1143

KONFIRMATORISTEN FAKTORIMALLIEN JA FAKTOREIDEN RAKENNEYHTÄLÖMALLIEN RYHMÄ- JA TASOVERTAILUISTA:

15-vuotiaiden suomalaisnuorten aamuvireisyyden, unilaadun ja mielialan sukupuolierot ja perheen taloudellisella tilanteella selitettävät erot

Esko Levälahti

Pro gradu -tutkielma
7. lokakuuta 1998

Jyväskylän yliopisto

tilastotieteen laitos

KONFIRMATORISTEN FAKTORIMALLIEN JA FAKTOREIDEN RAKENNEYHTÄLÖMALLIEN RYHMÄ- JA TASOVERTAILUISTA:

15-vuotiaiden suomalaisnuorten aamuvireisyyden, unilaadun ja mielialan sukupuolierot ja perheen taloudellisella tilanteella selitettävät erot

Levälahti, Esko Pasi Johannes

Tilastotiede

7. lokakuuta 1998

Jyväskylän yliopisto

63 sivua ja 4 sivua liitteitä

Tutkielmassa tarkastellaan ryhmä- ja tasovertailuja konfirmatorisissa faktorimalleissa ja faktoreiden rakenneyhtälömalleissa. Siinä keskitytään tarkastelemaan usean ryhmän tapauksissa tasoparametrisia faktorimalleja ja tasoparametrisia faktoreiden rakenneyhtälömalleja sekä erityisesti tasoparametrista faktorimallia tilanteessa, jossa ryhmitteleviä tekijöitä on kaksi. Malli nimetään kaksisuuntaiseksi tasovertailumalliksi. Silloin tulee kyseeseen faktoriin liittyvän yhdysvaikutuksen tutkiminen ja suuntien päävaikutusten tutkiminen, jos yhdysvaikutusta ei ole. Tutkielman empiirisessä osassa tarkastellaan 15-vuotiaiden suomalaisnuorten aamuvireisyyden, unilaatun ja mielialan mittaamista sekä näiden faktoreiden välisiä yhteyksiä. Lisäksi selvitetään näihin faktoreihin ja niiden välisiin yhteyksiin liittyviä sukupuolieroja ja perheen taloudellisen tilanteen (lyhyesti PTT:n) mukaisia eroja. Tutkielma on tehty aikaväleillä syyskuu 1995 - marraskuu 1996 ja maaliskuu 1998 - lokakuu 1998. Se perustuu suurelta osin lähdeteokseen Leskinen, E. (1987): *Faktorianalyysi. Konfirmatoristen faktorimallien teoria ja rakentaminen.* Jyväskylän yliopiston tilastotieteen laitoksen julkaisuja 1/1987. Tutkielman teoreettisissa osassa havaitaan, että faktoriin liittyvää yhdysvaikutusta kaksisuuntaisessa tilanteessa voidaan tutkia asettamalla keskiarvoprofiilit yhdensuuntaisiksi rajoittava hypoteesi. Empiirisen osan faktorimallien perusteella voidaan sanoa, että pojilla, joiden PTT on hyvä ja huono, voidaan aamuvireisyyttä, unilaatua ja mielialaa mitata tietyillä niitä kuvaavilla muuttujilla. Polkumallit osoittavat, että mieliala selittää unilaatua ja unilaatu selittää aamuvireisyyttä. Faktoreiden keskiarvot ovat selvästi alemmat pojilla, joiden PTT on huono. Estimoitun kaksisuuntaisen tasovertailumallin perusteella havaitaan selkeät PPT:n mukaiset erot, mutta ei sukupuolieroja.

SISÄLLYSLUETTELO

1 JOHDANTO

Konfirmatorinen faktorimalli on malli, jossa faktoreiden latausrakenteeseen asetetaan etukäteen rajoituksia (Jöreskog 1969, Leskinen 1987). Nämä rajoitukset ovat käytännön faktorianalyyttisen tutkimuksen kannalta kuitenkin enemmän faktorimallin tulkintaa helpottavia kuin rajoittavia tekijöitä. Eksploratiivisessa faktorimallissa tällaisia rajoituksia ei tehdä: faktorimalli saadaan tulkinnalliseen muotoon rotatoimalla eli kertomalla faktorimallin latausrakenne (faktorilataukset sisältävä matriisi) sopivalla muunnosmatriisilla (Harman 1967, Mulaik 1972, Leskinen 1987). Muunnosmatriisi voidaan valita monella eri tavalla, joten mahdollisia faktoriratkaisujakin on useita. Siksi lopullinen faktoriratkaisu riippuu siitä, mikä rotaatiomenetelmä valitaan. Eksploratiivista faktorianalyysiä voidaan siten pitää hyvin mielivaltaisena analyysimenetelmänä. Konfirmatorisen faktorianalyysin avulla tämä ongelma on ratkaistavissa asettamalla latausrakenteeseen kohdistuvat rajoitukset siten, että latausrakennetta ei voida rotatoida rikkomatta asetettuja rajoituksia. Tällöin faktorimalli on rotaatioyksikäsitteinen.

Konfirmatoristen faktorimallien rakentamisessa pyritään yleensä malleihin, jotka ovat identifioituvia. Identifioituvan konfirmatorisen faktorimallin latausrakenne voidaan laskea yksikäsitteisesti havaittujen muuttujien kovarianssimatriisin (tai korrelaatiomatriisin) sisältämän informaation avulla (Jöreskog & Sörbom 1976, Leskinen 1987). Identifioituva faktorimalli on aina rotaatioyksikäsitteinen. Konfirmatorisen faktorimallin identifioituvuus mahdollistaa tehokkaiden tilastollisten mallien rakentamisperiaatteiden käytön konfirmatorisen faktorianalyysin teoriassa. Faktorimalli, jossa jokainen havaittu muuttuja mittaa vain yhtä faktoria, on ns. mittausmalli (Leskinen 1987). Mittausmalli on aina identifioituva.

Identifioituvan konfirmatorisen faktorimallin tapauksessa faktoreiden välille voidaan määritetään kausaalisuhteita. Tällöin puhutaan faktoreiden rakenneyhtälömalleista, joita ovat faktoreiden regressio-, polku- ja moniyhtälömallit. Polku- ja moniyhtälömallit ovat regressiomallin laajennuksia.

Identifioituvia konfirmatorisia faktorimalleja voidaan rakentaa yhtä aikaa usealle eri perusjoukolle eli ryhmälle ja suorittaa eri ryhmien faktorimalleille vertailuja (Jöreskog 1971).Tällöin kyseessä ovat ryhmävertailut. Usean ryhmän tapauksessa voidaan faktorimalli laajentaa käsittämään havaittujen muuttujien ja faktoreiden tasoparametrit sekä tasoparametreista muodostuvat odotusarvot, joille voidaan suorittaa ryhmien välisiä vertailuja. Tässä tapauksessa puhutaan faktoreiden tasovertailuista

(Jöreskog & Sörbom 1981). Ryhmävertailut ja tasovertailut voidaan edelleen laajentaa käsittämään faktoreiden rakenneyhtälömallit. Tasovertailuissa kausaalisuhdetta käytetään faktoreiden keskiarvoerojen selittämiseen.

Faktorimallit ja faktoreiden rakenneyhtälömallit laajennuksineen ovat erikoistapauksia kovarianssirakennemalleista. Tämän tutkielman teoreettiset tarkastelut perustuvat Karl Jöreskogin 1960-luvulla kehittämään kovarianssirakennemalliin LISREL (Linear Structural RELations) -malliin (Jöreskog 1967, Jöreskog 1969, Nummenmaan, Konttinen, Kuusinen & Leskinen 1997).

Luvussa 3 selvitetään LISREL-mallien teoriaa yhden ryhmän tapauksessa ensiksi konfirmatorisen faktorimallin osalta ja tämän jälkeen faktoreiden rakenneyhtälömallin osalta. Tämän jälkeen käydään läpi faktorimallien ja faktoreiden rakenneyhtälömallien rakentamisen vaiheet. Luvussa 4 tarkastellaan erikseen faktorimallin ja faktoreiden rakenneyhtälömallin ryhmävertailuja. Tarkoituksena on selvittää ryhmävertailujen yleiset periaatteet. Lukujen 3 ja 4 tarkastelut ovat pohjana luvussa 5 tarkasteltaville tasovertailuille.

Tässä tutkielmassa on tarkoitus keskittyä faktoreiden ja faktoreiden rakenneyhtälömallien tasovertailuihin. Tasoparametristen faktorimallien ja tasoparametristen faktoreiden rakenneyhtälömallien rakentaminen on tehokas tapa tutkia latenttien muuttujien keskiarvoeroja. Käytännössä faktoreiden keskiarvoeroja ei voida muilla menetelmillä yhtä tehokkaasti tutkia.

Tasovertailuja käsittelevässä luvussa 5 tarkastellaan aluksi tasoparametrista faktorimallia yleisesti ja odotusarvoparametrisoinnin identifioituvuutta. Tämän jälkeen selvitetään, kuinka tasoparametrista faktorimallia voidaan käyttää tilanteessa, jossa on mukana kaksi ryhmittelevää tekijää kuten kaksisuuntaisessa varianssianalyysissä. Tässä tapauksessa tulee erityisesti kyseeseen ryhmittelevien tekijöiden yhdysvaikutuksen tutkiminen. Luvun lopussa selvitetään, kuinka faktoreiden tasoeroja voidaan selittää tasoparametrisen faktoreiden rakenneyhtälöiden avulla.

Teoreettisia tarkasteluja havainnollistetaan käytännön tutkimuksella, jonka taustat ja aineisto esitellään luvussa 2. Aineiston pohjalta rakennetut mallit on estimoitu faktorimallien ja faktoreiden rakenneyhtälömallien sekä näiden ryhmävertailuihin ja tasovertailuiluihin sopivalla LISREL-ohjelmiston versiolla 8.12a (Jöreskog & Sörbom 1993a). Tämän rinnalla korrelaatioiden laskennassa on käytetty PRELIS-ohjelmiston versiota 2.12a (Jöreskog & Sörbom 1993b).

2 AINEISTO JA TUTKIMUSONGELMAT

Seuraavaksi esitellään tässä tutkielmassa käytettävä aineisto ja tutkimus, josta aineisto on peräisin. Tämän lisäksi esitellään tämän tutkielman empiirisessä osassa taustalla olevat tutkimusongelmat ja -hypoteesit.

WHO-koululaistutkimus on monivuotinen, kansainvälinen terveyskasvatustutkimus, jonka tarkoituksena on kuvata nuorten elämäntapaa terveyden näkökulmasta. Tässä tutkielmassa käytetään osaa WHO:n koululaistutkimuksen 1994 suomenkielisiltä 9. luokilta kerätystä kyselyaineistosta (Tynjälä & Törmäkangas 1994), jonka otoskoko on 1323 oppilasta. Aineiston tarkoituksena on lisätä tietoa nuorten sosiaalisista verkostoista, kouluympäristöstä, lepotottumuksista ja vireisyydestä.

Tarkastellaan seuraavaksi tässä tutkielmassa rakennettavissa malleissa esiintyviä havaittuja muuttujia ja indeksejä:

Aamuvireisyys -indeksi (1-3):

1 = ei ollenkaan vireä (summa15-32)
2 = kohtalaisen vireä (summa 33-42)
3 = hyvin vireä (summa 43-57)

väsymyksen useus kouluaamuisin (1-4):

1 = harvoin tai ei koskaan
2 = satunnaisesti
3 = 1-3 kertaa viikossa
4 = 4 kertaa viikossa tai useammin

Aamupirteyskokemus (1-4):

1 = harvoin tai ei koskaan
2 = silloin tällöin
3 = melko usein

Unilaatu -indeksi (1-3):

1 = huono (summa 19-37)
2 = kohtalainen (summa 38-44)
3 = hyvä (summa 45-48)
4 = usein tai joka aamu

Univaikeuksien ongelmallisuusaste, indeksi (1-3):

1 = pieni (summa 19-37) 45-48
2 = keskimääräinen (summa 38-44)
3 = suuri (summa 45-48)

Masentuneisuus, ärtyneisyys, hermostuneisuus (jokainen 1-3):

1 = harvemmin kuin kerran kuukaudessa tai ei koskaan
2 = kerran kuukaudessa-kerran viikossa
3 = useammin kuin kerran viikossa - lähes päivittäin

Edellä esitetyistä muuttujien ja indeksien luokitteluista huomataan, että aamuvireisyys -indeksi, unilaatu -indeksi ja univaikeuksien ongelmallisuusastetta kuvaava indeksi on uudelleenkoodattu 3-4 -luokkaisiksi. Tämä siksi, että keskiarvotarkasteluissa käytettävät tilastolliset testit toimivat paremmin, jos kaikki muuttujat samassa mallissa ovat asteikoltaan samanlaisia.

Aineiston muuttujista rakennettujen indeksien käyttö voidaan perustella aikaisempaan tutkimukseen perustuen. Aamuvireisyys -indeksinä käytetään ns. Smithin aamuvireisyysmittaria (Smith, Reilly & Midkif , 1989), joka on kehitetty kuvaamaan ihmisten vuorokausirytmiä ja siihen liittyviä eroja. Unen laatua kuvaava indeksi koostuu osittain samoista ja osittain muunnetusti osioita, joita on unilaatu -indeksissä, jota on käytetty tutkimuksissa Minors & Waterhouse (1987), Åkerstedt & Torsvall (1989) ja Åkerstedt, Hume, Minors & Waterhouse (1994). Vastaavalla tavalla univaikeuksien ongelmallisuutta kuvataan indeksillä, joka on muodostettu muuttamalla enemmän osioita sisältävä mittari (Cook & Burd, 1990) vähempiosioiseksi.

Tarkastellaan seuraavaksi edellä esiteltyyn aineistoon liittyviä tutkimusongelmia:

1. Minkälaisilla faktoreilla voidaan mitata 15-vuotiaiden nuorteen aamuvireisyyttä, unilaatua ja mielialaa?
2. Minkälaisilla rakenneyhtälöillä voidaan kuvata 15-vuotiaiden nuorteen aamuvireisyyden, unilaadun ja mielialan välisiä yhteyksiä?
3. Minkälaisia ovat 15-vuotiaiden nuorten aamuvireisyyteen, unilaatuun ja mielialaan ja niiden välisiin yhteyksiin liittyvät sukupuolierot ja perheen taloudellisen tilanteen mukanaan tuomat erot?

Edellä esitettyjen tutkimusongelmien 1 ja 2 perusteella voidaan muodostaa seuraavat tutkimus-hypoteesit:

1. Aamuvireisyys -indeksi, väsymyksen useus kouluaamuisin, ja aamupirteyskokemus mittaavat aamuvireisyys -faktoria.
2. Unilaatu -indeksi ja univaikeuksien ongelmallisuusaste -indeksi mittaavat unilaatu -faktoria.
3. Masentuneisuus, ärtyneisyys ja hermostuneisuus mittaavat mieliala -faktoria.
4. Unilaatu selittää aamuvireisyyttä, mieliala selittää unilaatua, ja siten mieliala vaikuttaa aamuvireisyyteen unilaadun kautta.

Tutkimusongelmaan 3 liittyvät hypoteesit esitellään kappaleissa 4 ja 5.

Tutkimusongelman 3 perusteella huomataan, että aineistossa kaksi ryhmittelevää tekijää: sukupuoli ja oppilaan perheen taloudellinen tilanne (lyhyesti PTT), tarkemmin sanoen oppilaan oma käsitys siitä. Tarkastelu kohteena on pääosin kaksi osaryhmää: pojat, joiden mielestä oman perheen taloudellinen tilanne on hyvä ja pojat, joiden mielestä se on huono. Poikkeuksena on kappale 5.3, jossa tarkastellaan aamuvireisyys -faktoriin liittyviä tasoeroja. Tarkasteluissa on mukana myös tyttöjen taloudellisesti hyvin toimeen tulevien ja huonosti toimeen tulevien perheiden tytöt.

Muuttujien kovarianssimatriisit ja keskiarvot on esitetty liitteessä 1 erikseen taloudellisesti hyvin ja huonosti toimeen tulevien perheiden poikien osaryhmissä. Aamuvireisyys -indeksin, väsymyksen useus kouluaamuisin -muuttujan ja aamupirteyskokemus -muuttujan osalta kovarianssit ja keskiarvot on esitelty tyttöjen vastaavissa osaryhmissä liitteessä 2.

Tässä tutkielmassa esitettäviä tuloksia on aiemmin esitetty tutkimushypoteesien 1 - 4 osalta tutkimuksessa Tynjälä, Kannas ja Levälahti (1995) erikseen tytöillä ja pojilla.

3 KONFIRMATORINEN FAKTORIMALLI

Faktorianalyysin perusmalli (Jöreskog & Sörbom 1979) on muotoa

$$\mathbf{y} = \Lambda\eta + \varepsilon\,, \tag{3.1}$$

jossa p x 1 -vektori **y** sisältää havaitut muuttujat. Vektori η sisältää faktorit, joita on k kappaletta (k < p) eli η on k x 1 -vektori. Λ on p x k -latausmatriisi. Faktorimallin jäännösvektori ε (p x 1 -vektori) sisältää vektorin **y** vaihtelusta sen osan, jota faktorimalli ei selitä.

Faktorianalyysin perusmallissa (3.1) oletetaan, että faktoreiden ja jäännösten odotusarvot ovat nollia eli $E(\eta) = 0$ ja $E(\varepsilon) = 0$. Faktorit ja jäännökset oletetaan korreloimattomiksi eli $E(\eta\varepsilon^T) = E(\varepsilon\eta^T) = 0$. Faktoreiden kovarianssimatriisia merkitään Ω:llä eli $\mathrm{cov}(\eta) = E(\eta\eta^T) = \Omega$ on k x k -matriisi. Jäännösten (ε) jäännöskovarianssimatriisia merkitään Θ:lla. Θ on p x p -matriisi, joka yleensä oletetaan diagonaaliseksi.

Malli (3.1) on konfirmatorinen faktorimalli, mikäli latausmatriisissa Λ on rajoituksia. Se on identifioituva konfirmatorinen faktori, jos rajoitukset Λ:ssa valitaan sopivalla tavalla. Rajoituksia voidaan tarvittaessa tehdä myös faktoreiden kovarianssimatriisiin Ω.

Konfirmatorisen faktorimallin tapauksessa voidaan tarvittaessa luopua jäännöskovarianssimatriisin diagonaalisuusoletuksesta ja antaa osan ε-jäännöksistä korreloida. Vaihtoehto jäännösten kovarianssivapautuksille on ns. spesififaktoreita sisältävän mallin rakentaminen (Leskinen 1989). Jotta tällainen malli olisi tällöin identifioituva ja sisällöllisesti tulkinnallinen, mikään spesififaktori ei saa korreloida muiden faktoreiden ja spesififaktoreiden kanssa eikä myöskään jäännösten kanssa.

