Suomalaismiesten kognitiivisen kykyprofiilin muutokset 1988–2001.

Flynnin efektiä suomalaisessa aineistossa?

Salla Koivunen

Pro Gradu-tutkielma
Jyväskylän yliopisto
Psykologian laitos
Kevät 2007

TIIVISTELMÄ

Suomalaismiesten kognitiivisen kykyprofiilin muutokset 1988–2001. Flynnin efektiä suomalaisessa aineistossa?

Tekijä: Salla Koivunen
Ohjaaja: Professori Heikki Lyytinen
Psykologian pro-gradu -tutkielma
Kevät 2007
Jyväskylä yliopisto
29 s.

Tässä tutkimuksessa tarkasteltiin suomalaismiesten kognitiivisen kykyprofiilin muutoksia vuosina 1988–2001. Pitkittäisaineisto on Puolustusvoimien keräämä ja se sisältää tiedot varusmiesten suoriutumisesta Peruskoe 1 -testissä, joka on kognitiivisia kykyjä mittaava testi. Peruskoe on Puolustusvoimien kehittämä. Tutkimuksen kohteena on koko populaatio suomalaisista miehistä ikäluokittain, yhteensä 382 022 tutkittavaa. Varusmiespalvelus on kaikille suomalaisille miehille pakollinen, ja vuosittain noin 12 % jää pois varusmiespalveluksesta. Peruskoe 1 koostuu kolmesta osatestistä, joiden avulla voidaan laskea P1-luku, joka vastaa g-faktoria. Osatestit ovat visuaalisen päättelyn testi, matemaattisen päättelyn testi ja kielellisen päättelyn testi. Tässä tutkimuksessa tarkastellaan osatestien raakapistemäärien ja niiden kehitystä. Flynnin efektillä tarkoitetaan sellaista ilmiötä, jossa kognitiivisia kykyjä mittaavien testien pistemäärät kohoavat tasaisesti vuosittain. Flynnin efektiä on selitetty mm. ympäristöstä, koulutuksesta ja ravinnosta johtuvilla tekijöillä, sillä pistemäärien kehittyminen on liian nopeaa ollakseen periytyvää. Tämän tutkimuksen tärkein kysymys on se, onko Flynnin efektiä suomalaisessa populaatiossa. Lisäksi tarkastellaan, voiko aineistossa olevalla koulutus -muuttujan avulla selittää mahdollista muutosta. Tutkimuksen riippuvat muuttujat ovat osatestien raakapistemäärät, riippumattomia muuttujia kokeen suoritusvuosi (aika) ja koulutus.

Tutkimuksessa todettiin, että Flynnin efektiä on suomalaisessa aineistossa vuodesta 1988 vuoteen 1996, mutta tämän jälkeen Flynnin efekti päättyy ja on joidenkin osatestien osalta jopa käänteinen. Vuosina 1988–1996 testipistemäärät nousevat, mutta vuosina 1997–2001 ne laskevat osittain jopa vuotta 1988 alemmalle tasolle. Koulutus oli yhteydessä pistemääriin siten, että korkeammin koulutetut (lukion tai ylioppilastutkinnon suorittaneet) suoriutuivat paremmin kuin matalammin koulutetut (peruskoulun tai ammattikoulun suorittaneet), toisaalta myös korkeammin koulutettujen pisteet laskivat rajuimmin. Jatkossa olisi mielenkiintoista tarkastella, onko Flynnin efektin pysähtyminen suomalaisessa aineistossa pysyvää, ja tarkastella tämän aineiston ulkopuolelta laajemmin erilaisten yhteiskunnallisten tekijöiden vaikutusta pistemäärien kehitykseen. Myös koulutuksen vaikutusta tulisi tarkastella lähemmin, tässä tutkimuksessa ongelmana olivat osittain koulutustietojen puuttuminen aineistosta ennen vuotta 1993. Tutkimus ei myöskään vastaa siihen, koskeeko pistemäärien muutos kaikkia suomalaisia, tutkimuksen kohteena ovat olleet ainoastaan miehet. Tätä tutkimusta voi kuitenkin pitää tärkeänä lisänä kognitiivisten kykyjen muutosten tutkimuskentässä laajan pitkittäisaineiston vuoksi.

Avainsanat: *Flynnin efekti, kognitiivinen kykyprofiili, g-faktori.*

Sisällysluettelo

1. JOHDANTO 4

1.1. G-faktori 4
1.1.2 Spearmanin faktorianalyysi 5
1.1.3. Cattelin teoria 6
1.1.4 Flynnin efekti 6
1.1.5 Koulutuksen vaikutus kognitiiviseen kykyyn ja Flynnin efektiin 7

2. TUTKIMUS 8

2.1. Tutkimusongelma 8
2.2. Aineisto 8
2.3. Tutkimuksen toteuttaminen 9
2.3.1. Tutkittavat 9
2.3.2. Menetelmät ja muuttujat 9
2.3.3. Aineiston analysointi 11

3. TULOKSET 12

3.1. Raakapistemäärien kehitys 12
3.2. Koulutukseen hakeutuminen ja sen muutokset 14
3.3. Koulutuksen vaikutus raakapisteisiin 15

4. POHDINTA 19

4.1. Flynnin efekti suomalaismiesten aineistossa 19
4.2. Koulutus pistemuutosten selittäjänä 20
4.3. Muut selitysmallit 21
4.3.1. Tulokset Norjasta ja Tanskasta 21
4.3.2. Ympäristön muutokset 22
4.3.3. Motivaatiotekijät 23
4.4. Tutkimuksen saavutukset ja rajoitukset sekä tulevaisuuden haasteet 24

Kiitokset 25

Lähteet 26

1. JOHDANTO

Tässä tutkimuksessa tarkastellaan suomalaismiesten kognitiivisen kykyprofiilin muutosta vuosina 1988–2001. Tarkastelun kohteena on erityisesti se, miten testien pistemäärät muuttuvat suhteessa aikaan. Tutkimuksessa tarkastellaan myös, onko tässä aineistossa havaittavissa ns. Flynnin efektiä ja jos niin on, voiko aineistossa olevalla tiedolla koulutuksesta selittää havaittuja muutoksia.

Tutkimuksen aineisto on Puolustusvoimien keräämä. Se koostuu Peruskoe 1 -kykytestin tuloksista ja muista tiedoista, jotka on kerätty asevelvollisuutta suorittavista miehistä heidän varusmiespalveluksensa ensimmäisillä viikoilla. Varusmiespalveluksensa Puolustusvoimissa suorittaa suurin osa suomalaisista miehistä, noin 12 % jää vuosittain pois terveydellisistä, uskonnollisista tai vakaumuksellisista syistä (Puolustusvoimien Esikunta, puhelinkontakti 25.8.2006).

Puolustusvoimissa tarkastellaan kognitiivisten kykytestien avulla soveltuvuutta koulutukseen ja testit vaikuttavat myös asevelvollisen sijoituspäätöksiin erityisesti erikoiskoulutuksiin hakeuduttaessa. Peruskoe koostuu kolmesta osatestistä; visuaalisen päättelyn testistä, matemaattisen päättelyn testistä ja kielellisen päättelyn testistä. Peruskokeen osatestien avulla on mahdollista laskea yhteinen yleinen älykkyystekijä, ns. g-faktori. Tässä tutkimuksessa tarkastelen Peruskokeen eri osatestien raakapistemäärien kehitystä vuodesta 1988 vuoteen 2001 sekä koulutuksen ja pistemäärien suhdetta toisiinsa kyseisinä vuosina.

1.1. G-faktori

G-faktorilla tarkoitetaan yleistä älykkyyden tekijää (general intelligence). Se on tärkein psykometrinen mittari ihmisen kognitiivisia kykyjä tarkasteltaessa (Jensen, 1998). G-faktori on korkeimman asteen faktori joka voidaan eristää hierarkisessa faktorianalyysissa erilaisia kognitiivisia kykyjä mittaavista testeistä. Tämän lähestymistavan tausta-ajatuksena on faktorianalyysi, jonka alun perin kehitti Spearman (Jensen, 1998). Älykkyystesteistä voidaan yleensä eristää seitsemän faktoria faktorianalyysin avulla: kielellinen sisältö, kielellinen taito, abstrakti päättely, matemaattinen päättely, visuospaattinen päättely, muisti ja havainnon nopeus. (Jensen, 1998)

Yleisesti oletetaan, että älykkyys on jossain määrin periytyvää (Jensen, 1998; Neisser, 1998). G-faktorilla ajatellaan olevan useita biologisia korrelaatteja. Neurobiologisilla tekijöillä kuten informaation prosessoinnin nopeudella on yhteyttä g-faktoriin (Jensen 1998). Neuraalisen johtamiskyvyn ja glukoosiaineenvaihdunnan on myös ajateltu olevan yhteydessä älykkyyteen (Reed & Jensen 1992; Haier et al. 1988).

Täysin puhtaita g:n mittareita ei ole, mutta erilaiset kognitiivisia kykyjä mittaavat testit ovat eri tavoin g-latautuneita (Chabris, 1998). Eniten g-latautuneina testeinä, eli testeinä joista g-faktori voidaan eristää helpommin, pidetään mm. Ravenin Progressiivisia matriiseja, Wechslerin WAIS:a, sekä WISC:a ja Stanford-Binet'n testiä (Jensen 1998; Raven, 2000; Neisser, 1998). Yleisesti kielellistä kykyä mittaavia testejä voidaan pitää g-latautuneina (Ree & Earles 1993). Myös testit, joissa tutkittava joutuu huomioimaan ja määrittelemään asioiden välisiä suhteita, ovat g-latautuneita (Spearman, 1923). Lisäksi sellaiset testit, joissa tutkittava joutuu tekemään valintoja, päätöksiä ja olemaan ratkaisuissaan kekseliäs, ovat voimakkaasti g-latautuneita (Jensen, 1998). G-faktorin on havaittu olevan lisäksi invariantti paitsi testien kesken, myös kulttuurista toiseen (Jensen, 1998; Carroll, 1993; Reeve & Hakel 2002). Testitehtävien monimutkaisuuden kasvaessa kasvaa yleensä myös testin g-painotteisuus (Jensen, 1998). Tästä näkökulmasta yleisälykkyyttä eli g-faktoria voisi ajatella lähinnä kykynä käsitellä ja hallita menestyksekkäästi monimutkaisia asioita (Jensen, 1998).

