

Pitkittäisaineistojen faktorimalleista ja niiden
rakentamisesta sekä sovellus ”Vanhemmat, opettajat ja
lapsen oppiminen” -tutkimukseen

Eija Räikkönen

Tilastotieteen pro gradu -tutkielma
16. toukokuuta 2007

Jyväskylän yliopisto
Matematiikan ja tilastotieteen laitos

Tiivistelmä

Eija Räikkönen: *Pitkittäisaineistojen faktorimalleista ja niiden rakentamisesta sekä sovellus "Vanhemmat, opettajat ja lapsen oppiminen" -tutkimukseen*

Tilastotieteen pro gradu -tutkielma, Jyväskylän yliopisto,

16. toukokuuta 2007.

Sivuja 53, liitteitä 2.

Pitkittäistutkimukset ovat lisääntyneet nopeasti, sillä ne mahdollistavat monipuoliset tutkittavien ilmiöiden tarkastelut. Tutkielman teoriaosan tavoitteena oli tarkastella konfirmatorisen faktorimallin rakentamista pitkittäisasetelmassa. Asetelman erikoistapauksina tarkasteltiin stationaarisia ja invariantteja pitkittäisfaktorimalleja sekä pitkittäisfaktorimallin rakentamista ryhmävertailutilanteessa. Lisäksi tarkasteltiin pitkittäisfaktorimallien rakenneyhtälömallinnusta. Sovellusosan tavoitteena oli mallintaa ensimmäisellä luokalla olevien lasten mielialoja konfirmatorisen pitkittäisfaktorimallin avulla sekä tutkia mielialojen mittaamisen pysyvyyttä. Tutkimusaineistona oli Vanhemmat, opettajat ja lapsen oppiminen -tutkimuksessa kerätty emootio-pitkittäisaineisto. Mielialamuuttujia käytettiin kolmelta mittauspäivältä: maanantailta, keskiviikkolta ja perjantailta.

Mielialamuuttujien faktorirakenteiden etsimisessä alustavana menetelmänä oli eksploratiivinen faktorianalyysi, jonka tulosten perusteella saatujen faktorirakenteiden sopivuutta testattiin konfirmatorisen faktorianalyysin avulla. Eksploratiivisen faktorianalyysin faktorointivaiheessa ilmeni kuitenkin ongelmia, joiden takia mallinnettiin vain negatiivisia mielialoja. Konfirmatorisessa faktorianalyysissä aineistoon sovitettiin ensin konfirmatorinen pitkittäisfaktorimalli ja sen erikoistapaus, stationaarinen pitkittäisfaktorimalli. Sen jälkeen tutkittiin mielialojen mittaamisen pysyvyyttä faktoreiden rakenneyhtälömallin avulla. Mallien erikoistapauksena tutkittiin rakenneparametrien aikainvarianttisuutta. Konfirmatoristen faktorianalyysien tulosten perusteella voidaan lasten negatiivisten mielialojen mittaamista pitää aineistossa melko pysyvänä. Parhaiten aineistoon sopi aikainvarianttien rakenneparametrien rakenneyhtälömalli. Stationaarinen pitkittäisfaktorimalli sopi aineistoon lähes yhtä hyvin.

Avainsanoja: konfirmatorinen faktorianalyysi, pitkittäisaineisto, pitkittäisfaktorimalli

Sisältö

1 Johdanto	1
2 Pitkittäisaineiston konfirmatorinen faktorianalyysi	3
2.1 Pitkittäisfaktorimalli	3
2.2 Pitkittäisfaktorimallin rakentaminen	8
2.2.1 Faktorimallin rakentaminen mittauskerroille	8
2.2.1.1 Mallin spesifointi	9
2.2.1.2 Mallin identifoituvuus	10
2.2.1.3 Parametrien estimointi	11
2.2.1.4 Mallin hypoteesien testaus	13
2.2.1.5 Mallin riittävyystarkastelut	17
2.2.2 Poikkileikkausfaktorimallien yhdistäminen pitkittäisfaktorimalliksi sekä mallin rakentaminen	23
2.3 Pitkittäisfaktorimallin ryhmävertailuista	25
2.4 Pitkittäisfaktorimallin erikoistapauksia	26
2.5 Faktoreiden rakenneyhtälömallit	27
2.5.1 Mallin rakentaminen	30
3 Aineisto ja asetelma	33
3.1 Vanhemmat, opettajat ja lapsen oppiminen-tutkimus	33
3.2 Tutkittavat	34
3.3 Menetelmä ja muuttujat	34
4 Pitkittäisfaktorimallin rakentaminen sovellusaineistolle	36
4.1 Eksploraatiivinen faktorianalyysi	36
4.2 Konfirmatorinen faktorianalyysi	42
4.2.1 Kolmen mittauskerran pitkittäisfaktorimalli	42
4.2.2 Kolmen mittauskerran stationaarinen pitkittäisfaktorimalli	45
4.2.3 Kolmen faktorin rakenneyhtälömalli	47
4.2.4 Rakenneyhtälömallin rakenneparametrien aikainvarianttisuus	48
5 Yhteenveto	50
Lähdeluettelo	52
Liitteet	54
Liite A: Vanhempien päiväkirjan kysymyssarja 8.2	54
Liite B: LISREL-komentojonot	55

1 Johdanto

Pitkittäistutkimukset ovat lisääntyneet nopeasti eri tieteenaloilla viimeisten vuosikymmenten aikana, sillä pitkittäistutkimukset mahdollistavat monipuoliset tutkittavien ilmiöiden muutosten ja pysyvyyksien tarkastelut. Toisaalta pitkittäisaineistojen analysoiminen on tutkimusmetodisesti vaativaa ja siksi avuksi tarvitaan kehittyneitä tilastollisia menetelmiä ja malleja. Tässä tutkielmassa tarkastellaan pitkittäisaineiston analysointia konfirmatorisen pitkittäisfaktorimallin avulla. Tutkielmassa esitetään keskeinen konfirmatorisen pitkittäisfaktorimallin teoria sekä sovelletaan sitä käytännössä pitkittäisaineiston analysointiin.

Pitkittäisaineisto (*Longitudinal data*) on määritelmänsä mukaan tutkimusasetelma, jossa samoja havaintoyksiköitä mitataan toistuvasti eri mittauskerroilla, Leskinen (1995). Pitkittäisaineistosta käytetään myös useita muita termejä tieteenalasta riippuen. Pitkittäisaineisto-termin rinnalla käytettyjä termejä ovat mm. paneeliaineisto (*Panel data*), seuranta- tai seuruuaineisto (*Follow-up data*) sekä toistomittausaineisto (*Repeated measurement data*), Leskinen (2005). Kaikille näille asetelmille on yhteistä se, että mielenkiinnon kohteena ovat muuttujien arvoissa ajan kuluessa tapahtuvat muutokset.

Pitkittäisaineistoissa on oleellista, että tutkittavat ovat koko ajan samoja, mutta mitattavat muuttujat voivat vaihdella eri mittauskerroilla. Tässä tutkielmassa tarkastellaan sellaista pitkittäisaineistoa, jossa samoja kohdehenkilöitä on mitattu kolme kertaa ja mitatut muuttujat ovat olleet samoja kaikilla mittauskerroilla. Silloin uusintamittaukset lisäävät mitattujen muuttujien lukumäärää, eivät tutkittavien henkilöiden määrää. Pitkittäisaineistot ovat siis luonteeltaan aina monimuuttujaisia. Koska mittaukset tehdään aina samoista havaintoyksiköistä, ovat mittauskerrat keskenään korreloituneita, Leskinen (2005).

Tämän tutkielman tavoitteet voidaan jakaa kahteen osaan. Tutkielman teoriaosan tarkoituksena on tutkia konfirmatorisen pitkittäisfaktorimallin teoriaa ja rakentamisen vaiheita. Pitkittäisfaktorimallin rakentaminen on jaettu kahteen osaan. Ensimmäisen osan tavoitteena on konfirmatorisen faktorimallin rakentaminen yksittäiselle mittauskerralle. Konfirmatorisen faktorimallin päävaiheina käsitellään mallin spesifiointi, identifioituvuustarkastelut, parametrien estimointi, hypoteesien testaus ja riittävyystarkastelut. Toisen vaiheen tavoitteena on rakentaa eri mittauskertojen konfirmato-

risista faktorimalleista mittauskertojen välisten kovarianssirakenteiden avulla pitkittäisfaktorimalli. Teoriaosan kolmantena tavoitteena on perehtyä pitkittäisfaktorimallin erikoistapauksiin: pitkittäisfaktorimallin rakentamiseen ryhmävertailutilanteessa, ajan suhteen stationaariseen pitkittäisfaktorimalliin sekä ryhmien suhteen invarianttiin pitkittäisfaktorimalliin. Viimeisenä tavoitteena teoriaosassa on kuvata pitkittäisfaktorimallin faktoreiden mittauskertojen väliset yhteydet faktoreiden rakenneyhtälöiden avulla.

Tutkielman soveltavassa osassa on tarkoitus tutkia peruskoulun ensimmäisellä luokalla olevien lasten mielialaa mittaavien muuttujien faktorirakenteita ja yhteyksiä. Tavoitteena on ensinnäkin mallintaa ensimmäisellä luokalla olevien lasten mielialo- ja konfirmatorisen pitkittäisfaktorimallin avulla sekä tutkia, ovatko löydetty faktorirakenteet stationaarisia mittauskerrasta toiseen. Tutkielmassa analysoidaan pitkittäisaineistoa, jossa mukana on kolme mittauspäivää ja kultakin mittauspäivältä on mukaan valittu kolme mielialaa. Tarkoituksena on saada aikaan kolmen faktorin pitkittäisfaktorimalli, jossa kukin faktori edustaa yhtä mittauspäivää. Toisena tavoitteena on tutkia lasten mielialojen mittaamisen pysyvyyttä. Tämä tapahtuu tutkimalla faktorien keskinäisiä suhteita faktoreiden rakenneyhtälöiden avulla.

Tutkielman alkuosa sisältää tutkielman kannalta olennaisen tilastotieteen teorian. Luvussa 2 käsitellään konfirmatorisen pitkittäisfaktorimallin teoriaa keskittyen mallin rakentamisen vaiheisiin. Tutkielman soveltavan osan aloittaa luku 3, jossa esitellään käytettävä tutkimusaineisto. Aineisto on Jyväskylän yliopistossa toimivan Suomen Akatemian Oppimisen ja motivaation huippututkimusyksikön tutkimusohjelmaan kuuluvan Vanhemmat, opettajat ja lapsen oppiminen -pitkittäistutkimuksen osa-aineisto. Tutkimuksen vastuullisena tutkijana toimii dosentti Kaisa Aunola. Tutkielman neljännessä luvussa tarkastellaan tutkimusaineiston eksploratiivista ja konfirmatorista faktorianalyysia. Tutkimusaineiston analysoinnissa on käytetty SPSS 14.0- ja LISREL 8.7 -ohjelmia.

2 Pitkittäisaineiston konfirmatorinen faktorianalyysi

Konfirmatorinen faktorimalli on faktorimalli, jossa faktoreiden latausmatriisiin Λ kohdistuu rajoituksia. Yleisimmät latausten kiinnitykset ovat 0 ja 1, Leskinen (1987). Kun konfirmatorista faktorimallia sovelletaan pitkittäisaineiston analyysissä, kutsutaan mallia *pitkittäisfaktorimalliksi*, Tisak & Meredith (1990); Leskinen (1995).

Tässä luvussa tarkastellaan konfirmatorisen pitkittäisfaktorimallin teoriaa ja rakentamisen vaiheita. Kappaleessa 2.1 esitellään yleinen pitkittäisfaktorimalli. Seuraavaksi kappaleessa 2.2 käsitellään pitkittäisfaktorimallin rakentamisen vaiheita. Kappaleessa 2.3 tarkastellaan pitkittäisfaktorimallin käyttöä ryhmävertailutilanteessa ja kappaleessa 2.4 käsitellään lyhyesti pitkittäisfaktorimallin erikoistapauksia. Teoriaosuuden päättää kappale 2.5, johon on sisällytetty pitkittäisfaktorimallien rakenneyhtälömallit. Teoriaosuutta on pyritty selkiyttämään lyhyiden esimerkkien avulla. Pääasiallisina lähteinä luvussa on käytetty artikkelia Tisak & Meredith (1990) ja Bollenin vuonna 1989 julkaistua teosta *Structural Equations with Latent Variables* sekä Leskinen vuonna 1987 julkaistua teosta *Faktorianalyysi. Konfirmatoristen faktorimallien teoria ja rakentaminen*. Tarkemmat viitteet löytyvät tekstistä.

2.1 Pitkittäisfaktorimalli

Oletetaan, että tutkittavasta perusjoukosta on poimittu satunnaisesti aineisto, jossa on yhteensä N havaintoa, p havaittua muuttujaa mittauskerralta t sekä T mittausajankohtaa, merkitään Y_{ijt} , $i = 1, \dots, N$; $j = 1, \dots, p$; $t = 1, \dots, T$. Pitkittäisaineiston konfirmatorinen faktorimalli eli pitkittäisfaktorimalli voidaan esittää yleisen konfirmatorisen faktorimallin muodossa

$$\mathbf{y} = \Lambda\boldsymbol{\eta} + \boldsymbol{\varepsilon}, \quad (2.1)$$

missä

\mathbf{y} on havaittujen muuttujien $pT \times 1$ -vektori,

$\boldsymbol{\eta}$ on latenttien muuttujien eli faktoreiden $kT \times 1$ -vektori,

$\boldsymbol{\varepsilon}$ on havaittujen muuttujien mittausvirheet eli jäännökset sisältävä $pT \times 1$ -satunnaisvektori ja

$\boldsymbol{\Lambda}$ on havaittujen muuttujien ja faktoreiden välisten latausten $pT \times kT$ -matriisi.