3.1 Konfirmatorisen faktorimallien rakentamisen vaiheet

Tarkastellaan seuraavaksi konfirmatorisen faktorimallin rakentamisen vaiheita. Tarkastelu perustuu pääosin julkaisuun Leskinen (1987): Faktorianalyysi. Konfirmatoristen faktorimallien teoria ja rakentaminen (sivut 66 - 143).

Konfirmatorisen faktorimallin rakentamisessa järjestyksessä läpikäytävät viisi päävaihetta ovat

1) mallin spesifiointi

2) mallin identifioituvuustarkastelut

3) mallin parametrien estimointi

4) mallia koskevien hypoteesien testaus

5) mallin riittävyystarkastelut .

Mikäli malli ei jossakin vaiheessa täytä sille asetettuja vaatimuksia, joudutaan palaamaan takaisin johonkin sitä edeltäneeseen vaiheeseen.

Mallin spesifiointivaiheessa täsmennetään tutkimushypoteesien perusteella mallin muoto sellaiseksi, kuin se halutaan estimoida. Identifioituvuustarkastelujen tarkoituksena on tarkistaa, voidaan spesifioitu malli estimoida siten, että tulokseksi saadaan kullekin tuntemattomalle parametrille yksikäsitteinen ratkaisu. Kun mallin spesifiointi ja identifioituvuustarkastelut on suoritettu huolellisesti, on mallin estimointi yleensä helppo toteuttaa. Mallin estimoinnin jälkeen testataan, selittääkö estimoitu malli tutkittavia ilmiöitä. Tässä vaiheessa, mallin riittävyystarkasteluissa, tutkitaan miltä osin malli sopii havaintoaineistoon. Mallin riittämättömiä kohtia voidaan modifioida havaintoaineiston kanssa paremmin yhteen sopiviksi. Mallin rakentamisen eri vaiheita tarkastellaan tarkemmin seuraavissa kappaleissa.

3.1.1 Mallin spesifiointi

Konfirmatorisen faktorimallin rakentamisen ensimmäisessä vaiheessa spesifioidaan eli täsmennetään mallin muoto. Mallin spesifiointi tapahtuu erityisesti sisällöllisiin tutkimushypoteeseihin perustuen. Toisaalta malli pyritään spesifioimaan siten, että se olisi identifioituva. Jokaiselle parametrille parametrimatriiseissa Λ, Ω ja Θ valitaan jokin seuraavista vaihtoehdoista: parametri estimoidaan vapaasti, parametri estimoidaan jonkun toisen parametrin kanssa yhtä suurena tai parametri kiinnitetään vakioksi. Latausmatriisin Λ spesifiointi riippuu havaittujen muuttujien ja faktoreiden välisistä yhteyksistä, kovarianssimatriisin Ω spesifiointi riippuu faktoreiden välisistä yhteyksistä ja kovarianssimatriisin Θ spesifiointi riippuu jäännösvariansseista ja niiden välisistä yhteyksistä.

Jokainen tietyllä tavalla spesifioitu faktorimalli tuottaa havaituille muuttujille tietynlaisen teoreettisen kovarianssirakenteen Σ. Havaittujen muuttujien teoreettinen kovarianssimatriisi on faktorianalyysin perusmallin (3.1) perusteella muotoa

$$\Sigma = \mathrm{E}\,\mathbf{y}\mathbf{y}^{\mathrm{T}} = \Lambda\Omega\Lambda^{\mathrm{T}} + \Theta\,, \tag{3.2}$$

koska mallissa oletetaan faktorit ja jäännökset keskenään korreloimattomiksi ja koska jäännösten ja faktoreiden odotusarvot ovat nollia.

Konfirmatorinen faktorimalli pyritään spesifioimaan siten, että faktoreiden mitta-asteikko eli latausmatriisi Λ on yksikäsitteinen. Faktoreiden latausrakenne saadaan etumerkkiä vaille yksikäsitteiseksi kiinnittämällä faktoreiden varianssit vakioiksi, yleensä ykkösiksi, jolloin faktoreiden yhteyksiä kuvaava matriisi Ω on korrelaatiomatriisi. Faktorimallin rakentamisen ja sisällöllisen tulkinnan kannalta ei ole oleellista se, onko latausrakenne etumerkin suhteen yksikäsitteinen vai ei. Yksikäsitteisyys saadaan myös siten, että kunkin faktorin yksi lataus, yleensä ensimmäinen lataus, kiinnitetään vakioksi, yleensä ykköseksi. Tällöin latausrakenne on myös etumerkin suhteen yksikäsitteinen.

Faktoreiden mitta-asteikon yksikäsitteisyys ei takaa mallin identifioituvuutta. Tämän vuoksi kannattaa yleensä spesifioida ns. mittausmalleja , joissa latausmatriisin Λ kullakin rivillä on vain yksi nollasta poikkeava lataus ja joissa on mitta-asteikon yksikäsitteisyyden vaatimat kiinnitykset. Mittausmalleja voidaan tarvittaessa laajentaa spesifioimalla jäännösten kovarianssivapautuksia tai spesififaktoreita, jolloin identifioituvuustarkastelut tarvitsee kohdistaan ainoastaan jäännöskovarians-seihin tai spesififaktoreihin liittyviin parametreihin.

3.1.2 Mallin identifioituvuustarkastelut

Konfirmatorisen faktorimallin parametri on identifioituva eli yksilöityvä, mikäli se voidaan ratkaista teoreettisen kovarianssimatriisin Σ määrittämien yhtälöiden avulla. Jos kaikki faktorimallin parametrit ovat identifioituvia, koko malli on identifioituva (Jöreskog 1981). Mallin parametri on yli-identifioituva, jos se voidaan ratkaista Σ:sta useammalla kuin yhdellä tavalla. Yli-identifioituvuus ei kuitenkaan ole ristiriidassa identifioituvuuden kanssa, sillä kaikki eri ratkaisutavat antavat samat arvot mallin parametreille. Identifioitumattomasta parametrista voidaan vastaavasti käyttää nimitystä ali-identifioituva parametri.

Faktorimallin identifioituvuudelle on olemassa kaksi välttämätöntä ehtoa. Jotta faktorimallin parametrit voitaisiin ratkaista teoreettisen kovarianssimatriisin Σ avulla, pitää olla voimassa

$$[pk + (1/2)\, k\, (k + 1) + p] - r_1 \leq (1/2)\, p(p+1), \tag{3.3}$$

jossa pk mahdollisten parametrien määrä latausmatriisissa Λ, $(1/2)\, k\, (k+1)$ on mahdollisten parametrien määrä faktoreiden kovarianssimatriisissa Ω ja p mahdollisten parametrien määrä jäännöskovarianssimatriisissa Θ. Kovariansseja ja variansseja on teoreettisessa kovarianssimatriisissa Σ yhteensä $(1/2)\, p\, (p+1)$ kappaletta. r_1 on rajoitusten määrä parametrimatriiseissa Λ, Ω ja Θ. Rotaatioyksikäsitteisyys on toinen välttämätön ehto faktorimallin identifioituvuudelle. Jotta faktorimalli olisi rotaatioyksikäsitteinen, tarvitaan vähintään k^2 kappaletta rajoituksia parametrimatriiseihin Λ ja Ω eli pitää olla voimassa ehto

$$r_2 \geq k^2 . \tag{3.4}$$

Yhdistämällä ehdot (3.3) ja (3.4), saadaan faktorimallin identifioituvuudelle yksi välttämätön ehto r*,

$$r^* \geq \max(r_1, r_2)\,. \tag{3.5}$$

Riittäviä ehtoja faktorimallin identifioituvuudelle ei voida määrittää yleisesti. Identifioituvuustarkastelut voidaan tietenkin suorittaa identifioituvuuden määritelmään perustuen, mikä on kuitenkin erittäin työlästä, mikäli kyseessä on suuri ja suhteellisen monimutkaisen rakenteen omaava malli. Identifioituvuustarkastelut voidaan vaihtoehtoisesti suorittaa siten, että estimoidaan LISREL-ohjelmalla spesifioitu malli, jolloin ohjelma tarkistaa samalla sen identifioituvuuden. Tällöin ongelmaksi jää ei-identifioituvien parametrien paikallistaminen.

3.1.3 Mallin parametrien estimointi

Faktorimallin parametrien estimoinnissa pyritään spesifioidun mallin tuntemattomille parametreille löytämään arvot siten, että kovarianssimatriisin Σ sovite

$$\hat{\Sigma} = \hat{\Lambda}\hat{\Omega}\hat{\Lambda}^T + \hat{\Theta} \tag{3.6}$$

olisi mahdollisimman lähellä otoskovarianssimatriisia **S**.

Konfirmatoristen faktorimallien estimointiin sopivia menetelmiä on monia, joista tarkastellaan suurimman uskottavuuden menetelmää, yleistettyä pienimmän neliösumman menetelmää, yleistä painotettua pienimmän neliösumman menetelmää ja tavallista pienimmän neliösumman menetelmää. Havaittujen muuttujien noudattaessa moniulotteista normaalijakaumaa, voidaan estimoinnissa käyttää suurimman uskottavuuden menetelmää. Kun multinormaalijakautuneisuus ei ole täysin voimassa, voidaan käyttää yleistettyä pienimmän neliösumman menetelmää. Jos havaitut muuttujat eivät ole lähellekään multinormaalijakautuneita, voidaan parametrit estimoida yleisellä painotetulla pienimmän neliösumman menetelmällä. Menetelmä vaatii kuitenkin havaittujen muuttujien jakaumien symmetrisyyden ja kohtalaisen suuren otoskoon. Pienimmän neliösumman menetelmää käytetään silloin, kun mitään muuta estimointimenetelmää ei voida käyttää.

Estimointimenetelmissä periaatteena on minimoida kohdefunktio F estimoitavien parametrien suhteen. Pienimmän neliösumman menetelmässä minimoitava kohdefunktio on

$$F(\Lambda,\Omega,\Theta) = (1/2)\mathrm{tr}\,[\Sigma(\Lambda,\Omega,\Theta)-\mathbf{S}]^2$$

$$= (1/2)\sum_{i=1}^{p}\sum_{j=1}^{p}[\sigma_{ij}(\Lambda,\Omega,\Theta)-s_{ij}]^2. \tag{3.7}$$

Tavallisessa eli painottamattomassa pienimmän neliösumman menetelmässä ei käytetä jakaumaoletuksia. Se on suhteellisen tehoton estimointimenetelmä ja sitä käytetään vain silloin, kun havaitut muuttujat eivät ole lähellekään normaalijakautuneita.

Pienimmän neliösumman menetelmää tehokkaampi estimointimenetelmä on yleistetty pienimmän neliösumman menetelmä. Yleistettyä pienimmän neliösumman menetelmää voidaan käyttää silloin, kun havaitut muuttujat noudattavat likimain normaalijakaumaa. Yleistetyssä pienimmän neliösumman menetelmässä minimoitava kohdefunktio on

$$F(\Lambda,\Omega,\Theta) = (1/2)\mathrm{tr}\left[\left(\mathbf{I}-\mathbf{S}^{-1}\Sigma\right)^2\right]. \tag{3.8}$$

Suurimman uskottavuuden menetelmällä on etuna sen hyvät estimointiteoreettiset ominaisuudet: suurimman uskottavuuden estimaattorit ovat tarkentuvia, niillä on minivarianssiominaisuudet, ja ne ovat asymptoottisesti normaalijakautuneita. Minimoitava kohdefunktio suurimman uskottavuuden menetelmässä on

$$F(\Lambda,\Omega,\Theta) = \ln|\Sigma(\Lambda,\Omega,\Theta)| - \ln|\Sigma| + \mathrm{tr}[\Sigma^{-1}(\Lambda,\Omega,\Theta)\mathbf{S}] - p\ . \tag{3.9}$$

Yleistetty painotettu pienimmän neliösumman menetelmä (Browne 1984) tuottaa asymptoottisesti harhattomia ja jakaumasta riippumatomia parhaita pienimmän neliösumman estimaatteja, joilla on minimivarianssi ominaisuudet. Normaalijakauman tapauksessa se tuottaa samat estimointitulokset kuin suurimman uskottavuuden menetelmä. Yleisessä painotetussa pienimmän neliösumman menetelmässä minimoitava kohdefunktio (Jöreskog & Sörbom 1993) on

$$F(\Lambda,\Omega,\Theta) = (s - \sigma)^T W^{-1}(s - \sigma)$$

$$= \sum_{g=1}^{k}\sum_{h=1}^{g}\sum_{i=1}^{k}\sum_{j=1}^{i} w^{gh,ij}\left(s_{gh} - \sigma_{gh}\right)\left(s_{gh} - \sigma_{gh}\right) , \tag{3.10}$$

missä $s^T = (s_{11}, s_{21}, s_{22}, s_{31}, \ldots ,s_{pp})$ eli s on havaittujen muuttujien kovarianssimatriisin **S** alakolmiosta muodostettu vektori. $\sigma^T = (\sigma_{11}, \sigma_{21}, \sigma_{22}, \sigma_{31}, \ldots , \sigma_{pp})$ on vastaavasti teoreettisen kovarianssimatriisin **Σ** alakolmiosta muodostettu vektori. Kohdefunktiossa $\mathbf{W}^{-1}$ on ns. painomatriisi. Painot $w_{gh,ij}$ lasketaan kaavalla

$$w_{gh,ij} = m_{ghij} - s_{gh}\, s_{ij} , \tag{3.11}$$

missä

$$m_{gh,ij} = (1 / N)\sum_{a=1}^{N}\left(y_{ag}-\bar{y}_g\right)\left(y_{ah}-\bar{y}_h\right)\left(y_{ai}-\bar{y}_i\right)\left(y_{aj}-\bar{y}_j\right) , \tag{3.12}$$

eli m_{ghhi}:t ovat havaittujen muuttujien 4. keskusmomentteja. N on havaintojen lukumäärä.

3.1.4 Hypoteesien testaaminen

Konfirmatoristen faktorimallin estimoinnin jälkeen testataan, onko estimoitu malli voimassa eli onko nollahypoteesi

$$H_0: \Sigma = \Lambda\Omega\Lambda^T + \Theta .$$

voimassa. Vastahypoteesina on

H_1 : kovarianssimatriisissa Σ ei ole rajoituksia.

H_0 -hypoteesin voimassaoloa testataan χ^2 -testillä, joka saadaan kohdefunktion F minimin avulla seuraavasti:

$$\chi^2 = (N-1)F_{min}(\Sigma) \ . \tag{3.12}$$

Hypoteesien testaamisessa χ^2 -testisuuretta verrataan vapausasteisiin df, jotka voidaan laskea kaavasta

$$df = (1/2)p(p+1) - t_0 , \tag{3.13}$$

missä (1/2)p(p+1) on kovarianssimatriisissa Σ olevien kovarianssien ja varianssien lukumäärä ja t_0 on estimoitavien parametrien lukumäärä H_0 -mallissa. Päätös H_0 -hypoteesin hyväksymisestä tai hylkäämisestä tehdään yleensä estimoitua χ^2 -arvoa vastaavan p-arvon eli todennäköisyyden

$$p = p(\chi 2(df) \geq \hat{\chi}^2(df)) \tag{3.14}$$

perusteella. Ho -hypoteesi hyväksytään, jos p-arvo on suuri ja hylätään, jos $p < .05$ (0.05:n merkitsevyystasolla). Kuitenkaan estimoitua mallia ei pidä hylätä tai hyväksyä pelkästään p-arvon perusteella etenkään silloin, kun otoskoko on hyvin suuri tai hyvin pieni. Tällöin mallin sopivuutta havaintoaineistoon arvioidaan jollakin muulla tavalla, esimerkiksi vertaamalla χ^2 -arvon suuruutta suhteessa vapausasteisiin. Yleensä mallia voidaan pitää vähintään kohtalaisesti aineistoon sopivana, jos $\chi^2(df) < 2df$.

Faktorimalleille voidaan laskea ns. normeerattu yhteensopivuusindeksi (Tucker & Lewis 1973, Bentler & Bonett 1980), jonka avulla voidaan arvioida suuren otoskoon vaikutusta χ^2 -testiin ja mallin sopivuutta aineistoon. Normeerattu yhteensopivuusindeksi (Normed Fit Index, NFI) vertaa estimoitua mallia ns. nollamalliin, jossa havaitut muuttujat oletetaan korreloimattomiksi. Normeerattu yhteensopivuusindeksi lasketaan nollamallin ja estimoidun mallin χ^2 -arvojen avulla:

$$\mathrm{NFI} = \frac{\chi_0^2 - \chi_1^2}{\chi_0^2} = \frac{(N-1)F_0 - (N-1)F_1}{(N-1)F_0} = \frac{F_0 - F_1}{F_0} \qquad (3.15)$$

Kaavasta nähdään, että normeerattu yhteensopivuusindeksi ei riipu otoskoosta ja se voi saada arvoja 0:n ja 1:n väliltä. Malli on sitä parempi, mitä lähempänä 1:stä normeerattu yhteensopivuusindeksi on.

Critical N –indeksin (CN) perusteella voidaan arvioida χ^2 -testin sopivuutta mallin riittävyysmittana. Jos kriittinen N on pienempi kuin otoksen suuruus, χ^2 -testi ei sovellu mahdollisesti liian suuren otoksen vuoksi mallin riittävyyden mittaamiseen.

Hypoteesien testaamisessa on mahdollisuus verrata kahta estimoitua mallia toisiinsa ja testata kumpi malli voidaan hyväksyä. Tämä edellyttää, että kyseessä on ns. sisäkkäiset mallit eli mallit ovat toistensa erikoistapauksia. Tällaisia hypoteeseja voidaan testata χ^2 -peräkkäistestillä (ks. luku 4).

3.1.5 Mallin riittävyystarkastelut

Kun hypoteesien testaamisvaiheessa on todettu, että estimoitu malli sopii kokonaisuudessaan aineistoon (χ^2 -testin mielessä tai muulla tavalla arvioituna), tarkastellaan mallin riittävyyttä tarkemmin ja paikallistetaan mallin mahdolliset riittämättömät kohdat. Riittävyystarkastelut jaetaan neljään ryhmään, koko mallia koskeviin tarkasteluihin, muuttujakohtaisiin tarkasteluihin, parametrikohtaisiin tarkasteluihin ja havaintokohtaisiin tarkasteluihin.

