuid vähemal määral, kui on leitud siiani.

Meetod

Valim

Valimiks käesoleva uurimuse puhul on Eesti koolide abituriendid aastatest 2001 ja 2012, lisaks on vahevõrdluseks kaasatud abituriendid aastast 2005. Valimi võrreldavuse tagamiseks kaasati valimisse ainult 18-19 aastased õpilased, osalejate keskmine vanus kõigi osalejate peale kokku 18.19 eluaastat.

Osalejaid kõikide uurimust läbiviidud 3 aastakäigu peale kokku oli 1328, neist meessoost 544 ja naissoost 784 õpilast. Valimite kirjeldus aastate lõikes on järgmine: 2001. aasta valimis oli osalejaid 573 (247 noormeest ja 326 neidu) 27 koolist, 2005. aasta valimis oli osalejaid 417 (155 noormeest ja 262 neidu) 17 koolist ning 2012. aasta valimis oli osalejaid 338 (142 noormeest ja 196 neidu) 14 koolist. Algne 2012. aasta valim koosnes 347 testitust, kuid sarnaselt varasematele töödele (Must et al., 2009) jäeti analüüsist välja vastajad ($n = 9$), kelle üldvõimekuse testiskoor oli madalam kui 2 standardhälvet eakaaslaste keskmisest, et hoida ära katseisiku madalast motivatsioonitasemest või muudest faktoritest mõjutatud madalate skooride võimalikku mõju valimi testisoorituse keskmisele tasemele.

Et saada võimalikult parem ülevaade uuringus käsitlevate probleemide olemusest Eestis ning usaldusväärsem võimalus teha järeldusi kogu Eesti kohordi kohta, koostati valimid üle Eesti paiknevate koolide õpilastest, kuhu koolidest olid võrdväärsetl kaasaatud nii suuremate linnade, maakonnakeskuste kui ka väiksemate administratiivüksuste koolid kõigist Eesti piirkondadest. Täpsem info koolide administratiivse jaotuse kohta aastate lõikes on esitatud joonisel 3.

Koolide tasemete hindamiseks kasutati Eesti koolide pingerida (Postimees, 2012), mille alusel omistati igale uuringus osalenud koolile perioodi 2010-2012 keskmine koht Eesti üldhariduskoolide pingereas. Koolid jagati seejärel koha alusel viide gruppi, täpsem info on esitatud joonisel 2.

Mõõtevahend

Mõõtvahendina on käesolevas uuringus kasutatud Raveni Standardiseeritud Progressiivseid Maatrikseid (SPM; Raven, 2000), mis on kirjandusele toetudes kultuurierinevustest, rassist ja rahvusest sõltumatu ning parim mitteverbaalne mõõtevahend üldintelligentsuse g mõõtmiseks (Jensen, 1998, J. Flynn et al., 2012). Test

on välja töötatud psühholoog John Carlyle Raveni (1902-1970) poolt eesmärgiga kõrvaldada mõõtmise tulemuste sõltuvus inimese keelelisest arengust. Test algab lihtsamate ülesannetega ja ülesanded muutuvad progresseeruvalt järjest keerukamaks. Raveni SPM testi sooritajal tuleb ette antud kaheksa kujundi põhjal leida seaduspärasus või loogiline järgnevus ja valida kuuest variandist kujundiseeria jätkamiseks või mustri lõpetamiseks sobivaim kujund. Test koosneb 60 ülesandest ja 5-st alatestist (alatestide omavahelised seosed esitatud tabelis 1, valimite keskmised tasemed tabelites 2 ja 3). Kokku on Raveni SPM testiga võimalik koguda summaarselt 60 punkti. SPM on Eestis normeeritud ajapiiranguta 2001. aastal 4874 õpilasest koosneva valimi peal vanuses 7-19 eluaastat (Pullmann et al., 2004). Skaala sisereliaabluse koefitsient oli antud valimil $\alpha = 0.82$ ja testi täideti ajapiiranguta.

Tabel 1

Raven SPM testi alaskaalade omavahelised seosed

2012/2001	A	B	C	D	E	Koguskoor
Alatest A		.05	.12*	.05	.07	.21***
Alatest B	.05		.20***	.19***	.24***	.49***
Alatest C	.10	.19**		.23***	.45***	.71***
Alatest D	.04	.12*	.26***		.25***	.54***
Alatest E	.11	.17**	.57***	.20**		.85***
Koguskoor	.26***	.42***	.77***	.52***	.85***	

Märkused. * $p < .05$, ** $p < .01$, *** $p < .001$. Seosed on leitud Pearsoni korrelatsioonkordajat (r) kasutades. Diagonaalist üleval on 2001. aasta tulemused ($n = 552$) ja allpool 2012. aasta tulemused ($n = 304$).

Protseduur

Kõigi kolme testimislaine puhul toimus õpilaste testimine grupitestimisena koolitundide ajal. Enne testimist täitsid õpilased nõusolekulehe uuringus osalemaks. Testimisele kulus kaks koolitundi, kus esimese tunni ajal täitsid õpilased SPM testi ja teise tunni vältel vastasid õpilased enesekohastele küsimustikele ja täiendavale ankeedile. Peale andmete sisestamist ja töötlemist edastati elektroonselt igale uurimuses osalejale soovi korral individuaalne tagasisideleht, kus nende tulemused olid lahti seletatud võrreldes samast soost ja vanusest Eesti abiturientide keskmiste tulemustega. Töö autor osales nii andmekogumisel kui nende sisestamisel.

Andmeanaliüs

Kogutud andmete statistiliseks analüüsimiseks kasutati SPSS 20.0 (IBM) ja Statistika 8.0 andmetöötluspaketti (StatSoft). Kasutati kirjeldavaid statistikuid ning keskmiste erinevuse hindamise meetodeid (*t*-test ja *Anova*) nii sooliste kui kohortide keskmiste ning teiste grupikeskmiste erinevuste hindamiseks.