G-faktorin hyödyllisyyttä käytännössä on pohdittu paljon. Yleisesti ollaan kaiketi sitä mieltä, että g:n ennustavuus on melko hyvä, esimerkiksi työssä tai koulutuksessa menestymisessä. (Jensen 1998; Gottfredson 1997, Thorndike 1985). Jotkut sosiaaliset ongelmat kuten syrjäytyminen, rikollisuus ja alhainen sosioekonominen status liittyvät negatiiviseen korrelaatioon g-faktorin kanssa (Jensen 1998).

1.1.2 Spearmanin faktorianalyysi

Spearman kehitti faktorianalyysin. Faktorianalyysin lähtökohta menetelmänä oli se, että älykkyystestien eri osa-alueiden tulokset korreloivat positiivisesti keskenään (Spearman, 1923; Jensen, 1998). Spearman myös määritteli ensimmäisenä g-faktorin ja termin yleisälykkyys faktorianalyysinsä avulla vuonna 1904 (Spearman, 1923). Spearmanin kahden faktorin teoriassa älykkyystestit mittaavat kahta faktoria, g-faktoria (yleistä tekijää) ja s-faktoria (erityistä tekijää) (Spearman, 1923). Yleisälykkyys on samalle yksilölle sama kaikissa älykkyyttä vaativissa

tehtävissä, ja erityiset tekijät puolestaan vaihtelevat tehtävästä toiseen ja ovat eriytyneitä kulloisellekin tehtävätyypille (Jensen, 1998). Spearmanin kuvasi g:tä henkiseksi energiaksi, jota voi käyttää erilaisiin tehtäviin(Spearman, 1923). Yksilöiden väliset erot johtuivat Spearmanin mielestä henkisen energian määrän eroista ja kyvyistä käyttää tätä energiaa hyödyksi. (Jensen 1998, Spearman, 1923).

1.1.3. Cattelin teoria

Cattell kehitti Spearmanin näkemystä. Hän tutki laajassa faktorianalyyttisessa tutkimuksessaan faktoreiden välisiä suhteita, jotka hän eristi kolmeen kerrokseen hierarkisesti (Cattell, 1971). Spearmanin tavoin Cattell on sitä mieltä, että on olemassa yleisiä kykyjä ja näiden lisäksi erityisiä kykyjä. Cattell kehitti teorian joustavasta älykkyydestä (Gf, fluid intelligence) ja kiteytyneestä älykkyydestä (Gc, chrystallized intelligence), jotka sijoittuvat ylemmille tasoille faktoreiden hierarkiassa (Cattell, 1971). Joustava älykkyys (Gf) koostuu lähinnä päättelykyvystä ja informaation prosessointikyvystä ja kiteytynyt älykkyys (Gc) puolestaan koostuu hankituista, tilanneriippuvaisista tiedoista ja taidoista (Cattell, 1971; Jensen, 1998). Cattellin joustavan älykkyyden faktori (Gf) vastaa g-faktoria eli yleistä älykkyyden faktoria. Joustavan älykkyyden (Gf) on osoitettu olevan vähemmän riippuvainen koulutuksesta kuin kiteytyneen älykkyyden (Gc) (esim. Roberts et al. 2000).

Carrollin (1993) faktorianalyyttinen malli on samantapainen kuin Cattellin teoria: Carroll katsoo älykkyyden olevan hierarkinen järjestelmä, mutta kolmessa tasossa. Ensimmäinen kerros (stratum I) sisältää kapeat erityiskyvyt, toinen kerros (stratum II) puolestaan koostuu sekä joustavasta älykkyydestä (Gf) että kiteytyneestä älykkyydestä (Gc) ja kolmas kerros (stratum III) puolestaan vastaa yleisälykkyyttä, g-faktoria. (Carroll, 1993).

1.1.4 Flynnin efekti

Flynnin efektillä tarkoitetaan älykkyysosamäärien kehitystä kohti korkeampia pistemääriä. Ilmiö on havaittavissa viime vuosikymmeniltä. Nousu on ollut keskimäärin .3 pistettä vuosittain. (Jensen 1998; Flynn, 1984, 1987, 1999). Nousua on havaittu erityisesti joustavaa älykkyyttä (Gf) mittaavissa testeissä, erityisesti kulttuurista riippumattomissa testeissä, kuten Ravenin progressiivisissa matriiseissa tai sen kaltaisissa kognitiivisten kykyjen testeissä (Flynn, 1987,

1998b). Ravenin progressiiviset matriisit, ja tämän testin kaltaiset visuaalisen päättelyn testit, on myös todettu hyvin vähän riippuvaisiksi kulttuurin lisäksi koulutuksesta (Raven 2000). Flynnin efekti on ollut havaittavissa useissa valtioissa erityisesti länsimaissa (Flynn, 1984, 1987, 1999, 2000), mutta on julkaistu tutkimuksia myös kehittyvistä maista (esim. Daley et al., 2003).

Vaikka kognitiivisten kykyjen voidaan olettaa olevan periytyviä, on Flynnin efektissä tapahtuva muutos liian nopeaa ollakseen yksinomaan perinnöllisten tekijöiden aiheuttamaa (Flynn, 1987, 1998b; Jensen, 1998). Erilaisia ehdotuksia, kuten koulutusta, ravitsemuksen muutosta, yhteiskunnan rakennemuutoksia ja ärsykkeiden lisääntymistä ympäristössä on tarjottu selittäjäksi pistemäärien nousulle (Chabris 1998; Neisser, 1998). Flynn on itse pyrkinyt selittämään muutosta perinnöllisen ja ympäristöllisen tekijän yhteisvaikutuksella siten, että tekijät vaikuttavat toisiaan vahvistaen. Ihminen on muokannut omaa ympäristöään omille kognitiivisille kyvyilleen sopivammaksi ja kehittävämmäksi jo olemassa olleiden kognitiivisten lähtökohtiensa perusteella (Dickens & Flynn, 2001). Ympäristön asettamat vaatimukset eivät ole ylittäneet yksilön kognitiivisia kykyjä, vaan ympäristö ja kyvyt ovat vastanneet toisiaan (Dickens & Flynn, 2001).

1.1.5 Koulutuksen vaikutus kognitiiviseen kykyyn ja Flynnin efektiin

Tiedetään, että g:n ennustavuus on hyvä sille, miten yksilö pärjää koulutuksessa tai työelämässä (Thorndike, 1985; Jensen, 1998). Vähemmän on kuitenkin tutkimuksia siitä, miten koulutus vaikuttaa älykkyystestien pistemääriin ja voiko koulutuksella selittää Flynnin efektiä. On kuitenkin todettu, että älykkyydellä ja koulutuksella on kaksisuuntainen yhteys suhteessa toisiinsa (Cesi & Williams, 1997). Lisäksi on todettu älykkyydellä ja koulutuksen määrällä olevan yhteys siten, että mitä korkeammin yksilö on kouluttautunut, sitä paremmin hän on suoriutunut kognitiivisia kykyjä mittaavissa testeissä (Cesi, 1991).

Koulutusjärjestelmän muutokset ja Flynnin efekti on myös havaittu ajoittuneen samalle ajanjaksolle, eli laadulliset ja määrälliset muutokset yhdessä voisivat selittää Flynnin efektiä osaltaan (Meisenberg et al., 2006). Myös esikouluun osallistumisen on havaittu selittävän älykkyystestien pistemäärien nousua ja Flynnin efektiä (Daley et al., 2003).

Toisaalta on myös raportoitu päinvastaisiakin tuloksia. Kokeellisesti on yritetty tutkia sitä, miten interventioiden avulla on pystytty nostamaan älykkyystestien tulosten pistemääriä. Nousua on

havaittu, mutta se ei kuitenkaan ole ollut pysyvää tai kestävää. On myös väitetty, että ainoa koulutuksen tuoma vaikutus kognitiivisissa kykytesteissä pärjäämiseen on harjoitteluvaikutus eli tottuminen erilaisten testien tekemiseen kouluaikana. (Jensen 1998)

2. TUTKIMUS

2.1. Tutkimusongelma

Tämän tutkimuksen tarkoituksena on tarkastella, miten Puolustusvoimien Peruskokeen tehneet suomalaismiehet ovat muuttuneet kognitiiviselta kykyprofiililtaan vuosina 1988 - 2001. Tutkimuksen hypoteesina on, että kykytestien raakapistemäärät ovat nousevia suhteessa aikaan Flynnin efektin mukaisesti. Tarkastelun kohteena on myös, miten koulutus voisi selittää mahdollisesti havaittavaa muutosta. Hypoteesina on, että korkeammin kouluttautuneet suoriutuvat kykytesteissä paremmin kuin vähemmän kouluttautuneet.

2.2. Aineisto

Aineistona tutkimuksessa oli Puolustusvoimien keräämä peruskokeista ja muista tiedoista koostuva pitkittäisaineisto, jossa havaintoyksiköitä oli 436 508 kpl. Peruskokeita on tehty Puolustusvoimissa vuodesta 1955 lähtien. Peruskoe 1 on viimeksi muuttunut vuonna 1983, joten kokeet ovat säilyneet täysin muuttumattomina tässä tutkimuksessa tarkastelussa olevan ajankohdan ajan.