Koko mallia koskevat tarkastelut perustuvat etupäässä χ^2 -testiin. χ^2 -testiä käytetäänkin usein mallin rakentamisessa enemmän riittävyyden arviointiin kuin hypoteesien testaamiseen. Se on erityisen hyödyllinen riittävyyden mitta silloin, kun mallia modifioidaan aineiston antaman informaation avulla. Koko mallin riittävyyttä voidaan arvioida myös normeeratun yhteensopivuusindeksin avulla tai toisella otoskoosta riippumattomalla yleisellä yhteensopivuusindeksillä, GFI:llä (Goodness-of-Fit-Index). Yleistetyn painotetun pienimmän neliösumman estimoinnin tapauksessa voidaan määrittää yleisesti GFI:n laskentakaava:

$$\mathrm{GFI} = 1 - \frac{(\mathbf{s}-\hat{\boldsymbol{\sigma}})^{T}\, \mathbf{W}^{-1} (\mathbf{s}-\hat{\boldsymbol{\sigma}})}{\mathbf{s}^{T}\, \mathbf{W}^{-1}\, \mathbf{s}}, \qquad (3.14)$$

missä jakolausekkeen osoittajassa on sovitetun mallin kohdefunktion minimi ja nimittäjässä on painotettu otoskovarianssimatriisi (Tanaka & Huba 1984, Jöreskog & Sörbom 1993b) . Muiden estimointimenetelmien tapauksessa GFI voidaan johtaa (3.14):sta. Esimerkiksi suurimman uskotavuuden estimoinnissa GFI on

$$\mathrm{GFI} = 1 - \frac{\mathrm{tr}(\hat{\Sigma}^{-1}\mathbf{S} - \mathbf{I})^2}{\mathrm{tr}(\hat{\Sigma}^{-1}\mathbf{S})^2} \qquad (3.15)$$

(Jöreskog & Sörbom 1981). GFI saa arvoja 0:n ja 1:n väliltä. Jos GFI on suurempi kuin 0.90, on malli GFI:n mielessä riittävä.

Mallin riittävyyttä voidaan mitata myös RMR:llä (Root Mean square Residual) , joka mittaa keskimääräistä jäännöskovarianssia ja jäännösvarianssia. RMR (Jöreskog & Sörbom 1981) lasketaan kaavalla

$$\mathrm{RMR} = \sqrt{2 \sum_{j=1}^{p} \sum_{i=1}^{j} \left(s_{ij} - \hat{\sigma}_{ij}\right)^2 / [p(p+1)]}\;, \qquad (3.16)$$

missä s_{ij}:t ovat otoskovarianssimatriisin **S** alkioita ja $\hat{\sigma}_{ij}$:t kovarianssimatriisin $\hat{\Sigma}$ alkioita.

Koko mallin riittävyyttä mittaavia indeksejä on muitakin edellä esitettyjen indeksien lisäksi. Näistä mainittakoon vielä RMSEA (Root Mean Square Error of Approximation). Jos estimoidulle mallille RMSEA < .05, se on RMSEA:n mielessä tarpeeksi hyvä, eikä sitä tarvitse tarkentaa. Jos taas RMSEA on välillä .05 ja .08, on estimoitu malli kohtalainen eikä sitä RMSEA:n mielessä ole välttämätöntä tarkentaa.

Faktorimallin muuttujakohtaisissa riittävyystarkasteluissa lasketaan havaituille muuttujille muuttujakohtainen riittävyysindeksi (squared multiple correlation)

$$\hat{R}_i^2 = 1 - \hat{\theta}_i / s_{ii}, \tag{3.16}$$

jossa i = 1,...,p. Indeksit $\hat{R}_i^2$ voidaan tulkita havaittujen muuttujien reliabiliteetti- tai kommunaliteettikertoimiksi. Ne kuvaavat havaittujen muuttujien kykyä mitata faktoreita. Teoriassa $\hat{R}_i^2$:t saavat arvoja 0:n ja 1:n väliltä. Mitä lähempänä 1:stä nämä indeksit ovat, sitä paremmin ne mittaavat faktoreita.

Parametrikohtaisten riittävyystarkastelujen tarkoituksena on varmistaa, että parametrien estimaatit ovat sisällöllisesti tulkinnallisia ja tilastotieteellisesti kelvollisia: faktoreiden varianssien estimaattien $\hat{\psi}_{jj}$, j = 1,...,k ja jäännösvarianssien estimaattien $\hat{\theta}_i$, i = 1,..., p pitää olla positiivisia. Estimoitujen kovarianssimatriisien $\hat{\Omega}$ ja $\hat{\Theta}$ pitää olla positiivisesti definiittejä. Jos on $\hat{\Omega}$ standardoitu korrelaatiomatriisiksi, pitää faktoreiden välisten korrelaatioiden olla -1:n ja 1:n välillä.
Parametrikohtaisissa riittävyystarkasteluissa tarkastellaan myös parametrien estimointitarkkuutta keskivirheiden avulla. Suuret keskivirheet ovat merkki estimoinnin epätarkkuudesta. Parametrien nollasta eroavuutta voidaan arvioida t-arvon avulla, joka saadaan parametrin estimaatin ja keskivirheen suhteena. Jos $|t - \text{arvo}| > 1.96$, voidaan parametria pitää nollasta eroavana. Jos kuitenkin mallin parametrien estimaattorit korreloivat voimakkaasti, eivät estimaateille lasketut t-arvot ole luotettavia. Tällöin parametrien nollastaeroavuutta pitää arvioida myös estimaattoreiden korrelaatioiden avulla.

Faktorimallissa kiinnitettyjen ja yhtäsuurina estimoitujen parametrien tarkastelut tehdään modifikaatioindeksien avulla. Suuri modifikaatioindeksin arvo osoittaa, että kiinnitys tai yhtäsuurena estimointi ei ole sopiva. Modifikaatioindeksit ovat estimointimenetelmän kohdefunktion ensimmäisiä ja toisia derivaattoja kiinnitettyjen yhtäsuurina estimoitujen parametrien suhteen (Jöreskog & Sörbom 1981, Sörbom 1986). Modifikaatioindeksi ilmaisee, kuinka paljon χ^2 -arvo vähintään laskee, jos kiinnitetty tai yhtäsuurena estimoitu parametri vapautetaan.

Faktorimallin havaintokohtaiset tarkastelut perustuvat jäännöskovarianssimatriisiin **R**.

$$\mathbf{R} = \mathbf{S} - \hat{\Sigma} = [s_{ij} - \hat{\sigma}_{ij}] = [R_{ij}] \ , \tag{3.17}$$

jossa i,j = 1,..., p. Suuret jäännökset R_{ij} osoittavat, miltä osin estimoitu malli ei sovi aineistoon. Jäännöskovarianssimatriisin **R** perusteella voi kuitenkin olla vaikeata arvioida jäännösten suuruutta, jos otoskovarianssimatriisi **S** sisältää suuruudeltaan toisistaan poikkeavia variansseja. Tämän takia jäännökset normalisoidaan. Normalisoidut jäännökset saadaan kaavalla

$$NR_{ij} = \frac{s_{ij} - \hat{\sigma}_{ij}}{\sqrt{\left(\hat{\sigma}_{ii}\hat{\sigma}_{jj} + \hat{\sigma}_{ij}^2\right)/N}}, \tag{3.18}$$

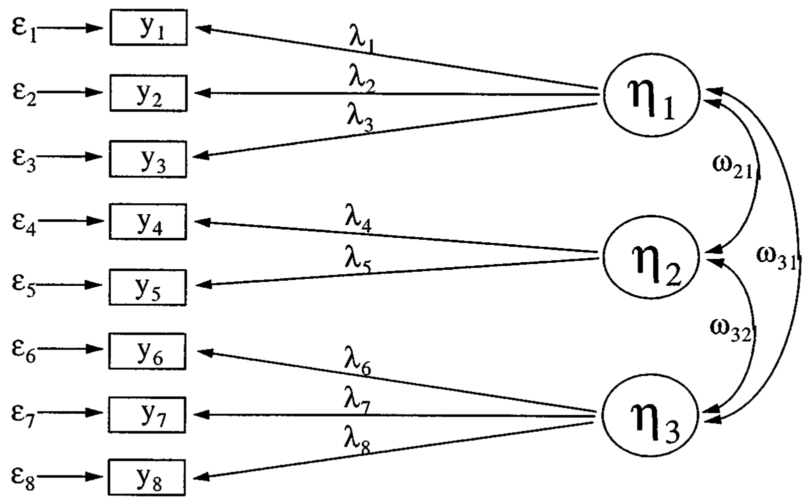
jossa i,j = 1,..., p. Normalisoidut jäännökset, joille on voimassa

$$|NR_{ij}| > 2,$$

ovat merkki mallin riittämättömyydestä. Kuitenkin rakenteeltaan suuren ja monimutkaisen faktorimallin tapauksessa voidaan mallia pitää riittävänä, vaikka mallille saataisiinkin muutamia tämän ehdon täyttäviä normalisoituja jäännöksiä .

3.2 Aamuvireisyyden, unilaadun ja mielialan faktorimalli

Tarkastellaan seuraavaksi, kuinka konfirmatorisen faktorimallin rakentaminen etenee käytännössä kappaleissa 3.1.1 - 3.1.5 esitettyjen vaiheiden mukaisesti. Rakennettava malli spesifioidaan luvussa 2 esitettyihin tutkimushypoteeseihin 1-3 perustuen. Niiden perusteella käsiteltävä muuttujajoukko on esitettävissä konfirmatorisen faktorimallin avulla, tarkemmin sanoen kolmen faktorin mittausmallin avulla. Tällainen malli on esitettynä kuviossa 1.



y_1 = Aamupirteyskokemus	y_5 = Unilaatu -indeksi
y_2 = Väsymyksen useus kouluaamuisin	y_6 = Hermostuneisuus
y_3 = Aamuvireisyys -indeksi	y_7 = Ärtyneisyys
y_4 = Univaikeuksien ongelmallisuusaste -indeksi	y_8 = Masentuneisuus

η_1 = Aamuvireisyys	η_3 = Mieliala
η_2 = Unilaatu	

Kuvio 3.A Kolmen faktorin malli aamuvireisyys-, unilaatu- ja mielialamuuttujille

Jotta tämä malli olisi mitta-asteikoltaan yksikäsitteinen, pitää jokaisen faktorin varianssi tai yksi lataus asettaa vakioksi. Tässä tapauksessa valitaan näistä jälkimmäinen asetus. Siten rakennettava faktorimalli voidaan esittää parametrimatriisien avulla seuraavasti:

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & \lambda_{52} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda_{73} \\ 0 & 0 & \lambda_{83} \end{bmatrix}, \quad \Omega = \begin{bmatrix} \omega_{11} & & \\ \omega_{21} & \omega_{22} & \\ \omega_{31} & \omega_{32} & \omega_{33} \end{bmatrix}, \quad \Theta = \mathrm{diag}(\theta_1\, \theta_2\, \theta_3\, \theta_4\, \theta_5\, \theta_6\, \theta_7\, \theta_8)\ .$$

Malli tuottaa havaituille muuttujille kovarianssirakenteen

$$\Sigma = \begin{bmatrix} \sigma_{11} & & & & & & & \\ \sigma_{21} & \sigma_{22} & & & & & & \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & & & & & \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} & & & & \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_{55} & & & \\ \sigma_{61} & \sigma_{62} & \sigma_{63} & \sigma_{64} & \sigma_{65} & \sigma_{66} & & \\ \sigma_{71} & \sigma_{72} & \sigma_{73} & \sigma_{74} & \sigma_{75} & \sigma_{76} & \sigma_{77} & \\ \sigma_{81} & \sigma_{82} & \sigma_{83} & \sigma_{84} & \sigma_{85} & \sigma_{86} & \sigma_{87} & \sigma_{88} \end{bmatrix} =$$

$$\begin{bmatrix} \omega_{11}+\theta_1 & & & & & & & \\ \omega_{11}\lambda_{21} & \omega_{11}\lambda_{21}^2+\theta_2 & & & & & & \\ \omega_{11}\lambda_{31} & \omega_{11}\lambda_{21}\lambda_{31} & \omega_{11}\lambda_{31}^2+\theta_3 & & & & & \\ -\omega_{21} & -\omega_{22}\lambda_{21} & -\omega_{21}\lambda_{31} & \omega_{22}+\theta_4 & & & & \\ \omega_{21}\lambda_{52} & \omega_{21}\lambda_{21}\lambda_{52} & \omega_{21}\lambda_{31}\lambda_{52} & -\omega_{22}\lambda_{52} & \omega_{22}\lambda_{52}^2+\theta_5 & & & \\ \omega_{31} & \omega_{31}\lambda_{21} & \omega_{31}\lambda_{31} & -\omega_{32} & \omega_{32}\lambda_{52} & \omega_{33}+\theta_6 & & \\ \omega_{31}\lambda_{73} & \omega_{31}\lambda_{21}\lambda_{73} & \omega_{31}\lambda_{31}\lambda_{73} & -\omega_{32}\lambda_{73} & \omega_{32}\lambda_{52}\lambda_{73} & \omega_{33}\lambda_{73} & \omega_{22}\lambda_{73}^2+\theta_7 & \\ \omega_{31}\lambda_{83} & \omega_{31}\lambda_{21}\lambda_{83} & \omega_{31}\lambda_{31}\lambda_{83} & -\omega_{32}\lambda_{83} & \omega_{32}\lambda_{52}\lambda_{83} & \omega_{33}\lambda_{83} & \omega_{33}\lambda_{73}\lambda_{83} & \omega_{33}\lambda_{83}^2+\theta_8 \end{bmatrix}$$

Kovarianssirakenteeseen perustuen voidaan aloittaa spesifioidun mallin identifioituvuustarkastelut. Koska spesifioitu malli on mittausmalli, tiedetään, että se on identifioituva. Havainnollistetaan kuitenkin seuraavassa identifioituvuustarkasteluita tarkastelemalla spesifioidun mallin identifioituvuutta.

Aluksi voidaan todeta, että välttämättömät ehdot spesifioidulle mallille ovat voimassa. Spesifioidussa mallissa on 19 vapaasti estimoitavaa parametria, kun taas kovarianssimatriisissa Σ kovariansseja on yhteensä 36 kappaletta. Malli on myös rotaatioyksikäsitteinen, sillä siihen on asetettu enemmän kuin $k^2 = 9$ kappaletta rajoituksia.

Spesifioidun mallin identifioituvuus voidaan osoittaa identifioituvuuden määritelmän perusteella. Identifioituvuuden osoittaminen aloitetaan kirjoittamalla auki kovarianssimatriisin Σ kovarianssit ja hakemalla tätä kautta mallin tuntemattomille parametreille esitysmuodot, joissa ne saadaan esitettyä kovarianssien σ avulla.

Kirjoitetaan auki aluksi kovarianssit σ_{21}, σ_{31}, σ_{32} ja σ_{11}, joiden avulla saadaan esitysmuodot tuntemattomille parametreille ω_{11}, λ_{31}, λ_{21} ja θ_1 seuraavasti:

$$\sigma_{21} = \lambda_{21}\,\omega_{11} \;\Leftrightarrow\; \omega_{11} = \frac{\sigma_{21}}{\lambda_{21}} \tag{3.19}$$

$$\sigma_{31} = \lambda_{31}\,\omega_{11} \;\Leftrightarrow\; \lambda_{31} = \frac{\sigma_{31}}{\omega_{11}} \tag{3.20}$$

$$\sigma_{32} = \lambda_{31}\,\lambda_{21}\omega_{11} \;\Leftrightarrow\; \lambda_{21} = \frac{\sigma_{32}}{\lambda_{31}\omega_{11}} \tag{3.21}$$

$$\sigma_{11} = \omega_{11} + \theta_1 \;\Leftrightarrow\; \theta_1 = \sigma_{11} + \omega_{11} \tag{3.22}$$

Kun yhtälö (3.21) sijoitetaan yhtälöön (3.19), saadaan

$$\omega_{11} = \frac{\lambda_{21}}{\dfrac{\sigma_{32}}{\lambda_{31}\,\omega_{11}}}\,. \tag{3.23}$$

Kun tähän sijoitetaan yhtälö (3.20), saadaan

$$\omega_{11} = \frac{\sigma_{21}\,\sigma_{31}}{\sigma_{32}}\,. \tag{3.24}$$

Tässä ω_{11} on esitettynä tunnettujen kovarianssien σ avulla, joten ω_{11} on identifioituva. Kun vastaavasti sijoitetaan (3.24) (3.22):een, havaitaan, että myös θ_1 on identifioituva. Sama havainto voidaan tehdä parametrin λ_{31} osalta, kun yhtälö (3.24) sijoitetaan yhtälöön (3.20). Kun näin saatu yhtälö sijoitetaan yhtälöön (3.21), havaitaan, että myös λ_{31} on identifioituva. Vastaavalla tavalla kirjoittamalla auki muita kovarianssimatriisin Σ kovariansseja ja tekemällä sijoituksia, voidaan osoittaa mallin muutkin tuntemattomat parametrit identifioituviksi. Siten spesifioitu faktorimalli on kokonaisuudessaan identifioituva.

Seuraavassa vaiheessa estimoidaan identifioituvaksi todettu malli. Mallin estimoinnissa on käytetty yleistä pienemmän neliösumman menetelmään. Estimointitulokset pojilla, joiden PTT on huono, on esitetty taulukossa 3.1. Perusestimointitulosten lisäksi taulukossa on esitetty myös standardoitu kovarianssimatriisi kuvaa faktoreiden välisiä korrelaatioita. Korrelaatioarvioiden perusteella voidaan päätellä, minkälaisia rakenneyhtälöitä faktoreiden välillä on.

Suoritetaan seuraavana estimoidun mallin estimointitulosten perusteella riittävyystarkastelut ja hypoteesien testaaminen. Kun faktorimallin tilastollista yhteensopivuutta testataan yleistä H_1 -hypoteesia vastaan, saadaan χ^2-testin tulos (17) = 18.27, P = 0.37. Estimoitu malli sopii siten hyvin aineistoon. Myös muiden koko mallin riittävyyttä kuvaavien indeksien, GFI:n, RMSEA:n, ja RMR:n, perusteella voidaan todeta, että malli on riittävä. Tulos RMR = .021 tarkoittaa, että keskimäärin kaksi sadasosaa havaituista otoskovariansseista on jäänyt selittämättä. Havaittujen muuttujien reliabiliteettikertoimet $\hat{R}^2_{y(i)}$ ovat kohtalaisen korkeita lukuunottamatta unilaadun reliabiliteettikerrointa, joka on pieni. Masentuneisuuden ja ärtyneisyyden lataukset ovat suuria. Kun hermostuneisuuden lataus on kiinnitetty ykköseksi, voidaan η_3 tulkita mielialafaktoriksi. Aamuvireisyysfaktorin osalta voidaan tehdä samat tulkinnat, vaikka aamuvireisyysindeksin lataus on suuruusluokaltaan hieman pienempi. Unilaatufaktorilla unilaatu -muuttujan saama lataus on kohtalaisen pieni. Tämä ja unilaatu -muuttujan saama pieni reliabiliteettikerroin osoittavat, että unilaadulla heikko mittauskyky unilaatufaktorin mittaamisessa. Mittausvirheiden varianssit ovat kohtalaisen pieniä lukuunottamatta uniongelmat-, unilaatu-, ja aamuväsymys -muuttujan virhevariansseja, jotka ovat hieman suurempia. Joidenkin parametrien keskivirheiden osalta voidaan todeta, että mallin estimoinnissa on hieman epätarkkuutta. Erityisesti näin voidaan sanoa niiden parametrien osalta, joiden keskivirhe on suurempi kuin .10. Modifikaatioindeksien maksimiarvo MI_{max} osoittaa, että uniongelmat -muuttujan ja masentuneisuus -muuttujan jäännökset olisivat korreloituneita. Jäännösten välisen kovarianssin vapauttamiselle ei ole kuitenkaan mitään sisällöllisiä perusteita. Normalisoitujen jäännösten maksimiarvo $|NR_{max}| = 2.53$ perusteella voidaan sanoa, että estimoitu malli ei ole selittänyt kovin hyvin kaikkia otoskovariansseja. Mallin modifioimiselle ei kuitenkaan löydy sisällisiä perusteita tai muita tilastollisia perusteita.