Merkitään seuraavaksi faktoreiden kovarianssimatriisia merkinnällä

$$\text{cov}(\boldsymbol{\eta}) = \boldsymbol{\Omega}$$

ja mittausvirheiden kovarianssimatriisia merkinnällä

$$\text{cov}(\boldsymbol{\varepsilon}) = \boldsymbol{\Theta},$$

jolloin $\boldsymbol{\Omega}$ on faktoreiden $kT \times kT$ -kovarianssimatriisi ja $\boldsymbol{\Theta}$ on jäännösten $pT \times pT$ -kovarianssimatriisi. Silloin havaittujen \mathbf{y} -muuttujien teoreettinen $pT \times pT$ -kovarianssimatriisi on muotoa

$$\text{cov}(\mathbf{y}) = \boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Omega}\boldsymbol{\Lambda}' + \boldsymbol{\Theta},$$

missä

$$\boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Lambda}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{\Lambda}_T \end{pmatrix},$$

$$\boldsymbol{\Omega} = \text{cov}(\boldsymbol{\eta}) = \begin{pmatrix} \boldsymbol{\Omega}_{11} & & & \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} & & \\ \vdots & \vdots & \ddots & \\ \boldsymbol{\Omega}_{T1} & \boldsymbol{\Omega}_{T2} & \dots & \boldsymbol{\Omega}_{TT} \end{pmatrix}$$

ja

$$\Theta = \text{cov}(\varepsilon) = \begin{pmatrix} \Theta_{11} & & & \\ \Theta_{21} & \Theta_{22} & & \\ \vdots & \vdots & \ddots & \\ \Theta_{T1} & \Theta_{T2} & \dots & \Theta_{TT} \end{pmatrix}.$$

Olkoon mittauskerran t sisäinen poikkileikkausfaktorimalli muotoa

$$\mathbf{y}_t = \mathbf{\Lambda}_t \boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \dots, T,$$

missä

\mathbf{y}_t on havaittujen muuttujien $p \times 1$ -vektori,

$\boldsymbol{\eta}_t$ on $k \times 1$ -vektori, joka sisältää latentit muuttujat eli faktorit $\eta_1, \eta_2, \dots, \eta_k$,

$\mathbf{\Lambda}_t$ on tuntemattomat latausparametrit sisältävä $p \times k$ -matriisi ja

$\boldsymbol{\varepsilon}_t$ on havaittuihin y -muuttujiin liittyvät mittausvirheet eli jäännökset sisältävä $p \times 1$ -vektori.

Pitkittäisfaktorimalli on silloin konfirmatorinen faktorimalli, jossa mittauskertojen sisäiset poikkileikkausfaktorimallit on yhdistetty mittauskertojen välisten faktoreiden kovarianssimatriisien avulla yhdeksi konfirmatoriseksi faktorimalliksi.

Olkoon mittauskerran t faktoreiden kovarianssimatriisi muotoa

$$\text{cov}(\boldsymbol{\eta}_t) = \mathbf{\Omega}_{tt}, \quad t = 1, \dots, T,$$

ja jäännösten kovarianssimatriisi puolestaan muotoa

$$\text{cov}(\boldsymbol{\varepsilon}_t) = \mathbf{\Theta}_{tt}, \quad t = 1, \dots, T.$$

Silloin mittauskertojen sisäiset kovarianssimatriisit $\boldsymbol{\Sigma}_{tt}$ ovat yleistä muotoa

$$\boldsymbol{\Sigma}_{tt} = \mathbf{\Lambda}_t \mathbf{\Omega}_{tt} \mathbf{\Lambda}_t' + \mathbf{\Theta}_{tt}, \quad t = 1, \dots, T.$$

Poikkileikkausfaktorimalleja yhdistävät mittauskertojen s ja t väliset kovarianssimatriisit Σ_{st} , jossa $s \neq t$, ovat yleistä muotoa

$$\Sigma_{st} = \Lambda_s \Omega_{st} \Lambda_t' + \Theta_{st}, \quad s, t = 1, \dots, T, \quad (2.2)$$

missä

Λ_s ja Λ_t ovat mittauskertojen s ja t $p \times k$ -latausmatriiseja,

Ω_{st} on mittauskertojen s ja t latenttien muuttujien eli faktoreiden $\boldsymbol{\eta}_s$ ja $\boldsymbol{\eta}_t$ välinen $k \times k$ -kovarianssimatriisi ja

Θ_{st} on mittausvirheiden eli jäännösten $\boldsymbol{\varepsilon}_s$ ja $\boldsymbol{\varepsilon}_t$ välinen $p \times p$ -autokovarianssimatriisi.

Pitkittäisfaktorimallissa oletetaan, että faktorit ja mittausvirheet ovat korreloimattomia sekä mittausten sisällä että mittausten välillä, $E(\boldsymbol{\eta}\boldsymbol{\varepsilon}') = \mathbf{0}$. Edelleen oletetaan, että jäännökset eivät korreloi keskenään mittausten t sisällä, $E(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_j) = 0$, $i \neq j$. Mittauskertojen välillä jäännökset voivat korreloida keskenään. Yleensä oletetaan myös, että jäännösten autokovarianssimatriisit Θ_{st} ovat diagonaalisia. (Tisak & Meredith 1990).

Esimerkki 2.1. Tarkastellaan konfirmatorista pitkittäisfaktorimallia, jossa on mukana kolme mittauskertaa. Kultakin mittauskerralta on mukana yksi faktori ja kutakin faktoria on valittu mittaamaan 3 havaittua y -muuttujaa, yhteensä siis 9 havaittua muuttujaa. Pitkittäisfaktorimalli on silloin muotoa

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

missä

\mathbf{y} on havaittujen muuttujien 9×1 -vektori,

$\boldsymbol{\eta}$ on faktoreiden 3×1 -vektori,

$\boldsymbol{\varepsilon}$ on mittausvirheet sisältävä 9×1 -vektori ja

Λ on tuntemattomat latausparametrit sisältävä 9×3 -vektori.

Pitkittäisfaktorimallin graafinen esitys on kuvassa 1. Mallin havaittujen muuttujien teoreettinen 9×9 -kovarianssimatriisi on muotoa

$$\text{cov}(\mathbf{y}) = \Sigma = \Lambda\Omega\Lambda' + \Theta,$$

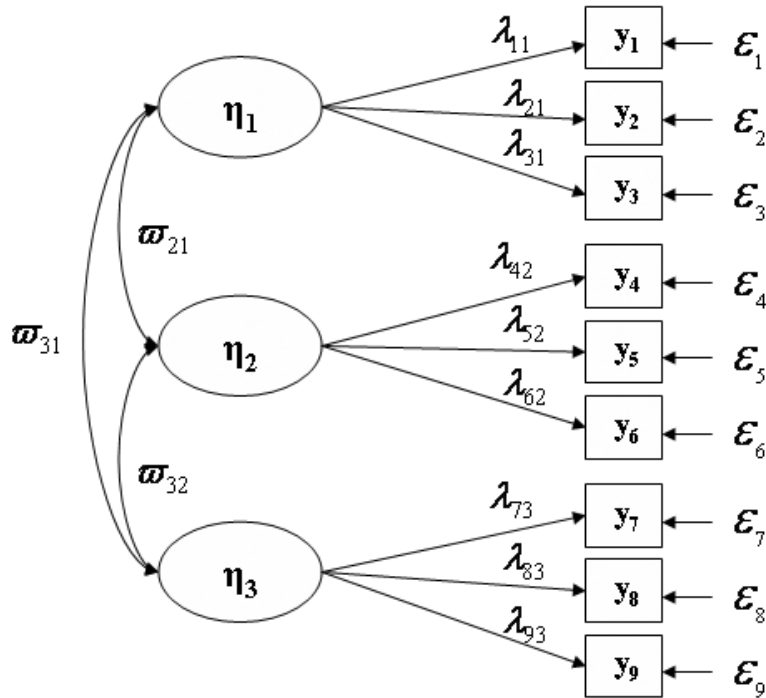
missä

$$\Lambda = \begin{pmatrix} \Lambda_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Lambda_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Lambda_3 \end{pmatrix} = \begin{pmatrix} \lambda_{11} & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & 0 & 0 \\ 0 & \lambda_{42} & 0 \\ 0 & \lambda_{52} & 0 \\ 0 & \lambda_{62} & 0 \\ 0 & 0 & \lambda_{73} \\ 0 & 0 & \lambda_{83} \\ 0 & 0 & \lambda_{93} \end{pmatrix},$$

$$\Omega = cov(\boldsymbol{\eta}) = \begin{pmatrix} \Omega_{11} & & \\ \Omega_{21} & \Omega_{22} & \\ \Omega_{31} & \Omega_{32} & \Omega_{33} \end{pmatrix} = \begin{pmatrix} \omega_{11} & & \\ \omega_{21} & \omega_{22} & \\ \omega_{31} & \omega_{32} & \omega_{33} \end{pmatrix}$$

ja

$$\Theta = cov(\boldsymbol{\varepsilon}) = \begin{pmatrix} \Theta_{11} & & & & & & & & \\ \Theta_{21} & \Theta_{22} & & & & & & & \\ \Theta_{31} & \Theta_{32} & \Theta_{33} & & & & & & \\ & & & & & & & & \\ & & & & & & & & \\ & & & & & & & & \\ & & & & & & & & \\ & & & & & & & & \\ & & & & & & & & \end{pmatrix} = \begin{pmatrix} \theta_{11} & & & & & & & & \\ 0 & \theta_{22} & & & & & & & \\ 0 & 0 & \theta_{33} & & & & & & \\ 0 & 0 & 0 & \theta_{44} & & & & & \\ 0 & 0 & 0 & 0 & \theta_{55} & & & & \\ 0 & 0 & 0 & 0 & 0 & \theta_{66} & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \theta_{77} & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_{88} & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \theta_{99} \end{pmatrix}.$$



Kuva 1: Hypoteettinen kolmen mittauskerran konfirmatorinen pitkittäisfaktorimalli.

2.2 Pitkittäisfaktorimallin rakentaminen

Seuraavassa tarkastellaan lähemmin pitkittäisfaktorimallin rakentamista. Rakentaminen tapahtuu kahdessa vaiheessa. Ensin jokaiselle mittauskerralle rakennetaan konfirmatorinen faktorimalli eli poikkileikkausfaktorimalli. Sen jälkeen yksittäisten mittauskertojen poikkileikkausfaktorimalleista rakennetaan pitkittäisfaktorimalli yhdistämällä poikkileikkausfaktorimallit eri mittauskertojen faktoreiden välisillä kovarianssimatriiseilla toisiinsa, Tisak & Meredith (1990). Kappale perustuu pääosin teokseen Leskinen (1987). *Faktorianalyysi. Konfirmatoristen faktorimallien teoria ja rakentaminen*.

2.2.1 Faktorimallin rakentaminen mittauskerroille

Tarkastellaan seuraavaksi poikkileikkausfaktorimallin rakentamista yksittäiselle mittauskerralle t , $t = 1, \dots, T$. Merkitään mittauskerran t poikkileikkausfaktorimallia seuraavasti:

$$\mathbf{y} = \mathbf{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon}, \quad (2.3)$$

missä

\mathbf{y} on havaittujen muuttujien $p \times 1$ -vektori,

$\boldsymbol{\eta}$ on latenttien muuttujien eli faktoreiden $k \times 1$ -vektori,

$\boldsymbol{\varepsilon}$ on havaittujen muuttujien mittausvirheet eli jäännökset sisältävä $p \times 1$ -satunnaisvektori ja

$\mathbf{\Lambda}$ on havaittujen muuttujien ja faktoreiden välisten latausten $p \times k$ -matriisi.

Silloin yksittäisen mittauskerran havaittujen \mathbf{y} -muuttujien teoreettinen kovarianssimatriisi on muotoa

$$\boldsymbol{\Sigma} = \mathbf{\Lambda}\boldsymbol{\Omega}\boldsymbol{\Lambda}' + \boldsymbol{\Theta}. \quad (2.4)$$

Poikkileikkausfaktorimallin (2.3) rakentamisessa on erotettavissa viisi vaihetta, jotka suoritetaan järjestyksessä. Ensin spesifoidaan malli tutkimushypoteesien perusteella ja sen jälkeen suoritetaan identifioituvuustarkastelut. Seuraavaksi estimoidaan mallin parametrit ja suoritetaan mallia koskevien hypoteesien testaus. Lopuksi tehdään riittävyystarkastelut. Jos mallissa ilmenee ongelmia jonkin vaiheen yhteydessä, palataan sopivaan sitä edeltävään vaiheeseen.