Peruskoe on kykytesti, jota käytetään Puolustusvoimissa selvittämään alokkaan kykyprofiilia ja siten soveltuvuutta jatkokoulutukseen, kuten aliupseeri- tai reserviupseerikoulutukseen. Aineiston voidaan katsoa edustavan koko populaatiota. Asepalvelus on suomessa pakollinen kaikille miehille, ja vain noin 12 % vuosittain jää pois palveluksesta joko terveydellisten, uskonnollisten tai eettisten syiden vuoksi (Puolustusvoimien Esikunta, puhelinkontakti 25.8.2006). Palvelukseen astuneita ja peruskokeeseen osallistuneita varusmiehiä on tutkimuksessa keskimäärin 25 000 kpl vuosittain. Peruskokeiden tulokset löytyivät suurimmalta osalta tutkittavia (N = 384 022), puuttuvia tietoja oli noin 13 prosentilla populaatiosta.

Peruskoe 1 sisältää on kolme osatestiä, visuaalisen päättelyn testin, kielellisen päättelyn testin ja matemaattisen päättelyn testin. Testit on muokattu yksinomaan Puolustusvoimien käyttöön Puolustusvoimien Koulutuksen Kehittämiskeskuksessa. Kuviotestistä, sanatestistä ja numerotestistä on kustakin joka vuosi suoritettu standardointi, ja laskettu yhteinen p-arvo, jonka voidaan katsoa vastaavan yleistä kognitiivisen kyvykkyyden profiilia, eli ns. G-faktoria (henkilökohtainen kontakti, Puolustusvoimien koulutuksen kehittämisyksikkö). Peruskokeen avulla ennustetaan kognitiivista suoriutumiskykyä.

P1-luku on kolmen osatestin normeerattujen pistemäärien summan normeerattu arvo, jonka keskiarvo on 5 ja jonka arvot vaihtelevat välillä 1 ja 9 (Peruskoe 1, henkilökohtainen kontakti, Puolustusvoimien Koulutuksen kehittämisyksikkö). P1-luku perustuu g-faktoriteoriaan, ja Puolustusvoimat on valinnut nämä kolme perusominaisuutta omiin tarpeisiinsa vastaavaksi (henkilökohtainen kontakti, Puolustusvoimien Koulutuksen Kehittämisyksikkö). Erityisesti kuviopäättelytestiä voidaan pitää vahvasti g-faktorilatautuneeksi (Raven 2000). Peruskoe 1 on pääsääntöisesti tehty heti asepalveluksen alettua pääesikunnan yhtenäisten määräysten mukaisesti kussakin asepalveluksen suorittamisvaruskunnassa ja se on pakollinen kaikille alokkaille. Kokeentekopäiviä on vuosina 1988–1996 aineistossa enintään viisi vuodessa, vuodesta 1997 eteenpäin kokeentekopäiviä aineistossa oli kuitenkin jopa yli sata vuosittain.

2.3. Tutkimuksen toteuttaminen

2.3.1. Tutkittavat

Tutkittavia oli yhteensä 384 022 kappaletta. Tutkittavat olivat aloittaneet asepalveluksen ja siten osallistuneet Puolustusvoimien peruskokeisiin vuosina 1988–2001. Tutkittavat olivat tutkimushetkellä keskimäärin 19–20 -vuotiaita suomalaisia miehiä, jotka olivat syntyneet vuosina 1968–1982.

2.3.2. Menetelmät ja muuttujat

Kaikkien osatestien maksimipistemäärä oli 40. Tehtävät ovat monivalintatehtäviä alkaen helposta ja päättyen vaikeaan. Kolmesta testiosiosta Puolustusvoimat laskee standardoinnin avulla ns. P1-luvun, jonka voidaan katsoa vastaavan g-faktoria, eli yleistä kykytekijää. Tässä tutkimuksessa

keskityn eri testien raakapisteiden tarkasteluun. Tavoitteena on tehdä havaintoja pisteiden kehityssuunnasta suhteessa aikaan ja standardoimattomien pisteiden avulla pienetkin muutokset tulevat paremmin esiin.

Kielellisen päättelyn testi (sanatesti) koostuu neljästä eri kysymystyypistä: Tutkittavan tulee valita annetulle sanalle synonyymi tai antonyymi, valita samaan kategoriaan kuuluvat sanat, löytää kategoriaan kuulumaton sana ja keksiä sanapareille oikeat suhteet. Kielellisen päättelyn testi edustaa kiteytynyttä älykkyyttä (Gc) mittaavia testejä.

Matemaattisen päättelyn testi (numerotesti) muodostuu neljästä eri ongelmatyypistä: Tutkittavan tulee järjestää numeroita tiettyyn järjestykseen, ratkaista sanallisesti esitettyjä matemaattisia ongelmia, ratkaista yksinkertaisia laskutehtäviä ja keksiä lukuparien välille oikeat suhteet. Matemaattisen päättelyn testi edustaa kiteytynyttä älykkyyttä (Gc) mittaavia testejä.

Visuaalisen päättelyn testi (kuviotesti) muodostuu matriiseista, joissa kuvista on poistettu pala, ja tutkittavan tulee päätellä, mikä annetuista vaihtoehdoista sopii puuttuvan palan paikalle. Tämä testi on Puolustusvoimissa Ravenin progressiivisen matriisit -testi. Ravenin progressiiviset matriisit edustaa joustavaa älykkyyttä (Gf) mittaavia testejä.

Tarkastelun kohteena olivat siis peruskokeiden osioiden raakapisteet ja niissä tapahtuva muutos suhteessa aikaan. Aineiston perusteella oli mahdollista tarkastella pistemäärien suhteita kokeen suorittamisen ajankohtaan ja tutkittavan koulutukseen. Koulutus on aineistossa luokiteltu neljään luokkaan: peruskoulu, ammattikoulu, lukio ja ylioppilas. Ongelmana koulutusmuuttujien valinnassa oli se, että tiedot puuttuivat vuosilta 1988–1992. Koulutustiedot oli tallennettu järjestelmällisesti vasta vuodesta 1993 lähtien, joten tarkasteluja koulutuksen suhteen saattoi tehdä vain 52 prosentille populaatiosta (N = 202 715).

Tutkimuksen tärkeimpiä riippumattomia muuttujia olivat siis kokeen suoritusvuosi eli aika ja koulusivistys. Koulutus oli selittävänä muuttujana. Tarkasteltiin siis erityisesti, onko koulutukseen hakeutumisessa tapahtunut muutoksia, ja miten pistemäärät kehittyvät eri koulutuksen saaneilla. Muutokset koulutukseen hakeutumisessa voisivat selittää mahdollisia oletettuja muutoksia pistemäärissä. Riippuvana muuttujana olivat eri osatestien raakapisteet.

2.3.3. Aineiston analysointi

Aineiston analysoinnissa käytettiin varianssianalyysiä (ANOVA), t-testiä ja lineaarista mallia. Kaikki tilastolliset analyysit suoritettiin SPSS 11 for Mac OS X -ohjelmistolla ja SPSS 13.1 for Windows -ohjelmistolla.

Varianssianalyysissä tarkasteltiin ensin kokeen suorittamisvuoden vaikutusta raakapisteisiin kunkin testin osalta erikseen, jotta saataisiin selville, miten suorittamisvuosi vaikuttaa kokeen tulokseen. Varianssianalyysin avulla oli mahdollista tarkastella myös onko suorittamisvuodella ja koulusivistyksellä yhteyttä raakapisteisiin kunkin testin osalta. Tässä nähtäisiin mahdolliset omavaikutukset ja yhteisvaikutukset raakapisteisiin muuttujien osalta.

Suoritettiin varianssianalyysi siten, että riippuvina muuttujina olivat peruskokeen osatestien raakapisteet ja riippumattomana muuttujana eli selittävinä muuttujina peruskokeen suoritusvuosi. Koulutus oli myös selittävänä muuttujana niiltä osin kuin se löytyi eli vuodesta 1993 eteenpäin ja riippuvana muuttujana tällöinkin peruskokeiden osatestien raakapisteet.

Koulusivistyksen vaikutusta raakapistemääriin tarkasteltiin varianssianalyysin ja lineaarisen mallin avulla. Lisäksi tutkittiin t-testin avulla, onko varusmiesten tuloksissa eroja riippuen koulutuksesta. Koulusivistys jaettiin kahteen luokkaan siten, että peruskoulu ja ammattikoulu muodostivat ryhmän “alempi koulutus” ja lukion käyneet ja ylioppilaaksi kirjoittaneet muodostivat ryhmän “korkeampi koulutus”.

Lisäksi tehtiin erilliset analyysit vuosien 1997–2001 kokeentekopäiville. Oletuksena oli, että sellaisina päivinä, jolloin kokeen oli tehnyt suuri ryhmä, poikkeaisi tuloksiltaan niistä tutkittavista, jotka olivat tehneet kokeen eri aikaan kuin suurin osa samana vuonna asepalveluksen aloittaneista ja samassa erässä aloittaneista. Tämä siitä syystä, että kokeentekopäivien lukumäärä muuttui vuodesta 1996 eteenpäin. Aikaisempina vuosina kokeentekopäiviä oli muutamia vuodessa, ja 1996 vuodesta eteenpäin niitä oli parhaimmillaan yli sata. Oletus oli se, että kokeen tekotilanne saattaisi olla erilainen. Tarkoituksena oli kontrolloida testiolosuhteiden mahdollista vaihtelua ja vaikutusta suoriutumiseen.

Jaettiin kokeentekopäivät kahteen ryhmään. Ensimmäisessä ryhmässä oli massatestaukseen osallistuneet alokkaat, ja toisessa ryhmässä massatestausten ulkopuolella kokeen tehneet.