Koska unilaatu -muuttujalla on unilaatufaktorin mittaamisessa ja unilaatufaktorilla kohdassa 3.4 rakennettavassa polkumallissa on sisällöllisesti tärkeä osa, hyväksytään estimoitu malli, vaikka se unilaatu -muuttujan mittauskyvyn heikkouden vuoksi osoittautuu hieman riittämättömäksi.

Kun spesifioitu malli estimoidaan pojille, joiden PTT on hyvä, saadaan seuraavat yleisen pienimmän neliösumman estimointitulosten riittävyysindeksit:

$\chi^2(17) = 31.04$, $p = .020$, $N = 184$, $CN = 197.98$

$NFI = .96$, $GFI = .96$, $RMSEA = .067$, $RMR = .031$

$|NR_{max}| = |NR_{86}| = 2.60$, $MI_{MAX} = MI(\lambda_{72}) = 12.66$, $MI(\theta_{86}) = 12.46$

Mallille yhteensopivuustestin tulokseksi yleistä H_0 -hypoteesia vastaan saadaan $\chi^2(17) = 31.04$, $p = .020$. Testin perusteella malli todetaan riittämättömäksi. Vaikka NFI:n, GFI:n, RMSEA:n ja RMR:n arvoissa ei mallin riittämättömyys näy, ei otoskoko ole vaikuttanut χ^2 -testin tulokseen. Tämän voi todeta kriittisen N:n arvon perusteella (CN > N) ja erityisesti normalisoitujen jäännösten maksimiarvon ja suurien modifikaatioindeksien perusteella voidaan todeta, että estimoitu malli on tietyiltä osin riittämätön. Modifikaatioindeksien maksimiarvon perusteella voitaisiin vapauttaa ärtyneisyys -muuttujan lataus unilaatufaktorilla, mihin ei kuitenkaan ole sisällöllisiä perusteita. Normalisoitujen jäännösten maksimiarvon ja modifikaatioindeksin $MI(\theta_{86}) = 12.46$, voidaan päätellä, että mallin riittämättömyys voi johtua hermostuneisuuden ja masentuneisuuden jäännösten korreloituneisuudesta.

Taulukko 3.1 Aamuvireisyyden, unilaadun ja mielialan faktorimallin estimointitulokset pojilla, joiden perheen taloudellinen tilanne on huono

Aamuvireisyyden η_1, unilaadun η_2 ja mielialan η_3 faktorimallin latausten Λ, jäännösvarianssien θ ja reliabiliteettien $R^2_{y(i)}$ estimaatit (suluissa keskivirheet):

	$\hat{\Lambda}$ (s.e.)			$\hat{\Theta}$ (s.e.)	$\hat{R}^2_{y(i)}$
	$\eta 1$	$\eta 2$	$\eta 3$		
Aamupirteys	1*	0*	0*	.26(.04)	.61
Aamuväsymys	-.94(.11)	0*	0*	.42(.05)	.47
Aamuvireisyys	.71(.08)	0*	0*	.21(.03)	.50
Uniongelmat	0*	-1*	0*	.40(.08)	.49
Unilaatu	0*	.56(.11)	0*	.38(.04)	.24
Hermostuneisuus	0*	0*	1*	.15(.02)	.54
Ärtyneisyys	0*	0*	.91(.12)	.17(.02)	.47
Masentuneisuus	0*	0*	.90(.12)	.24(.03)	.37

*) kiinnitetty

Faktoreiden kovarianssimatriisi $\left(\hat{\Psi}\right)$ ja standardoitu kovarianssimatriisi (suluissa keskivirheet):

	$\hat{\Psi}$ (s.e.)			standardoitu $\hat{\Psi}$		
	$\eta 1$	$\eta 2$	$\eta 3$	$\eta 1$	$\eta 2$	$\eta 3$
$\eta 1$	.41(.07)			1		
$\eta 2$	.24(.04)	.38(.10)		.61	1	
$\eta 3$	-.07(.02)	-.14(.03)	.18(.03)	-.26	-.53	1

Riittävyysindeksit:

$\chi^2(17) = 18.27$, $P = 0.37$, $N = 265$

GFI = .98 RMSEA = .017 RMR = .021

$|NR_{max}| = 2.53$, $MI_{MAX} = MI(\theta_{65}) = 4.78$

Taulukossa 3.2 on esitetty faktorimalli, jossa hermostuneisuuden ja masentuneisuuden jäännösten välinen kovarianssi on vapautettu. Estimaatiksi jäännösten väliselle kovarianssille saadaan -.11ja keskivirheeksi .04. χ^2-testin tulokseksi $\chi^2(16) = 18.12$, $p = .32$, jonka perusteella voidaan malli hyväksyä. Mallin riittävyydestä voidaan tehdä samat päätelmät kuin taulukon 3.1 mallista edellä.

<u>Taulukko 3.2</u> Aamuvireisyyden, unilaadun ja mielialan faktorimallin estimointitulokset pojilla, joiden perheen taloudellinen tilanne on hyvä.

Aamuvireisyyden η_1, unilaadun η_2, ja mielialan η_3 faktorimallin latausten Λ, jäännösvarianssien Θ ja reliabiliteettien estimaatit $\hat{R}^2_{y(i)}$ estimaatit (suluissa keskivirheet):

	$\hat{\Lambda}$ (s.e.)			$\hat{\Theta}$ (s.e.)	$\hat{R}^2_{y(i)}$
	η_1	η_2	η_3		
Aamupirteys	1*	0*	0*	.31(.05)	.63
Aamuväsymys	-1.13(.11)	0*	0*	.25(.05)	.73
Aamuvireisyys	.73(.07)	0*	0*	.22(.03)	.56
Uniongelmat	0*	-1*	0*	.42(.10)	.46
Unilaatu	0*	.50(.13)	0*	.38(.05)	.20
Hermostuneisuus	0*	0*	1*	.18(.05)	.51
Ärtyneisyys	0*	0*	.71(.18)	.20(.03)	.31
Masentuneisuus	0*	0*	.93(.20)	.20(.05)	.44

*) kiinnitetty

Vapautettu $\theta_{86} = -.11(.04)$.

Faktoreiden kovarianssimatriisi ja standardoitu kovarianssimatriisi:

	$\hat{\Psi}$ (s.e.)			Standardoitu $\hat{\Psi}$		
	η_1	η_2	η_3	η_1	η_2	η_3
η_1	.53(.09)			1		
η_2	..29(.06)	.37(.11)		.65	1	
η_3	-.14(.03)	-.16(.04)	.19(.06)	-.45	-.62	1

Riittävyysindeksit:

$\chi^2(16) = 18.12$, $p = .32$, N= 184,

GFI = .98, RMSEA = .027, RMR = .021,

$|NR_{max}| = 2.09$, $MI_{MAX} = MI(\lambda_{23}) = 4.63$

3.3 Faktoreiden rakenneyhtälömallit

Konfirmatorista faktorimallia voidaan laajentaa siten, että faktoreiden välille määritellään rakenneyhtälöitä. Tällöin oletetaan, että faktorimallissa (3.1) faktoreiden η väliset korrelaatiot johtuvat siitä, että ne selittävät tai ennustavat toisiaan. Jotta rakennettava rakenneyhtälömalli olisi sisällöllisesti tulkinnallinen, pitää rakennettu faktorimalli olla identifioituva. Toisaalta sisällöllisesti tulkinnallisia rakenneyhtälömalleja voidaan rakentaa ainoastaan faktorimalleille, jotka ovat mittausmalleja.

Tarkastellaan seuraavaksi faktoreiden rakenneyhtälöihin liittyvää teoriaa ja rakenneyhtälömallien rakentamista. Tarkastelu perustuu teokseen Leskinen (1987) (sivut 167 - 186).

Faktoreiden η rakenneyhtälömalli on muotoa

$$\begin{cases} y = \Lambda\eta + \varepsilon \\ \eta = B\eta + \zeta \end{cases} \tag{3.25}$$

missä ylempi rivi on kuten (3.1):ssä. B on rakenneyhtälömallin rakenneparametrit sisältävä k x k -matriisi ja ζ on k x 1 -jäännösvektori. Jäännösten kovarianssimatriisia merkitään ψ:llä. ψ on siten k x k -matriisi, mutta yleensä oletetaan, että ψ on diagonaalinen. Tarvittaessa jäännösten ζ kovariansseja voidaan kuitenkin vapauttaa.

Rakenneyhtälömallin rakenneyhtälöosalla $\eta = B\eta + \zeta$ on olemassa ns. supistettu muoto, jossa faktorit sisältävä vektori η esiintyy vain yhtälön vasemmalla puolella. Supistettu muoto voidaan johtaa seuraavasti:

$$\begin{aligned} & \eta = B\eta + \zeta \\ \Leftrightarrow\ & (\mathbf{I} - B)\eta = \zeta \\ \Leftrightarrow\ & \eta = (\mathbf{I} - B)^{-1}\zeta \end{aligned} \tag{3.26}$$

Faktoreiden rakenneyhtälömallin rakenneyhtälöosa voidaan aina esittää supistetussa muodossa (3.26), koska matriisi **I** - B on epäsingulaarinen.

Yksinkertaisin tapaus faktoreiden rakenneyhtälömallista kahden faktorin regressiomalli, jossa ns. eksogeeninen eli selittävä faktori selittää tai ennustaa ns. endogeenista eli selitettävää faktoria. Voidaan sanoa myös, että eksogeeninen faktori vaikuttaa endogeeniseen faktoriin. Monimutkaisempi tapaus tästä on useamman kuin yhden selittävän faktorin regressiomalli (englanniksi multiple regression). Jos myös selitettäviä faktoreita on enemmän kuin yksi, on kyseessä faktoreiden moniyhtälömalli. Faktoreiden polkumalli on malli, jossa kahden faktorin regressiomalleja on useita siten, että ne muodostavat polun, jolloin mallissa on faktoreita, jotka ovat sekä selitettäviä että selittäviä.

Faktoreiden rakenneyhtälömallit voidaan jakaa rekursiivisiin ja simultaanisiin malleihin. Edellä mainittu kahden faktorin regressiomalli on esimerkki rekursiivisesta rakenneyhtälömallista. Simultaanisissa malleissa faktoreiden välillä on vuorovaikutteisia yhteyksiä. Esimerkiksi kahden faktorin regressiomalli on simultaaninen, jos kumpikin faktori selittää toistaan. Rekursiivisissa rakenneyhtälömalleissa täsmennetään tarkasti faktoreiden väliset vaikutussuunnat. Simultaanisissa malleissa oletetaan, että faktorit ovat vuorovaikutuksellisesti toisiinsa yhteydessä tai toisaalta simultaanisia yhteyksiä täsmennetään myös silloin, kun faktoreiden välisiä vaikutussuuntia ei pystytä varmuudella täsmentämään.

3.3.1 Rakenneyhtälömallien rakentaminen

Rakenneyhtälömallien rakentamisessa kannattaa yleensä menetellä siten, että ensin rakennetaan faktorimalli ja vasta tämän jälkeen tarkennetaan mallia faktoreiden välisten yhteyksien osalta. Faktoreiden rakenneyhtälömallien rakentamisessa käydään läpi samat vaiheet kuin konfirmatorisen faktorimallin rakentamisessa.

Faktoreiden rakenneyhtälömallin spesifioinnin tuloksena saadaan faktoreiden kovarianssimatriisille teoreettinen rakenne, joka on muotoa

$$\operatorname{cov}(\eta) = E\eta\eta^{T} = (\mathbf{I} - \mathbf{B})^{-1}\Psi(\mathbf{I} - \mathbf{B}^{T})^{-1} \tag{3.27}$$

Tämä saadaan siis sijoittamalla η:n paikalle supistettu muoto (3.18). Havaittujen muuttujien teoreettinen kovarianssimatriisi on siten muotoa

$$\begin{aligned} \mathrm{cov}(\mathbf{y}) &= \Lambda\Omega\Lambda^T + \Theta \\ &= \Lambda\left[(I-B)^{-1}\Psi(I-B^T)^{-1}\right]\Lambda^T + \Theta \end{aligned} \tag{3.28}$$

Rakenneyhtälön identifioituvuustarkastelut on yksinkertaisinta tehdä faktoreiden kovarianssimatriisin perusteella. Yleisesti rakenneyhtälömallien identifioituvuudesta voidaan sanoa, että malli, joka on rekursiivinen ja jonka jäännökset ovat korreloimattomia, on aina identifioituva. Simultaanisten rakenneyhtälömallien ja ei-diagonaalisen jäännöskovarianssimatriisin omaavien rakenneyhtälömallien tapauksessa identifioituvuus joudutaan varmentamaan erikseen.

Rakenneyhtälömallin muuttujakohtaisista riittävyystarkasteluista huomautettakoon, että kullekin rakenneyhtälömallin selitettävälle muuttujalle voidaan laskea selitysasteet, joiden avulla voidaan arvioida selittävien muuttujien selityskykyä tai ennustuskykyä ts. minkäsuuruinen osa kunkin selitettävän faktorin η_i vaihtelusta on pystytty selittämään. Selitysaste i:nnelle selitettävälle on muotoa

$$R^2_{\eta(i)} = 1 - \psi_i / \mathrm{var}(\eta_i), \tag{3.29}$$

missä ψ_i on i:nnen rakenneyhtälön jäännöksen ζ_i varianssi.

3.4 Aamuvireisyyden, unilaadun ja mielialan polkumalli

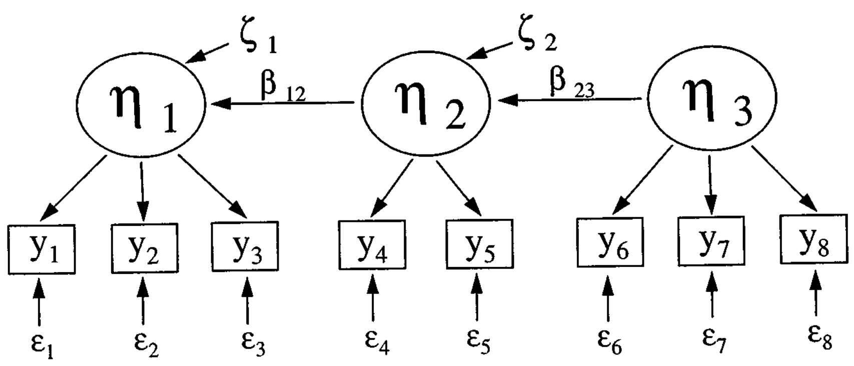
Tarkastellaaan taulukossa 3.1 esitettyjä faktorimallin estimointituloksia pojilla, joiden PTT on huono. Faktoreiden kovarianssimatriisin ja standardoidun kovarianssimatriisin perusteella voidaan olettaa, että faktoreiden välillä on olemassa kausaalisuhteita. Lähdetään seuraavaksi tarkastelemaan luvussa 2 esitettyjen faktoreiden välisiin yhteyksiin liittyvien tutkimushypoteesin 3 pohjalta polkumallin rakentamista vaiheittain taloudellisesti hyvin toimeen tulevien perheiden poikien osaryhmälle. Rakennettava malli kuvata kuvion 3.2 graafin avulla. Malli koostuu rakenneyhtälöistä, jotka ovat muotoa

$$\begin{cases} \eta_1 = \beta_{12}\eta_2 + \zeta_1, \\ \eta_2 = \beta_{23}\eta_3 + \zeta_2, \\ \eta_3 = \zeta_3. \end{cases}$$

Yhtälöiden faktorit η_i, i = 1,2,3, määritellään kuten kappaleessa 3.2.

Mallin rakenneparametrimatriisi B on muotoa

$$B = \begin{bmatrix} 0 & \beta_{12} & 0 \\ 0 & 0 & \beta_{23} \\ 0 & 0 & 0 \end{bmatrix}.$$



y_1 = Aamuvireisyys -indeksi
y_2 = Väsymyksen useus kouluaamuisin
y_3 = Aamupirteyskokemus
y_4 = Unilaatu –indeksi
y_5 = Univaikeuksien ongelmallisuusaste -indeksi
y_6 = Masentuneisuus
y_7 = Ärtyneisyys
y_8 = Hermostuneisuus

η_1 = Aamuvireisyys
η_2 = Unilaatu
η_3 = Mieliala

<u>**Kuvio 3.2**</u> Kolmen faktorin polkumalli

Malli tuottaa faktoreille kovarianssirakenteen

$$\boldsymbol{\Omega} = \begin{bmatrix} \omega_{11} & & \\ \omega_{21} & \omega_{22} & \\ \omega_{31} & \omega_{32} & \omega_{33} \end{bmatrix} = \begin{bmatrix} \psi_{11} + \beta_{12}^2\psi_{22} + \beta_{12}^2\beta_{23}^2\psi_{33} & & \\ \beta_{12}\psi_{22} + \beta_{12}\beta_{23}^2\psi_{33} & \psi_{22} + \beta_{23}^2\psi_{33} & \\ \beta_{21}\beta_{23}\psi_{33} & \beta_{23}\psi_{33} & \psi_{33} \end{bmatrix}, \tag{3.22}$$

koska $\boldsymbol{\Omega} = (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I} - \mathbf{B}^T)^{-1}$.

Kovarianssirakenteen perusteella voidaan lähteä tarkastelemaan spesifioidun faktoreiden polkumallin identifioituvuutta mallin faktoreille tuottaman teoreettisen kovarianssimatriisin perusteella. Mallin faktorirakenteen identifioituvuus on todettu edellä kappaleessa 3.2. Huomattakoon, että mallin rakenneyhtälöosassa on ainoastaan rekursiivisia rakenneyhtälöitä ja jäännösten kovarianssi matriisi $\boldsymbol{\Psi}$ on diagonaalinen (jäännökset ζ_i, i=1,2,3, eivät korreloi). Tämän vuoksi tiedetään ilman identifioituvuustarkasteluita, että myös rakenneyhtälöosa on identifioituva. Tarkastellaan seuraavaksi lyhyesti, kuinka spesifioitu malli voidaan osoittaa identifioituvaksi.

Todetaan aluksi, että välttämättömät ehdot spesifioidulle mallille ovat voimassa. Spesifioidussa mallissa on 5 vapaasti estimoitavaa parametria, kun taas kovarianssimatriisissa $\boldsymbol{\Omega}$ kovariansseja ja varriansseja on yhteensä 6 kappaletta.