2.2.1.1 Mallin spesifointi

Poikkileikkausfaktorimallin *spesifointi*- eli *täsmentämisvaiheessa* määritellään mallin muuttujien väliset yhteydet valitsemalla λ -, ω - ja θ -parametreille sopiva käsittelytapa tutkimushypoteesin pohjalta. Parametri voidaan estimoida vapaasti, yhtäsuurena jonkun toisen parametrin kanssa tai parametri voidaan kiinnittää vakioksi. Se, mikä tapa kullekin parametrille valitaan, riippuu tutkimushypoteesissa oletetuista havaittujen muuttujien ja faktoreiden välisistä yhteyksistä, faktoreiden keskinäisistä yhteyksistä sekä mittausvirheiden variansseista.

Jotta poikkileikkausfaktorimallin parametrisoinnista tulisi yksikäsitteinen, pitää mallin jokaisen faktorin skaala kiinnittää. Faktoreiden skaalakiinnitykset koskevat joko faktoreiden variansseja tai faktoreiden ja havaittujen muuttujien välisiä latausparametreja. Kunkin faktorin skaalayksikäsitteisyys saadaan kiinnittämällä joko fakto-

rin varianssi ykköseksi tai yksi lataus ykköseksi.

Esimerkki 2.2. Esimerkin 2.1 pitkittäisfaktorimallissa on mukana kolme mittauskerrtaa eli kolme poikkileikkausfaktorimallia. Jotta poikkileikkausfaktorimallit olisivat skaalaysikäsitteisiä, asetetaan kunkin mittauskerran poikkileikkausfaktorimallista ensimmäinen lataus ykköseksi:

$$\lambda_{11} = 1, \lambda_{42} = 1, \lambda_{73} = 1.$$

Muut lataukset estimoidaan vapaasti. Lisäksi kunkin mittauskerran faktorin varianssi sekä havaittujen muuttujien mittausvirheiden varianssit estimoidaan vapaasti. Mittauskerralla 1 vapaasti estimoitavia parametreja ovat siis parametrit λ_{21} , λ_{31} , ω_{11} , θ_{11} , θ_{22} ja θ_{33} . Mittauskerralla 2 vapaasti estimoitavia parametreja ovat parametrit λ_{52} , λ_{62} , ω_{22} , θ_{44} , θ_{55} sekä θ_{66} ja mittauskerralla 3 vapaasti estimoitavia parametreja ovat parametrit λ_{83} , λ_{93} , ω_{33} , θ_{77} , θ_{88} ja θ_{99} .

2.2.1.2 Mallin identifioituvuus

Poikkileikkausfaktorimallin *identifioituvuudella eli yksilöityvyydellä* on keskeinen asema poikkileikkausfaktorimallin rakentamisessa. Jotta valitun poikkileikkausfaktorimallin parametrit voidaan estimoida, on mallin oltava identifioituva. Identifioituvuus on faktorimallin teoreettinen ominaisuus, joten se ei riipu aineistosta eikä estimointimenetelmistä. Poikkileikkausfaktorimallin identifioituvuus liittyy kovarianssimatriisiin (2.4) rakenteeseen.

Identifioituvuus voidaan määritellä eri tavoin. Sovellusten kannalta käyttökelpoisiin määritelmä on seuraava: Poikkileikkausfaktorimallin yksittäinen parametri on identifioituva, jos se on ratkaistavissa kovarianssimatriisin Σ avulla. Jos poikkileikkausfaktorimallin kaikki parametrit ovat ratkaistavissa Σ :n avulla, on poikkileikkausfaktorimalli identifioituva.

Jos parametri voidaan ratkaista kovarianssimatriisista Σ usealla tavalla käyttäen kovarianssiesityksen (2.4) eri yhtälöitä, kyseinen parametri on *yli-identifioituva*. Eri ratkaisut kuitenkin tuottavat saman numeerisen arvon yli-identifioituvalla paramet-

rille, joten yli-identifioituva poikkileikkausfaktorimalli on identifioituvuuden määritelmän mukainen. Yli-identifioituvuustilanteessa kaikkia kovarianssimatriisin Σ yhtälöitä ei tarvitse käyttää parametrien ratkaisemiseen. Tällöin käyttämättömistä yhtälöistä saadaan vapausasteita mallin riittävyystestaukseen χ^2 -testillä.

Faktorimallin (2.3)

$$\mathbf{y} = \mathbf{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

identifioituvuudelle on olemassa välttämättömiä ehtoja, jotka ovat voimassa kun tarkasteltavat parametrit ovat ratkaistavissa kovarianssiesityksen (2.4) avulla. Symmetrisessä kovarianssimatriisissa Σ on $(1/2)p(p+1)$ yhtälöä. Toisaalta kovarianssimatriisissa Σ on pk kappaletta mahdollisia λ -parametreja, $(1/2)k(k+1)$ kappaletta mahdollisia ω -parametreja ja p kappaletta mahdollisia θ -parametreja eli yhteensä s parametria,

$$s = pk + (1/2)k(k+1) + p = (1/2)(2p+k)(k+1).$$

Merkitään seuraavaksi mallin parametreille tehtävien rajoitusten lukumäärää r :llä. Tällöin välttämätön ehto λ -, ω - ja θ -parametrien ratkaisemiseksi kovarianssimatriisista Σ on muotoa

$$s - r \leq (1/2)p(p+1).$$

Käytännössä toimivia riittävyysehtoja poikkileikkausfaktorimallin (2.3) identifioituvuudelle ei voida yleisessä tapauksessa johtaa. Identifioituvuuden määritelmään perustuvat identifioituvuustarkastelut soveltuvat pienille ja melko yksinkertaisille poikkileikkausfaktorimalleille, mutta yleisessä tapauksessa tarkastelut ovat työläisiä.

Konfirmatoristen faktorimallien käytössä keskeisiä malleja ovat ns. *mittausmallit*. Mittausmallissa kukin havaittu y -muuttuja mittaa vain yhtä η -faktoria. Kun mittausmallista kiinnitetään kaikkien faktoreiden varianssit tai jokaisen faktorin yksilataus nolasta poikkeavaksi, yleensä ykköseksi, on malli aina identifioituva. Silloin mittausmallit ovat hyvä lähtökohta poikkileikkausfaktorimallin rakentamiselle.

2.2.1.3 Parametrien estimointi

Olkoon havaintoaineisto kuten kappaleessa 2.1. Oletetaan, että aineistosta on estimoitu odotusarvovektori $\hat{\boldsymbol{\mu}}$ sekä otoskovarianssimatriisi \mathbf{S} , joihin parametrien estimointi voidaan perustaa kun havaitut y -muuttujat ovat jatkuvia. Jos y -muuttujat ovat järjestystason muuttujia, käytetään otoskovarianssimatriisin \mathbf{S} tilalla polykorista otoskorrelaatiomatriisia \mathbf{R} . Estimaattori \mathbf{S} on teoreettisen kovarianssimatriisin $\boldsymbol{\Sigma}$ tyhjentävä estimaattori.

Oletetaan seuraavaksi, että on spesifioitu identifioituva poikkileikkausfaktorimalli. Mallin tuntemattomat parametrit $\boldsymbol{\Lambda}$ -, $\boldsymbol{\Omega}$ - ja $\boldsymbol{\Theta}$ -matriiseissa estimoidaan siten, että niiden tuottaman kovarianssimatriisin $\boldsymbol{\Sigma}$ sovite

$$\hat{\boldsymbol{\Sigma}} = \hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Omega}}\hat{\boldsymbol{\Lambda}}' + \hat{\boldsymbol{\Theta}}$$

olisi mahdollisimman lähellä otoskovarianssimatriisia \mathbf{S} , Leskinen (1987).

Konfirmatorisen poikkileikkausfaktorimallin estimaattoreita ei voida johtaa yleisessä tapauksessa analyttisesti vaan parametrien estimointi suoritetaan iteratiivisesti tietokoneohjelman, esimerkiksi LISREL -ohjelman avulla, Bollen (1989). Parametrit voidaan estimoida joko vapaasti, yhtäsuurena jonkun toisen parametrin kanssa tai parametri voidaan kiinnittää vakioksi, Leskinen (1987). Parametrien estimointimenetelmiä on useita erilaisia. Seuraavassa esitellään suurimman uskottavuuden estimointimenetelmä sekä yleistetty pienimmän neliösumman menetelmä.

Suurimman uskottavuuden menetelmässä (Maximum Likelihood, lyhyesti ML-menetelmä) maksimoidaan otoksen uskottavuusfunktioita $\boldsymbol{\Lambda}$ -, $\boldsymbol{\Omega}$ - ja $\boldsymbol{\Theta}$ -parametrimatriiseissa olevien parametrien suhteen. Yhtäpitävää uskottavuusfunktion maksimoinnin kanssa on minimoida kohdefunktio

$$F_{ML} = \ln |\boldsymbol{\Sigma}(\boldsymbol{\Lambda}, \boldsymbol{\Omega}, \boldsymbol{\Theta})| + tr(\mathbf{S}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Lambda}, \boldsymbol{\Omega}, \boldsymbol{\Theta})) - \ln |\mathbf{S}| - \mathbf{p}, \quad (2.5)$$

sillä kohdefunktio saavuttaa miniminsä samassa pisteessä kuin uskottavuusfunktio saavuttaa maksiminsa, Leskinen (1987). Matriisit $\boldsymbol{\Sigma}$ ja \mathbf{S} oletetaan positiivisesti definiiteiksi, Bollen (1989). ML -menetelmässä oletetaan, että mittauskerran t havaitut y -muuttujat noudattavat p -ulotteista normaalijakaumaa. Lisäksi otoskoon on oltava riittävän suuri, yleensä vähintään 100.

Kun otoskoko on riittävän suuri, parametrien estimoinnin taustalla olevan otoskovarianssi- tai otoskorrelaatiomatriisin estimointi on luotettavaa. Tällöin myös ML-menetelmän hyvät estimointiteoreettiset ominaisuudet ovat voimassa. Menetelmän tuottama estimaattori on tarkentuva, asympotoottisesti tehokas sekä asympotoottisesti harhaton. Pienille otoksille estimaattori on kuitenkin harhainen. Estimaattorin asympotoottinen varianssi on tarkentuvien estimaattorien joukosta pienin. ML-menetelmä on riippumaton havaittujen y -muuttujien skaalauksista eli se on skaalainvariantti. Lisäksi menetelmä on skaalavapaa.

Yleistetyssä pienimmän neliösumman menetelmässä (engl. Generalized Least Squares, lyhyesti GLS) minimoidaan otoskovarianssimatriisin käänteismatriisilla \mathbf{S}^{-1} painotettuja jäännösneliösummia. Menetelmän kohdefunktio on muotoa

$$F_{GLS} = \frac{1}{2} tr \left\{ [\mathbf{I} - \mathbf{S}^{-1} \boldsymbol{\Sigma}(\boldsymbol{\Lambda}, \boldsymbol{\Omega}, \boldsymbol{\Theta})]^2 \right\}.$$

GLS -estimointimenetelmä on asympotoottisesti tehokas. Myös GLS -menetelmä on sekä skaalainvariantti että skaalavapaa. GLS -menetelmässä oletetaan, että havaitut muuttujat ovat joko normaalisti jakautuneita tai ainakin, etteivät jakaumat ole kovin huipukkaita. Jos havaittujen muuttujien jakaumat ovat hyvin paksu- tai ohuthäntäisiä, eivät asympotoottiset keskivirheet ole tarkkoja eikä χ^2 -merkitsevyytestaus anna tarkkoja tuloksia, Bollen (1989).

2.2.1.4 Mallin hypoteesien testaus

Tarkastellaan seuraavaksi poikkileikkausfaktorimallin suurimman uskottavuuden estimointia ja uskottavuussuhteen testiä. Uskottavuussuhteen testin avulla voidaan testata valitun poikkileikkausfaktorimallin sopivuutta mittauskerran t poikkileikkausaineistoon.