Testisoorituse kasvu hindamiseks kasutati standardhälbeühikuid (Cohen'i *d*), mida arvutatakse keskmise ja standardhälbe ühiku kaudu ning mille alusel tõlgendatakse keskmiste erinevuste efekti suurus järgnevalt: "väike erinevus, $d = 0.2$ ", "mõõdukas erinevus, $d = 0.5$ " ja "suur erinevus, $d = 0.8$ ". (Cohen, 1987). Muutuste esitamine standardhälbe ühikutes on mõistlik seepärast, et nii on eri alltestide soorituste muutused omavahel võrreldavad (Must, 2011). Tulemuste esitamiseks IQ-punktides korrutatakse standardhälbed läbi 15-ga.

Tulemused

Keskiste testisoorituse tasemete võrdlus

Esmalt võrreldi, millisel määral erinevad 2005. aastal testitud abiturientide Raveni SPM tulemused võrreldes neli aastat varem (2001. aastal) peamiselt samades koolides testitud õpilaste testisooritusest. Täpsemad tulemused on esitatud tabelis 2 nii koguvalimile kui eraldi neidudele ja noormeestele. Tulemused näitasid, et kogu valimi lõikes jäid tulemused samale tasemele ($d = -0.03$), kuid tegemist on keskmistatud tulemusega ja võrdlus sugude lõikes näitas, et kui noormeeste testisoorituses ilmnis nõrk, kuid statistiliselt oluline testisoorituse paranemine ($d = 0.23$, $p < .05$), siis neidudel leiti samas suurusjärgus oluline langus ($d = -0.25$, $p < .01$).

Tabel 2

Testisoorituse muutused: 2001. ja 2005. aasta valimite võrdlus

		2001 <i>N</i> = 552		2005 <i>N</i> = 411		Muutus	
		<i>m</i>	<i>SD</i>	<i>m</i>	<i>SD</i>	<i>d</i>	<i>t</i>
<i>Kogu valim</i>	Alatest A	11.9	0.4	11.8	0.5	-0.22	1.8
	Alatest B	11.4	0.8	11.4	0.8	0.00	-0.4
	Alatest C	10.7	1.2	10.6	1.3	-0.08	1.2
	Alatest D	10.6	1.0	10.6	1.0	0.00	0.1
	Alatest E	8.9	2.0	8.8	2.1	-0.05	0.4
	Koguskoor	53.4	3.6	53.3	4.1	-0.03	0.7
<i>Neiud</i>	Alatest A	11.9	0.4	11.8	0.5	-0.22	1.6
	Alatest B	11.5	0.8	11.4	0.8	-0.13	1.0
	Alatest C	10.6	1.2	10.4	1.3	-0.16	1.6
	Alatest D	10.7	1.0	10.6	1.0	-0.10	1.3
	Alatest E	8.8	2.0	8.4	2.1	-0.20	2.6**
	Koguskoor	53.4	3.5	52.5	3.8	-0.25	2.8**
<i>Noormehed</i>	Alatest A	11.8	0.4	11.8	0.5	0.00	1.0
	Alatest B	11.4	0.9	11.5	0.8	0.12	-1.7
	Alatest C	10.8	1.2	10.9	1.3	0.08	-0.5
	Alatest D	10.6	1.0	10.7	1.1	0.10	-1.4
	Alatest E	8.9	2.1	9.6	2.0	0.34	-3.0**
	Koguskoor	53.5	3.7	54.4	4.2	0.23	-2.3*

Märkused. * $p < .05$, ** $p < .01$. *m* = keskmine; *SD* = standardhälve, *d* = standardhälbeühikud (Cohen'i *d*), *t* = *t*-statistiku väärtus.

Teiseks võrreldi aastatel 2001 ja 2012 testitud abiturientide Raveni SPM skooore, mille tulemused on näha nii sugude kaupa kui koguvalimi lõikes tabelis 3. Tulemuseks saadi, et kogu valimi osas leiti koguskoori osas suur erinevus, täpsemalt oluline langus ($d = -0.58$, $p < .001$). Võrdlus sugude lõikes näitas, et noormeeste osas oli koguskooride muutus standardhälbeühikutes mõõdukalt langev ($d = -0.31$, $p < .01$), kuid neidude tulemuste osas toimus suur langus ($d = -0.82$, $p < .001$).