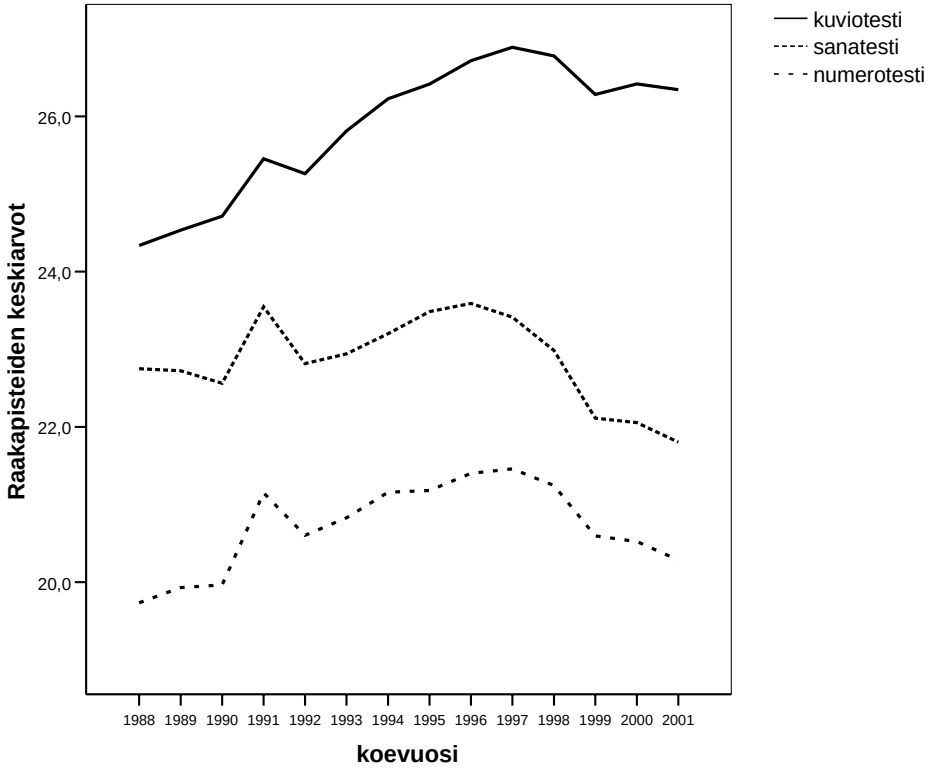
Massatestauksen rajaksi asetettiin 5000 henkilöä. Ryhmien eroa tutkittiin t-testillä suhteessa kaikkien osatestien raakapistemääriin. Vuodesta 1996 lähtien massatestauksiin oli osallistunut yhteensä 56 937 tutkittavaa ja 99 991 tutkittavaa oli tehnyt kokeen muuna ajankohtana.

3. TULOKSET

3.1. Raakapistemäärien kehitys

Kaikkien tutkittavien raakapisteiden keskiarvo *(hajonta)* oli kuviotestissä 25.832 *(sd=5.6023)* ja sana- ja numerotesteissä 22.854, *(sd=7.1071)* ja 20.691 *(sd= 7.7893)*.

Kuva 1. Testien raakapisteiden kehitys



Pistemäärien kehitys vaihteli suhteessa aikaan. ANOVA:n tulokset osoittivat, että pistemäärät nousivat ensin vuoteen 1997 asti ja sitten laskivat tilastollisesti erittäin merkitsevästi vuoteen 2001 saakka.(kuviotesti *F= 702.734, df=1/13/375010, p< .000*; sanatesti *F= 156.521, df= 1/13/375010, p< .000*; numerotesti *F= 157.189, df=1/13!375010, p< .000*).

Kuvasta 1 käy ilmi miten vuosina 1988–1996 kaikkien osatestien raakapistemäärät olivat yleisarviona tasaisessa nousussa (ks. myös taulukko 1). Kuviotestin raakapistemäärien lisäys oli keskimäärin .29 pistettä, sanatestin raakapistemäärien lisäys keskimäärin .10 pistettä ja numerotestin raakapistemäärien lisäys keskimäärin .20 pistettä vuosittain. Kuviotestin osalta siis kasvu pistemäärissä on lähes samaa suuruusluokkaa kuin on raportoitu näyttönä Flynnin efektistä (.30 pistettä vuosittain). Sanatestin ja numerotestin pistemäärien nousut eivät aivan yllä tähän samaan kasvuun vuositasolla. Vuosina 1990–1991 on kaikissa osatestien pistemäärissä yksittäinen merkittävä nousu, joka kuitenkin seuraavana vuonna palaa yleiselle pistemäärien kehitystasolle. Tämän yllättävän heilahduksen syitä ei ole tiedossa. Kehitys jatkuu kuitenkin jälleen vuodesta 1992 muiden vuosien tapaan tasaisesti.

Vuosina 1988–1996 kaikkien osatestien raakapistemäärät olivat yleisarviona tasaisessa nousussa (kuva 1, taulukko 1). Kuviotestin raakapistemäärien lisäys oli keskimäärin .29 pistettä, sanatestin raakapistemäärien lisäys keskimäärin .10 pistettä ja numerotestin raakapistemäärien lisäys keskimäärin .20 pistettä vuosittain. Kuviotestin osalta siis kasvu pistemäärissä on lähes samaa suuruusluokkaa kuin Flynnin efektissä (.30 pistettä vuosittain). Sanatestin ja numerotestin pistemäärien nousut eivät aivan yllä tähän samaan kasvuun vuositasolla. Vuosina 1990–1991 on kaikissa osatestien pistemäärissä merkittävä nousu, jonka syitä ei ole tiedossa. Kehitys jatkuu kuitenkin jälleen vuodesta 1992 muiden vuosien tapaan tasaisesti.

Nousun jälkeen pistemäärät muuttuvat siten, että vuosina 1997–2001 kaikkien testin raakapistemäärät alkavat laskea (taulukko 1). Kuviotestin raakapistemäärät laskivat keskimäärin .073 pistettä vuodessa, sanatesti .357 pistettä vuodessa ja numerotesti .223 pistettä vuodessa. Sanatestin pistemäärät laskivat siis keskimäärin nopeinta tahtia, nopeammin kuin sanatestin pisteet nousivat aikaisempina vuosina. Numerotestin pistemäärät laskivat myös enemmän kuin vuosina 1988–1996 nousivat. Kuviotestin pistemäärät sen sijaan osoittivat vuositasolla pienintä laskua muihin osatesteihin verrattuna ja vuoden 2001 pistemäärä jäi korkeammalle kuin vuoden 1988. Sanatestin pistemäärät olivat vuonna 2001 matalammat kuin vuonna 1988. Numerotestin pistemäärät olivat vuonna 2001 hieman vuotta 1988 korkeammalla. Laskeva muutos ilmenee selvänä pistemäärien nousuna tutkittavien suuren lukumäärän vuoksi. Muutokset ovat suuresta aineistosta johtuen tilastollisesti merkitseviä vuositasojen välillä, vaikka keskiarvoissa ilmenevät muutokset ovat pieniä.

Vuosina 1997–2001 sen sijaan kaikkien testin raakapistemäärät osoittivat laskusuuntaa (taulukko 1). Kuviotestin raakapistemäärät laskivat keskimäärin .073 pistettä vuodessa, sanatesti .357 pistettä vuodessa ja numerotesti .223 pistettä vuodessa. Sanatestin pistemäärät laskivat siis keskimäärin nopeinta tahtia, nopeammin kuin sanatestin pisteet nousivat aikaisempina vuosina. Numerotestin pistemäärät laskivat myös enemmän kuin vuosina 1988–1996 nousivat. Kuviotestin pistemäärät sen sijaan osoittivat vuositasolla pienintä laskua muihin osatesteihin verrattuna ja vuoden 2001 pistemäärä jäi korkeammalle kuin vuoden 1988. Sanatestin pistemäärät olivat vuonna 2001 matalammat kuin vuonna 1988. Numerotestin pistemäärät olivat vuonna 2001 hieman vuotta 1988 korkeammalla.

Taulukko 1. Raakapisteiden keskiarvot koevuosittain

Vuosi	**Kuviotesti RP *(sd)***	**Sanatesti RP *(sd)***	**Numerotesti RP *(sd)***
1988	24.337 *(5.81)*	22.750 *(7.46)*	19.734 *(8.08)*
1989	24.535 *(5.70)*	22.724 *(7.44)*	19.932 *(7.98)*
1990	24.715 *(5.68)*	22.563 *(7.32)*	19.966 *(7.95)*
1991	25.454 *(5.48)*	23.549 *(7.14)*	21.251 *(7.84)*
1992	25.262 *(5.55)*	22.816 *(7.18)*	20.605 *(7.91)*
1993	25.813 *(5.44)*	22.942 *(7.01)*	20.830 *(7.79)*
1994	26.226 *(5.38)*	23.201 *(6.98)*	21.159 *(7.79)*
1995	26.417 *(5.32)*	23.485 *(6.92)*	21.181 *(7.69)*
1996	26.718 *(5.27)*	23.591 *(6.85)*	21.402 *(7.63)*
1997	26.891 *(5.34)*	23.413 *(6.90)*	21.460 *(7.63)*
1998	26.780 *(5.38)*	22.986 *(6.85)*	21.245 *(7.54)*
1999	26.284 *(5.70)*	22.113 *(6.97)*	20.595 *(7.61)*
2000	26.419 *(5.67)*	22.055 *(6.98)*	20.525 *(7.63)*
2001	26.345 *(5.67)*	21.804 *(6.99)*	20.285 *(7.56)*

3.2. Koulutukseen hakeutuminen ja sen muutokset

Koulutukseen hakeutuminen vaihteli hieman (taulukko 2). Lukion käyneiden tai ylioppilastutkinnon suorittaneiden määrä osoitti yleisesti kasvua 1993–2001. Lukion käyneiden määrä oli nousussa 1993–1997, mutta laski hieman 1998–1999, noustakseen taas 2000–2001. Ylioppilastutkinnon

suorittaneiden määrä kasvoi vuosina 1996–1999 ja laski hieman vuosina 2000–2001. Vuosina 1993–1996 ylioppilastutkinnon suorittaneiden määrää koskevissa tiedoissa on aineistossa puutteita, joiden takia tiedoista ei ole mahdollista tehdä päätelmiä. On mahdollista, että aluksi ylioppilaat on koodattu lukion käyneiden joukkoon.