Minimoimalla kohdefunktio (2.5)

$$F_{ML} = \ln |\boldsymbol{\Sigma}(\boldsymbol{\Lambda}, \boldsymbol{\Omega}, \boldsymbol{\Theta})| + tr(\mathbf{S}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\Lambda}, \boldsymbol{\Omega}, \boldsymbol{\Theta})) - \ln |\mathbf{S}| - p$$

estimoitavien parametrien suhteen on saatu suurimman uskottavuuden estimaatit $\hat{\boldsymbol{\Lambda}}$, $\hat{\boldsymbol{\Omega}}$ ja $\hat{\boldsymbol{\Theta}}$. Silloin kovarianssimatriisin $\boldsymbol{\Sigma}$ sovitteeksi saadaan

$$\hat{\Sigma} = \hat{\Lambda}\hat{\Omega}\hat{\Lambda}' + \hat{\Theta}.$$

Kun nyt nollahypoteesiksi ja vastahypoteesiksi valitaan hypoteesipari

$$H_0 : \Sigma = \Lambda\Omega\Lambda' + \Theta$$

H_1 : Kovarianssimatriisissa Σ ei ole rajoituksia,

saadaan poikkileikkausfaktorimallin aineistoon sopivuudelle uskottavuussuhteen testi

$$\lambda = \frac{\max L_{H_0} \{\Sigma(\Lambda, \Omega, \Theta); \mathcal{S}\}}{\max L_{H_0 \cup H_1} \{\Sigma(\Lambda, \Omega, \Theta); \mathcal{S}\}}.$$

Nollahypoteesin H_0 ollessa voimassa maksimi saavutetaan kun

$$\hat{\Sigma} = \hat{\Lambda}\hat{\Omega}\hat{\Lambda}' + \hat{\Theta},$$

ja vastahypoteesin H_1 vallitessa maksimi saavutetaan kun

$$\hat{\Sigma} = \mathcal{S}.$$

Nollahypoteesin ollessa voimassa testisuure $-2 \ln \lambda$ noudattaa asympotoottisesti χ^2 -jakaumaa vapausastein

$$df = t_1 - t_0 = \frac{1}{2}p(p+1) - t_0,$$

missä $t_1 = \frac{1}{2}p(p+1)$ on vastahypoteesin H_1 mukaisen rajoittamattoman poikkileikkausfaktorimallin estimoitavien parametrien lukumäärä ja t_0 on hypoteesin H_0 määräämän poikkileikkausfaktorimallin estimoitavien parametrien lukumäärä. Jotta testi olisi käyttökelpoinen, on ehdon $t_1 > t_0$ oltava voimassa. Silloin uskottavuussuhteen testifunktioksi saadaan

$$\begin{aligned}
-2 \ln \lambda &= -2 \ln \left(\frac{\max L_{H_0}}{\max L_{H_0 \cup H_1}} \right) \\
&= -2 \ln (\max L_{H_0}) + 2 \ln (\max L_{H_0 \cup H_1}) \\
&= (N - 1) \left[\ln |\hat{\Sigma}| + \text{tr}(\hat{\Sigma}^{-1} \mathbf{S}) \right] - (N - 1) [\ln |\mathbf{S}| + p] \\
&= (N - 1) \left[\ln |\hat{\Sigma}| + \text{tr}(\hat{\Sigma}^{-1} \mathbf{S}) - \ln |\mathbf{S}| - p \right].
\end{aligned}$$

Testifunktio voidaan yhtäpitävästi esittää muodossa

$$-2 \ln \lambda = (N - 1) F_{ML}(\hat{\Lambda}, \hat{\Omega}, \hat{\Theta}),$$

jolloin asymptoottisen χ^2 -testin testisuuren arvo saadaan suurimman uskottavuuden estimoinnissa käytettävän kohdefunktion F_{ML} minimin avulla. Estimoidulle χ^2 -arvolle voidaan laskea myös sitä vastaava p -arvo, joka ilmaisee todennäköisyyden

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

eli estimoidun poikkileikkausfaktorimallin empiirisen merkitsevyydystason, kun H_0 -hypoteesi saa tukea. Valitaan merkitsevyydystaso $\alpha = 0.05$. Jos $p \geq 0.05$, saa H_0 -hypoteesi tukea eli sanotaan, että poikkileikkausfaktorimalli sopii aineistoon. Jos sen sijaan

$p < 0.05$, ei H_0 -hypoteesi ole voimassa eikä poikkileikkausfaktorimalli sovi aineistoon.

Kun tutkitaan estimoidun poikkileikkausfaktorimallin yhteensopivuutta mittauskerran t poikkileikkausaineistoon, käytetään χ^2 -testiä yhteensopivuusmittana. Päätöstä poikkileikkausfaktorimallin hyväksymisestä tai hylkäämisestä ei kuitenkaan pitäisi tehdä ainoastaan χ^2 -testin perusteella vaan testisuuretta tulisi arvioida myös suhteessa vapausasteisiin: Jos estimoitu χ^2 -arvo on suuri suhteessa testin vapausasteisiin ($\hat{\chi}^2 > 2df$), on estimoitu poikkileikkausfaktorimalli liian yksinkertainen ja sitä tulisi modifioida. Jos taas estimoitu χ^2 -arvo on pieni vapausasteisiin verrattaessa, voi malli olla liian yksityiskohtainen ja tulosten yleistettävyyden siten kyseenalaista.

Koska χ^2 -testisuure

$$-2 \ln \lambda = (N - 1) F_{ML}(\hat{\Lambda}, \hat{\Omega}, \hat{\Theta})$$

on suoraan verrannollinen otoksen havaintojen lukumäärään N , saattaa käytännössä syntyä ongelmia poikkileikkausfaktorimallia koskevia hypoteeseja testattaessa otoskoon ollessa hyvin suuri ($N > 500$). Silloin χ^2 -yhteensopivuustesti hylkää estimoidun faktorimallin erittäin herkästi, vaikka kovarianssimatriisin sovitetun $\hat{\Sigma}$ ja otoskovarianssimatriisin S välinen poikkeama olisikin hyvin pieni. Tästä johtuen on kehitetty erilaisia χ^2 -testisuureeseen perustuvia yhteensopivuusindeksejä, mm. Bentlerin ja Bonettin (1980) kehittämä normeerattu yhteensopivuusindeksi (Normed Fit Index, lyhyesti NFI), joissa suuren otoskoon vaikutus χ^2 -yhteensopivuustestiin on eliminoitu. NFI -indeksiin palataan uudelleen poikkileikkausfaktorimallin riittävyystarkastelujen yhteydessä.

Toisaalta myös liian pieni otoskoko ($N < 100$) saattaa aiheuttaa ongelmia mallia koskevia hypoteeseja testattaessa. Koska suurelta otoskovarianssimatriisin S ja sovitetun $\hat{\Sigma}$ väliset erot eivät ole välttämättä tilastollisesti merkitseviä kun N on pieni, saattaa χ^2 -yhteensopivuustesti hyväksyä useita erilaisia mallivaihtoehtoja aineistoon sopivina. Tällöin χ^2 -testi ei ole kovinkaan tehokas ja sen käyttö mallin valintakriteerinä ei ole perusteltua.

Kun tutkimushypoteesia halutaan tarkentaa testaamalla sisäkkäisiä hypoteeseja, voidaan käyttää χ^2 -peräkkäistestiä. Oletetaan, että valittu H_0 -hypoteesin mukainen poikkileikkausfaktorimalli M_1 on oikea. Kun mallia M_1 testattiin yleistä vastahypoteesia H_1 vastaan, saatiin testisuureen arvo $\hat{\chi}_1^2$ sekä vapausasteet df_1 . Haluttaessa tarkentaa tutkimushypoteesia tutkitaan poikkileikkausfaktorimallin M_1 erikoistapausta, poikkileikkausfaktorimallia M_2 , joka kuuluu mallin M_1 määrittämien mallien joukkoon. Silloin nollahypoteesia

$$H_0 : \text{Malli } M_2 \text{ on oikea}$$

voidaan testata aiempaa nollahypoteesimallia M_1 vastaan. Vastahypoteesi on nyt muotoa

$$H_1 : \text{Malli } M_1 \text{ on oikea,}$$

missä $H_0 \subset H_1$. Testisuureen arvoksi saadaan $\hat{\chi}_2^2$ ja vapausasteiksi df_2 . Jos uusi H_0 -hypoteesi on voimassa, noudattaa peräkkäistestien testisuure

$$D = \hat{\chi}_2^2 - \hat{\chi}_1^2$$

asymptoottista χ^2 -jakaumaa vapausastein

$$df = df_2 - df_1.$$

2.2.1.5 Mallin riittävyystarkastelut

Estimoidun poikkileikkausfaktorimallien riittävyttä voidaan arvioida myös muilla tavoin kuin χ^2 -yhteensopivuustestin avulla. Riittävyystarkasteluja voidaan suorittaa koko faktorimallille, yksittäisille muuttujille, yksittäisille parametreille sekä yksittäisille havainnoille, Leskinen (1987).

Koko poikkileikkausfaktorimallia koskevia riittävyysindeksejä ovat χ^2 -yhteensopivuustestin lisäksi mm. approksimointivirhe RMSEA (Root Mean Square Error of Approximation, Browne & Cudeck (1993)), normeerattu yhteensopivuuksindeksi NFI (Normed Fit Index, Bentler & Bonett (1980)), suhteellinen yhteensopivuuksindeksi CFI (Comparative Fit Index, Tanaka (1993)) sekä keskimääräinen standardoitu jäännös SRMR (Standardized Root Mean Square Residual, Bollen (1989)).

Approksimointivirhe RMSEA on epäkeskiseen eli vastahypoteesin mukaiseen χ^2 -jakaumaan perustuva riittävyysindeksi, joka kertoo faktorimallin liiallisesta yksinkertaisuudesta. RMSEA on muotoa

$$RMSEA = \sqrt{\frac{d}{df}},$$

missä

$$d = \max \left\{ \frac{\hat{\chi}^2 - df}{N - 1}, 0 \right\}.$$

Estimoitu malli sopii RMSEA -indeksin mielessä hyvin aineistoon, jos indeksin arvo on enintään 0.05. Jos $0.05 < RMSEA \leq 0.08$, voidaan mallin sopivuutta aineistoon

pitää vielä tyydyttävänä. Approksimointivirheelle on olemassa jakauma, joten approksimointivirheelle voidaan laskea myös todennäköisyysarvio $P(RMSEA \leq 0.05)$ eli p -arvo. Kun p -arvo on vähintään 0.05, voidaan RMSEA -indeksin perusteella sanoa, että malli sopii aineistoon hyvin, Browne & Cudeck (1993).

Normeeratun yhteensopivuusindeksin NFI avulla voidaan eliminoida suuren otokoon vaikutusta χ^2 -yhteensopivuustestiin. Indeksissä testattavan poikkileikkausfaktorimallin estimoitua χ^2 -arvoa verrataan riippumattomuusmallin estimoituun χ^2 -arvoon. NFI -indeksi on silloin muotoa

$$NFI = \frac{\hat{\chi}_0^2 - \hat{\chi}_1^2}{\hat{\chi}_0^2} = \frac{(N-1)\hat{F}_0 - (N-1)\hat{F}_1}{(N-1)\hat{F}_0} = \frac{\hat{F}_0 - \hat{F}_1}{\hat{F}_0},$$

missä $\hat{\chi}_0^2$ on χ^2 -testin arvo riippumattomuusmallin tilanteessa ja $\hat{\chi}_1^2$ on χ^2 -testin arvo estimoidun mallin tilanteessa, Bentler & Bonett (1980). Riippumattomuusmalli on rajoitetuin mahdollinen malli. Konfirmatorisen faktorianalyysin tapauksessa riippumattomuusmallissa oletetaan, että havaittujen y -muuttujien taustalla ei ole faktorirakenteita ja että havaitut muuttujat ovat keskenään korreloimattomia, Bollen (1990). NFI -indeksin arvoalue on

$$0 \leq NFI \leq 1.$$

Jos indeksin arvo on vähintään 0.90, voidaan sovitettavaa mallia pitää riittävänä; toisaalta jos $NFI < 0.90$, ei sovitettava malli ole NFI -indeksin mielessä riittävä, Bentler & Bonett (1980).

Suhteellinen yhteensopivuusindeksi CFI on approksimointivirheen tavoin epäkeskisyysjakaumaan perustuva riittävyysindeksi, joka rankaisee suuresta estimoitavien parametrien lukumäärästä. CFI määritellään muodossa

$$CFI = \frac{(\hat{\chi}_0^2 - df_0) - (\hat{\chi}_1^2 - df_1)}{\hat{\chi}_0^2 - df_0},$$

missä alaindeksi 0 viittaa riippumattomuusmalliin ja alaindeksi 1 testattavaan malliin. Myös CFI -indeksin arvoalue on

$$0 \leq CFI \leq 1.$$

Indeksin ollessa vähintään 0.90, on sovitettava malli CFI -indeksin mielessä riittävä. Jos CFI -indeksi on alle 0.90, ei sovitettavaa mallia voida pitää riittävänä kuvaamaan aineistoa, Tanaka (1993).

Keskimääräinen jäännös RMR on jäännösmatriisiin $\mathbf{S} - \hat{\Sigma}$ perustuva koko mallia arvioiva riittävyysindeksi, joka mittaa keskimääräistä jäännöskovarianssia ja -varianssia. RMR -indeksi on muotoa

$$RMR = \sqrt{\frac{2 \sum_{i=1}^p \sum_{j=1}^i (s_{ij} - \hat{\sigma}_{ij})^2}{p(p+1)}},$$

missä alkiotason jäännöskovarianssit ovat muotoa $(s_{ij} - \hat{\sigma}_{ij})$. Jos jäännöskovarianssi on suuri ja positiivinen, poikkileikkausfaktorimalli aliestimoi muuttujien y_i ja y_j välistä kovarianssia; jos sen sijaan jäännöskovarianssi on suuri ja negatiivinen, muuttujien y_i ja y_j välinen estimoitu kovarianssi on liian suuri. Koska RMR -indeksiä pitäisi tutkia suhteessa havaittuihin kovariansseihin ja variansseihin, on helpompaa käyttää standardoitua keskimääräistä jäännöstä SRMR, Bollen (1989).

Muuttujakohtaisia tarkasteluja voidaan tehdä, kun jokaiselle havaitulle y -muuttujalle määritellään muuttujakohtainen riittävyysindeksi \hat{R}_i^2 (*squared multiple correlation*):

$$\hat{R}_i^2 = 1 - \frac{\hat{\theta}_i}{\hat{\sigma}_{ii}}, \quad i = 1, \dots, p.$$

Indeksi \hat{R}_i^2 kertoo, kuinka hyvin yksittäinen y_i -muuttuja mittaa faktorirakennetta. Jos indeksin arvo on lähellä ykköstä, vastaava y_i -muuttuja mittaa hyvin faktorirakennetta. Jos indeksin arvo on lähellä nollaa, ei kyseinen y_i -muuttuja toimi lainkaan mittarina ja se on joko poistettava mallista tai mallia on modifioitava muuttujan suhteen. Indeksi \hat{R}_i^2 voidaan tulkita havaitun muuttujan kommunaliteetti- tai reliabiliteetikertoimeksi, Leskinen (1987).