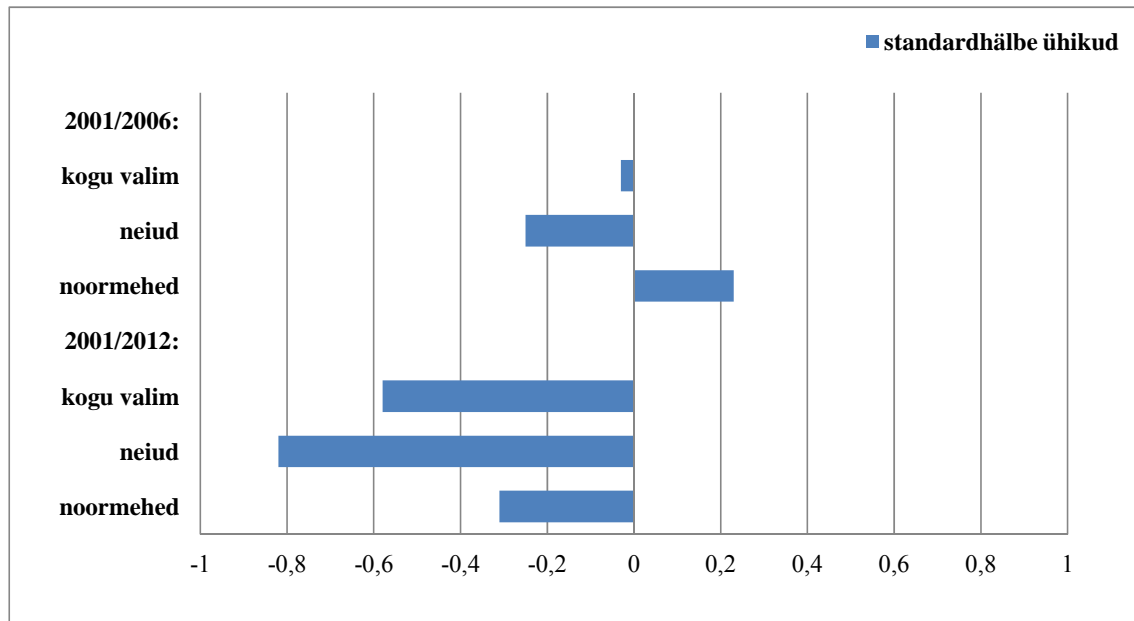
Tabel 3

Testisoorituse muutused: 2001. ja 2012. aasta valimite võrdlus

		2001 <i>N</i> = 552		2012 <i>N</i> = 304		Muutus	
		<i>m</i>	<i>SD</i>	<i>m</i>	<i>SD</i>	<i>d</i>	<i>t</i>
<i>Kogu valim</i>	Alatest A	11.9	0.4	11.7	0.6	-0.40	5.3***
	Alatest B	11.4	0.8	11.3	0.9	-0.12	2.9**
	Alatest C	10.7	1.2	10.3	1.3	-0.32	4.6***
	Alatest D	10.6	1.0	10.1	1.2	-0.45	6.9***
	Alatest E	8.9	2.0	7.9	2.2	-0.48	6.3***
	Koguskoor	53.4	3.6	51.2	4.0	-0.58	8.3***
<i>Neiud</i>	Alatest A	11.9	0.4	11.6	0.6	-0.60	5.0***
	Alatest B	11.5	0.8	11.3	0.8	-0.25	3.0**
	Alatest C	10.6	1.2	10.0	1.3	-0.48	4.8***
	Alatest D	10.7	1.0	10.0	1.3	-0.61	6.5***
	Alatest E	8.8	2.0	7.5	2.2	-0.62	6.6***
	Koguskoor	53.4	3.5	50.4	3.8	-0.82	8.7***
<i>Noormehed</i>	Alatest A	11.8	0.4	11.7	0.5	-0.22	2.2*
	Alatest B	11.4	0.9	11.3	0.9	-0.11	1.1
	Alatest C	10.8	1.2	10.6	1.3	-0.16	1.5
	Alatest D	10.6	1.0	10.2	1.0	-0.40	2.8**
	Alatest E	8.9	2.1	8.5	2.2	-0.19	1.9
	Koguskoor	53.5	3.7	52.3	4.0	-0.31	2.7**

Märkus. * $p < .05$, ** $p < .01$, *** $p < .001$. *m* = keskmine; *SD* = standardhälve, *d* = standardhälbeühikud (Cohen'i *d*), *t* = *t*-statistiku väärtus.

Joonisel 1 on ülevaatlilikult esitatud abiturientide SPM testi üldskooride muutused aastatel 2001-2005 ja 2001-2012 nii kogu valimite kui sugude lõikes.



Joonis 1. Abiturientide üldvõimekuse testi soorituse muutused aastatel 2001-2005 ja 2001-2012.

Märkus. Joonisel on kujutatud 2005. ja 2012. aastal testitud abiturientide Raveni SPM koguskoori testisoorituse vahe standardhälbe ühikutes (Cohen'i d) võrreldes 2001. aasta eakaaslaste sooritusega (0 – erinevust pole, positiivne suund – tulemus on parem, negatiivne suund – tulemus on halvem). Tinglikult defineeritakse Coheni efekti suurust järgnevalt: "väike, $d = 0.2$ ", "mõõdukas, $d = 0.5$ " ja "suur, $d = 0.8$ " (Cohen, 1987).

Testisoorituse muutus koolide taseme ja asukoha lõikes

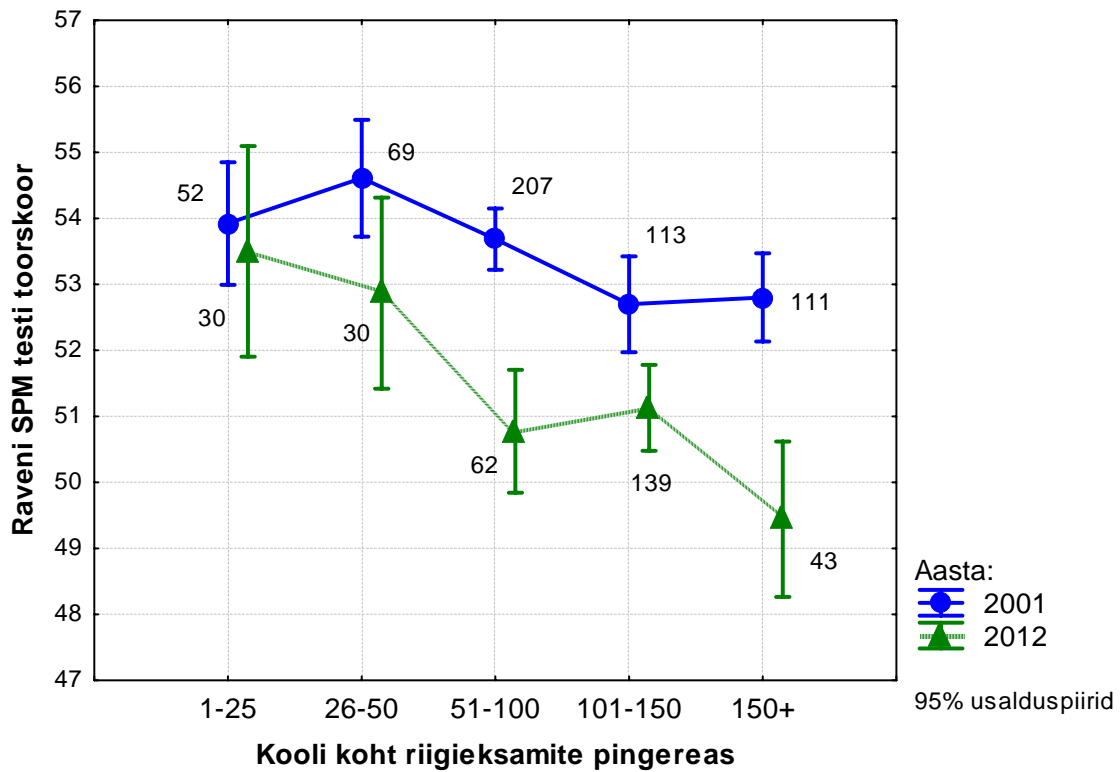
Kuivõrd kõige olulisem langus keskmistes testisooritustes oli toimunud 2001. ja 2012. aasta võrdluses, siis on oluline kontrollida nende valimite võrreldavust, sest valimites oli otsene kattuvus vaid kahe kooli puhul.