Vain peruskoulun tai ammattikoulun suorittaneiden lukumäärä laski pääsääntöisesti koko tarkasteluajan ajalla. Määrä oli laskeva vuosina 1993–1996, mutta nousi vuonna 1997. Kuitenkin jälleen vuosina 1998–2000 suunta oli laskeva, ja vuonna 2001 määrässä tapahtui pieni nousu. Ammattikoulun suorittaneiden määrä oli kasvussa vuosina 1993–1995, mutta laskussa vuosina 1996–1997. Määrä nousi taas hieman vuonna 1998 laskeakseen uudestaan vuosina 1999–2001 (taulukko 2).

Taulukko 2. Asepalveluksen alkamiseen mennessä suoritetut koulutukset

Vuosi	Peruskoulu (%)	Amm.koulu (%)	Lukio (%)	Ylioppilas (%)	Yhteensä
1993	5272 *(18.8 %)*	11920 *(42.4 %)*	10923 *(38.9 %)*	- *(- %)*	28115
1994	3381 *(15.9 %)*	9239 *(43.4 %)*	8681 *(40.7 %)*	2 *(0.0 %)*	21303
1995	2945 *(12.2 %)*	10937 *(45.5 %)*	10168 *(42.3 %)*	2 *(0.0 %)*	24052
1996	2601 *(10.9 %)*	10705 *(44.9 %)*	10518 *(44.1 %)*	4 *(0.0 %)*	23828
1997	2787 *(11.6 %)*	10157 *(42.3 %)*	10862 *(45.3 %)*	187 *(0.1 %)*	23993
1998	2459 *(11.0 %)*	9621 *(43.1 %)*	9267 *(41.5 %)*	876 *(3.9 %)*	22323
1999	2091 *(8.9 %)*	10091 *(43.2 %)*	9057 *(38.7 %)*	2140 *(9.1 %)*	23379
2000	1496 *(8.6 %)*	7258 *(41.9 %)*	7080 *(40.9 %)*	1452 *(8.4 %)*	17286
2001	1209 *(9.2 %)*	5324 *(40.8 %)*	5706 *(43.7 %)*	826 *(6.3 %)*	13065

3.3. Koulutuksen vaikutus raakapisteisiin

Pistemäärien kehityssuuntaan vaikutti paitsi aika, myös koulutus: Koulutuksella oli omavaikutusta (ANOVA: kuviotesti $F=57.009$, $df=1/3/202144$, $p< .000$; sanatesti $F= 101.378$, $df=1/3/202144$, $p< .000$; numerotesti , $F= 103.674$, $df=1/3/202144$, $p< .000$), ja sillä on merkitsevä yhdysvaikutusta testin suoritusvuoden kanssa (ANOVA: kuviotesti $F= 1.680$, $df=1/33/202144$, $p< .01$; sanatesti $F= 1.935$, $df=1/33/202144$, $p< .001$; numerotesti $F= 1.744$, $df=1/33/202144$, $p< .005$). Koulusivistys vaikutti siten, että korkeamman koulutuksen (lukio ja ylioppilastutkinto) saaneet pärjäsivät peruskokeessa paremmin kuin matalamman koulutuksen (peruskoulu ja ammattikoulu) saaneet kaikissa osatesteissä (Lineaarinen malli: kuviotesti $F= 196.662$, $df=1/3/202716$, $p< .000$;

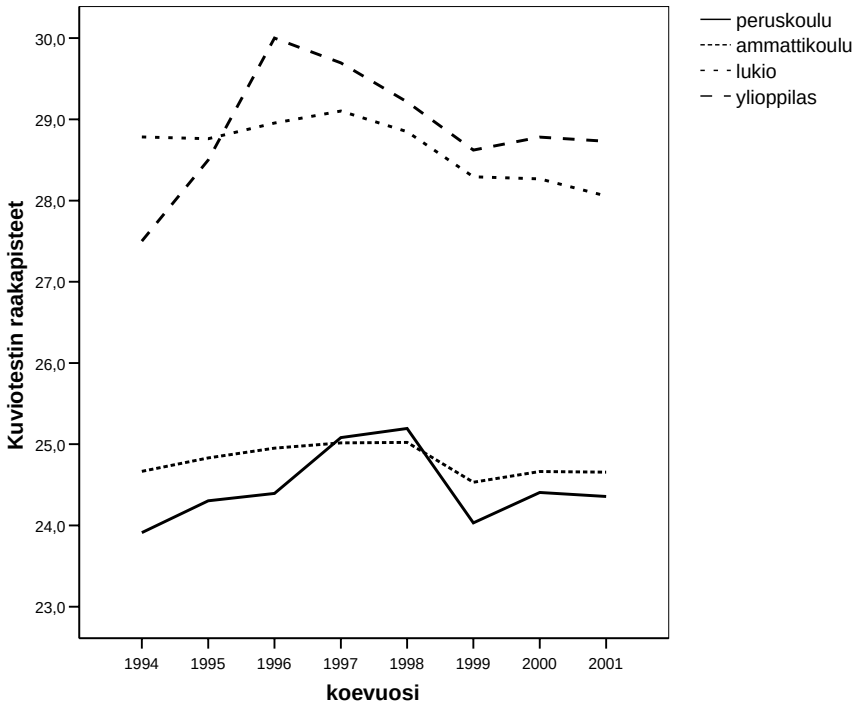
sanatesti *F= 405.389, df=1/3/202716, p< .000*; numerotesti *F= 369.759, df=1/3/202716, p< .000*). Suoriutumisessa kielellisen päättelyn testissä ja matemaattisen päättelyn testissä korkeammin ja matalammin kouluttautuneiden välillä on myös selkeä ero korkeammin kouluttautuneiden eduksi (taulukko 3). Yhdysvaikutukset ovat monimutkaisia, kuten kuvista 2-4 ilmenee.

Taulukko 3. Raakapisteiden keskiarvot koulutuksittain

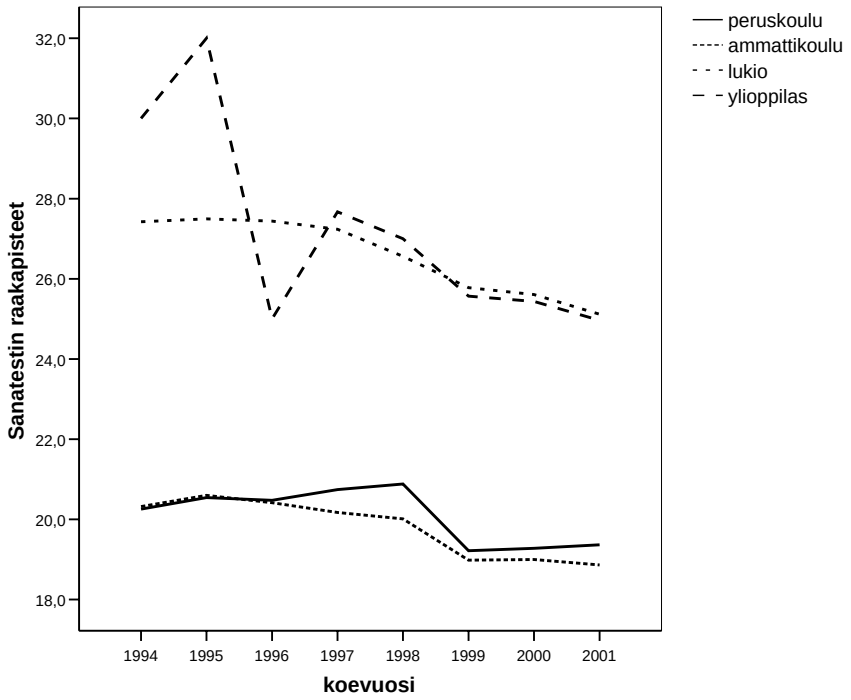
Koulutus	**Kuviotesti RP *(sd)***	**Sanatesti RP *(sd)***	**Numerotesti RP *(sd)***
Peruskoulu	24.239 *(5.84)*	20.138 *(6.92)*	17.577 *(7.77)*
Amm.koulu	24.727 *(5.29)*	19.910 *(5.96)*	17.733 *(6.71)*
Lukio	28.659 *(4.55)*	26.792 *(5.79)*	25.226 *(6.23)*
Ylioppilas	28.812 *(4.70)*	25.679 *(5.99)*	24.600 *(6.37)*

Kuvissa 2, 3 ja 4 voidaan nähdä, että suoriutuminen kaikissa kolmessa osatestissä on selkeästi parempaa korkeammin koulutetuilla. Toisaalta näyttäisi kuitenkin siltä, että korkeammin koulutettujen pistemäärät laskevat selvimmin ja aiheuttavat jopa kokonaan ns. käänteisen Flynnin efektin eli pisteet laskevat yhtä paljon kuin ovat Flynnin efektissä nousseet. Tarkemmin katsottuna peruskoulun käyneillä pistemäärät eivät laske vuosina 1996–1998 lainkaan. Ammattikoulun käyneillä sanatestin pistemäärät lähtevät loivaan laskuun vuodesta 1996, mutta muiden testien osalta laskua ei ole havaittavissa. Lukion käyneillä laskua tapahtuu kaikissa osatesteissä; sanatestin ja numerotestin pisteet laskevat vuodesta 1996, ja kuviotestin pisteet laskevat vuodesta 1997. Ylioppilailla pisteiden kehityssuunata on sama kuin lukion suorittaneilla, joskin ylioppilaiden lukumäärät vuosittain aineistossa ovat vähäiset. Eniten laskevat kuitenkin sanatestin ja numerotesti raakapisteet. (Kuva 2,3 ja 4).