Spesifioidun mallin identifioituvuus voidaan osoittaa identifioituvuuden määritelmän perusteella. Identifioituvuuden osoittaminen aloitetaan kirjoittamalla auki kovarianssimatriisin $\boldsymbol{\Omega}$ kovarianssit ja hakemalla tätä kautta mallin tuntemattomille parametreille esitysmuodot, joissa ne saadaan esitettyä kovarianssien ω avulla.

Matriisiyhtälön 3.22 perusteella huomataan, että mallin estimoitavat parametrit voidaan esittää kovarianssimatriisin $\boldsymbol{\Omega}$ alkioiden avulla ja siten voidaan todeta identifioituvuuden määritelmän perusteella, että spesifioitu malli on identifioituva.

Taulukko 3.3 Aamuvireisyyden, unilaadun ja mielialan faktoreiden polkumallin estimointitulokset, pojilla joidenperheen taloudellinen tilanne on huono

Aamuvireisyyden η_1, unilaadun η_2 ja mielialan η_3 faktoreiden polkumallin latausten Λ, jäännösvarianssien ja kovarianssien θ, reliabiliteettien $\hat{R}^2_{y(i)}$ estimaatit (suluissa keskivirheet).

	$\hat{\Lambda}$ (s.e.)			Θ (s.e.)	$\hat{R}^2_{y(i)}$
	η_1	η_2	η_3		
Aamupirteys	1*	0*	0*	.26(.04)	.61
Aamuväsymys	-.94(.11)	0*	0*	.42(.05)	.47
Aamuvireisyys	.71(.08)	0*	0*	.21(.03)	.50
Uniongelmat	0*	-1*	0*	.38(.08)	.52
Unilaatu	0*	.54(.11)	0*	.38(.04)	.24
Hermostuneisuus	0*	0*	1*	.15(.02)	.53
Ärtyneisyys	0*	0*	.93(.12)	.16(.02)	.47
Masentuneisuus	0*	0*	.91(.13)	.24(.03)	.37

*) kiinnitetty

Rakenneparametrit B, jäännösvarianssit Ψ ja selitysasteet $\hat{R}^2_{\eta(i)}$:

$\hat{\beta}_{12}$ =.59(.12), $\hat{\beta}_{23}$ =-.77(.16)

$\hat{\psi}_{11}$ =.27(.06), $\hat{\psi}_{22}$ =.31(.08)

$\hat{R}^2_{\eta_1}$ =.35, $\hat{R}^2_{\eta_2}$ =.25

Riittävyysindeksit:

$\chi^2(18) = 18.86$, P = 0.40, N =265 , CN = 488.15

NFI = .97 GFI = .98 RMSEA = .013 RMR = .022

$|NR_{max}| = 2.55$, $MI_{max} = MI(\theta_{31}) = 4.61$

Yleisen pienimmän neliösumman estimointitulokset mallille pojilla, joiden PTT on huono, on esitetty taulukossa 3.3. Malli voidaan todeta siinä määrin riittäväksi kuin ryhmälle rakennettu faktorimalli. Mallin rakenneparametrit ovat erittäin merkitseviä. Kuitenkin selitysasteet ovat suhteellisen pieniä: unilaatu selittää aamuvireisyyden vaihtelusta 35 % ja mieliala unilaadun vaihtelusta 25 %. Siten faktoreiden jäännösvarianssit suhteessa faktoreiden variansseihin eivät ole kovin pieniä. Estimointitulosten perusteella voidaan tutkia mielialan yhteyttä aamuvireisyyteen laskemalla mielialan epäsuora yhteys aamuvireisyyteen, joka on tulo

$$\hat{\beta}_{12}\,\hat{\beta}_{23} = -.46, \text{ keskivirhe s.e} = .11 \,.$$

Tämän perusteella mielialan yhteys aamuvireisyyteen unilaadun kautta on erittäin merkitsevä.

4 RYHMÄVERTAILUT

Konfirmatorisia identifioituvia faktorimalleja ja niiden rakenneyhtälömalleja voidaan rakentaa yhtä aikaa usealle ryhmälle eli perusjoukolle ja suorittaa ryhmien välisiä mallivertailuja (Jöreskog 1971). Vertailut tehdään siten, että ryhmien välille tehdään yhtäsuuruuskiinnityksiä, jotka voivat kohdistua kokonaisiin parametrimatriiseihin tai yksittäisiin parametreihin. Yleisesti yhtä aikaa eri ryhmille rakennettujen mallien voimassaolo testataan χ^2 -testillä kuten yhden ryhmän tapauksessakin. Yhtäsuuruuskiinnitysten voimassaolo testataan tavallisesta χ^2 -testistä muodostetulla χ^2-peräkkäis-testillä.

Ryhmävertailuja koskevat teoreettiset tarkastelut tässä luvussa perustuvat julkaisuun Leskinen (1987) (sivut 66 - 143).

Rakennettaessa faktorimalleja usealle ryhmälle yhtä aikaa oletetaan, että kaikista perusjoukoista $g = 1,...,G$ on poimittu riippumattomat otokset ja jokaisessa ryhmässä on voimassa faktorimalliesitykset

$$y_g = \Lambda_g \eta_g + \varepsilon_g , \; g = 1,...G. \tag{4.1}$$

Tällöin havaittujen muuttujien teoreettiset kovarianssimatriisit ovat muotoa

$$\Sigma_g = \Lambda_g \operatorname{cov}(\eta)_g \Lambda_g^T + \Theta_g , \; g = 1,...G. \tag{4.2}$$

Faktoreiden rakenneyhtälömalleja rakennettaessa usealle ryhmälle yhtä aikaa oletetaan (4.1):n lisäksi, että kaikille ryhmille on voimassa rakenneyhtälöesitykset

$$\eta_g = B_g \eta_g + \zeta_g \tag{4.3}$$

Tällöin oletetaan myös, että ryhmille, joille on voimassa rakenneyhtälöesitykset (4.3), on voimassa rakenneyhtälöesitysten supistetut muodot. Tällöin (4.2):ssa näiden ryhmien osalta faktoreiden kovarianssimatriisit ovat muotoa

$$\mathrm{cov}(\eta)_g = \left(I - B_g\right)^{-1} \Psi_g \left(I - B_g^T\right)^{-1}. \tag{4.4}$$

4.1 Konfirmatoristen faktorimallien rakentaminen usealle ryhmälle samanaikaisesti

Konfirmatoristen faktorimallien rakentaminen samanaikaisesti usealle ryhmälle kannattaa yleensä suorittaa seuraavassa järjestyksessä: ensin rakennetaan kullekin ryhmälle erikseen faktorimallit. Yleensä pyritään rakentamaan kaikille ryhmille rakenteeltaan samanlainen faktorimalli, mikä mahdollistaa eri ryhmien faktorimallien parametrien yhtäsuuruusvertailut. Tämän jälkeen estimoidaan rakennetut mallit yhtä aikaa kaikille ryhmille.

Rakennettaessa faktorimalleja yhtäaikaa usealle ryhmälle, käydään läpi samoja vaiheita kuin yhden ryhmän tapauksessa. Identifioituvuustarkasteluissa riittää yhden ryhmän identifioituvuuden tarkastelu, jos kaikille ryhmille on spesifioitu rakenteeltaan samanlainen faktorimalli, muussa tapauksessa identifioituvuustarkastelut on tehtävä kunkin ryhmän osalta erikseen.

Parametrien estimointi tapahtuu samojen periaatteiden mukaan kuin yhden ryhmän tapauksessa. Minimoitava kohdefunktio on usean ryhmän tapauksessa muotoa

$$F = (N_1/N)F_1 + (N_2/N)F_2 + \ldots + (N_G/N)F_G, \tag{4.5}$$

missä F_1, F_2, ... , F_G ovat valitun estimointimenetelmän minimoitavat kohdefunktiot kussakin ryhmässä. N_1, N_2, ... , N_G ovat eri ryhmien otoskoot ja N on näiden otoskokojen summa. Jos ryhmien välillä ei ole yhtäsuuruuskiinnityksiä, tuottaa kohdefunktion minimointi kullekin ryhmälle samat estimointitulokset kuin yhden ryhmän tapauksessa.

Ryhmävertailuissa tulisi välttää tilanteita, joissa jonkun ryhmän otoskoko eroaa hyvin paljon muiden ryhmien otoskoista, koska tällöin suuremman otoskoon omaavat ryhmät dominoivat estimointia kun taas pienempien ryhmien mukanaolo ei vaikuta kovin paljon estimointiin.

Samanaikaisesti estimoitujen mallien yhteistä tilastollista riittävyyttä χ^2 -testillä, jonka testisuure on

$$\chi^2 = (N-1)F_{min}, \tag{4.6}$$

missä F_{min} on kohdefunktion minimi. Testisuure noudattaa asymptoottisesti χ^2 -jakaumaa vapausastein Gp(p+1)/2 - t, missä t on vapaasti estimoitavien parametrien lukumäärä.

Yhtäsuuruushypoteeseja testataan χ^2 -testistä muodostetulla χ^2 -peräkkäistestillä. Oletetaan, että malli M_1 on estimoitu siten, että siinä on testattava yhtäsuuruuskiinnitys mukana ja malli M_2 on samanlainen malli kuin M_1, mutta siinä ei ole testattavaa yhtäsuuruushypoteesia mukana. M_1 on siten mallin M_2 erikoistapaus. Tällöin testattava hypoteesi on muotoa

H_0: Malli M_1 on voimassa.

Vastahypoteesina on siten

H_1: Malli M_2 on voimassa.

Tällöin χ^2 -peräkkäistestin testisuure saadaan silloin laskettua kaavalla

$$\chi^2_{21} = \chi^2_2 - \chi^2_1 \tag{4.7}$$

missä χ^2_2 on mallin M_2 χ^2 -testisuure ja χ^2_1 on mallin M_1 testisuure. Vastaavat vapausasteet ovat df_2 ja df_1. Tällöin χ^2 -peräkkäistestin testisuure noudattaa asymptoottisesti χ^2 -jakaumaa vapausastein df = df_2 - df_1.

Päätös H_0-hypoteesin hyväksymisestä tai hylkäämisestä tehdään testisuuretta vastaavan p-arvon perusteella, kuten tavallisessa χ^2 -testissä. Silloin kun havaintomäärä on niin suuri (tai pieni), että yhtäsuuruushypoteesien testausta ei voida suorittaa peräkkäistestien, voidaan yhtäsuuruushypoteesien voimassaoloa arvioida ns. informaatiokriteerien avulla, esimerkiksi Akaiken informaatiokriteerin (AIC:n) avulla (Akaike, 1987).

4.1.1 Faktorimallien yhtäsuuruusvertailut

Ennen faktorimallin rakentamista samanaikaisesti eri ryhmille voidaan testata, ovat kovarianssimatriisit yhtä suuria eri perusjoukoissa. Testattava H_0 -hypoteesi on muotoa

$$H_0: \Sigma_1 = \Sigma_2 = \ldots = \Sigma_G ,$$

jolloin vastahypoteesi on

H_1: Kovarianssimatriisit eivät ole yhtäsuuria.

Mikäli H_0 -hypoteesi hyväksytään, voidaan usean ryhmän faktorimallien rakentaminen palauttaa yhden perusjoukon faktorimallin rakentamiseen, jolloin faktorimallin rakentaminen perustuu ryhmien yhdistettyyn kovarianssimatriisiin. Näin ei kuitenkaan välttämättä kannata menetellä, sillä kovarianssimatriisien yhtäsuuruus ei merkitse sitä, että eri ryhmien faktorimallit olisivat täsmälleen yhtäsuuret.

Faktorimallien vertaileminen aloitetaan yleensä testaamalla, ovatko latausmatriisit yhtäsuuria eri ryhmissä, jolloin H_0 -hypoteesi on

$$H_{0\Lambda}: \Lambda_1 = \Lambda_2 = \ldots = \Lambda_G.$$

Nollahypoteesin mukaisen faktorimallin voimassaoloa voidaan testata tavallisella χ^2 -tekstillä kuten yhden ryhmän faktorimallin tapauksessakin, jolloin vastahypoteesina on, että kovarianssimatriisissa Σ ei ole rajoituksia. Latausmatriisien yhtäsuuruutta testattaessa vastahypoteesi on

$$H_{1\Lambda}: y_g = \Lambda_g \eta_g + \varepsilon_g .$$

Latausmatriisien yhtäsuuruus on ehtona sille, että faktorimallin vertailut olisivat muilta osin mielekkäitä. Jos hypoteesi $H_{0\Lambda}$ hyväksytään, voidaan testata esimerkiksi jäännösten kovarianssimatriisien yhtäsuuruutta eli hypoteesin

$$H_{0\Theta}: \Theta_1 = \Theta_2 = ... = \Theta_G$$

voimassaoloa. Myös faktoreiden kovarianssimatriisien yhtäsuuruutta voidaan testata, jolloin testattava nollahypoteesi on

$$H_{0\Omega}: \Omega_1 = \Omega_2 = ... = \Omega_G.$$

Mikäli eri ryhmille samanaikaisesti rakennettujen faktorimallien faktoreiden välille rakennetaan rakenneyhtälöitä, toimii faktoreiden kovarianssimatriisien yhtäsuuruuden testaus esitestinä rakenneyhtälömallien yhtäsuuruusvertailuille.

Yhtäsuuruuden testaus voidaan tarvittaessa kohdistaa myös mallin yksittäisiin parametreihin. Hypoteesien $H_{0\Theta}$, $H_{0\Omega}$ ja matriisien Θ ja Λ yksittäisiin parametreihin kohdistuvien yhtäsuuruushypoteesien testaamisjärjestyksellä ei ole väliä.

4.1.2 Faktoreiden rakenneyhtälömallien yhtäsuuruusvertailut

Faktoreiden rakenneyhtälömallien rakentaminen samanaikaisesti eri ryhmille tapahtuu samojen periaatteiden mukaan kuin faktorimallienkin samanaikaisesti rakentaminen. Yleensä rakenneyhtälöt kannattaa rakentaa faktorimallien yhtäsuuruusvertailujen jälkeen.

Rakenneyhtälömallien yhtäsuuruusvertailujen ehtona on latausmatriisien yhtäsuuruus eli hypoteesin H_0 voimassaolo. Rakenneyhtälömallien ryhmävertailuissa keskeisin vertailutilanne on rakenneparametrimatriisien B yhtäsuuruuden testaaminen, jolloin testattava nollahypoteesi on

$$H_{0B}: B_1 = B_2 = ... = B_G.$$

Vastahypoteesina on malli, jossa parametrimatriisit ovat erisuuruisia. Tämän lisäksi voidaan testata nollahypoteesia

$$H_{0\Psi}: \Psi_1 = \Psi_2 = \ldots = \Psi_G$$

eli rakenneyhtälömallin jäännösten kovarianssimatriisin yhtäsuuruutta.

4.2 Aamuvireisyyden, unilaadun ja mielialan faktorimallien ja polkumallien ryhmävertailut

Tarkastellaan seuraavaksi, kuinka konfirmatoristen faktorimallien ja faktoreiden rakenneyhtälömallien ryhmävertailut etenevät käytännössä. Valitaan vertailtaviksi ryhmiksi kaksi ryhmää: pojat, joiden PTT on hyvä ja pojat, joiden PTT on huono. Pohjana näille tarkasteluille ovat otoskovarianssimatriisit liitteessä 1, ryhmille erikseen rakennetut faktorimallit kappaleessa 3.2 ja kappaleessa 3.4 rakennettu polkumalli pojilla, joiden PTT on huono.

Ennen faktorimallia ja polkumallia koskevia ryhmävertailuja testataan, ovatko tarkasteltavien ryhmien otoskovarianssimatriisit samansuuruiset. Tällöin testattava H_0 -hypoteesi on

$$H_0: \Sigma_1 = \Sigma_2 .$$

Kun otoskovarianssimatriisien yhtäsuuruutta olettavaa hypoteesia testataan yleistä H_1 -hypoteesia

$$H_1: \Sigma_1 \neq \Sigma_2$$

vastaan, saadaan χ^2-testin tulos $\chi^2(36) = 45.69$, $p = .129$, minkä perusteella hypoteesi otoskovarianssimatriisien yhtäsuuruudesta voidaan hyväksyä. Kuitenkin $MI_{MAX} = MI(\sigma_{86}^{(1)}) = MI(\sigma_{86}^{(2)}) = 10.18$, mikä osoittaa, että hermostuneisuuden ja masentuneisuuden välisten kovarianssien $\sigma_{86}^{(1)}$ ja $\sigma_{86}^{(2)}$ osalta hypoteesi kovarianssimatriisien yhtäsuuruudesta ei välttämättä sovi. Tämä näkyy jo eroavuutena kappaleessa 3.2 rakennettujen faktorimallien jäännösten kovarianssirakenteessa.

Tämän tuloksen perusteella voidaan olettaa, että tarkasteltaville ryhmille saadaan rakenteeltaan suhteellisen samansuuruiset faktori- ja polkumallit, lukuunottamatta havaittujen muuttujien jäännösten kovarianssirakennetta.

Taulukossa 4.1 on esitetty edellä tarkasteltujen otoskovarianssimatriisien pohjalta ryhmille yhtäaikaa rakennettu kolmen faktorin malli. Estimointitulokset tilastollisessa mielessä vastaavat ryhmille kohdassa 3.2 erikseen rakennettujen faktorimallien estimointituloksia. Testattaessa mallia yleistä H_1 -hypoteesia vastaan, saadaan χ^2 -testin tulokseksi $\chi^2(33) = 36.39$, $p = .31$. Tuloksen perusteella voidaan estimoitu malli hyväksyä.

Aloitetaan ryhmävertailut testaamalla, ovatko latausrakenteet yhtäsuuret. Testattava hypoteesi on

$$H_{0\Lambda}\colon \Lambda_1 = \Lambda_2 .$$

Kun valittua hypoteesia testataan yleistä vastahypoteesia H_1 vastaan, saadaan tulos $\chi^2(38) = 38.96$, $p = .43$. Asettamalla vastahypoteesi

$$H_{1\Lambda}\colon y_g = \Lambda_g \eta_g + \varepsilon_g ,\ g = 1,2$$

eli ryhmille yhtäaikaa rakennettu malli ilman yhtäsuuruussidoksia, voidaan testata, onko yhtäsuuruussidos $\Lambda_1 = \Lambda_2$ voimassa. Peräkkäistestin tulokseksi saadaan $\chi^2_\Lambda = \chi^2_{0\Lambda} - \chi^2_{1\Lambda} = \chi^2(5) = 2.57$, $p = .766$. Testin tuloksen perusteella voidaan hyväksyä asetettu yhtäsuuruushypoteesi.