Ühe võimaliku seletusena kontrolliti, kas oluline tulemuste erinevus võis olla tuleneda valimitesse kaasatud koolide erinevast tasemest. Selleks jagati nii 2001. kui 2012. aastal uuringus osalenud koolid vastavalt riigieksamite pingereale viide gruppi ning võrreldi Raveni SPM testi keskmisi nende gruppide lõikes.

Tulemused näitasid, et mõlema valimi kooli tasemete ja kooli keskmise testisoorituse vahel ilmnes oluline seos nii koolide riigieksamite tulemuste järjestuse ($r = -.23$, $p < .001$) kui gruppidesse jagamise puhul [$F(9,846) = 13.0$, $p < .001$], kus riigieksamil edukamaid tulemusi näidanud koolide õpilased said keskmiselt ka paremaid tulemusi Raveni üldvõimekuse testis.

Vaadates tulemusi eraldi valimite lõikes ilmnes koolide tasemegruppide ja testimisaastate vahel oluline interaktsioon, $F(9,846) = 13.0$, $p < .001$. Kui esimese poolsaja kooli lõikes ei

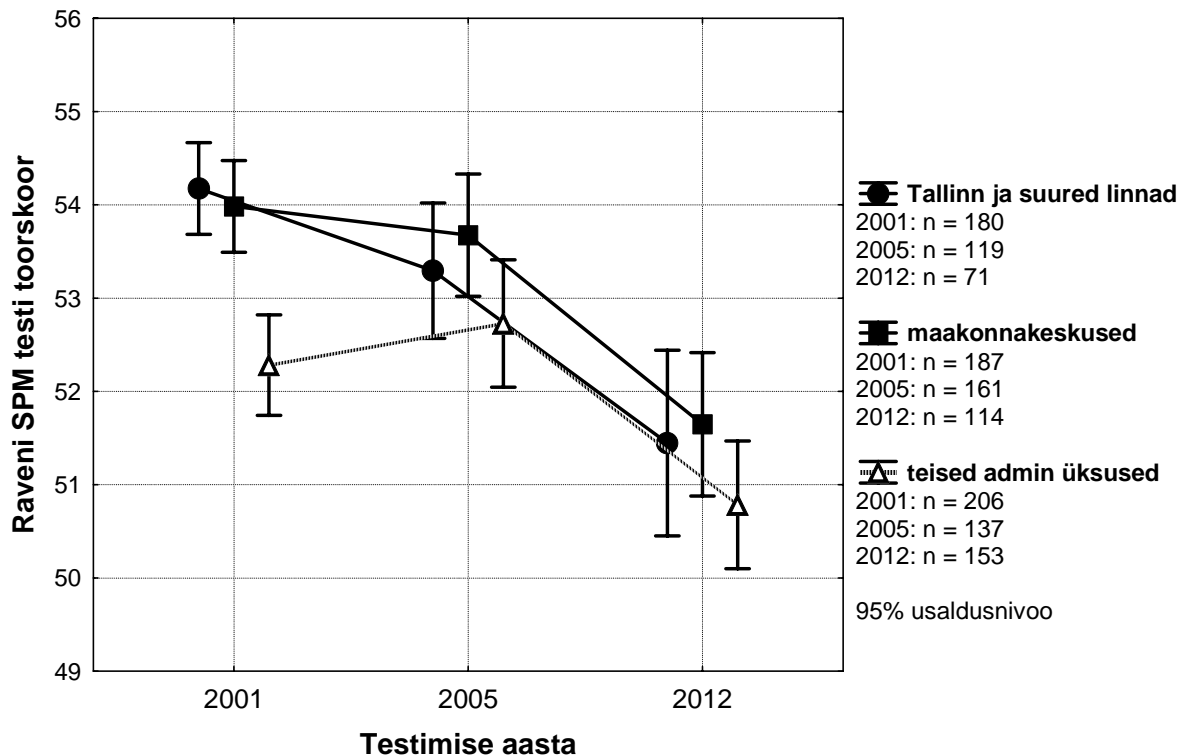
ilmnenud keskmiste soorituste vahel erinevatel testimisaastatel erinevusi, siis 2012. aasta valimis pingereas madalamal kohal olnud koolide õpilaste keskmised Raveni testisooritused olid oluliselt madalamad võrreldes sama taseme koolide õpilaste testitulemustega 11 aastat varem. Täpsemad tulemused on graafiliselt esitatud joonisel 2.



Joonis 2. Koolide keskmiste testisoorituste jaotus koolide pingereade lõikes

Teiseks analüüsiti, millisel määral on valimites esindatud erinevate administratiivsete piirkondade koolid ja kas 2012. aasta valimi oluliselt madalam sooritus ilmneb ka selle näitaja lõikes.

Tulemused näitasid, et Raveni testi keskmine tulemus oli kolme valimi koolide administratiivse asukohaga oluliselt seotud, $F(2,1264) = 18.26, p < .001$. Erinevate aastate valimite keskmiste sooritustasemete võrdlus näitas, et 2012. aastal testitud abiturientide keskmine sooritustase Raveni testis oli kõigi administratiivsete üksuste lõikes oluliselt madalam 2001. ja 2005. aasta vastavatest keskmistest. Tulemused on graafiliselt esitatud joonisel 3.