Parametrikohteisilla tarkasteluilla on tärkeä rooli tutkittaessa poikkileikkausfaktorimallin riittävyttä. Parametrikohteiset tarkastelut tehdään kaikille parametreille riippumatta siitä, onko parametri kiinnitetty vakioksi, estimoitu yhtäsuurena jonkun toisen parametrin kanssa tai estimoitu vapaasti.

Estimoitujen parametrien arvojen tulee ensinnäkin olla sisällöllisesti tulkittavissa. Lisäksi estimaattien on oltava tilastollisesti kelvollisia: faktoreiden varianssien esti-

maattien $\hat{\omega}$ sekä jäännösvarianssien estimaattien $\hat{\theta}$ on oltava positiivisia, faktoreiden välisten korrelaatiokertoimien estimaattien on oltava välillä $[-1, 1]$ ja estimoitujen kovarianssimatriisien $\hat{\Omega}$ ja $\hat{\Theta}$ on oltava positiivisesti definiittejä. Mikäli poikkileikkausfaktorimallissa on tilastollisesti kelvottomia parametrien estimaatteja, ei mallia saa käyttää vaan sitä on modifioitava.

Oletetaan, että estimoitava poikkileikkausfaktorimalli on identifioituva ja että mallin parametrien estimoinnissa on käytetty suurimman uskottavuuden menetelmää. Silloin parametrien estimointitarkkuutta ja parametrien tilastollista merkitsevyyttä voidaan tarkastella estimaattoreiden keskivirheiden sekä estimaattoreiden välisten korrelaatioiden avulla. Jos parametrin estimaattorin keskivirhe on pieni, on kyseisen parametrin estimointi tilastollisesti tarkkaa. Keskivirheiden avulla voidaan parametrien estimaateille laskea edelleen luottamusväli ja t -arvot. Luottamusväli, joka sisältää 95% prosentoin todennäköisyydellä tuntemattoman parametrin, voidaan laskea kaavan

$$\text{parametrin estimaatti} \pm 1.96 \times \text{keskivirhe}$$

avulla. Parametrien tilastollista merkitsevyyttä voidaan arvoida vertaamalla parametrin estimaattia keskivirheeseensä, jolloin saadaan

$$t\text{-arvo} = \frac{\text{parametrin estimaatti}}{\text{keskivirhe}}.$$

Parametri tulkitaan nolasta eroavaksi, kun $|t\text{-arvo}| > 1.96$. Tämä vastaa t -testin käyttöä likimain 5% merkitsevyytasolla.

Estimaattoreiden väliset voimakkaat korrelaatiot merkitsevät sitä, että kyseisten parametrien estimoinnit riippuvat voimakkaasti toisistaan. Silloin poikkileikkausfaktorimalli voidaan tulkita kyseisten parametrien suhteen lähes ei-identifioituvaksi. Ongelmaa voidaan korjata vähentämällä poikkileikkausfaktorimallista estimoitavien parametrien määrää.

Konfirmatorisen faktorimallin parametrien nolakiinnitysten tai muiden rajoitusten sopivuutta voidaan arvioida *modifikaatioindeksien* (lyhyesti MI) avulla. Indeksien laskeminen perustuu estimointimenetelmän kohdefunktiosta kiinnitettyjen tai rajoitettujen parametrien suhteen laskettuihin ensimmäisiin ja toisiin derivaattoihin. Modifikaatioindeksin arvo kertoo, kuinka paljon estimoidun poikkileikkausfaktorimallin

χ^2 -testisuure vähintään laskee, jos kyseinen mallin parametri vapautetaan estimoitavaksi, Leskinen (1987). Oletetaan, että on estimoitu yli-identifioituva poikkileikkausfaktori malli. Merkitään mallin parametreja vektorilla $\boldsymbol{\tau}$,

$$\boldsymbol{\tau} = \text{vec}(\boldsymbol{\Lambda}, \boldsymbol{\Omega}, \boldsymbol{\Theta}).$$

Oletetaan, että parametri τ_i on kiinnitetty nolaksi. Modifikaatioindeksi on silloin testisuure nollahypoteesille

$$H_0 : \tau_i = 0.$$

Testisuure noudattaa χ^2 -jakaumaa yhdellä vapausasteella,

$$MI(\tau_i) \sim \chi^2(1).$$

Vastahypoteesi on muotoa

$$H_1 : \tau_i \neq 0.$$

Yleensä testin merkitsevyystasoksi valitaan joko $\alpha = 0.05$ tai $\alpha = 0.005$. Merkitsevyystasolla $\alpha = 0.05$ χ^2 -vertailuarvoksi saadaan

$$\chi_{0.95}^2 = 3.84.$$

Jos $MI(\tau_i) < 3.84$, on H_0 voimassa eli nolaksi kiinnitettyä parametria ei ole järkevää vapauttaa. Jos sen sijaan $MI(\tau_i) \geq 3.84$, H_0 hylätään ja kyseinen parametri vapautetaan estimoitavaksi. Parametrien vapauttamisessa on kuitenkin huomioitava, että mallia voidaan modifioida vain sellaisten parametrien osalta, jotka ovat tutkimusongelman kannalta mielekkäitä. Merkitsevyystasolla $\alpha = 0.005$ χ^2 -vertailuarvo puolestaan on

$$\chi_{0.995}^2 = 7.85.$$

Jos $MI(\tau_i) < 7.85$, on nollahypoteesi voimassa eikä nollassa kiinnitettyä parametria ei ole järkevää vapauttaa. Jos sen sijaan $MI(\tau_i) \geq 7.85$, nollahypoteesi hylätään ja kyseisen parametrin voi vapauttaa estimoitavaksi, jos se on tulkinntallisesti järkevää, Bollen (1989).

Havaintokohtaiset tarkastelut perustuvat estimoidun poikkileikkausfaktorimallin tuottamiin jäännöksiin kovariansseille ja variansseille:

$$\mathbf{S} - \hat{\Sigma} = [s_{ij} - \hat{\sigma}_{ij}], \quad i, j = 1, \dots, p.$$

Suorittamalla jäännösten standardointi saadaan jäännökset helpommin vertailtaviksi. Asymptoottisesti on voimassa, että

$$E(s_{ij}) \approx \sigma_{ij}$$

ja

$$Var(s_{ij}) \approx (\sigma_{ii}\sigma_{jj} + \sigma_{ij}^2)/N, \quad i, j = 1, \dots, p.$$

Silloin jäännösten standardointi suoritetaan kaavalla

$$NR_{ij} = \frac{s_{ij} - \hat{\sigma}_{ij}}{\sqrt{(\hat{\sigma}_{ii}\hat{\sigma}_{jj} + \hat{\sigma}_{ij}^2)/N}}, \quad i, j = 1, \dots, p.$$

Estimoidun poikkileikkausfaktorimallin ollessa oikea saadaan standardoitujen jäännösten odotusarvoksi $E(NR_{ij}) \approx 0$ ja varianssiksi $Var(NR_{ij}) \approx 1$ kaikilla i, j . Poikkileikkausfaktorimalli on riittämätön, jos

$$|NR_{ij}| > 1.96.$$

Tällöin malli ei kykene selittämään kyseistä jäännöstä vastaavien havaittujen muuttujien y_i ja y_j välistä kovarianssia tai korrelaatiota, joten malli tulisi spesifoida uudelleen, Leskinen (1987).

2.2.2 Poikkileikkausfaktorimallien yhdistäminen pitkittäisfaktorimalliksi sekä mallin rakentaminen

Tarkastellaan seuraavaksi eri mittauskertojen poikkileikkausfaktorimallien yhdistämistä pitkittäisfaktorimalliksi. Oletetaan, että jokaiselle mittauskerralle t on rakennettu konfirmatorinen faktorimalli (2.3). Yksittäisten mittauskertojen konfirmatoriset faktorimallit ovat muotoa

$$\mathbf{y}_t = \mathbf{\Lambda}_t \boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \dots, T.$$

Silloin havaittujen muuttujien mittauskertojen väliset kovarianssimatriisit (2.2) ovat muotoa

$$\boldsymbol{\Sigma}_{st} = \mathbf{\Lambda}_s \boldsymbol{\Omega}_{st} \mathbf{\Lambda}'_t + \boldsymbol{\Theta}_{st}, \quad s, t = 1, \dots, T; s \neq t.$$

Yleensä oletetaan, että mittauskertojen välisten jäännösten kovarianssimatriisit $\boldsymbol{\Theta}_{st}$ ovat nollamatriiseja. Yhdistämällä eri mittauskertojen poikkileikkausfaktorimallit mittauskertojen välisillä kovarianssimetriiseilla toisiinsa saadaan pitkittäisfaktorimalli (2.1)

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

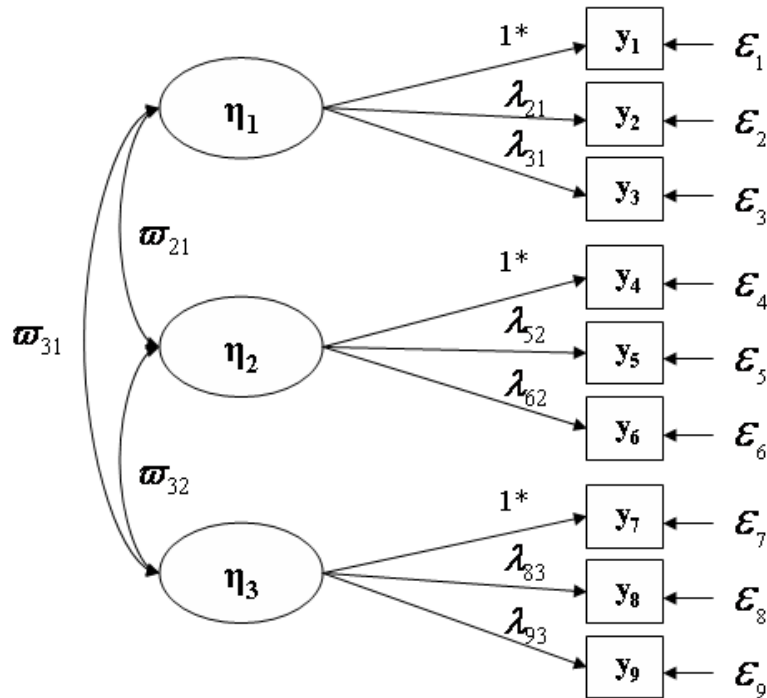
Tisak & Meredith (1990).

Pitkittäisfaktorimallin rakentamisessa on erotettavissa samat viisi vaihetta kuin poikkileikkausfaktorimallin rakentamisessakin: mallin spesifointi, identifioituvuustarkastelut, parametrien estimointi, mallin hypoteesien testaus ja riittävyystarkastelut. Oletetaan, että jokaiselle mittauskerralle t , jossa $t = 1, \dots, T$, on rakennettu oma konfirmatorinen faktorimallinsa. Tarkastellaan mallinrakennuksen vaiheita vain mitausten välisille kovarianssirakenteille, Leskinen (1995).

Pitkittäisfaktorimallin rakentaminen aloitetaan spesifioimalla eri mittauskertojen faktoreiden väliset ω -kovarianssiparametrit tutkimushypoteesin pohjalta. Seuraavaksi suoritetaan spesifioitujen kovarianssiparametrien identifioituvuustarkastelut. Koska eri mittauskertojen poikkileikkausfaktorimallit ovat identifioituvia, myös mittauskertojen väliset faktoreiden kovarianssit ovat identifioituvia. Identifioituvuustarkastelujen jälkeen suoritetaan kovarianssiparametrien estimointi, mallin hypoteesien testaus

sekä riittävyystarkastelut.

Esimerkki 2.3. Spesifioidaan seuraavaksi esimerkin 2.1 kolmen mittauskerran pitkittäisfaktorimalli. Yksittäisten mittauskertojen poikkileikkausfaktorimallit spesifioitiin jo esimerkissä 2.2, joten riittää, kun spesifioidaan eri mittauskertoja yhdistävät kovarianssit sekä mittauskertojen väliset jäännöskovarianssit. Oletetaan, että esimerkin pitkittäisfaktorimalli on rekursiivinen malli, joten spesifioitavat parametrit ovat mittauskertojen 1 ja 2 faktoreiden välinen kovarianssi ω_{21} , mittauskertojen 2 ja 3 faktoreiden välinen kovarianssi ω_{32} sekä mittauskertojen 1 ja 3 faktoreiden välinen kovarianssi ω_{31} . Estimoidaan nämä kovarianssit vapaasti. Mittauskertojen väliset jäännöskovarianssimatriisit oletettiin nollamatriiseiksi, joten kiinnitetään kaikki mittauskertojen väliset jäännöskovarianssit θ_{st} , jossa $s \neq t$; $s, t = 1, 2, 3$, nolliksi. Spesifioitu pitkittäisfaktorimalli on esitetty graafisesti kuvassa 2.



Kuva 2: Hypoteettisen kolmen mittauskerran pitkittäisfaktorimallin spesifointi. Tähdellä (*) merkityt parametrit kiinnitetty.