Seuraavaksi voidaan testata mallin jäännösvarianssien yhtäsuuruutta asettamalla hypoteesi

$$H_{0\Theta}\colon \Theta_1 = \Theta_2 .$$

Kun vastahypoteesiksi asetetaan aikaisemmin hyväksytty $H_{0\Lambda}$ ja χ^2 -testin tulokseksi yleistä vastahypoteesia H_1 vastaan testattaessa $\chi^2(46) = 46.46$, $p = .39$, saadaan jäännöskovarianssit yhtäsuuriksi olettavalle $H_{0\Theta}$ -hypoteesille laskettua peräkkäistestitulos $\chi^2_{\Theta\Lambda} = \chi^2_{0\Theta} - \chi^2_{0\Lambda} = \chi^2(8) = 6.60$, $p = .580$. Tuloksen perusteella voidaan todeta jäännöskovarianssien olevan yhtäsuuria ryhmien välillä.

Taulukko 4.1 Aamuvireisyyden, unilaadun ja mielialan faktorimallin estimointitulokset pojilla.

Aamuvireisyyden η_1, unilaadun η_2 ja mielialan η_3 faktorimallin latausten Λ, jäännösvarianssien ja -kovarianssien θ estimaatit (suluissa keskivirheet).

	Λ						$\Theta^{(1)}$ **	$\Theta^{(2)}$
	$\lambda_1^{(1)}$	$\lambda_1^{(2)}$	$\lambda_2^{(1)}$	$\lambda_2^{(2)}$	$\lambda_3^{(1)}$	$\lambda_3^{(2)}$		
Aamupirteys	1*	1*	0*	0*	0*	0*	31(.05).	.26(.04)
Aamuväsymys	-1.13(.11)	-.94(.11)	0*	0*	0*	0*	.25(.05)	.42(.05)
Aamuvireisyys	.73(.07)	.71(.08)	0*	0*	0*	0*	.22(.03)	.21(.03)
Uniongelmat	0*	0*	-1*	-1*	0*	0*	.42(.10)	.40(.08)
Unilaatu	0*	0*	.50(.13)	.56(.11)	0*	0*	.38(.05)	.38(.04)
Hermostuneisuus	0*	0*	0*	0*	1*	1*	.18(.05)	.15(.02)
Ärtyneisyys	0*	0*	0*	0*	.71(.18)	.91(.12)	.20(.03)	.17(.02)
Masentuneisuus	0*	0*	0*	0*	.93(.20)	.90(.12)	.20(.05)	.24(.03)

*) kiinnitetty

**) Vapautettu $\theta_{86}^{(1)} = -.11(.04)$

Faktoreiden kovarianssit:

	$\hat{\Psi}^{(1)}$ (s.e.)			$\hat{\Psi}^{(2)}$ (s.e.)		
	$\eta 1$	$\eta 2$	$\eta 3$	$\eta 1$	$\eta 2$	$\eta 3$
$\eta 1$	.53(.09)			.41(.07)		
$\eta 2$	.29(.06)	.37(.11)		.24(.04)	.38(.10)	
$\eta 3$	-.14(.03)	-.16(.04)	.19(.06)	-.07(.02)	-.14(.03)	.18(.03)

Riittävyysindeksit:

$\chi^2(33) = 36.39$, $P = .31$, $N = 449$,

GFI = .98, RMSEA = .015, RMR = .021

Vastaavasti voidaan testata faktoreiden kovarianssimatriisien yhtäsuuruutta asettamalla hypoteesi

$$H_{0\Omega}: \mathbf{\Omega}_1 = \mathbf{\Omega}_2 .$$

Testitulokseksi yleistä H_1 -hypoteesia vastaan testattaessa saadaan $\chi^2(44) = 47.66$, $p = .33$. Kun vastahypoteesiksi asetettaan edellä hyväksytty hypoteesi $H_{1\Lambda}$, voidaan hypoteesin $H_{0\Omega}$ mukaiselle yhtäsuuruusoletukselle laskea peräkkäistestitulos $\chi^2_{\Omega\Lambda} = \chi^2_{0\Omega} - \chi^2_{1\Lambda} = \chi^2(6) = 8.70$, $p = .191$. Tämän perusteella voidaan hyväksyä hypoteesi faktoreiden kovarianssimatriisien yhtäsuuruudesta.

Tarkastellaan seuraavaksi ryhmille yhtäaikaa rakennettu polkumallia yhtäsuuruuskiinnityksellä $\Lambda_1 = \Lambda_2$ taulukossa 4.2 ja sen ryhmävertailuja. Edellä todettiin, että faktoreiden kovarianssimatriisit ovat yhtäsuuret. Siksi on odotettavissa, että kolmen faktorin polkumallin rakenneparametrit ovat yhtäsuuret.

Testattaessa mallia yleistä vastahypoteesia H_1 vastaan, saadaan χ^2 -testin tulokseksi $\chi^2(40) = 39.58$, $p = .49$. Tämän perusteella voidaan todeta, että $H_{0\Lambda}$ -hypoteesin mukainen malli on voimassa yhtäaikaa molemmissa ryhmissä.

Rakenneparametrien yhtäsuuruus voidaan testata asettamalla hypoteesi

$$H_{0B}: \mathbf{B}_1 = \mathbf{B}_2 .$$

Testattaessa H_{0B} -hypoteesin mukaista mallia yleistä H_1 -hypoteesia vastaan, saadaan testitulos $\chi^2(42) = 44.12$, $p = .38$. Kun vastahypoteesiksi valitaan edellä hyväksytyn mallin mukainen hypoteesi eli

$H_{0\Lambda}$: kolmen faktorin polkumalli, jossa kiinnitys $\Lambda_1 = \Lambda_2$,

voidaan laskea peräkkäistestitulos $\chi^2_{B\Lambda} = \chi^2_{0B} - \chi^2_{1\Lambda} = \chi^2(2) = 4.54$, $p = .103$, minkä perusteella voidaan hyväksyä hypoteesi rakenneparametrien yhtäsuuruudesta.

Seuraavaksi voidaan testata jäännösten varianssien yhtäsuuruutta asettamalla hypoteesi

$$H_{0\Psi}: \Psi_1 = \Psi_2 .$$

Testitulokseksi yleistä H_1 -hypoteesia vastaan saadaan $\chi^2(44) = 47.58$, $p = .33$. Peräkkäistestin tulokseksi H_{0B} -hypoteesia vastaan saadaan $\chi^2_{\Psi B} = \chi^2_{0\Psi} - \chi^2_{1B} = \chi^2(2) = 3.46$, $p = .177$. Siten hypoteesi faktoreiden jäännösten varianssien yhtäsuuruudesta voidaan hyväksyä.

Taulukko 4.2 Aamuvireisyyden η_1, unilaadun η_2 ja mielialan η_3 faktorimallin latausten Λ ja kovarianssien θ sekä reliabiliteettien estimaatit (suluissa keskivirheet).

	$\hat{\Lambda}^{(1)} = \hat{\Lambda}^{(2)}$			$\hat{R}^{2(1)}_{y(i)}$	$\hat{R}^{2(2)}_{y(i)}$
	η_1	η_2	η_3		
Aamupirteys	1*	0*	0*	.66	.59
Aamuväsymys	-1.02(.07)	0*	0*	.69	.51
Aamuvireys	.71(.05)	0*	0*	.57	.48
Uniongelmat	0*	-1*	0*	.42	.53
Unilaatu	0*	.53(.08)	0*	.19	.24
Hermostuneisuus	0*	0*	1*	.43	.55
Ärtyneisyys	0*	0*	.88(.10)	.39	.45
Masentuneisuus	0*	0*	.90(.11)	.35	.38

*) kiinnitetty

Vapautettu $\theta^{(1)}_{86} = -.09(.03)$

$\hat{\beta}^{(1)}_{12} = .89(.18)$, $\hat{\beta}^{(1)}_{23} = -.96(.20)$, $\psi^{(1)}_{11} = .32(.07)$, $\psi^{(1)}_{22} = .18(.06)$

$\hat{\beta}^{(2)}_{12} = .57(.11)$, $\hat{\beta}^{(2)}_{23} = -.77(.15)$, $\psi^{(2)}_{11} = .25(.05)$, $\psi^{(2)}_{22} = .32(.08)$

Mallin riittävyysindeksit:

$\chi 2(40) = 39.58$, $P = .49$, $N = 449$, $CN = 720.37$

NFI = .97 GFI = .98 RMSEA = .0 RMR = .021

5 FAKTOREIDEN TASOVERTAILUMALLIT

Tähän asti konfirmatoristen faktorimallien tarkastelu on kohdistunut faktorimalleihin, joissa havaittujen muuttujien ja faktoreiden odotusarvoja ei ole parametrisointu. Tarkastellaan nyt lähinnä ryhmävertailujen näkökulmasta faktorimalleja, joissa on mukana havaittujen muuttujien ja faktoreiden odotusarvoparametrisoinnit. Teoreettiset tarkastelut kohdissa 5, 5.1 ja 5.4 perustuvat julkaisuun Leskinen (1987) (sivut 203 - 225).

Tasoparametrinen faktorimalli on muotoa

$$\mathbf{y} = \tau + \Lambda\eta + \varepsilon, \tag{5.1}$$

jossa

$$E\eta = \kappa \text{ ja } E\varepsilon = \mathbf{0}$$

sekä

$$\mathrm{cov}(\eta) = \Omega \text{ ja } \mathrm{cov}(\varepsilon) = \Theta.$$

Silloin havaittujen muuttujien odotusarvovektori on muotoa

$$E\mathbf{y} = \mu = \tau + \Lambda\kappa \tag{5.2}$$

ja havaittujen muuttujien kovarianssimatriisiesitys on muotoa

$$\Sigma = E(\mathbf{y} - \mu)(\mathbf{y} - \mu)^T = \Lambda\,\Omega\,\Lambda^T + \Theta. \tag{5.3}$$

Odotusarvovektorin μ parametrisoinnissa τ -parametrit kuvaavat havaittujen muuttujien konstruoinnista aiheutunutta mittaamisen yleistasoa tasoa ja komponentti $\Lambda\kappa$ faktorirakenteen aiheuttamaa tasoa. Odotusarvovektori μ sisältää p yleistasoa τ_i (i = 1,...,p) ja m faktoreiden odotusarvoa κ_j (j = 1,...,m). Siten odotusarvoparametrisointi ei ole identifioituva ilman rajoituksia.

Yleensä kahden tai useamman ryhmän tapauksessa tai tilanteissa, joissa yhdestä muuttujasta on havaintoja usealta ajankohdalta, voidaan tasoparametrisointia pitää käyttökelpoisena laajennuksena. Silloin odotusarvoparametrisointi saadaan identifioituvaksi asettamalla malliin tulkinnan kannalta sopivia rajoituksia.

Seuraavaksi tarkastellaan kappaleessa 5.1 odotusarvojen parametrisoinnin identifioituvuutta yhden faktorin mallin ja neljän ryhmän tapauksessa. Tämän jälkeen kappaleissa 5.4 ja 5.5 tarkastellaan tilannetta, jossa neljä ryhmää muodostavat kaksisuuntaisen varianssianalyysin tapaan kaksisuuntaisen tasovertailutilanteen. Kappaleissa 5.4 ja 5.5 tarkastellaan tasoparametrista faktoreiden rakenneyhtälömallia ja sitä, kuinka rakenneparametrisoinnin avulla voidaan selittää faktoreiden tasoeroja.

5.1 Odotusarvoparametrisoinnin identifioituvuus

Tarkastellaan neljän ryhmän tapauksessa yhden faktorin mallia

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} + \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} \eta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix},$$

jossa $E\eta = \kappa$.

Odotusarvovektorit ovat silloin muotoa

$$\eta^{(g)} = \tau^{(g)} + \Lambda^{(g)} \kappa^{(g)}, \quad g = 1,2,3,4 .$$

Yleisin tapa saada τ- ja κ- parametrit identifioituviksi on asettaa yhtäsuuruusrajoitus

$$\tau^{(1)} = \tau^{(2)} = \tau^{(3)} = \tau^{(4)} =: \tau$$

ja asettaa jossakin ryhmässä $\kappa^{(g)} = 0$, esimerkiksi kun g = 1. Silloin odotusarvovektori $\eta^{(g)}$ on muotoa

$$\mu^{(1)} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} \quad \text{ja} \quad \mu^{(g)} = \begin{bmatrix} \tau_1 + \lambda_1^{(g)} \kappa^{(g)} \\ \tau_2 + \lambda_2^{(g)} \kappa^{(g)} \\ \tau_3 + \lambda_3^{(g)} \kappa^{(g)} \end{bmatrix}, \text{ kun } g = 2,3,4.$$

Lisäksi asetetaan rajoitus

$$\Lambda^{(1)} = \Lambda^{(2)} = \Lambda^{(3)} = \Lambda^{(4)}.$$

Nyt parametrit τ_1, τ_2, τ_3, $\kappa^{(2)}$, $\kappa^{(3)}$ ja $\kappa^{(4)}$ ovat identifioituvia, sillä ne voidaan ratkaista odotusarvovektoreiden $\mu^{(1)}$, $\mu^{(2)}$, $\mu^{(3)}$, $\mu^{(4)}$ (eli 12 odotusarvon) avulla.

Tällaisessa odotusarvoparametrisoinnissa parametrit $\kappa^{(2)}$, $\kappa^{(3)}$ ja $\kappa^{(4)}$ tulkitaan tasoeroiksi suhteessa ryhmään 1, jonka taso on kiinnitetty kiinteäksi vertailutasoksi. Mahdolliset odotusarvoerot ryhmien välillä oletetaan johtuvan faktorin tasosta κ.

Asetettujen rajoitusten voimassaoloa voidaan arvioida modifikaatioindeksien ja χ^2 -peräkkäistestien avulla.

5.2 Faktoreiden kaksisuuntainen tasovertailumalli

Tarkastellaan kappaleessa 5.1 esitettyä tasoparametrista yhden faktorin mallia. Laajennetaan mallin ryhmien tasoparametrien vertailu vastaamaan kaksisuuntaista varianssianalyysia. Siksi tässä kappaleessa mallista käytetään nimitystä kaksisuuntainen tasovertailumalli.

Käytetään nyt tasoparametreista $\kappa^{(g)}$ merkintää α_{AB} (A = 1,2 ja B = 1,2) eli

$$\kappa^{(1)} = \alpha_{11},\ \kappa^{(2)} = \alpha_{12},\ \kappa^{(3)} = \alpha_{21},\ \kappa^{(4)} = \alpha_{22} \tag{5.4}$$

Silloin kyseessä on kaksisuuntainen tasovertailutilanne, jossa voidaan testata suuntien A ja B yhdysvaikutuksen olemassaoloa ja mahdollisesti suuntien A ja B päävaikutuksia, mikäli tilastollisesti merkitsevää yhdysvaikutusta ei ole. Suuntien A ja B vaikutuksia voidaan arvioida tarkastelemalla graafisesti α -tasoja (ns. keskiarvoprofiilit).

Yhdysvaikutuksen olemassaoloa voidaan testata yhdensuuntaisuusrajoitteen

$$\alpha_{21} - \alpha_{11} = \alpha_{22} - \alpha_{12} \tag{5.5}$$

avulla. Tapauksessa, jossa on asetettu rajoite $\alpha_{11} = 0$, testattava nollahypoteesi on

$$H_0 : \alpha_{21} = \alpha_{22} - \alpha_{12}$$

Vastahypoteesi on silloin

H_1: $\alpha_{21} \neq \alpha_{22} - \alpha_{12}$ eli suunnilla A:lla ja B:llä on merkitsevä yhdysvaikutus.

Tilanteessa, jossa hypoteesi keskiarvoprofiilien yhdensuuntaisuudesta voidaan hyväksyä, voidaan testata A:n ja B:n päävaikutuksen olemassaoloa. Silloin tasoparametri α_{12} voidaan tulkita suunnan A päävaikutukseksi ja tasoparametri α_{21} suunnan B päävaikutukseksi. Päävaikutusten tilastollista merkitsevyyttä voidaan arvioida t - testien avulla. Tarkemmin tilastollinen merkitsevyys voidaan testata testaamalla χ^2 -peräkkäistestin avulla rajoitteen

$$\alpha_{12} = 0 \tag{5.6}$$

ja rajoitteen

$$\alpha_{21} = 0 \tag{5.7}$$

voimassaoloa, kun on asetettu yhdensuuntaisuusrajoite $\alpha_{21} = \alpha_{22} - \alpha_{12}$. Siten testattava nollahypoteesi suunnan A päävaikutukselle voidaan esittää muodossa

$$H_{0A}: \alpha_{21} = \alpha_{22}.$$

Vastahypoteesi on silloin

$H_{1A}: \alpha_{21} = \alpha_{22} - \alpha_{12}$ eli suunnalla A on merkitsevä päävaikutus.

Vastaavasti suunnan B päävaikutukselle testattava nollahypoteesi on silloin

$$H_{0A}: \alpha_{12} = \alpha_{22}.$$

Vastahypoteesi on silloin

$H_{1A}: \alpha_{21} = \alpha_{22} - \alpha_{12}$ eli suunnalla B on merkitsevä päävaikutus.

5.3 Aamuvireisyyden tasovertailut

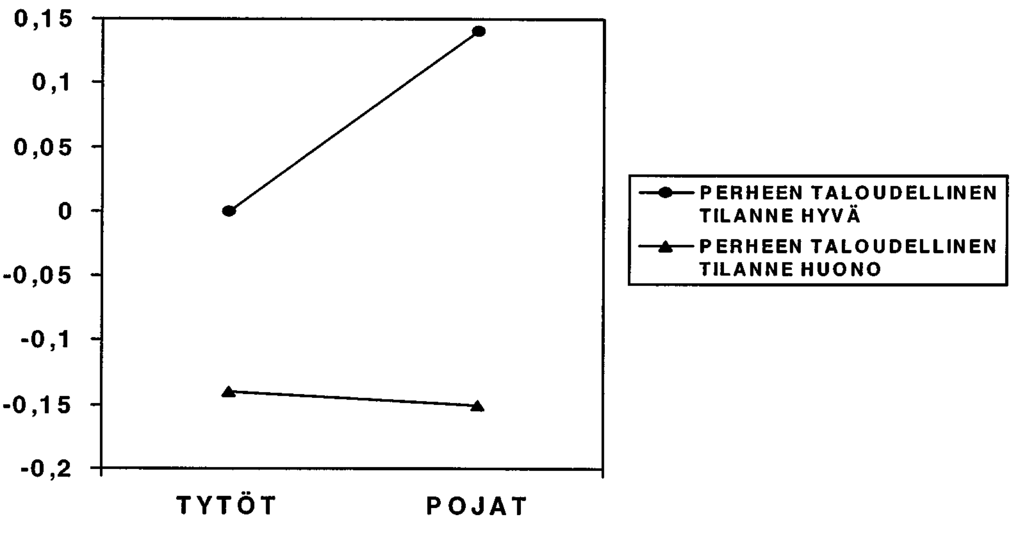
Valitaan tarkasteltavaksi neljää ryhmää: tytöt, joiden PTT on hyvä (ryhmä 1) ja huono (ryhmä 2) sekä pojat, joiden PTT on hyvä (ryhmä 3) ja huono (ryhmä 4). Rakennetaan aluksi kappaleen 5.1 parametrisoinnin mukainen tasovertailumalli aamuvireisyydelle. Tämän jälkeen tarkastellaan sukupuolen ja perheen taloudellisen tilanteen mukaisia tasoeroja kaksisuuntaisen tasovertailumallin avulla. Mallien testitulokset on esitetty taulukossa 5.1. Taulukossa 5.2 on kaksisuuntaisen tasovertailumallin estimointitulokset.