Joonis 3. Koolide keskmiste testisoorituste jaotus koolide administratiivsete asukohtade lõikes.

Testitulemuste sõltuvus erinevatest teguritest

Järgnevalt uuriti dispersioonanalüüsiga (*Anova*) sarnaselt varasematele töödele (nt Must et al., 2003), kuidas sõltub abiturientide Raveni testi sooritustase testimise aastast (2002/2005/2012), soost (neiud/noormehed), vanusest (18/19) ja kooli asukohast (suurlinn/maakonnakeskus/muu) ning nende tunnuste omavahelisest interaktsioonist.

Tulemused näitasid, et Raveni testi tulemus sõltus oluliselt testimise aastast, abiturientide soost ja kooli administratiivsest asukohast. Seejuures osutus statistiliselt oluliseks ka kaks interaktsiooni, täpsemalt interaktsioonid testimise aasta ja soo [$F(2,1231) = 6.86, p < .01$] ning testimise aasta ja kooli asukoha [$F(4,1231) = 2.50, p < .05$] vahel.

Täpsemad tulemused on esitatud tabelis 4.

Tabel 4*Testisoorituse sõltuvus soost, vanusest, testimise aastast ja kooli asukohast.*

	<i>Vabadus- astmete arv</i>	<i>F-statistik</i>	<i>Olulisusnivoo p-väärtus</i>
Sugu	1	33.83	.000***
Vanus	1	0.9	.344
Aasta	2	25.33	.000***
Asukoht	2	6.59	.001**
Aasta x Sugu	2	6.86	.001**
Aasta x Vanus	2	2.54	.079
Sugu x Vanus	1	0.47	.494
Aasta x Asukoht	4	2.5	.041*
Sugu x Asukoht	2	0.36	.699
Sugu x Asukoht	2	1.65	.192
Aasta x Sugu x Vanus	2	0.05	.955
Aasta x Sugu x Asukoht	4	2.12	.077
Aasta x Vanus x Asukoht	4	1.27	.279
Sugu x Vanus x Asukoht	2	0.02	.979
Aasta x Sugu x Vanus x Asukoht	4	0.31	.875

Märkus. * $p < .05$, ** $p < .01$, *** $p < .001$.

Arutelu

Käesolev töö otsis vastust küsimusele, kas ja kuidas on viimase tosina aasta vältel muutunud Eesti koolide abiturientide vaimse võimekuse sooritustase Raveni testiga hinnatuna eeldatava Flynnni efekti kontrollimiseks. Tulemustest ilmnes, et hüpotees Eesti abiturientide testiskooride tõusu kohta ei leidnud kinnitust - pigem vastupidi, erinevate aastate testisooritusi võrreldes ilmnes hoopis märkimisväärne tulemuste langus 0.58 standardhälbe ühiku võrra. Millega seda seletada ja kas tulemust on võimalik varasemate Eesti uuringutega võrrelda?

Tulemuste võrreldavus varasemate Eesti uuringutega

Kuigi töö sissejuhatuses tutvustatud Musta-Torki uuringu tulemustest selgus Eestis Flynnni efekti märkimisväärne esinemine 0.44 SD ühikut (Must et al., 2003), ei leitud ka nende uuringus kõikide alatestide raames üldist tõusutrendi – skooride tõusutrend ilmnes ennekõike mitteverbaalsetes alaosades, kuid langustrend ilmnes alaosades (sh 1997/2006 valimite vahel on paaris alatestis negatiivsed tulemused), mis sisaldasid aritmeetika, informeerituse ning arvutamisalaseid ülesandeid (Must et al., 2009). Seega on ilmnenu ka Eesti andmetel langustrende, kuigi vastavad alatestid ei ole Raveni testiga sisuliselt võrreldavad.

FE tuvastamine eeldab, et võrreldakse võimalikult sarnaste inimgruppide testisooritusi eri ajahetkedel (Must, 2011). Must ja kolleegid (2003) viitavad regionaalsusele kui ühele oma uuringu nõrgale kohale, sest suurema osa valimist moodustasid Tartu ja Tartumaa ning Tallinna koolid. Võiks ühe seletusena oletada, et käesoleva uuringu testisoorituse langustrendi tõid kaasa valimis olevad keskmisest nõrgema tasemega väiksemad koolid, sest lisaks suurematele linnadele ja maakonnakeskustele oli valimi esinduslikkuse tagamiseks käesolevas töös kaastatud võimalikult erinevaid koole üle Eesti. Seda võimalust kontrolliti töös analüüsides andmeid koolide riigieksamite pingerea lõikes ja administratiivset esindatust silmas pidades. Mõlema tunnuse lõikes ilmnes siiski Raveni testi soorituses 2012. aastal märkimisväärne langus võrreldes 2001. aastal kogutud samalaadse võrreldava valimiga.

Anti-Flynni efekt

Kas mujal maailmas on samuti täheldatud vastupidiselt Flynnni efektile hoopis sooritustaseme langust? Skooride tõusu pidurdumist ja langust on avastatud mujalgi ja seda on nimetatud anti-Flynni efektiks (Shayer & Ginsburg, 2009). Eelkõige on seelaadseid tulemusi raporteeritud Skandinaaviamaades, näiteks Norras ja Taanis.