2.3 Pitkittäisfaktorimallin ryhmävertailuista

Seuraavaksi tarkastellaan pitkittäisfaktorimallin (2.1) yleistämistä kahden tai useamman perusjoukon (ryhmän) tilanteisiin. Keskeisenä tavoitteena on tällöin vertailla eri perusjoukoille rakennettuja pitkittäisfaktorimalleja toisiinsa. Vertailu suoritetaan estimoimalla eri perusjoukkojen faktorimallit samanaikaisesti asettamalla mallien parametrien välille sopivia yhtäsuuruussidoksia ja testaamalla saatuja ratkaisuja χ^2 -yhteensopivuustestillä, Leskinen (1995); Tisak & Meredith (1990).

Olkoon tarkastelun kohteena G perusjoukkoa eli ryhmää. Oletetaan, että jokaisesta perusjoukosta on poimittu toisistaan riippumattomasti otokset, joiden koot ovat N_1, N_2, \dots, N_G . Havaintoaineisto on silloin muotoa

$$Y_{ijt}^{(g)}, \quad i = 1, \dots, N_g; j = 1, \dots, p; t = 1, \dots, T; g = 1, \dots, G,$$

missä

N_g on ryhmän g otoskoko,

p on havaittujen muuttujien lukumäärä mittauskerralla t ja

T on mittauskertojen lukumäärä.

Oletetaan edelleen, että kullekin ryhmälle on estimoitu $pT \times pT$ -otoskovarianssimatriisi $\mathbf{S}^{(g)}$, jossa $g = 1, \dots, G$. Ryhmän g pitkittäisfaktorimalliesitykseksi saadaan tällöin

$$\mathbf{y}^{(g)} = \mathbf{\Lambda}^{(g)}\boldsymbol{\eta}^{(g)} + \boldsymbol{\varepsilon}^{(g)}, \quad g = 1, \dots, G,$$

jolloin havaittujen y -muuttujien teoreettinen $pT \times pT$ -kovarianssimatriisi ryhmälle g on muotoa

$$\text{cov}(\mathbf{y}^{(g)}) = \boldsymbol{\Sigma}^{(g)} = \mathbf{\Lambda}^{(g)}\boldsymbol{\Omega}^{(g)}\mathbf{\Lambda}^{(g)'} + \boldsymbol{\Theta}^{(g)}, \quad g = 1, \dots, G,$$

Leskinen (1995); Tisak & Meredith (1990).

Pitkittäisfaktorimallin rakentaminen usealle perusjoukolle samanaikaisesti voidaan suorittaa samoin kuin yhden perusjoukon tapauksessakin. Kaikille ryhmille g , jossa $g = 1, \dots, G$, rakennetaan ensin poikkileikkausfaktorimallit jokaiselle mittauskerralle t , jossa $t = 1, \dots, T$. Poikkileikkausfaktorimallien rakentaminen tapahtuu kuten

kappaleessa 2.2.1. Ensin spesifioidaan malli tutkimushypoteesin pohjalta. Yleensä jokaiselle ryhmälle spesifioidaan rakenteeltaan samanlainen konfirmatorinen faktorimalli. Tämä mahdollistaa faktorimallien parametrien yhtäsuuruusvertailut. Spesifioinnin jälkeen suoritetaan mallin identifioituvuustarkastelut, parametrien estimointi, hypoteesien testaus ja riittävyystarkastelut. (Leskinen 1995). Poikkileikkausfaktorimallien rakentamisen jälkeen muodostetaan jokaiselle ryhmälle g oma pitkittäisfaktorimalli yhdistämällä poikkileikkausfaktorimallit mittausten välisillä kovarianssirakenteilla toisiinsa, Tisak & Meredith (1990).

2.4 Pitkittäisfaktorimallin erikoistapauksia

Seuraavassa tarkastellaan pitkittäisfaktorimallin erikoistapauksina ajan suhteen stationaarista faktorimallia, ryhmien suhteen invarianttia faktorimallia sekä näiden yhdistelmää, stationaarista ja invarianttia faktorimallia. Näiden mallien voimassaoloa voidaan testata esimerkiksi χ^2 -peräkkäistestien avulla. Lähteenä on käytetty artikkeleita Tisak & Meredith (1990).

Pitkittäisfaktorimalli (2.1)

$$\mathbf{y} = \mathbf{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

on ajan suhteen *stationaarinen*, jos sille on voimassa

$$\mathbf{\Lambda}_{11} = \mathbf{\Lambda}_{22} = \dots = \mathbf{\Lambda}_{TT}.$$

Stationaarisuuden voimassaolo on tärkeää, sillä se ilmaisee pitkittäisfaktorimallin avulla kuvatun ilmiön samanlaisuutta eri mittauskerroilla. Jos malli ei ole stationaarinen, aiheuttaa se ongelmia mittausten välisissä vertailuissa.

Jos stationaarisuus on voimassa, voidaan edelleen tutkia ovatko eri mittauskertojen faktoreiden kovarianssimatriisit yhtäsuuria,

$$\mathbf{\Omega}_{11} = \mathbf{\Omega}_{22} = \dots = \mathbf{\Omega}_{TT},$$

tai ovatko eri mittauskertojen jäännösten kovarianssimatriisit yhtäsuuret,

$$\Theta_{11} = \Theta_{22} = \dots = \Theta_{TT}.$$

Pitkittäisfaktorimallin erikoistapauksena ryhmävertailutilanteessa voidaan tutkia ryhmien invarianttisuutta. Ryhmien suhteen invariantille mallille on voimassa

$$\Lambda_{tt}^{(1)} = \Lambda_{tt}^{(2)} = \dots = \Lambda_{tt}^{(G)}, \quad t = 1, \dots, T.$$

eli ryhmien lataukset ovat yhtäsuuret mittauskerralla t .

Kolmas pitkittäisfaktorimallin erikoistapaus on *stationaarinen ja invariantti* pitkittäisfaktorimalli, jolle on voimassa

$$\Lambda_{tt}^{(g)} = \Lambda, \quad t = 1, \dots, T; \quad g = 1, \dots, G.$$

2.5 Faktoreiden rakenneyhtälömallit

Tässä kappaleessa käsitellään pitkittäisfaktorimallin faktoreiden eri mittauskertojen välisten rakenneyhtälöiden teoriaa ja rakentamisen vaiheita. Lähteenä on käytetty teosta Leskinen (1987).

Tarkastellaan edelleen pitkittäisfaktorimallia (2.1)

$$\mathbf{y} = \Lambda \boldsymbol{\eta} + \boldsymbol{\varepsilon}.$$

Pitkittäisfaktorimallissa faktoreiden mittauskertojen väliset yhteydet voidaan esittää myös faktoreiden rakenneyhtälöiden avulla. Tällöin faktoreiden oletetaan olevan ajan suhteen rekursiivisia eli faktoreiden oletetaan selittävän tai ennustavan toisiaan.

Olkoon faktoreiden mittauskertojen väliset rakenneyhtälöt muotoa

$$\boldsymbol{\eta} = B\boldsymbol{\eta} + \boldsymbol{\zeta}, \tag{2.6}$$

missä

\mathbf{B} on β -rakenneparametrit sisältävä $kT \times kT$ -matriisi ja

ζ on η -faktoreiden välisten rakenneyhtälöiden jäännöstermit sisältävä $kT \times 1$ -vektori.

Koska rakenneyhtälöt oletetaan ajan suhteen rekursiivisiksi, voidaan rakenneparametrimatriisi \mathbf{B} esittää aina alakolmiomuodossa. Merkitään lisäksi rakenneyhtälöiden ζ -jäännösten kovarianssimatriisia seuraavasti:

$$\text{cov}(\zeta) = \Psi.$$

Silloin faktoreiden rakenneyhtälömalli on kokonaisuudessaan muotoa

$$\begin{cases} \mathbf{y} &= \mathbf{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\eta} &= \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\zeta}. \end{cases} \quad (2.7)$$

Oletetaan, että matriisi $\mathbf{I} - \mathbf{B}$ on epäsingulaarinen. Rakenneyhtälöt (2.6) voidaan esittää nyt muodossa

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

jolloin η -faktoreiden kovarianssimatriisi voidaan esittää faktoreiden välisten rakenneyhtälöiden avulla muodossa

$$\begin{aligned} \text{cov}(\boldsymbol{\eta}) = \boldsymbol{\Omega} &= E [[(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\zeta}][\boldsymbol{\zeta}'(\mathbf{I} - \mathbf{B}')^{-1}]] \\ &= (\mathbf{I} - \mathbf{B})^{-1}E(\boldsymbol{\zeta}\boldsymbol{\zeta}')(\mathbf{I} - \mathbf{B}')^{-1} \\ &= (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I} - \mathbf{B}')^{-1}. \end{aligned}$$

Havaittujen y -muuttujien teoreettinen kovarianssimatriisi on silloin muotoa

$$\begin{aligned} \text{cov}(\mathbf{y}) = \boldsymbol{\Sigma} &= \mathbf{\Lambda}\boldsymbol{\Psi}\boldsymbol{\Lambda}' + \boldsymbol{\Theta} \\ &= \mathbf{\Lambda} [(\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I} - \mathbf{B}')^{-1}] \boldsymbol{\Lambda}' + \boldsymbol{\Theta}. \end{aligned}$$

Faktoreiden mittauskertojen välisillä rakenneyhtälöille voidaan johtaa selityskertoimia, joiden avulla voidaan arvioida selittävien faktoreiden selitys- tai ennustuskykyä. Selityskertoimet ovat muotoa

$$R_{\eta_t}^2 = 1 - \frac{\psi_t}{\text{var}(\eta_t)}, \quad t = 2, \dots, T,$$

missä ψ_t on t :nneen rakenneyhtälön jäännöksen ζ_t varianssi. Estimoitujen $R_{\eta_t}^2$ -kertoimien avulla voidaan rakenneyhtälömallin riittävyystarkastelujen yhteydessä arvioida, kuinka hyvin selittävillä faktoreilla on pystytty selittämään kunkin selitettävän η_t -faktorin vaihtelua.

Faktoreiden rakenneyhtälömallien selityskertoimista saadaan johdettua rakenneyhtälömallien pysyvyys- eli *stabiilisuuskertoimet*. Jos kutakin faktoria selittää vain edellisen mittauskerran faktori, ovat selityskertoimet muotoa

$$R(\eta_t) = \sqrt{R_{\eta_t}^2} = \rho(\eta_t, \eta_{t-1}), \quad t = 2, \dots, T,$$

missä $\rho(\eta_t, \eta_{t-1})$ on peräkkäisten mittauskertojen faktoreiden välinen korrelaatiokerroin.

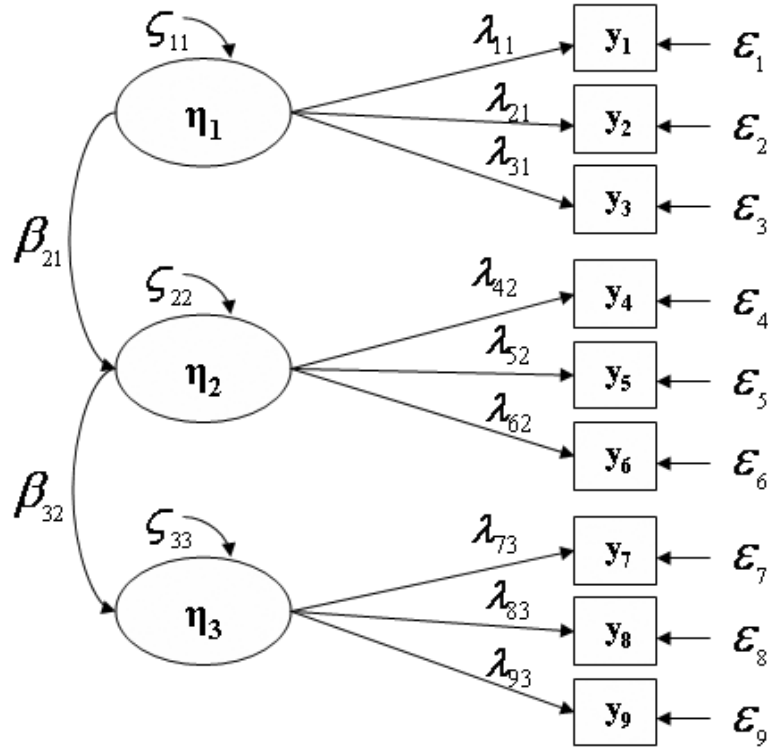
Esimerkki 2.4. Jatketaan esimerkin 2.1 pitkittäisfaktorimallin tarkastelua. Kuvaetaan seuraavaksi mallin faktoreiden mittauskertojen väliset yhteydet rakenneyhtälöiden avulla. Rakenneyhtälömallin graafinen esitys on kuvassa 3. Mallin faktoreiden rakenneyhtälöt ovat muotoa

$$\begin{cases} \eta_1 &= \zeta_1 \\ \eta_2 &= \beta_{21}\eta_1 + \zeta_2 \\ \eta_3 &= \beta_{32}\eta_2 + \zeta_3. \end{cases}$$

Silloin rakenneparametrimatriisi \mathbf{B} on muotoa

$$\mathbf{B} = \begin{pmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ 0 & \beta_{32} & 0 \end{pmatrix}.$$

Faktori η_1 on nyt selittävä (ennustava) faktori, faktori η_2 on sekä selittävä että selitettävä faktori ja faktori η_3 on selitettävä faktori. Siten faktori η_1 selittää faktoria η_2 ja faktori η_2 selittää faktoria η_3 . Esimerkin rakenneyhtälöt ovat tällöin rekursiivisia yhtälöitä eli faktoreiden välillä ei esiinny vuorovaikutteisia relaatioita.