Yksinkertaisin tapa rakentaa tasoparametrinen faktorimalli aamuvireisyydelle on käyttää kappaleen 5.1 mukaista parametrisointia eli kiinnitetään aamuvireisyyden taso taloudellisesti hyvin toimeentulevien perheiden tyttöjen ryhmässä nollaksi ja havaittujen muuttujien yleistasot ryhmien välillä yhtäsuuriksi. Silloin saadaan seuraavat estimaatit κ -parametreille:

$$\kappa^{(1)} = 0, \hat{\kappa}^{(2)} = -.14\,(.07), \hat{\kappa}^{(3)} = -.14(.08), \hat{\kappa}^{(4)} = -.15\;(.07)$$

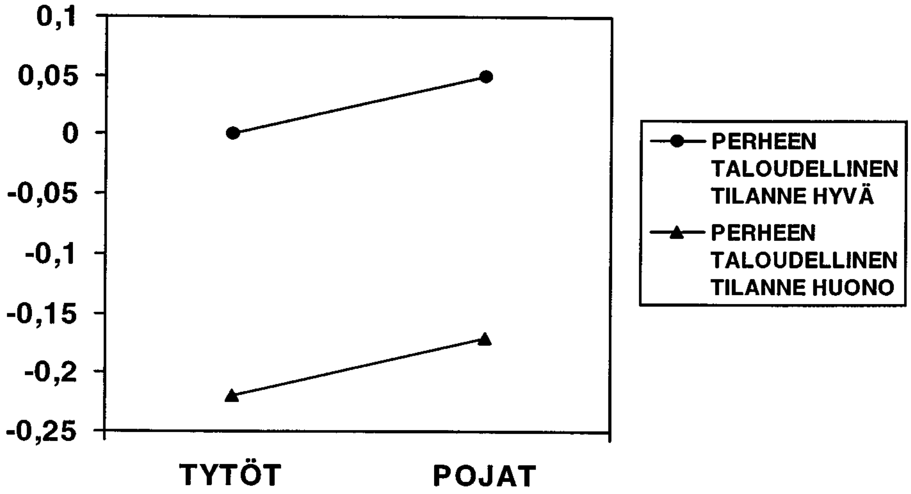
Tässä $\kappa^{(1)}$ on kiinnitetty. Estimaateista huomataan, että ryhmien 2, 3 ja 4 aamuvireisyyden tasot eroavat taloudellisesti hyvin toimeentulevien perheiden tyttöjen aamuvireisyyden tasosta. Koska kyseessä on selkeä kaksisuuntainen tasovertailutilanne, eivät tasoparametrien estimaatit edellä anna kovin selkeää kuvaa aamuvireisyyden tasoeroista.

Tutkitaan seuraavaksi aamuvireisyyden tasoeroja kaksisuuntaisen tasovertailumallin avulla, joten jatkossa käytetään tasoparametreista (5.3):n mukaisia merkintöjä. Aluksi testataan, onko sukupuolella (suunta A) ja taloudellisella tilanteella (suunta B) tilastollisesti merkitsevää yhdysvaikutusta. Kuviossa 5.1 aamuvireisyyden tasoerot on kuvattu graafisesti. Keskiarvoprofiilit eivät näytä samansuuntaisilta. Testi yhdensuuntaisuusrajoitteen $\alpha_{21} = \alpha_{22} - \alpha_{12}$ voimassaololle kuitenkin osoittaa, että sukupuolella ja perheen taloudellisella tilanteella ei ole tilastollisesti merkitsevää yhdysvaikutusta ($\chi^2(1) = 2.31,\ p = 0.13$).



Kuvio 5.1 Aamuvireisyysfaktorin sukupuolierot ja perheen taloudellisen tilanteen mukaiset tasoerot tasoparametrisen faktorimallin estimaattien $\hat{\kappa}$ avulla kuvattuna.

Estimaattien ja niiden keskivirheiden perusteella näyttäisi siltä, että sukupuolella ei ole tilastollisesti merkitsevää päävaikutusta ($\hat{\alpha}_{21}$ = .05, keskivirhe = .05), mutta perheen taloudellisella tilanteella on tilastollisesti erittäin merkitsevä päävaikutus ($\hat{\alpha}_{12}$ = -.22, keskivirhe = .05). Peräkkäistestien perusteella voidaan tehdä samat tulkinnat (ks. taulukko 5.1). Kaksisuuntaisen tasovertailumallin estimointitulokset on esitetty tarkemmin taulukossa 5.2.



Kuvio 5.2 Aamuvireisyysfaktorin sukupuolierot ja perheen taloudellisen tilanteen mukaiset erot kaksisuuntaisen tasovertailumallin estimaattien $\hat{\alpha}$ avulla kuvattuna.

Taulukko 5.1 Taloudellisesti hyvin (1) ja huonosti (2) toimeentulevien perheiden tyttöjen ja taloudellisesti hyvin (3) ja huonosti (4) toimeentulevien perheiden poikien aamuvireisyyden tasovertailujen testitulokset.

H_0 -hypoteesi	χ^2 -testi		Peräkkäistesti	
$\Lambda^{(1)} = \Lambda^{(2)} = \Lambda^{(3)} = \Lambda^{(4)}$	$\chi2(6) = 7.99$	p = .24		
Tasoparametrinen faktorimalli voimassa	$\chi2(12) = 17.75$	p = .12	$\chi2(6) = 9.76$	p = .14
$\alpha_{21} = \alpha_{22} - \alpha_{12}$	$\chi2(13) = 20.06$	p = .09	$\chi2(1) = 2.31$	p = .13
$\alpha_{21} = 0$ 1)	$\chi2(14) = 20.93$	p = .10	$\chi2(1) = 0.87$	p = .35
$\alpha_{12} = 0$ 1)	$\chi2(14) = 38.22$	p = .00	$\chi2(1) = 18.16$	p = .00

1) Verrattu malliin, jossa on yhdensuuntaisuusrajoite $\alpha_{21} = \alpha_{22} - \alpha_{12}$.

Taulukko 5.2 Taloudellisesti hyvin (1) ja huonosti (2) toimeentulevien perheiden tyttöjen ja taloudellisesti hyvin (3) ja huonosti (4) toimeentulevien perheiden poikien aamuvireisyyden (η) kaksisuuntaisen tasovertailumallin estimointitulokset, kun asetettu yhdensuuntaisuusrajoite $\alpha_{21} = \alpha_{22} - \alpha_{12}$.

	$\hat{\Lambda}^{(1)} = \hat{\Lambda}^{(2)} = \hat{\Lambda}^{(3)} = \hat{\Lambda}^{(4)}$	$\hat{\tau}^{(1)} = \hat{\tau}^{(2)} = \hat{\tau}^{(3)} = \hat{\tau}^{(4)}$
	η	
Aamupirteys	1*	2.39 (.05)
Aamuväsymys	-.94 (.05)	2.63 (.05)
Aamuvireisyys	.71(.03)	2.03(.04)

$\alpha_{11} = 0^*, \hat{\alpha}_{12} = -.22\,(.05), \hat{\alpha}_{21} = .05(.05), \hat{\alpha}_{22} = -.17\,(.07)$

*) kiinnitetty

Verrataan seuraavaksi kaksisuuntaisen tasovertailumallin tuloksia liitteessä 3 esitettyihin monimuuttujaisen varianssianalyysin (SPSS:n MANOVA, Norusis 1994) tuloksiin. Monimuuttujaiset testit osoittavat, että aamuvireisyyteen, väsymykseen kouluaamuisin ja aamupirteyteen sukupuolella ja PTT:llä ei ole tilastollisesti merkitsevää yhdysvaikutusta (p = .446). Yksimuuttujaiset testit osoittavat, että sukupuolella on melkein merkitsevä päävaikutus väsymyksen kouluaamuisin (p = .027) ja aamupirteyteen (p = .043), mutta aamuvireisyyteen ei vastaavaa päävaikutusta ole (p = .799). Monimuuttujaisten testien tulokset sukupuolen päävaikutukselle ovat melkein merkitseviä (p = .012). Yksimuuttujaiset PTT:n päävaikutukselle aamuvireisyyteen ja aamupirteyteen ovat erittäin merkitseviä. Väsymykseen kouluaamuisin PTT:n päävaikutus on tilastollisesti merkitsevä (p = .001). Monimuuttujaisen testin tulos PTT:n päävaikutukselle on melkein merkitsevä (p = .012). Tulokset ovat siten tulkinnallisesti melkein vastaavia kuin tasovertailutulokset. Pieni ero kuitenkin on sukupuolen päävaikutuksen tulkinnassa.

5.4 Tasoparametrinen faktoreiden rakenneyhtälömalli

Tarkastellaan seuraavaksi tasoparametrista faktoreiden rakenneyhtälömallia, joka on muotoa

$$\begin{cases} y = \tau + \Lambda\eta + \varepsilon \\ \eta = \alpha + B\eta + \zeta \end{cases} \tag{5.8}$$

Malli tuottaa havaittujen muuttujien ja faktoreiden odotusarvoille rakenneparametrisoinnit

$$Ey = \mu = \tau + \Lambda E\eta = \tau + \Lambda\kappa \tag{5.9}$$

jossa

$$E\eta = \kappa = (\mathbf{I}\text{-}\mathbf{B})^{-1}\alpha\,. \tag{5.10}$$

Tarkastellaan kahden ryhmän tapauksessa rakenneyhtälömallia

$$\begin{cases} \eta_1 = \alpha_1 + \beta_{12}\eta_2 + \zeta_1 \\ \eta_2 = \alpha_2 + \beta_{23}\eta_3 + \zeta_2 \\ \eta_3 = \alpha_3 + \zeta_3 \end{cases} \tag{5.11}$$

Malli tuottaa η-faktoreille odotusarvorakenteen

$$\begin{cases} E\eta_1 = \kappa_1 = \alpha_1 + \beta_{12}(\alpha_2 + \beta_{23}\alpha_3) \\ E\eta_2 = \kappa_2 = \alpha_2 + \beta_{23}\alpha_3 \\ E\eta_3 = \kappa_3 = \alpha_3 \end{cases} \tag{5.12}$$

Huomattakoon, että tämä tasoparametrinen faktoreiden rakenneyhtälömalli ei ole identifioituva ilman lisäehtoja odotusarvorakenteessa. Kahden tai useamman ryhmän tapauksessa saadaan malli identifioituvaksi asettamalla kappaleen 5.1 mukaiset lisärajoitukset odotusarvorakenteeseen. Kahden ryhmän tapauksessa asetetaan havaittujen muuttujien tasoille rajoitus

$$\tau^{(1)} = \tau^{(2)} =: \tau$$

ja asetetaan jossakin ryhmässä, esimerkiksi ensimmäisessä ryhmässä rajoitus

$$\kappa_1^{(1)} = \kappa_2^{(1)} = \kappa_3^{(1)} = 0 \Leftrightarrow \alpha_1^{(1)} = \alpha_2^{(1)} = \alpha_3^{(1)} = 0 \ .$$

Tällöin saadaan havaittujen muuttujien ja faktoreiden odotusarvoille seuraavat parametrisoinnit:

	1. ryhmä	2. ryhmä
$Ey = \mu$	τ	$\tau + \Lambda^{(2)}\kappa^{(2)}$
$E\eta_1 = \kappa_1$	0	$\alpha_1 + \beta_{12}(\alpha_2 + \beta_{23}\alpha_3)$
$E\eta_2 = \kappa_2$	0	$\alpha_2 + \beta_{23}\alpha_3$
$E\eta_3 = \kappa_3$	0	α_3

Tämän parametrisoinnin tulkinta ja ryhmävertailu voidaan suorittaa seuraavasti: aluksi oletetaan, että havaitut muuttujat mittaavat samalla tavoin η-faktoreita, jolloin niiden τ-yleistasoparametrit asetetaan yhtäsuuriksi molemmissa ryhmissä. Jos havaittujen muuttujien odotusarvoissa on ryhmien välillä eroja, kuvataan se komponentilla $\tau + \Lambda^{(2)}\kappa^{(2)}$ toisen ryhmän odotusarvoissa. Tämän eron tilastollinen merkitsevyys arvioidaan κ-odotusarvojen estimaattien ja niiden keskivirheiden avulla. Tällöin α-parametrien avulla voidaan kuvata ja tutkia mistä ero faktoreiden odotusarvoissa johtuu.

5.5 Tasoparametrinen faktorimalli ja rakenneyhtälömalli pojilla

Tarkastellaan aluksi tasoparametrisen faktorimallin estimointituloksia pojilla, joiden PTT on hyvä ja huono taulukossa 5.3 ja tehdään aamuvireisyyden, unilaadun sekä mielialan tasovertailut.

Taulukko 5.3 Taloudellisesti hyvin (1) ja huonosti (2) toimeentulevien perheiden poikien aavireisyyden (η_1), unilaadun (η_2) ja mielialan (η_3) tasoparametrisen faktorimallin estimointitulokset (suluissa keskivirheet).

		$\hat{\Lambda}^{(1)} = \hat{\Lambda}^{(2)}$		$\hat{\tau}^{(1)} = \hat{\tau}^{(2)}$
	η_1	η_2	η_3	
Aamupirteys	1*	0*	0*	2.46 (.06)
Aamuväsymys	-1.01 (.07)	0*	0*	2.56 (.07)
Aamuvireisyys	.72(.05)	0*	0*	2.07(.05)
Uniongelmat	0*	1*	0*	1.98 (.06)
Unilaatu	0*	-.48(.07)	0*	2.14(.04)
Masentuneisuus	0*	0*	1*	1.69 (.04)
Ärtyneisyys	0*	0*	0.85(.09)	1.86 (.04)
Hermostuneisuus	0*	0*	.90(.10)	1.52(.04)

$\kappa_1^{(1)} = 0* \ \kappa_2^{(1)} = 0* \ \kappa_3^{(1)} = 0*$

$\hat{\kappa}_1^{(2)} = -.29\,(.08),\ \hat{\kappa}_2^{(2)} = -.43(.08),\ \hat{\kappa}_3^{(2)} = .23(.05)$

*)kiinnitetty

$\chi^2(43) = 44.61$ p = .40

χ^2-testin perusteella voidaan sanoa, että tasoparametrinen faktorimalli sopii aineistoon erittäin hyvin. Vapaasti estimoidut κ-parametrit pojilla, joiden PTT on huono, ovat kaikki tilastollisesti erittäin merkitseviä. Aamuvireisyyden keskiarvo on tilastollisesti erittäin merkitsevästi pienempi pojilla, joiden PTT on huono, sillä $\hat{\kappa}_1^{(2)}$ on negatiivinen. Vielä selvempi vastaava keskiarvoero on unilaadussa. Myös mieliala on tilastollisesti merkitsevästi alemmalla tasolla pojilla, joiden PTT on huono, koska mielialafaktori kuvaa tässä alhaista mielialaa.

Näiden tulosten pohjalta voidaan lähteä analysoimaan tarkemmin keskiarvoeroja tasoparametrisen faktoreiden rakenneyhtälömallin avulla ja mahdollisesti tekemään päätelmiä siitä, mistä faktoreiden keskiarvoerot johtuvat. Mallin estimointitulokset on esitetty taulukossa 5.4.

Taulukko 5.4 Taloudellisesti hyvin (1) ja huonosti (2) toimeentulevien perheiden poikien aavireisyyden (η_1), unilaadun (η_2) ja mielialan (η_3) tasoparametrisen faktoreiden polkumallin estimointitulokset (suluissa keskivirheet).

	$\hat{\Lambda}^{(1)} = \hat{\Lambda}^{(2)}$			$\hat{\tau}^{(1)} = \hat{\tau}^{(2)}$
	η_1	η_2	η_3	
Aamupirteys	1*	0*	0*	2.47 (.06)
Aamuväsymys	-.98 (.07)	0*	0*	2.56 (.07)
Aamuvireisyys	.71(.05)	0*	0*	2.07(.05)
Unilaatu	0*	1*	0*	1.98 (.06)
Uniongelmat	0*	-.48(.07)	0*	2.14(.04)
Masentuneisuus	0*	0*	1*	1.69 (.04)
Ärtyneisyys	0*	0*	.84(.09)	1.86 (.04)
Hermostuneisuus	0*	0*	.90(.10)	1.52(.04)

$\hat{\beta}_{12}^{(1)} = \hat{\beta}_{12}^{(2)} = .65(.10)$, $\hat{\beta}_{23}^{(1)} = \hat{\beta}_{23}^{(2)} = -.83(.13)$

$\alpha_1^{(1)} = \alpha_2^{(1)} = \alpha_3^{(1)} = 0*$

$\alpha_1^{(2)} = -.01\,(.08)$, $\alpha_2^{(2)} = -.24(.08)$, $\alpha_3^{(2)} = .23(.05)$

*)kiinnitetty

$\chi^2(47) = 49.51$ p = .37

χ^2-testin tuloksen perusteella voidaan todeta, että tasoparametrinen faktoreiden polkumalli on voimassa molemmissa ryhmissä. Mielialan keskiarvo pojilla, joiden PTT on huono, on $\alpha_3^{(2)} = .23(.05)$, mikä on tilastollisesti erittäin merkitsevästi alempi kuin pojilla, joiden PTT on huono ($\alpha_3^{(1)} = 0$), kun otetaan huomioon, että mielialafaktori kuvaa alhaista mielialaa. Mallissa on vakioitu mielialan vaikutus unilaadun keskiarvoeroon: ero on silti tilastollisesti erittäin merkitsevä ($\alpha_2^{(2)} = -.24$, s.e = .08). Kun aamuvireisyyden keskiarvoerosta vakioidaan unilaadun ja mielialan tasoeron vaikutus, ei tilastollisesti merkitsevää keskiarvoeroa ole ($\alpha_1^{(2)} = -.01$, s.e. = .08).