Täpsemalt on Norras alates 1990.-ndaist aastaist toimunud skooride tõusu aeglustumine ja peatumine - skooride tõusu põhjusteks peeti toitumise ja arstiabi paranemist 1950.-ndaist kuni 1970 lõpuni ning peale stabiilsuse saavutamist ning inimeste eluolu paranemist skooride tõus ajas väidetavalt peatus (Sundet et al., 2004). Taanis avastati samuti peale 1990.-ndaid aastaid skooride tõusus seisak - peatumise trendi põhjuseks arvati olevat Taanis haridussüsteemi eripärana noorte 16-18 aastasel 3-aastasessse õppeprogrammi siirdumine, kuhu vastuvõtt toimub õppetöös edasijõudmise alusel (Teasdale & Owen, 2005). Mõlemad Põhjamaade uuringud, kus skooride tõusu ajas peatumine raporteeriti, tehti meessoost, täpsemalt ajateenijatest, koosneva valimi peal.

Lisaks pidurdustrendile on aga Suurbritannias registreeritud 14-aastaste teismeliste hulgas langustrend (Correspondent, 2009) ja ka 16-aastastel Hollandis (J. te Nijenhuis, 2012), mis annab kinnitust sellele, et IQ skooride langust esineb ka mujal maailmas. Suurbritannias J.R. Flynni enda poolt läbi viidud Raveni SPM uuringu põhjal 1980. aastal ning kordustestimise käigus 2008. aastal avastati 14-aastaste briti teismeliste noormeeste hulgas ca 6 IQ punkti madalamad skoorid kui 1980. aasta samavanustel ning selle põhjusteks peetakse arvutimängude eelistamist raamatulugemisele ning õpetamismeetodeid, mis õpetavad pigem valmistuma testideks kui loogiliselt mõtlema ja teadmisi seostama (Correspondent, 2009).

Flynni (J. Flynn et al., 2012) ning ka teiste uurijate (Kanaya, Scullin, et al., 2003) tähelepanek on olnud, et FE avaldub suurimal määral keskmisest madalama intelligentsustasemega kohordi hulgas (alaarenguga ning õpiraskustega lapsed) ning väikseim väga kõrge intelligentsusega kontingendi hulgas. See võib tähendada intelligentsuse ümberpaigumist või võrdsustumist. Samuti on olnud sotsiaalpoliitiline hoiak kogu arenenud Lääne maailmas ühtlustada IQ tasemeid, et seeläbi vähendada klassivahesid (Blackford, 2012). Kas saaksime järeldada siit, et Eesti ja Suurbritannia kooliõpilaste intelligentsusega on kõik kõige paremas korras, kuna nad on üldtasemelt keskmisest intelligentsemad ja sellest oli uuringutes tingitud ka suur langustrend, kuna Briti uuringus tuli langustrend esile just kõrgema intelligentsustasemega kontingendi hulgas?

Soolised erinevused langustrendis

Märkmisväärne on aga käesoleva uuringu puhul, et kui noormeeste puhul oli SPM koguskooris mõõdukas langus 0.31 SD ühikut, siis neidude puhul oli selleks lausa 0.82 SD ühikut, mis on tugev ja statistiliselt oluline efekt. Märkida tuleks ka seda, et kui võrrelda

tulemusi alatestide lõikes (vt tabel 3), saab neidude puhul kõikides alatestides ära tuua statistiliselt suure või mõõduka languse, samas kui noormeeste puhul leidis mõõdukalt olulise skoorilanguse vaid alatest D. Põhjalikum analüüs näitas (tulemuste osas raporteerimata), et 2012. aasta valimis olnud neid lahendasid Raveni testi ülesandeid oluliselt ($p < .001$) kauem võrreldes 2005. aasta neidudega ning samuti võrreldes mõlema aasta noormeestega. Kuivõrd testi lahendamise kiirus on oluliselt positiivselt seotud Raveni testi üldskoori tasemega, siis võib kaudselt oletada, et tütarlaste väga tugev testisoorituse langus 12 aastaga on vähemalt teatud määral tingitud nende kahanenud testile vastamise kiirusest ning seda tulemust võiks põhjalikumalt uurida.

Kahjuks ei ole siiski seda tulemust tütarlaste oluliselt madalama langustrendi kohta hetkel millegi samaväärsega võrrelda, kuna pidurdus- ja langustrendiga maades on uuringuraporteid ainult meessoost valimitelt kogutud tulemustelt. Samas võib olla, et uuringuid küll on tehtud, kuid raporteid ei ole veel jõutnud publitseerida.

Testide normeerimise vajalikkus

Levinud rusikareegli kohaselt on ühe intelligentsustesti normide taashindamist vajavaks perioodiks 1 põlvkond e 10-12 aastat (Kanaya, Ceci, et al., 2003). Seega võimaldab käesolev töö hinnata 2001. aastal rahvusliku Raveni SPM normeerimise (Pullmann et al., 2004) käigus kogutud testinormide ajakohastust ka tänapäeval tosin aastat hiljem. Kas saame Eesti tulemuste pinnalt järeldada, et Raveni SPM maatriksite tulemused võiksid olla ajas universaalsemad kui teised, kuna arvestades tsükli alguses ja lõpus kogutud andmete võrdlust ja tulemusi skooride tõusu ajas Flynn'i efektina siinpuhul ei toimunud?

Selle küsimuse vastus oleks käesoleva töö eesmärkides nimetatud praktilise väärtuse hindamisel oluline, kuid selleks, et teha otsust, kas Raveni SPM on normeerimistsükliks vastupidavam, kui verbaalsed IQ testid, on vaja taas koguda ja võrrelda andmeid ka teiste Eestis kasutusel olevate intelligentsustestidega sarnaselt Musta meeskonnale ning seejärel saadud tulemusi analüüsida. Käesoleva töö alusel ei ole siiski võimalik piisavalt täita seda eesmärki, et anda hinnangut koolipsühholoogidele ja karjäärinõustajatele, kas Raveni SPM normid Eestis on käesolevaks hetkeks aegunud ning vajavad uuendamist. Kuivõrd tulemustes ilmnes oodatud tõusu asemel langustrend, siis see vajab veel täiendavat uurimist enne normide leebemaks muutmist.