Kuva 3: Hypoteettinen kolmen mittauskerran konfirmatorinen pitkittäisfaktorimalli.

2.5.1 Mallin rakentaminen

Pitkittäisfaktoreiden rakenneyhtälömallien (2.7)

$$\begin{cases} \mathbf{y} = \mathbf{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\zeta}. \end{cases}$$

rakentaminen tapahtuu samojen vaiheiden mukaisesti kuin konfirmatorisen pitkittäisfaktorimallin (2.1) rakentaminenkin: ensin spesifoidaan malli, jonka jälkeen suoritetaan identifioituvuustarkastelut, estimoidaan mallin tuntemattomat parametrit, testataan mallia koskevia hypoteeseja sekä lopuksi suoritetaan riittävyystarkastelut.

Rakenneyhtälömallin (2.7) rakentaminen kannattaa aloittaa rakentamalla ensin konfirmatorinen pitkittäisfaktorimalli, jolloin saadaan selkeä kuva faktorimalliesityksen riittävydestä. Jos pitkittäisfaktorimalli osoittautuu riittäväksi, voidaan rakenneyhtälömallin rakentamista jatkaa spesifioimalla faktoreiden rakenneyhtälöt. Rakenneparametrimatriisi \mathbf{B} oletetaan rekursiiviseksi ja Ψ -matriisi diagonaaliseksi. Rakenneyhtälöiden spesifiointivaiheessa matriisien parametreille määritellään sopiva käsittelytapa tutkimushypoteesin pohjalta. Yksittäinen parametri voidaan estimoida joko vapaasti, yhtäsuurena jonkun toisen parametrin kanssa tai parametri voidaan kiinnittää vakioksi.

Spesifioinnin jälkeen faktoreiden rakenneyhtälömallille suoritetaan identifioituvuustarkastelut. Identifioituvuutta tarkastellaan ainoastaan rakenneparametrimatriisien \mathbf{B} ja Ψ osalta. Faktorimalliosan (2.1) parametrimatriisit $\mathbf{\Lambda}$, $\mathbf{\Omega}$ ja $\mathbf{\Theta}$ oletetaan identifioituviksi. Silloin rakenneparametrien identifioituvuustarkastelut voidaan suorittaa faktoreiden kovarianssimatriisiin $\mathbf{\Omega}$ sekä parametrimatriisien \mathbf{B} ja Ψ väliseen yhteyteen

$$\mathbf{\Omega} = (\mathbf{I} - \mathbf{B})^{-1} \Psi (\mathbf{I} - \mathbf{B}')^{-1} \quad (2.8)$$

perustuen. Jos estimoitavat parametrit \mathbf{B} - ja Ψ -matriiseissa ovat ratkaistavissa kovarianssimatriisiin $\mathbf{\Omega}$ määrittämien yhtälöiden avulla, ovat rakenneyhtälöt identifioituvia. Estimoitavia parametreja on silloin maksimissaan $(\frac{1}{2})(kT)(kT + 1)$, joka on $\mathbf{\Omega}$:n määrittämien yhtälöiden lukumäärä.

Rakenneparametrien identifioituvuutta tutkittaessa yhtälöiden (2.8) käyttö on kuitenkin yleisessä tapauksessa hankalaa. Näiden yhtälöiden sijaan voidaan käyttää myös yhtälöitä

$$(\mathbf{I} - \mathbf{B})\mathbf{\Omega}(\mathbf{I} - \mathbf{B}') = \Psi,$$

jolloin matriisia $\mathbf{I} - \mathbf{B}$ ei tarvitse kääntää. Rekursiiviset rakenneyhtälöt eli rakenneyhtälöt, joissa matriisi $\mathbf{I} - \mathbf{B}$ on alakolmiomatriisi ja Ψ -kovarianssimatriisi on diagonaalimatriisi, ovat aina identifioituvia.

Identifioituvuustarkastelujen jälkeen estimoidaan faktoreiden rakenneyhtälömallin (2.7) tuntemattomat parametrit matriiseissa $\mathbf{\Lambda}$, $\mathbf{\Omega}$ ja $\mathbf{\Theta}$ sekä \mathbf{B} ja Ψ samanaikaisesti. Parametrien estimoinnin jälkeen suoritetaan rakenneyhtälömallia koskevien hypoteesien testaus. Jos stationaarisuus on voimassa, voidaan rakenneyhtälömallin

erikoistapauksena testata β -rakenneparametrien aikainvarianttisuutta. Tällöin nollahypoteesi on muotoa

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_T.$$

Rakenneparametrien aikainvarianttisuustestaus voidaan suorittaa esimerkiksi χ^2 -peräkkäistestillä. Hypoteesien testauksen jälkeen suoritetaan vielä rakenneyhtälömallin riittävyystarkastelut kohdassa 2.2.1.5 esitettyjen yleisten periaatteiden mukaisesti.

Esimerkki 2.5. Tarkastellaan seuraavaksi esimerkin 2.4 rakenneyhtälömallin identifioituvuutta. Siinä faktorimalliosan λ -, ω - ja θ -parametrit ovat identifioituvia. Lisäksi oletetaan, että esimerkin 2.4 rekursiivisten rakenneyhtälöiden jäännökset ovat korreloimattomia. Rakenneyhtälöt tuottavat silloin seuraavan kovarianssimatriisin η -faktoreille:

$$\left\{ \begin{array}{l} \omega_{11} = \psi_{11} \\ \omega_{21} = \beta_{21}\psi_{11} \\ \omega_{22} = \beta_{21}^2\psi_{11} + \psi_{22} \\ \omega_{31} = \beta_{21}\beta_{32}\psi_{11} \\ \omega_{32} = \beta_{21}^2\beta_{32}\psi_{11} + \beta_{32}\psi_{22} \\ \omega_{33} = \beta_{21}^2\beta_{32}^2\psi_{11} + \beta_{32}^2\psi_{22} + \psi_{33} \end{array} \right.$$

Kovarianssimatriisin alkioiden ω_{11} , ω_{21} ja ω_{22} määrittämistä yhtälöistä havaitaan helposti, että parametrit ψ_{11} , β_{21} ja ψ_{22} ovat identifioituvia. Parametri β_{32} saadaan ratkaistua pienen laskutoimituksen avulla yhtälöstä ω_{31} , jonka jälkeen yhtälöstä ω_{33} saadaan ratkaistua ψ_{33} . Näin ollen kaikki esimerkin 2.4 rakenneyhtälöiden β - ja ψ -parametrit ovat identifioituvia ja siten myös koko esimerkkirakenneyhtälömalli on identifioituva.

3 Aineisto ja asetelma

Tässä luvussa esitellään tutkielmassa käytettävä tutkimusaineisto. Aluksi kappaleessa 3.1 esitellään lyhyesti Vanhemmat, opettajat ja lapsen oppiminen (VALO) -tutkimus, jonka osa-aineisto tämän tutkielman tutkimusaineisto on. Seuraavaksi kappaleessa 3.2 esitellään tutkielman tutkimusaineisto, ja luvun viimeisessä kappaleessa eli kappaleessa 3.3 perehdytään tutkimusaineiston keruumenetelmään sekä aineiston muuttujiin.

3.1 Vanhemmat, opettajat ja lapsen oppiminen-tutkimus

Tämän tutkielman tutkimusaineistona on Emootio -aineisto, joka on kerätty syksyllä 2006 osana VALO -tutkimusta. VALO -tutkimuksessa tarkoituksena on tutkia lasten oppimistilanteisiin liittämiä emootioita ja oppimismotivaatiota sekä niiden kehitystä perhe- ja kouluympäristöissä ensimmäisen kouluvuoden aikana.

VALO -tutkimus toteutettiin kaksivaiheisesti. Ensimmäisessä vaiheessa lukuvuonna 2005–2006 toteutettiin pienimuotoinen pilottitutkimus, jonka tarkoituksena oli testata eri mittarien, erityisesti emootiomittarien, ja asetelman toimivuutta sekä käytännön että analyysimenetelmien kannalta, Aunola & Nurmi (2005). Toisessa vaiheessa lukuvuosina 2006–2009 toteutetaan varsinainen VALO -pitkittäistutkimus. Tutkimus tullaan toteuttamaan samanlaisena kolmelle eri ikäkohortille kolmessa eri kaupungissa: Porissa, Mikkelissä ja Lappeenrannassa. Jokaiselle kohortille tehdään kaksi mitausta, yksi syyslukukaudella ja yksi kevätlukukaudella. Vuosittain kohdekaupunkien kultakin tutkimukseen osallistuvalla ensimmäiseltä luokalta valitaan satunnaisesti ja vanhempien luvalla yksi lapsi mukaan tutkimukseen. Mukana on sekä yleis- että erityisopetuksessa olevia lapsia. Tutkimukseen osallistuvat myös lasten opettajat sekä vanhemmat tai huoltajat. VALO -tutkimus on osa Suomen Akatemian Oppimisen ja motivaation huippututkimusyksikön tutkimusohjelmaa, Aunola & Nurmi (2006).

3.2 Tutkittavat

VALO -tutkimuksen ensimmäisessä aineistonkeruussa syksyllä 2006 tutkimukseen osallistui yhteensä 68 lasta (perhettä), 68 opettajaa sekä 121 vanhempaa, 53 isää ja 68 äitiä. Perheitä, joista tutkimukseen osallistui vain äiti, oli yhteensä 15 eli 22.06% kaikista perheistä.

Tutkimukseen osallistuneista henkilöistä koehenkilöiksi rajattiin VALO -tutkimukseen osallistuneet vanhemmat tai huoltajat. Vanhemmat, joilla oli puuttuvaa tietoa tähän tutkimukseen valittujen muuttujien osalta, jätettiin pois lopullisesta aineistosta. Tällöin koehenkilöiden kokonaismäärä oli 101 eli 40 isää ja 61 äitiä. Perheitä lopullisessa aineistossa oli yhteensä 63 eli 92.6% kaikista VALO -tutkimukseen osallistuneista perheistä. Lopullisen aineiston perheistä yhteensä 25 eli 36.8% kaikista VALO -tutkimuksen perheistä oli sellaisia, joista vastasi vain toinen vanhempi. Koska suurin osa vanhemmista edusti samoja perheitä, olivat näiden vanhempien vastaukset toisistaan riippuvia. Vastausten riippuvuus päätettiin kuitenkin jättää huomioimatta, jotta otoskoko saataisiin mahdollisimman suureksi.

3.3 Menetelmä ja muuttajat

VALO -tutkimukseen osallistuneet vanhemmat täyttivät viikon ajan joka päivä strukturoitua päiväkirjaa. Päiväkirja koostui erilaisista kyselylomakkeista, jotka käsitelivät päivän aikana tapahtuneita oppimistilanteita, oppimistilanteissa tapahtunutta vanhemman ja lapsen välistä vuorovaikutusta sekä lapsen toimintaa näissä tilanteissa. Vanhemmat täyttivät päiväkirjaa aina iltaisin päivän päätteeksi, Aunola & Nurmi (2006).

Päiväkirjasta valittiin mallinnettavaksi kysymyssarja 8.2. (*"Lapseni vaikutti tänään . . . "*, ks. Liite A), jossa vanhemmat arvioivat päivittäin lapsensa mielialoja eli emootioita. Kyseinen emootio -kysymyssarja on muokattu positiivisia ja negatiivisia mielialoja mittaavasta PANAS -mittarista (Positive and Negative Affect Schedule; Watson, Clark ja Tellegen (1988)). Ensin VALO -tutkimuksen pilottivaiheessa emootio-kysymyssarjaan valittiin mukaan alkuperäisestä PANAS -mittarista kuusi positiivista mielialaa (iloinen, ylpeä, onnellinen, tyytyväinen, innostunut, kiitollinen) sekä 14 negatiivista mielialaa (vihainen, hermostunut, syyllisen oloinen, ärtynyt, jän-

nittynyt, toivoton, turhautunut, levoton, surullinen, onneton, alakuloinen, ahdistunut, avuton, nolostunut/häpeilevä). Pilottivaiheessa mittarin toimivuutta tutkittiin eksploratiivisen faktorianalyysin avulla. Analyyseissa löytyi osioita, jotka eivät toimineet ollenkaan sekä osioita, jotka mittasivat täysin samaa mielialaa. Pilottivaiheen analyysien perusteella rakennettiin lopullinen kysymyssarja, josta poistettiin ne osiot, jotka eivät toimineet. Osiot, jotka mittasivat samaa mielialaa, yhdistettiin yhdeksi osioksi.