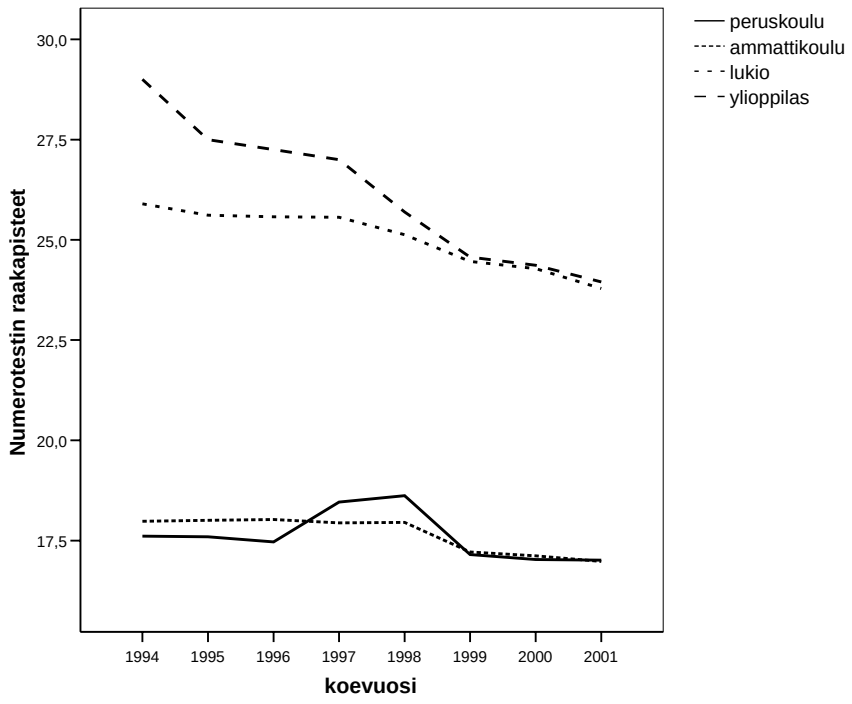
Kuva 2. Kuviotestin raakapisteet koulutuksittain



Kuva 3. Sanatestin pistemäärät koulutuksittain



Kuva 4. Numerotestin raakapisteet koulutuksittain



Massoissa peruskokeen tehneiden ja massatestausten ulkopuolella testiin osallistuneiden raakapistemäärät erosivat tilastollisesti merkitsevästi t-testin tulosten mukaan toisistaan kuviotestin *(t= 3.204/ 3.238, df=1/156926/ 122201, p < .001*) ja sanatestin (*t= -15.793/ -15.850, df=1/ 156926/ 119674, p < .000)* osalta vuodesta 1996 eteenpäin, jolloin testipäivien lukumäärissä tapahtui merkittävä muutos. Numerotestissä ryhmien välisiä eroja ei ollut havaittavissa *(t= .092/ .092, df= 1/ 156926/ 117854, p= .926)*. Paremmin peruskokeesta suoriutuivat ne tutkittavat, jotka olivat tehneet kokeen muuna kuin massatestausajankohtana.

4. POHDINTA

4.1. Flynnin efekti suomalaismiesten aineistossa

Tässä tutkimuksessa tuli esiin kaksi pisteiden kehityssuuntaa, joista ajallisesti aikaisempi vastasi asetettua hypoteesia, eli Flynnin efektin esiintymistä aineistossa. Vuosina 1988–1996 trendi oli siis kaikkien osatestien raakapisteissä nouseva, ja erityisesti visuaalisen päättelyn osatestissä (kuviotesti) tuli esiin Flynnin efektin suuruinen pistemäärien keskimääräinen nousu vuositasolla. Visuaalisen päättelyn osatesti oli Ravenin progressiiviset matriisit, joten tulokset ovat vertailukelpoiset testityypin osalta. Erityisesti Ravenin progressiivisissa matriiseissa on havaittu Flynnin efektiä, ja sen on todettu olevan vähiten riippuvainen kulttuurista ja koulutuksesta (Raven, 2000; Flynn, 1984, 1987). Kielellisen päättelyn testin (sanatesti) ja matemaattisen päättelyn testin (numerotesti) pistemäärät olivat kyseisinä vuosina myös nousevia, mutta keskimääräiset pistemäärien lisäykset vuositasolla eivät yltäneet yhtä suuriksi kuin visuaalisen päättelyn osatestissä.

Vuosina 1997–2001 sen sijaan pistemäärät laskivat rajusti erityisesti kielellisen päättelyn testin ja matemaattisen päättelyn testin osalta. Muutos oli niin suuri, että kielellisen päättelyn testin osalta vuoden 2001 pistemäärä oli alhaisempi kuin tutkimuksen lähtötasovuoden 1988. Myös numerotestin pistemäärät laskivat hyvin lähelle lähtötasovuotta. Pisteiden lasku muistuttaa kielellisen päättelyn testin osalta käänteistä Flynnin efektiä, eli pisteet laskevat saman verran kuin ovat ensin nousseet. Visuaalisen päättelyn testissä pisteiden lasku ei ollut kuitenkaan yhtä merkittävää.

Yhteenvetona voi siis todeta, että Flynnin efektiä on suomalaisessa aineistossa vuosina 1988–1996, mutta ei vuosina 1997–2001. Flynnin efektin osalta oli odotettavaa, että erityisesti visuaalisen päättelyn testi voisi tuoda odotetun kaltaisia muutoksia positiivisen kehityksen osalta, mutta mielenkiintoista on se, että kielellisen päättelyn testin pistemäärien lasku on niin suurta, että se vuositasolla yltää käänteiseen Flynnin efektiin vuodesta 1997 eteenpäin.

Flynnin efektiä on löytynyt myös muista Pohjoismaista (Flynn, 1987), ja erityisen mielenkiintoista on se, että Norjassa ja Tanskassa on havaittu samanlaista pistemäärien muutosta ensin kasvavalla trendillä Flynnin efektin mukaisesti ja tämän jälkeen käänteisesti (Teasdale & Owen, 2005; Sundet et al., 2004). Pistemäärien kasvun pysähtyminen ajoittuu norjalaisessa aineistossa 1990-luvun

puolivälin jälkeiselle ajalle, ja tanskalaisessa aineistossa 1990-luvun loppupuolelle (Teasdale & Owen, 2005; Sundet et al., 2004). Molemmat tutkimukset ovat käyttäneet hyväkseen nimenomaan asepalvelusikäisistä miehistä armeijan tarkoituksiin kerättyä aineistoa, ja erityisen hyvin vertailukelpoiseksi tutkimukset tekee se, että molemmissa testipatteristot ovat olleet hyvin samankaltaiset kuin Peruskoe 1. Erityisesti Ravenin progressiiviset matriisit ja sille samanlaiset testit ovat olleet käytössä, ja näissä on havaittu samanlaista muutosta kuin tässä tutkimuksessa suomalaisaineistosta (Teasdale & Owen, 2005; Sundet et al. 2004). Kehitys näyttäisi siis olevan samansuuntainen paitsi Norjassa ja Tanskassa, myös Suomessa.

4.2. Koulutus pistemuutosten selittäjänä

Kielellisen päättelyn testi ja matemaattisen päättelyn testi ovat todennäköisesti riippuvampia koulutuksesta kuin visuaalisen päättelyn testi, jonka on todettu olevan riippumaton kulttuurista, sosioekonomisesta statuksesta ja koulutuksesta (Raven, 2000). Tästä herää kysymys siitä, millainen vaikutus koulutuksella on pistemäärien muutoksiin. Tässä tutkimuksessa todettiin, että asepalveluksen alkamiseen mennessä suoritetuissa koulutuksissa oli jonkun verran muutoksia vuosina 1993–2001. Pääsääntöisesti pelkän peruskoulun suorittaneiden osuus tutkittavista väheni, kun taas vastaavasti ylioppilastutkinnon suorittaneiden määrä aineistossa kasvoi jossain määrin tarkastelun ajan. Periaatteessa asepalveluksen alkamiseen mennessä hankitun koulutuksen määrä on siis lisääntynyt suhteessa aikaan. Pistemäärien kasvu pysähtyi ja kääntyi osittain laskuun vuosina 1996–1997. Kyseisinä vuosina tapahtui tuolloin pientä vaihtelua, joka ei kuitenkaan ollut pysyvää, ja voisi näin ollen selittää laskevaa kehityssuuntaa pistemäärissä.

Koulutuksella ja pistemäärillä todettiin olevan yhteys siten, että korkeammin kouluttautuneet (lukio tai ylioppilas) suoriutuivat testeissä paremmin kuin matalammin kouluttautuneet (peruskoulu ja ammattikoulu). Tulos vastaa asetettua hypoteesia. Mielenkiintoista on kuitenkin se, että erityisesti korkeammin kouluttautuneiden pisteet laskevat ja aiheuttavat käänteisen Flynnin efektin koko aineistoon: Näyttäisi siltä, että matalamman koulutuksen saaneiden pistemäärät eivät keskimäärin vuositasolla ennen vuotta 1999 juurikaan laske, eikätkä voi näin ollen aiheuttaa koko aineistossa pistemäärien laskevaa kehitystä. Koulutukset jakautuivat melko tasaisesti korkeaan ja matalaan koulutukseen, ja näin ollen korkeammin koulutettujen pistemäärien lasku on todella merkittävää laskeakseen koko populaation pistemääriä.