Estimointitulosten ja odotusarvorakenteen (5.10) perusteella voidaan laskea pojille, joiden PTT on huono, seuraavat keskiarvojen estimaatit:

$$\begin{aligned}
\hat{\kappa}_1 &= \hat{\alpha}_1 + \hat{\beta}_{12}(\hat{\alpha}_2 + \hat{\beta}_{23}\hat{\alpha}_3) = \hat{\alpha}_1 + \hat{\beta}_{12}\hat{\alpha}_2 + \hat{\beta}_{12}\hat{\beta}_{23}\hat{\alpha}_3 \\
&= -.01 + .65(-.24) + .65(-.83).23 \\
&= -.01 + (-.16) + (-.12) = -.29 \\
\hat{\kappa}_2 &= \hat{\alpha}_2 + \hat{\beta}_{23}\hat{\alpha}_3 = -.24 + (-.83).23 = -.24 + (-.19) = -.43 \\
\hat{\kappa}_3 &= \hat{\alpha}_3 = .23
\end{aligned}$$

Keskiarvojen estimaateiksi saadaan samat arvot kuin taulukon 5.3 tasoparametrisessa faktorimallissa. Aamuvireisyyden tasoero voidaan tasoparametrisen faktoreiden polkumallin avulla jakaa kolmeen osaan: mielialan keskiarvoerosta johtuvaan osaan ($\hat{\beta}_{12}\hat{\beta}_{23}\hat{\alpha}_3 = -.12$) ja unilaadun keskiarvoerosta johtuviin osaan ($\hat{\beta}_{12}\hat{\alpha}_2 = -.16$) ja osaan, joka ei selity muulla keskiarvoerolla ($\hat{\alpha}_1 = -.01$). Ottamalla huomioon tämän keskivirhe, voidaan todeta, että aamuvireisyyden keskiarvoero selittyy kokonaan unilaadun ja mielialan tasoeroilla. Unilaadun keskiarvoero on jaettu kahteen osaan: mielialan keskiarvoerosta johtuvaan osaan ($\hat{\beta}_{23}\hat{\alpha}_3 = -.19$) ja osaan, joka ei selity muilla keskiarvoeroilla ($\hat{\alpha}_2 = -.24$).

6 YHTEENVETO

Tässä tutkielmassa on tarkasteltu konfirmatorista faktorimallia ja faktoreiden rakenneyhtälömallia sekä niiden rakentamista vaiheittain. Lisäksi on tarkasteltu konfirmatoristen faktorimallien ja faktoreiden rakenneyhtälömallien rakentamista yhtäaikaa usealle ryhmälle sekä ryhmien välisiä vertailumahdollisuuksia ja vertailujen suorittamista. Nämä tarkastelut ovat lähtökohtana tasovertailuille ja tasovertailumalleille, joihin tässä tutkielmassa on keskitytty.

Tasovertailuja koskevassa luvussa on aluksi tarkasteltu ryhmävertailutilanteissa tasoparametrista faktorimallia ja sitä, minkälaisin rajoituksin se on identifioituva. Tämän jälkeen on lähdetty selvittämään, kuinka tasoparametriseen faktorimalliin nojautuen voidaan kaksisuuntaisessa ryhmävertailutilanteessa tutkia ryhmittelevien tekijöiden yhdysvaikutusta ja ryhmien päävaikutuksia, jos yhdysvaikutusta ei ole. Tasoparametrisen faktorimallin lisäksi on tarkasteltu tasoparametrista faktoreiden rakenneyhtälömallia ja sen tuomia mahdollisuuksia faktoreiden keskiarvotarkasteluihin.

Teoreettisissa tarkasteluissa saatiin seuraava päätulos: faktoriin liittyvää yhdysvaikutusta kaksisuuntaisessa tilanteessa voidaan tutkia asettamalla yhdensuuntaisuusrajoite, joka rajoittaa ryhmien keskiarvoprofiilit yhdensuuntaisiksi. Tämä tulos on perusteltu tämän tutkielman empiirisessä osassa graafisesti. Tulos voidaan yleistää monimutkaisempiin tilanteisiin, esimerkiksi tilanteisiin, joissa on useampia ryhmittelijöitä ja faktoreita. Tällaisia tilanteita ei ole kuitenkaan tarkasteltu tässä tutkielmassa, joten tällaisten mallien toimivuutta käytännösssä tulee tarkastella erikseen.

Tässä tutkielmassa on saatu myös joitakin empiirisiä tuloksia liittyen 15-vuotiaiden nuorten aamuvireisyyteen, unilaatuun ja mielialaan. Suomalaisille 15-vuotiaille pojille rakennettiin faktori- ja polkumallit taloudellisesti hyvin ja huonosti toimeentulevien perheiden osaryhmissä ja ryhmille suoritiin vertailuja vertailuja. Näitä malleja laajennettiin rakentamalla myös tasoparametrinen faktorimalli ja tasoparametrinen faktoreiden rakenneyhtälömalli. Tasovertailuja koskevassa luvussa esitettiin lisäksi kaksisuuntainen tasovertailumalli 15-vuotiaiden nuorten aamuvireisyydelle. Mallissa ryhmittelevinä tekijöinä olivat sukupuoli ja perheen taloudellinen tilannne.

Rakennettuihin malleihin liittyy joitakin näkökulmia, joista katsoen saatujen tulosten voimassaoloa voidaan epäillä. Mallien estimointitulokset perustuvat havaittujen muuttujien ja indeksien kovarianssimatriiseihin, vaikka kyseessä on ei-jatkuvat, tarkemmin sanoen järjestysasteikolliset muuttujat ja järjestysasteikollisiksi koodatut indeksit. Näin on kuitenkin meneteltävä, jotta muuttujia ja indeksejä voidaan käyttää keskiarvotarkasteluissa. Lisäksi estimoitujen faktori- ja polkumallien tulosten osalta pojilla voidaan todeta, että ne ovat tulkinnallisesti lähes vastaavia kuin tutkimuksessa Tynjälä et al. (1995) pojille rakennettu polkumalli, jonka estimointi perustuu järjestysasteikollisille muuttujille soveltuvaan polykorisia korrelaatioita sisältävään matriisiin. Keskiarvotarkastelujen tulokset voidaan kuitenkin kyseenalaistaa havaittujen muuttujien ja indeksien järjestysasteikollisuuden vuoksi. Menettelyä voidaan perustella sillä, että havaitut järjestysasteikolliset muuttujien ja indeksit voidaan ajatella arvioiksi vastaavasta jatkuvasta muuttujasta ja rakennetut mallit ovat siten arvioita vastaavista malleista jatkuvilla muuttujilla ja toisaalta saadut tulokset ovat sisällöllisesti järkeviä.

Tutkimusongelmista 1 - 2 muodostetut tutkimushypoteesit 1 - 3 testataan pojilla, joiden PPT on hyvä ja huono. Tulokset tukevat niitä, lukuunottamatta pieniä yksityiskohtia. Unilaatumuuttujalla on heikko mittauskyky unilaatufaktorin mittaamisessa. Toisaalta mallissa pojilla, joiden PTT on hyvä, korreloivat hermostuneisuuden ja masentuneisuuden mittausvirheet. Myös tutkimushypoteesi 4 osoittautuu oikeaksi pojilla, joiden PTT on huono. Tosin selitysasteet eivät ole kovin suuria, mutta kohtuullisia. Sama voidaan todeta myös taloudellisesti hyvin toimeentulevien perheiden poikien osalta poikien ryhmille yhtäaikaa rakennetun mallin perusteella. Ryhmävertailujen päätulos on, että poikien ryhmissä faktorimallit ja faktoreiden polkumallit samanlaisia lukuunottamatta mittausvirheiden korreleituneisuutta hermostuneisuuden ja masentuneisuuden osalta pojilla, joiden PTT on hyvä.

Tasoparametrisen faktorimallin perusteella voidaan todeta, että pojilla, joiden PTT on huono, on aamuvireisyys, unilaatu ja mieliala selkeästi alemmalla tasolla. Tasoparametrisen faktoreiden rakenneyhtälömallin perusteella todettiin, että aamuvireisyyden keskiarvoero selittyy kokonaan unilaadun ja mielialan keskiarvoeroilla. Unilaadun tasoerosta lähes puolet selittyy mielialan tasoerolla. Tasoparametrisen faktoreiden rakenneyhtälömallin ja sen estimointitulosten perusteella laskettiin keskiarvojen estimaatit pojilla, joiden PTT on huono. Tulokset samoja kuin tasoparametrisessa faktorimallissa. Tämä osoittaa, että tasoparametrisen faktoreiden rakenneyhtälö-mallin tulokset luotettavia. Tulosten perusteella voidaan todeta, että tasoparametrinen faktoreiden rakenneyhtälömalli on käyttökelpoinen faktoreiden keskiarvoerojen selittämisessä. Tasoparametrisen faktorimallin estimointitulosten perusteella havaituista keskiarvoeroista voidaan vakioida osuus, joka johtuu toisen faktorin tasoeroista.

Neljän ryhmän tapauksessa, tytöille ja pojille, joiden PTT on hyvä ja huono, rakennetun aamuvireisyyden tasoparametrisen faktorimallin estimointitulosten perusteella voidaan todeta, että sen avulla ei pystytä tulkitsemaan ryhmien välisiä keskiarvoeroja. Estimoimalla kaksisuuntainen tasovertailumalli pystytään tekemään aamuvireisyyden keskiarvoeroihin liittyviä tulkintoja. Mallin perusteella voidaan todeta, että sukupuolella ja PPT:llä ei ole tilastollisesti yhdysvaikutusta. PTT:llä havaittiin olevan selkeä päävaikutus, mutta sukupuolella päävaikutusta ei ole. Tuloksia on verrattu SPSS:n MANOVA:n tuloksiin aamuvireisyys -faktoria mittaaville aamupirteys -muuttujalle, väsymystä kouluaamuisin mittaavalle muuttujalle ja aamuvireisyys -muuttujalle. Tulokset ovat samansuuntaisia. Erona on kuitenkin se, että SPSS:n MANOVA:n monimuuttujaiset testit osoittavat, että sukupuolella on melkein merkitsevä päävaikutus. Yksimuuttujaiset testit antavat saman tuloksen väsymykselle kouluaamuisin ja aamupirteydelle.

Kaksisuuntainen tasovertailumallin saatiin tulkittua aamuvireisyyteen liittyvät keskiarvoerot ryhmien välillä. Kaksisuuntainen tasovertailumalli näyttäisi siten saatujen tulosten perusteella olevan tehokas menetelmä faktoreiden tasoerojen tutkimisessa.

LÄHDEVIITTEET

Akaike, H. (1987): Factor analysis and AIC. *Psychometrica 52(3),* 317 - 322.

Bentler, P. M. ja Bonett, D. G. (1980): Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin 88*, 588 - 606.

Browne, M. W. (1984): Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psycholog, 37*, 62 - 83.

Cook, J. ja Burd, L. (1990): Preliminary report on construction and validation of a pediatric sleep disturbance questionnaire. *Perceptual and Motor Skills 70*, 259 - 267.

Harman, H.H. (1967): *Modern Factor Analysis. Second Edition.* Chicago: University of Chicago Press.

Jöreskog, K. G. (1967): Some contributions to maximum likelihood factor analysis. *Psychometrika 34*, 183 - 202.

Jöreskog, K. G. (1969): A general approach to confirmatory factor analysis. *Psychometrika 36*, 409 - 426.

Jöreskog, K. G. (1971): Simultaneous factor analysis in several populations. *Psychometrika 34*, 183 - 202.

Jöreskog, K. G. ja Sörbom, D. (1976): *LISREL III. Estimation of linear Structural equation systems by maximum likelihood methods.* Chicago: National Educational Resources, Inc.

Jöreskog, K. G. ja Sörbom, D. (1979): *Advanced in Factor Analysis and Structural Equations Models.* Cambridge, mass.: Abt Books.

Jöreskog, K. G. ja Sörbom, D. (1981): *LISREL V. Analysis of linear relationships by maximum likelihood and least squares methods.* University of Uppsala, Department of Statistics. Research Report 81 - 8.

Jöreskog, K. G. ja Sörbom, D. (1993a): *LISREL 8 User's Reference Guide.* Chicago: Scientific Software International inc.

Jöreskog, K. G. ja Sörbom, D. (1993b): *PRELIS 2 User's Reference Guide.* Chicago: Scientific Software International inc.

Leskinen, E. (1987): *Faktorianalyysi. Konfirmatoristen faktorimallien teoria ja rakentaminen.* Jyväskylän yliopiston tilastotieteen laitoksen julkaisuja 1/1987.

Leskinen, E. (1989): *Spesififaktoreiden mallintamisesta ja identifioituvuudesta kovarianssi-rakennemalleissa.* Jyväskylän yliopiston tilastotieteen laitoksen julkaisuja 5/1989.

Norusis, M.J. (1994): *SPSS Professional Statistics 6.1.* Chicago IL: SPSS Inc.

Nummenmaa, T., Konttinen, R., Kuusinen, J. ja Leskinen, E. (1997): *Tutkimusaineiston analyysi.* Porvoo: WSOY

Minors, D. ja Waterhouse J.M. (1987): The role of naps in alleviating sleepiness during an irregular sleep-wake schedule. *Ergonomics, 30,* 1261 - 1273.

Mulaik, S.A. (1972): *The Foundations of Factor Analysis.* New York: McGraw-Hill.

Smith, S., Reilly, C. ja Midkif, K. (1989): Evaluation of three circadian rhythm questionnaires with suggestions for an improved measure of morningness. *Journal of Applied Psychology 9,* 74(5): 728 - 738.

Sörbom, D. (1986): *Model modification.* University of Uppsala . Department of Statistics. Reseach Report 86-3.

Tanaka, J. S. ja Huba, G. J. (1984): Confirmatory hierarchical factor analysis of psychological distress measures. *Journal of Personality and Social Psychology, 46,* 621 - 635.

Tucker, L.R. ja Lewis, C. (1973): A reliability coefficient for maximum likelihood factor analysis. *Psychometrika 38,* 1 -10.

Tynjälä, J. ja Törmäkangas, K.(1995): WHO-koululaistutkimusaineisto 1994. Kirjassa: Kannas, L. (toim.) *Koululaisten kokema terveys, hyvinvointi ja kouluviihtyvyys.* Opetushallituksen julkaisuja 1995. Helsinki: Hakapaino oy.

Tynjälä, J., Kannas, L. ja Levälahti, E.(1995): *Aamuvireisyys 15-vuotiailla suomalaisnuorilla* (Morning alertness among 15-year-old Finns). Sosiaalilääketieteen päivät. KELA:n tutkimus- ja kehitysyksikkö. Turku 2. - 3.11.1995. Posteri

Åkerstedt, T., Hume, K., Minors, D. ja Waterhouse, J. (1994): The subjective meaning of good sleep, an intraindividual approach using the Karoniska Sleep Diary. *Perceptual and Motor Skills 79,* 278 - 296.

Åkerstedt, T. ja Torsvall, L. (1989): Shift work and transient insomnia. Julkaisussa J. Horne (toim.) *Sleep '88* Stuttgart: Gustav Fisher Verlag. 30 -32.

LIITE 1

Havaittujen muuttujien ja indeksien kovarianssit sekä otoskeskiarvot pojilla

Taloudellisesti hyvin toimeen tulevien perheiden pojat:

Aamuvireisyys	.517							
Aamuväsymys	-.433	.934						
Aamupirteys	.396	-.582	.862					
Unilaatu	.073	-.144	.151	.493				
Uniongelmat	-.237	.359	-.253	-.171	.818			
Masentuneisuus	-.093	.194	-.106	-.083	.181	.382		
Ärtyneisyys	-.056	.126	-.121	-.038	.115	.113	.309	
Hermostuneisuus	-.077	.160	-.108	-.101	.140	.067	.140	.377

Taloudellisesti huonosti toimeen tulevien perheiden pojat:

Aamuvireisyys	.434							
Aamuväsymys	-.272	.798						
Aamupirteys	.311	-.373	.689					
Unilaatu	.057	-.156	.103	.533				
Uniongelmat	-.151	.277	-.247	-.224	.814			
Masentuneisuus	-.034	.067	-.078	-.105	.083	.400		
Ärtyneisyys	-.034	.089	-.054	-.079	.123	.152	.317	
Hermostuneisuus	-.024	.071	-.061	-.108	.149	.162	.162	.333

Otoskeskiarvot:

	Pojat, PPT hyvä	Pojat, PTT huono
Aamuvireisyys	2.087	1.857
Aamuväsymys	2.576	2.823
Aamupirteys	2.467	2.162
Unilaatu	2.098	1.974
Uniongelmat	1.957	2.426
Masentuneisuus	1.505	1.736
Ärtyneisyys	1.886	2.030
Hermostuneisuus	1.696	1.909

LIITE 2

Havaittujen muuttujien ja indeksien kovarianssit sekä otoskeskiarvot tytöillä

Taloudellisesti hyvin toimeen tulevien perheiden tytöt:

Aamuvireisyys	.416		
Aamuväsymys	-.253	.733	
Aamupirteys	.316	-.449	.788

Taloudellisesti huonosti toimeen tulevien perheiden tytöt:

Aamuvireisyys	.478		
Aamuväsymys	-.358	.737	
Aamupirteys	.376	-.459	.772

Otoskeskiarvot:

	tytöt, PTT hyvä	tytöt, PTT huono
aamuvireisyys	2.023	1.933
aamuväsymys	2.695	2.809
aamupirteys	2.328	2.177

LIITE 3

Monimuuttujaisen kaksisuuntaisen varianssianalyysin (SPSS MANOVA:n) tulokset aamuvireisyyttä mittaaville havaituille muuttujille

Monimuuttujaiset testit sukupuolen ja perheen taloudellisen tilanteet yhdysvaikutukselle:

		F	vapausasteet (H_0-hypoteesi)	vapausasteet (virhe)	p
Pillais	.003	.889	3.00	1029.00	.446
Hotellings	.003	.889	3.00	1029.00	.446
Wilks	.997	.889	3.00	1029.00	.446
Roys	.003				

Monimuuttujaiset testit sukupuolen päävaikutukselle:

		F	vapausasteet (H_0-hypoteesi)	vapausasteet (virhe)	p
Pillais	.011	3.64	3.00	1029.00	.012
Hotellings	.011	3.64	3.00	1029.00	.012
Wilks	.989	3.64	3.00	1029.00	.012
Roys	.011				

Monimuuttujaiset testit perheen taloudellisen tilanteen päävaikutukselle:

		F	vapausasteet (H_0-hypoteesi)	vapausasteet (virhe)	p
Pillais	.021	7.52	3.00	1029.00	.000
Hotellings	.022	7.52	3.00	1029.00	.000
Wilks	.979	7.52	3.00	1029.00	.000
Roys	.021				

LIITE 3 (jatkoa)

F -testit sukupuolen päävaikutukselle:

	Neliösumma	Virhe-neliösumma	Keski-neliösumma	Keskivirhe-neliösumma	F	p
Aamuvireisyys	,036	472,679	,036	,458	,078	,799
Väsymys kouluaamuina	3,974	832,141	3,974	,807	4,924	,027
Aamupirteys	3,088	779,048	3,088	,756	4,086	,043

F -testit perheen taloudellisen tilanteen päävaikutukselle:

	Neliösumma	Virhe-neliösumma	Keski-neliösumma	Keskivirhe-neliösumma	F	p
Aamuvireisyys	6,230	472,679	6,230	,458	13,588	,000
Väsymys kouluaamuina	9,409	832,141	9,409	,807	11,657	,001
Aamupirteys	15,531	779,048	15,531	,756	20,554	,000