Töö piirangud

Kindlasti on ka käesolevas töös mitmeid kitsaskohti, mis ei võimalda tulemuste üldistamist. Näiteks ei saa täiesti välistada, et 2012. aasta testi sooritajatel oli madalam

motivatsioonitase testi sooritamise hetkel võrreldes 2001. aastal uuringus osalejatega. Kuna motivatsioon mõjutab võimekuste testide puhul testisooritust olulisel määral (Silm, 2012), siis sellestki võib tuleneda osaliselt leitud langustrend. Kuid kogunud testimise läbiviijad ei täheldanud testimise käigus märkimisväärset õpilaste motivatsiooni puudust, mis võinuks tulemusi nii mõjutada - osalejad olid valdavalt vastutulelikud, koostööaltid ning täitsid teste innukalt. Lisaks motiveeriti kõik osalejaid personaalse tagasiside võimalusega, mida aktiivselt kasutati ja mis eeldatavasti hoidis nende motivatsiooni tasemel saavutamaks sooritusel oma maksimumi. Samuti kontrolliti 2012. aasta andmeid ja ootuspärasest oluliselt madalamad testitulemused (<2SD keskmisest) jäeti valimist välja, et alamotiveeritud testitajate või muul viisil poolikuks jäänud vastamise tulemusel väga madalad skoorid ei jääks üldist keskmist kallutama. Seega võimaluste piires püüti kontrollida erinevaid põhjuseid, mida oli võimalik ennetada või vältida, kuid sellele vaatamata ilmnis tulemustes selge langustrend Raveni üldvõimekuse testi sooritusel.

Kokkuvõte

Kokkuvõttes võib töö tulemusel nentida, et 2001-2012 Raven SPM läbilõikeuuringus Eesti abiturientide peal leiti erinevalt eelnevatele Eestis tehtud uuringutele üldvõimekuse testisooritusel ilmne langustrend, kuid see tulemus ei ole varasemate Eesti õpilaste Flynn'i efekti uuringutega võrreldavad, kuna eelnevalt on kasutusel olnud erinev mõõdik, regionaalsete piirangutega valim ning ka uuringute ajaline vahe on olnud käesolevast uuringust täiesti erinev.

Dotsent Olev Must on tõdenud, et kuigi Eestis on koolilaste IQ-testi tulemused aja jooksul tõusnud, pole siiski võimalik väita, et testitavad oleksid muutunud taibukamateks ja nutikamateks. Nimelt võivad IQ-testide keskmiste tulemuste muutused ajas olla oma olemuselt tehnilised, kõrvalistest asjaoludest tingitud ning nende alusel ei ole võimalik langetada otsustusi põlvkondade intelligentsuse erinevuste kohta (Must, 2011).

Sellele vaatamata pakub käesolev töö olulist infot ja mõtteainet edaspidiseks näidates, et ligikaudu tosina aastaga on Eesti abiturientide testisooritus mittesõnalise Raveni SPM testiga märkimisväärselt kahanenud, mitte paranenud, nagu oleks Flynn'i efekti seaduspära arvestades eeldatud. Kahjuks ei suuda käesolev töö olemasolevate võimaluste piires selgitada selle testisoorituse languse võimalikke põhjuseid, kuid fikseerib hetkeseisu ning motiveerib loodetavasti veelgi enam jälgima maailmas toimuvad trende ja innustab edasist uurimistööd ka Eesti andmete näitel.

Kirjanduse loetelu

- Ang, S., Rodgers, J. L., & Wanstrom, L. (2010). The Flynn Effect within Subgroups in the U.S.: Gender, Race, Income, Education, and Urbanization Differences in the NLSY-Children Data. *Intelligence*, 38(4), 367–384.
- Binet, A., & Simon, T. (1916). Upon the necessity of establishing a scientific diagnosis of inferior states of intelligence. (L'Année Psych., 1905, pp. 163-191). *The development of intelligence in children (The Binet-Simon Scale)*. (lk 9–36). Baltimore, MD US: Williams & Wilkins Co.
- Cohen, J. (1987). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates.
- Correspondent, B. R. G., Science. (2009, veebruar 7). British teenagers have lower IQs than their counterparts did 30 years ago. *Telegraph.co.uk*.
- Flynn, J. R. (1998). IQ gains over time: Toward finding the causes. U. Neisser (Toim), *The rising curve: Long-term gains in IQ and related measures*. (lk 25–66). Washington, DC US: American Psychological Association.
- Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, 54(1), 5–20.
- Flynn, J. R., & Rossi-Casé, L. (2012). IQ gains in Argentina between 1964 and 1998. *Intelligence*, 40(2), 145–150.
- Flynn, J., Shaughnessy, M. F., & Fulgham, S. W. (2012). An interview with Jim Flynn about the Flynn effect. *North American Journal of Psychology*, 14(1), 25–38.
- Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1), 13–23.
- Herrnstein, R. J., & Murray, C. A. (1994). *The bell curve: Intelligence and class structure in American life*. New York, NY US: Free Press.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT US: Praeger Publishers/Greenwood Publishing Group.
- Kanaya, T., & Ceci, S. J. (2011). The Flynn Effect in the WISC Subtests Among School Children Tested for Special Education Services. *Journal of Psychoeducational Assessment*, 29(2), 125–136.
- Kanaya, T., Ceci, S. J., & Scullin, M. H. (2003). The rise and fall of IQ in special ed: Historical trends and their implications. *Journal of School Psychology*, 41(6), 453–465.
- Kanaya, T., Scullin, M. H., & Ceci, S. J. (2003). The Flynn Effect and U.S. Policies: The Impact of Rising IQ Scores on American Society via Mental Retardation Diagnoses. *American Psychologist*, 58(10), 778–90.
- Kikas, E., Toomela, A., & Mõttus, E. (2006). Ability grouping in schools: A study of academic achievement in five schools in Estonia. *Trames*, (1), 32–43.