4.3. Muut selitysmallit

4.3.1. Tulokset Norjasta ja Tanskasta

Norjassa ja Tanskassa on havaittu samaan aikaan samantyyppistä kehitystä kognitiivisia kykyjä mittaavissa testeissä kuin Suomessa nyt tässä aineistossa. Tanskalainen ja norjalainen tutkimus ovat tärkeitä siitä syystä, että Pohjoismaissa sosiaalipoliittiset, koulutukselliset, terveydenhuollolliset palvelut ovat hyvin samankaltaisia keskenään, ja tästä syystä voisi olettaa, että useimmissa Pohjoismaissa näkyvät samanlaiset kehityssuunnat olisivat samoilla keinoilla selitettävissä. Tanskalainen ja norjalainen tutkimus ovat myös tutkimusasetelmansa suhteen hyvin samankaltaisia tämän tutkimuksen kanssa.

Norjalaisessa tutkimuksessa epäiltiin, että pistemäärien kasvun pysähtyminen saattaisi johtua ns. kattoefektistä, joka tarkoittaa sitä, että standardointivaiheessa yhä suurempi osa tutkittavista saa korkeampia standardipisteitä (Sundet et al., 2004). Tässä tutkimuksessa ei kuitenkaan tarkasteltu standardipisteitä, ja näin ollen kattoefekti ei voi selittää raakapisteiden kasvun pysähtymistä. Toisaalta erityisesti Ravenin progressiiviset matriisit -testissä pitäisi olla tilaa pistemäärien kasvulle (Raven, 2000).

Tanskalaisessa tutkimuksessa havaittiin, että koulutukseen hakeutumisessa oli tapahtunut muutoksia samaan aikaan kuin pistemäärien kehitys oli pysähtynyt. Tanskalaisaineisto antoi viitteitä siitä, että korkeampaan koulutukseen hakeutumisen muutoksilla ja Flynnin efektin päättymisellä on ajallinen yhteys. (Teasdale & Owen, 2005).

Teasdale ja Owen nostavat esiin mahdollisen, joskin tutkimattoman selityksen pistemäärien laskulle. Laskua saattaisi selittää se, että yhä enenemässä määrin käytetään muita kuin perinteisiä kynä-paperi -tehtäviä ja testejä, sillä muutos on ollut enemmän tietokoneavusteisen opiskelun lisäämisen suuntaan viime vuosien aikana. He esittävät, että testaustapa siis jo itsessään voisi vaikuttaa suoriutumiseen. (Teasdale & Owen, 2005).

Norjalaistutkijat puolestaan arvelevat, että yleisestikin g-faktorin ja Flynnin efektin vaikutus on vähenemään päin erityisesti kehittyneissä maissa, sillä g-faktoriteorian taustalla olevat laskentatavat perustuvat tiettyyn aikakauteen ja tietylle yhteisölle ominaisille kykyjen kovariaatioille. Jos Flynnin

efektiä on pyritty selittämään elämäntyylin ja -tavan muutoksilla, näiden tekijöiden aiheuttaman muutokset näkyisivät nyt kehittyvissä maissa, kuten on todettukin (Daley et al., 2003). (Sundet et al., 2004).

4.3.2. Ympäristön muutokset

Flynnin efektiä on selitetty eniten ympäristössämme tapahtuvilla muutoksilla, ja ravinnon sekä terveydenhuollon muutoksilla (esim. Cocordia et al., 2003; Flynn, 2003). Lisäksi myös ihmisten vaikutukselle omaan ympäristöönsä on annettu painoarvoa selittävänä tekijänä, eli ihmiset muokkaavat ympäristöään omia kognitiivisia kykyjään vastaaviksi, ja näin ympäristö ja kognitiiviset kyvyt muokkaavat toinen toistaan, kehittäen entisestään jälleen yksilöiden kognitiivisia kykyjä (Dickens & Flynn, 2001). Kehittyneissä maissa on tapahtunut esimerkiksi viihdeteollisuuden kehityksessä kasvua puhumattakaan esimerkiksi erilaisten teknisten laitteiden kuten tietokoneiden ja erilaisten kommunikaatiovälineiden arkipäiväistymistä.

Voisiko ajatella, että jos Flynnin efektiä voisi selittää osaltaan testien tekemisestä johtuva harjoitteluvaikutus (esim. Cocordia et al. 2003), voisiko myös Flynnin efektille käänteistä kehitystä selittää testien tekemistavan vieraudella? Jos muokkaamamme ympäristö (Dickens & Flynn, 2001) nyt on hyvin erilainen verrattuna testien kehittämishetken ympäristöön, niin voisiko pelkästään kynä-paperi -tehtävien tekemisen vieraus, verrattuna esimerkiksi tietokoneella tehtyihin tehtäviin, selittää osaltaan tässäkin tutkimuksessa havaittua pistemäärien kehittymisen pysähtymistä? Toisaalta, suomalaisessa koulutusjärjestelmässä on vasta viime vuosina ryhdytty laajemmin käyttämään tietokoneavusteisia opetusmuotoja. Esimerkiksi vuonna 1997 asepalvelusta suorittamassa ollut mies olisi ollut oletettavasti vuosina 1985–1994 peruskoulussa, ja tuolloin tietokoneet tekivät vasta tuloaan.

Voisi myös arvuutella, miten tietoyhteiskunnan muutokset ovat vaikuttaneet suoriutumiseen kognitiivisia kykyjä mittaavissa testeissä. Tietokoneiden yleistyminen kotitalouksissa alkoi 1980-luvun lopulta, mutta oikeastaan 1990-luvun puoliväliin voisi katsoa yleisesti ajoittuvan hetken, jolloin tietokoneet Internetin myötä arkipäiväistyivät ja olivat kaikkien saatavilla. Tämän tutkimuksen tiimoilta ei voi ottaa kantaa siihen, onko esim. tietokonepelien yleistymisellä jotain tekemistä käänteisen Flynnin efektin kanssa. Asiaa olisi varmasti hedelmällistä pohtia tarkemmin. Jokin yhteinen muutos Pohjoismaissa luulisi olevan jollain tavalla yhteydessä samanlaisiin

pistemäärien kehityssuuntiin. Voisi kuitenkin ajatella, että visuaalispohjaisen päättelytaidon kasvaessa (tietokonepelit ja muu tietotekniikka) voisi jäädä vähemmälle kielellispohjainen tai matemaattispohjainen päättelytaito ja näiden kahden harjaannuttaminen.

4.3.3. Motivaatiotekijät

Tämän tutkimuksen aineisto on Puolustusvoimien keräämä nimenomaisesti asepalvelustaan suorittavilta miehiltä. Tämä nostaa väistämättä kysymyksen siitä, voisiko esimerkiksi asepalveluksen keston muutokset olla yhteydessä motivaatiotekijöihin alokkaiden keskuudessa? Motivaatiota ei tässä aineistossa ole mitattu millään tavoin. Voisi kuitenkin leikkiä ajatuksella, että hyvinkin kyvykäs mutta motivoitumaton alokas voisi pyrkiä antamaan itsestään älyllisesti kyvyttömämmän kuvan, jos taustalla olisi ajatus siitä, että voisi suorittaa asepalveluksensa lyhyemmässä ajassa. Tarkastelin tätä ajatusta silmällä pitäen, miten asepalveluksen kesto on muuttunut tutkimuksessa tarkasteltavan ajan aikana eli vuosina 1988–2001.

Aseellinen varusmiespalvelus on 180, 270 tai 362 vuorokauden mittainen. Erityistehtävissä palvelevat varusmiehet suorittavat asepalvelustaan pääsääntöisesti 270 tai 362 vuorokautta (Taskutietoa Puolustusvoimista 2005). Varusmies ei voi itse valita palvelusaikaansa, mutta hänen mielipidettään toki kuullaan tässä asiassa. Aliupseereiksi ja reserviupseereiksi koulutettavat palvelevat pisimmän mahdollisen ajan, ja näihin koulutuksiin valitaan erityisesti asevelvolliset, jotka ovat kognitiivisilta kyvyiltään ja muilta ominaisuuksiltaan johtajatehtäviin soveltuvia. Nykyisen, voimassaolevan asepalveluksen kestosta määräävä laki on astunut voimaan helmikuussa 1998 (HE 249/1998). Tätä ennen palvelusajat olivat 240, 285 tai 330 päivää (HE 249/1998). Palveluksen kesto on siis lyhentynyt vuoden 1997 jälkeen. Ensimmäiset alokkaat jotka pääsivät suorittamaan asepalvelustaan lyhyemmillä palvelusajoilla, astuivat palvelukseen heinäkuussa 1998 (HE 249/1998).

Voisi siis olla olemassa pieni mahdollisuus sille, että palvelusaikojen lyhentyminen olisi voinut vaikuttaa motivaatiota laskevasti vuoden 1998 aikoihin. Uuden lain voimaan astuminen merkitsi kuitenkin merkittävää muutosta asepalveluksen kestossa. Motivaatiokaan ei kuitenkaan yksinään riitä selittämään osatestien pistemäärien laskusuuntaa tässä aineistossa, ensinnäkin siitä syystä, että pisteiden kasvusuunta tyrehtyi jo vuoden 1996 jälkeen, ja toiseksi, voisi olettaa että motivaation lasku olisi hyvin tilapäistä ja tämän pitäisi näkyä nousevana trendinä pisteissä hyvin lyhyen ajan

jälkeen. Toisaalta tulee muistaa, että suoritusmotivaatiota ei tässä aineistossa ole mitattu joten syy-seuraussuhteista ei voi tehdä ehdottomia johtopäätöksiä suoritusmotivaation suhteen. Jälkikäteen motivaation tutkiminen on mahdotonta.