- Mingroni, M. A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn effect and other trends. *Psychological Review*, *114*(3), 806–829.
- Must, O. (2011). Flynni efekt: intelligentsustestide skooride tõus ajas. R. Möttus, J. Allik, & A. Realo (Toim), *Intelligentsuse psühholoogia* (lk 144–157). Tartu: Tartu Ülikooli Kirjastus.
- Must, O., Must, A., & Raudik, V. (2003). The secular rise in IQs: In Estonia, the Flynn effect is not a Jensen effect. *Intelligence*, *31*(5), 461–471.
- Must, O., te Nijenhuis, J., Must, A., & van Vianen, A. E. M. (2009). Comparability of IQ scores over time. *Intelligence*, *37*(1), 25–33.
- Postimees. (2012). Eesti koolide pingerida 2012. Salvestatud <http://www.postimees.ee/export/riigieksamid/2012/>
- Pullmann, H., Allik, J., & Lynn, R. (2004). The growth of IQ among Estonian schoolchildren from ages 7 to 19. *Journal of Biosocial Science*, *36*(6), 735–740.
- Raven, J. (2000). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 3: The Standard Progressive Matrices*. San Antonio, TX: Harcourt Assessment.
- Resing, W. C. M., & Tunteler, E. (2007). Children Becoming More Intelligent: Can the Flynn Effect Be Generalized to Other Child Intelligence Tests? *International Journal of Testing*, *7*(2), 191–208.
- Shayer, M., & Ginsburg, D. (2009). Thirty Years on--A Large Anti-Flynn Effect? (II): 13- and 14-Year-Olds. Piagetian Tests of Formal Operations Norms 1976-2006/7. *British Journal of Educational Psychology*, *79*(3), 409–418.
- Silm, G. (2012). *Testi tulemuse seos testi täitmise motivatsiooniga Tartu Ülikooli akadeemilise testi näitel* (Seminaritöö). Tartu Ülikool.
- Sundet, J. M., Barlaug, D. G., & Torjussen, T. M. (2004). The End of the Flynn Effect? *Intelligence*, *32*(4), 349–362.
- Te Nijenhuis, J. (2012). The Flynn effect, group differences, and g loadings. *Personality and Individual Differences*.
- Te Nijenhuis, Jan, Cho, S. H., Murphy, R., & Lee, K. H. (2012). The Flynn effect in Korea: Large gains. *Personality and Individual Differences*, *53*(2), 147–151.
- Te Nijenhuis, Jan, Murphy, R., & van Eeden, R. (2011). The Flynn effect in South Africa. *Intelligence*, *39*(6), 456–467.
- Teasdale, T. W., & Owen, D. R. (2005). A long-term rise and recent decline in intelligence test performance: The Flynn Effect in reverse. *Personality and Individual Differences*, *39*(4), 837–843.
- Terman, L. M. (1921). Mental growth and the IQ. *Journal of Educational Psychology*, *12*(7), 401–407.
- Tork, J. (1940). *Eesti laste intelligents*. Koolivara.

- Trimble, A. (2011). Exploring the Flynn Effect: A Comprehensive Review of the Causal Debate. *CMC Senior Theses*.
- Williams, W. M. (1998). Are we raising smarter children today? School- and home-related influences on IQ. U. Neisser (Toim), *The rising curve: Long-term gains in IQ and related measures*. (lk 125–154). Washington, DC US: American Psychological Association.
- Woodley, M. A. (2011). Heterosis doesn't cause the Flynn effect: A critical examination of Mingroni (2007). *Psychological Review*, 118(4), 689–693.

Tänu sõnad

Käesolevaga tahaksin tänada kõiki Raven 2012 projektis osalenuid, tänu kellele see uurimus sündis:

Esmalt oma seminaritöö juhendajat Helle Pullmanni, kes on olnud minu jaoks käesoleva töö Mentor, Mõtetesuunaja ning Juhendaja suurte algustähtedega.

Samaväärt tänulik olen oma grupikaaslastele Helenale, Hellele, Merilyle, Taimile ja Tarmo Richardile, kes meie ühise Raven 2012 projekti jooksul tänu oma panusele andmete kogumisele ja sisestamisele minugi jaoks käesoleva uurimistöö võimalikuks tegid.

Iseäranis erilised tänud nendele grupikaaslastele: Helenale moraalse toe ja Raven 2012 projektigrupi kaasamise eest; Hellele järeleaitamistundide eest SPSS-iga.

Samuti soovin tänada kõiki uurimuses osalenud koole ja nende abituriente, kes olid lahkelt nõus meid vastu võtma ja testimises innukalt osalema.

Tänuavaldused Helle Pullmannile, Jüri Allikule ja nende tublidele meeskondadele, kes on Eesti laste intelligentsust juba aastaid uurinud ning Raveni SPM põhjal regulaarselt erinevates vanuserühmades andmeid kogunud. Tänu nende tööle oli minul täna dekaaditagust minevikku, millega tänapäevaseid leide võrrelda. Aitäh!

Samuti väga tänuväärt on minu jaoks olnud dots. Olev Musta ning tema tööühma tehtud töö Eestis FE uurimise vallas – sealt sain tõuke ja ajendi selle teemaga tegelemaks. Täna neid võimaluse eest nende töö kitsaskohti vältida ning loodan, et neidki innustab minu väike leid edaspidi veelgi Eesti FE uurimisse panustama – jälgima tõusu- või langustrende ka väljaspool Raveni maatrikseid ning koos analüüsima võimalikke põhjuseid.

Lõpetuseks: eriti tänulik olen oma kallitele tublidele lastele Karlile ja Laurale, kes viimase mõne kuu jooksul on oma teismeeas läbinud edukalt ellujäämisõppused iseseisvaks eluks tänu ema pingutustele ülikooli lõpetamise teel.

Käesolevaga kinnitan, et olen korrektselt viidanud kõigile oma töös kasutatud teiste autorite poolt loodud kirjalikele töödele, lausetele, mõtetele, ideedele või andmetele.

Olen nõus oma töö avaldamisega Tartu Ülikooli digitaalarhiivis DSpace.

Merle Kõrgesaar