4.4. Tutkimuksen saavutukset ja rajoitukset sekä tulevaisuuden haasteet

Tämän tutkimuksen tulosten yleistämisen rajoitteet ovat erityisesti siinä, että aineisto koostui miehistä. Flynnin efektin voimakkuudesta eri sukupuolilla on ristiriitaisia tutkimustuloksia. Toisaalta on todettu, etteivät miesten ja naisten väliset erot ole merkittäviä (Flynn, 1998a), mutta toisaalta eestiläisessä aineistossa on havaittu eroja tyttöjen ja poikien välillä (Must et al., 2003). Tämän tutkimuksen aineistossa havaittu Flynnin efekti ja sen päättyminen on siis yleistettävissä vain suomalaisten miesten keskuuteen, ei koko populaatioon. Kuitenkin aineisto on kattava, sen voi katsoa edustavan koko populaatiota miesten osalta Suomessa vuosina 1988–2001. Vuosittainen kato on ollut keskimäärin samansuuruista ja kadon syyt ovat olleet samat ja lakiin perustuvat (Puolustusvoimien esikunta, puhelinkontakti 25.8.2006).

Aineiston ongelmat liittyvät koulutus -muuttujan osalta siihen, että tietoja ei ollut kerätty järjestelmällisesti ennen vuotta 1993. Lisäksi ylioppilaiden määrässä on todennäköisesti virhe aineistossa, ylioppilaiden lukumäärä vaikuttaa liian pieneltä. Toisaalta tämä tutkimus ei myöskään pysty aukottomasti selittämään Peruskokeen osatestien raakapistemäärien ensin nousevaa ja sitten erityisesti laskevaa kehitystä. Tutkimuksesta ja sen tuloksista nousee kuitenkin esiin useita mielenkiintoisia kysymyksiä jatkotutkimuksia ajatellen: Miten raakapistemäärät ovat kehittyneet vuodesta 2001 eteenpäin, onko laskusuunta jatkunut tai kenties ovatko pisteet kääntyneet uudelleen nousuun? Onko suomalaisen koulutusjärjestelmän sisällä tapahtunut laadullisia muutoksia (esim. ryhmien koossa) silloin, kun vuosina 1996–1998 asevelvolliset ovat olleet koulutusjärjestelmän piirissä? Voiko tietokoneiden yleistyminen ja näin ollen visuaalispohjaisen päättelyn lisääntyminen arkielämässä vaikuttaa tuloksiin? Miten aineiston ulkopuolelta mahdollisesti saatava tieto, kuten sosioekonominen asema, perheen lapsiluku ja syntymäjärjestys, voisivat selittää pistemäärien muutoksia? Näihin kysymyksiin ei tässä tutkimuksessa voida ottaa kantaa. Lisäksi olisi mielenkiintoista tutkia varusmiesten suoritusmotivaatiota testitilanteessa.

Muiden pohjoismaisten tutkimusten tavoin on siis todettavissa, että Flynnin efektiä löytyy myös suomalaisessa aineistossa, joskin se näyttää päättyvän 1990-luvun puolivälin jälkeen.

Koulutuksessa asepalveluksen alkuun mennessä tapahtuu joitain muutoksia, jotka eivät kuitenkaan aukottomasti selitä osatestien raakapistemäärien muutoksia tässä aineistossa. Tässä tutkimuksessa pystyttiin kuitenkin poissulkemaan testausympäristön ja -tilanteen vaikutus niiltä osin kuin analyysi massoissa Peruskokeen tehneiden ja ei-massoissa Peruskokeen tehneiden suoriutuminen osoittaa. Tämän tutkimuksen etu on se, että testien pistemäärätiedot löytyvät joka vuodelta, ja kehitystä on mahdollista seurata tarkemmin. Tätä tutkimusta voi pitää merkittävänä kognitiivisten kykyjen muutosten tutkimuskentässä, ja valaisee omalta osaltaan Flynnin efektin esiintymistä kehittyneissä maissa ja Pohjoismaissa, sekä tuo mielenkiintoisen lisän kysymykseen siitä, millaiset mahdolliset ympäristötekijät saattaisivat vaikuttaa kognitiiviseen kykyprofiiliin.

Kiitokset

Juhani Sinivuolle ja Asko Tolvaselle asiantuntija-avusta.

Lähteet:

Carroll, J.B. (1993). Human Cognitive Abilities: A survey of Factor-Analytical Studies. New York: Cambridge University Press.

Cattell, R.B. (1971). Abilities. Their Structure, Growth, and Action. Boston. Houghton-Mifflin.

Cesi, S.J. (1991). How much does Schooling Influence General Intelligence and its Cognitive Components? A Reassessment of the Evidence. Developmental Psychology, 27(5), 703-722.

Cesi, S.J., Williams, W.M. (1997). Schooling, intelligence and Income. American Psychologist, 52(10), 1051-1058.

Chabris, C.F. (1998). IQ Since "The Bell Curve" Commentary, August.

Cocordia, E.A., Kim, J-S., Shin, H-S., Kim, J-W., Ee, J., Wee, M.S.W., Howard, R.W. (2003). Evidence that rising population intelligence is impacting formal education. Personality and Individual Differences, 35, 797-810.

Daley, T.C., Whaley, S.E., Sigman, M.D., Espinosa, M.P., Neumann, C. (2003). IQ on the Rise: The Flynn Effect in Rural Kenyan Children. Psychological Science, 14(3), 215-219.

Dickens, W.T., Flynn, J.R. (2001). Heritability Estimates vs. Large Environmental Effects: The IQ Paradox Resolved. Psychological Review, 108(2), 346-369.

Flynn, J.R. (1984).The Mean IQ of Americans - Massive Gains 1932 to 1978. Psychological Bulletin, 95, 29-51.

Flynn, J.R. (1987). Massive IQ Gains in 14 Nations. What IQ Test Really Measure. Psychological Bulletin, 101(2), 171-191.

Flynn, J.R. (1988a). Israeli Military IQ Tests: Gender differences small. Journal of Biosocial

Science 30, 541-553.

Flynn, J.R. (1998b) IQ Gains over Time: Toward Finding the Causes. Teoksessa Rising Curve: Long-Term Gains in IQ and Related Measures. Neisser Ulric (Ed.). Washington DC, US, American Psychologist Association.

Flynn, J.R. (1999) Searching for Justice: The Discovery of IQ Gains Over Time. American Psychologist, 54 (1), 5-20.

Flynn, J.R. (2000). IQ Gains: WISC Subtest and fluid g. The Nature of Intelligence. Wiley, Chichester, Novartis Foundation Symposium 233, 202-207.

Flynn, J.R. (2003). Movies about Intelligence: The Limitations of g. Current Directions in Psychological Science, 12, 95-99.

Gottfredson, L. (1997). Why g Matters: The Complexity of Everyday Life. Intelligence, 24 (1), 79-132.

Haier, R.J., Siegel, B.V., Nuechterlein, K.H., Hazlett, E. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography. Intelligence, 12(2), 199-217.

HE 249: Hallituksen esitys eduskunnalle laiksi asevelvollisuuslain muuttamisesta annetun lain voimaantulosäännöksen 2 momentin kumoamisesta. HE 249/1998.

Jensen, A.R. (1998). The g Factor: The Science of Mental Ability. Westport (Conn.): Praeger.

Meisenberg, G., Lawless, E., Lambert, E., Newton, A. (2006). Determinants of Mental Ability on a Caribbean Island, and the Mystery of Flynn Effect. Mankind Quarterly, 46(3), 273-312.

Must, O., Must, A., Raudik, V. (2003). The Flynn Effect for gains in literacy found in Estonia is not a Jensen Effect. Personality and Individual Differences, 34(7), 1287-1292.

Neisser, U. (1998). Rising Curve: Long-Term Gains in IQ and Related Measures. Washington DC,

US, American Psychologist Association.

Peruskoe 1. (1983). Puolustusvoimien Koulutuksen Kehittämiskeskus.

Puolustusvoimien Esikunta. (25.8.2006). Puhelinkontakti: Majuri Vesa Tohkanen.

Raven, J. (2000) The Raven's progressive Matrices: Change and stability over culture and time. Cognitive Psychology, 41, 1-48.

Ree, M.J, Earles, J.A. (1993). G is to Psychology what Carbon Is for Chemistry: A Reply to Sternberg and Wagner, McClelland and Calfee. Current Directions in Psychological Science, 2(1).

Reed, T.E., Jensen, A.R. (1992). Conduction Velocity in a Brain Nerve Pathway of Normal Adults Correlates with Intelligence Level. Intelligence, 16(3-4), 259-272.

Reeve C.L. Hakel, M.D. (2002) Asking The Right Questions About g. Human Performance, 15(1/2), 47-74.

Roberts, R.D., Goff, G.N., Anjoul, F., Kyllonen, P.C., Pallier, G., Stankov, L. (2000). The Armed Services Vocational Aptitude Battery (ASVAB). Individual Differences, 1(12), 1041-6080.

Spearman C. (1923) The Nature of 'Intelligence' and The Principles of Cognition. London: Macmillan.

Sundet J.M., Barlaug, D.G., Torjussen, T.M. (2004). The end of Flynn Effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. Intelligence, 32(4), 349-362.

Taskutietoa Puolustusvoimista. (2005). http://www.mil.fi haettu 3.1.2007.

Teasdale, T.W., Owen D.R. (2000). Forty-year Secular Trends in Cognitive Abilities. Intelligence 28, 115-120.

Thorndike, R.L. (1985). The Central Role of General Ability in Prediction. Multivariate

Behavioural Research, 20, 241-254.