Lopulliseen emootio -kysymyssarjaan jäi lopulta päivätasolta mukaan 12 osiota eli 12 erilaista positiivista tai negatiivista mielialaa (vihainen/ärtynyt, hermostunut/levoton, syyllisen oloinen, iloinen/onnellinen/innostunut, turhautunut, ylpeä, surullinen/alakuloinen, ahdistunut, kiitollinen, nolostunut/häpeilevä). Lisäksi kysymyssarjaan lisättiin yksi negatiivinen mieliala, väsynyt. Emootio -kysymyssarjassa vanhempien tehtävänä oli arvioida lastensa mielialoja 5 -portaisella asteikolla *1 (= ei ollenkaan) – 5 (= erittäin paljon)*. Korkeat arvot positiivisissa mielialoissa ilmaisevat lapsen olleen innostunut, vilkas ja reipas päivän aikana tapahtuneissa oppimistilanteissa sekä kokeneen mielihyvää tilanteissa. Matalat arvot positiivisissa mielialoissa puolestaan kertovat lapsen passiivisuudesta oppimistilanteissa. Korkeat arvot negatiivisissa mielialoissa ilmaisevat lapsen ilmentäneen ahdistusta sekä negatiivista suhtautumista oppimistilanteisiin. Matalat arvot negatiivisissa tunteissa kertovat lapsen olleen rauhallinen oppimistilanteissa, Watson, Clark ja Tellegen (1988).

4 Pitkittäisfaktorimallin rakentaminen sovellusaineistolle

Tässä luvussa esitellään sovellusaineistolle suoritettujen eksploratiivisten ja konfirmatoristen faktorianalyysien tuloksia. Koska analysoitavana on havaittuja y -muuttujia kolmelta eri mittauspäivältä, on mallinrakennuksen lähtökohtana kolmen faktorin mittausmalli, jossa kukin havaittu muuttuja latautuu yhdelle faktorille. Kaikki havaitut muuttujat ovat järjestysasteikollisia muuttujia, joten konfirmatoristen faktorianalyysien perustana on havaittujen muuttujien polykorinen korrelaatiomatriisi.

Pitkittäisfaktorimallin rakentaminen aloitetaan tutkimalla mielialamuuttujien faktorirakenteita eksploratiivisen faktorianalyysin avulla. Sovellusaineiston eksploratiivista faktorianalyysia käsitellään kappaleessa 4.1. Kappaleessa 4.2 tarkastellaan sovellusaineiston konfirmatorista faktorianalyysia. Ensin aineistoon sovitetaan konfirmatorinen pitkittäisfaktorimalli, jonka jälkeen tutkitaan pitkittäisfaktorimallin latausrakenteen stationaarisuutta. Kolmantena konfirmatorisena mallina aineistoon sovitetaan faktoreiden rakenneyhtälömallia ja lopuksi sovitetaan rakenneyhtälömallia, jonka rakenneparametrit ovat aikainvariantteja.

4.1 Eksploratiivinen faktorianalyysi

Päiväkirjan kysymyssarjan 8.2 alustavat tarkastelut päivätasolla suoritettiin tutkimalla mielialamuuttujien faktorirakenteita eksploratiivisen faktorianalyysin avulla SPSS 14.0 -ohjelmalla. Analyysien lähtökohtana oli 12 mielialamuuttujaa, joiden korrelaatiorakennetta haluttiin kuvata ja selittää faktorimallin avulla. Aineiston pienuuden takia ($N = 101$) tutkimusasetelmaan valittiin mukaan vain kolme mittauskertaa: maanantai, keskiviikko ja perjantai. Faktorointimenetelmäksi valittiin pääakselifaktorointi (engl. Principal Axis Factoring) ja faktoreiden rotaatio suoritettiin faktoreiden keskinäisen korrelaation sallivalla vinokulmaisella Promax -rotaatiolla, koska positiivisten ja negatiivisten mielialojen oletettiin korreloivan keskenään. Sisältöteorian perusteella päädyttiin kahden faktorin malliin kullakin mittauskerralla. Tällöin ensimmäisen faktorin oletettiin kuvaavan negatiivista mielialaa ja toisen faktorin positiivista mielialaa.

Faktoroinnin seurauksena jokaiselle mielialamuuttujalle y_i saatiin kommunaliteetti h_i^2 , joka kertoo, kuinka hyvin kyseinen muuttuja mittaa faktoreita. Mielialamuuttujien kommunaliteetit saatiin seuraavan kaavan avulla:

$$h_i^2 = \sum_{j=1}^2 \lambda_{ij}^2, \quad i = 1, \dots, 12,$$

jossa λ_{ij} on muuttujan y_i lataus j :nnellä faktorilla η_j . Kommunaliteettien arvot ovat tilastollisesti kelvollisia, kun ne ovat välillä $[0, 1]$. Mitä lähempänä ykköstä mielialamuuttujan kommunaliteetin arvo on, sitä paremmin muuttuja mittaa faktorirakennetta. Kommunaliteettien avulla voidaan lisäksi laskea yksittäisen havaitun y_i -muuttujan jäännösvaihtelu. Jäännösvaihtelu saadaan laskettua kaavalla

$$\theta_i = 1 - h_i^2.$$

Mielialamuuttujien estimoidut kommunaliteetit sekä mielialamuuttujien ja faktoroiden väliset lataukset kultakin mittauspäivältä on esitetty taulukossa 1. Taulukosta nähdään, että kaikki muut kommunaliteetit ovat vähintään kohtalaisia ($h_i^2 \geq .230$), paitsi Väsynyt -muuttujan perjantain mittauksen kommunaliteetti (.182), joka on melko pieni. Havaitut muuttujat näyttäisivät jokaisena mittauspäivänä mittaavan faktorirakennetta vähintäänkin kohtalaisesti. Maanantaina voimakkaimmin faktorirakennetta mittasi muuttuja Hermostunut/levoton (.645) ja heikoimmin Kiitollinen (.218). Keskiviikkona faktorirakennetta mittasi voimakkaimmin muuttuja Kiitollinen (.601) ja heikoimmin muuttuja Ylpeä (.309). Perjantaina voimakkaimmin faktorirakennetta mittasi muuttuja Surullinen/alakuloinen (.599) ja heikoimmin Väsynyt (.182).

Kaikkina kolmena mittauspäivänä negatiiviset mielialat (vihainen/ärtynyt, hermostunut/levoton, syyllisen oloinen, turhautunut, surullinen/alakuloinen, ahdistunut, nolostunut/häpeilevä, väsynyt) latautuivat positiivisesti ensimmäiseen faktoriin, ja kaikki positiiviset mielialat (iloinen/onnellinen/innostunut, ylpeä, kiitollinen) latautuivat positiivisesti toiseen faktoriin. Latausten perusteella faktorit voidaan siis nimetä Negatiivisen mielialan (lyhyesti NEG) ja Positiivisen mielialan (lyhyesti POS) faktoreiksi. Taulukossa 1 on lisäksi esitetty kunkin mittauspäivän NEG- ja POS-faktoreiden väliset korrelaatiot. Jokaisena mittauspäivänä faktoreiden välillä on kohtalaista positiivista korrelaatiota (.302–.426), joten faktoreiden korrelaation sallivan

vinokulmaisen Promax -rotaatiomenetelmän käyttö oli aiheellista.

Taulukko 1: Mielialamuuttujien lataukset, kommunaliteetit sekä faktoreiden väliset korrelaatiot mittauspäivittäin.

y	$\hat{\Lambda}, h^2$									
	Ma			Ke			Pe			
	NEG	POS	h_i^2	NEG	POS	h_i^2	NEG	POS	h_i^2	
Vihainen/ärtynyt	.545	-.401	.645	.565	-.262	.500	.487	-.171	.317	
Hermostunut/levoton	.811	-.080	.720	.695	-.065	.521	.579	.086	.312	
Syällisen oloinen	.659	.324	.358	.646	.207	.359	.650	.090	.396	
Iloinen/onnellinen/innostunut	-.420	.421	.504	-.303	.435	.380	-.328	.475	.427	
Turhautunut	.644	.077	.379	.605	-.069	.403	.499	.092	.230	
Ylpeä	.199	.599	.297	.116	.590	.309	.184	.715	.466	
Surullinen/alakuloinen	.537	-.114	.353	.728	-.009	.535	.711	-.159	.599	
Ahdistunut	.633	-.006	.405	.687	.000	.472	.750	.015	.556	
Kiitollinen	.000	.467	.218	.098	.807	.601	.069	.653	.404	
Nolostunut/häpeilevä	.578	.412	.301	.685	.292	.403	.584	.179	.310	
Väsynyt	.507	-.092	.306	.543	-.149	.378	.409	-.051	.182	

Faktoreiden väliset korrelaatiot mittauspäivittäin: $\omega_{ma} = -.426$, $\omega_{ke} = -.379$ ja $\omega_{pe} = -.302$

Koska kysymyssarjassa 8.2 oli mukana vain kolme positiivista mielialaa (Iloinen/onnellinen/innostunut), valittiin ne kaikki kuvaamaan Positiivisen mielialan faktoria. Eksploratiivisten faktorianalyysien perusteella Negatiivisen mielialan faktoria valittiin kultakin mittauspäivältä kuvaamaan kolme muuttujaa. Muuttujat valittiin täysin subjektiivisesti, sillä kaikki negatiivista mielialaa ilmaisevat muuttujat olivat kommunaliteettien ja latausten mielessä yhtä hyviä: kaikkien negatiivisten mielialojen muuttujien kommunaliteetit ja lataukset olivat yhtenä mittauspäivänä korkeat ja toisena mittauspäivänä melko matalat. Millään muuttujista ei ollut kaikkina kolmena mittauspäivänä korkeita kommunaliteetteja ja latauksia.

Valitut muuttujat ovat seuraavat:

VIHA = vihainen/ärtynyt

HERM = hermostunut/levoton

TURH = turhautunut

ILOI = iloinen/onnellinen/innostunut

YLPE = ylpeä

KIIT = kiitollinen

Havaittuja muuttujia on siten yhteensä 18, kuusi kultakin mittauspäivältä. Jokainen muuttuja vastaa yhtä mielialaa. Eri päivien muuttujat tullaan jatkossa erottelemaan toisistaan muuttujan nimen loppuun lisättävällä mittauspäivää kuvaavalla lyhenteellä ma, ke tai pe. Esimerkiksi VIHAma tarkoittaa maanantaina arvioitua vihainen/ärtynyt-osiota, VIHAke keskiviikkona arvioitua vihainen/ärtynyt -osiota ja VIHApe perjantaina arvioitua vihainen/ärtynyt -osiota.

Seuraavaksi havaittujen muuttujien faktorirakenteita tutkittiin uudelleen eksploratiivisen faktorianalyysin avulla. Faktorointimenetelmäksi valittiin jälleen pääakselifaktorointi ja rotaatiomenetelmäksi Promax. Eksploratiivinen faktorianalyysi suoritettiin ensin kullekin mittauspäivälle erikseen valitsemalla faktoreiden lukumääräksi kaksi.

Analyysin tulokset on esitetty taulukossa 2. Taulukosta havaitaan, että kaikkina kolmena mittauspäivinä negatiiviset mielialat latautuvat NEG -faktoriin ja positiiviset mielialat POS -faktoriin. Negatiivisten mielialojen lataukset ovat kaikkina mittauspäivinä korkeammat kuin positiivisten mielialojen lataukset. Lisäksi kaikki lataukset näyttäisivät olevan kaikkina mittauspäivinä suunnilleen samaa suuruusluokkaa. Valittu kahden faktorin ratkaisun latausrakenne näyttäisi siis säilyvän samanalaisena mittauspäivästä toiseen. Faktoreiden väliset korrelaatiot yksittäisinä mittauspäivinä ovat negatiivisia ja melko heikkoja (-.214– -.347).

Muuttujien kommunaliteetit ovat vähintään kohtalaisia ($h_i^2 \geq .278$) kaikkina mittauspäivinä. Maanantaina heikoin kommunaliteetti on KIIT -muuttujalla (.308) eli muuttuja KIIT kykenee selittämään vähiten (30.8%) maanantain kahden faktorin ratkaisun vaihtelusta. Muuttujalla HERM on maanantaina korkein kommunaliteetti (.931), joten kyseinen muuttuja kykenee selittämään eniten kahden faktorin ratkaisun vaihtelusta. Maanantaina muuttujien kommunaliteetit ovat yleisesti hieman

matalammat kuin muina päivinä. Keskiviikkona mielialamuuttujien kommunaliteetit ovat kohtalaisen korkeita (.319–.644), samoin perjantaina (.278–.525).

Taulukko 2: Havaittujen mielialamuuttujien lataukset, kommunaliteetit sekä fakto-
reiden väliset korrelaatiot mittauspäivittäin.

y	$\hat{\Lambda}, h^2$								
	Ma			Ke			Pe		
	NEG	POS	h_i^2	NEG	POS	h_i^2	NEG	POS	h_i^2
VIHA	.680	-.196	.592	.724	-.092	.579	.621	-.172	.462
HERM	1.013	.185	.931	.764	.066	.553	.741	.135	.525
TURH	.709	.026	.491	.782	.125	.559	.707	.046	.488
ILOI	-.339	.459	.431	.406	.277	.319	-.293	.380	.278
YLPE	.092	.612	.345	.119	.702	.448	.119	.710	.482
KIIT	.035	.566	.308	-.071	.775	.644	-.010	.645	.419

Faktoreiden väliset korrelaatiot mittauspäivittäin: $\omega_{ma} = -.342$, $\omega_{ke} = -.347$ ja $\omega_{pe} = -.214$

Havaittujen muuttujien faktorirakenteiden tarkastelua jatkettiin suorittamalla seuraavaksi eksploraatiivinen faktorianalyysi siten, että yksittäisten mittauspäivien faktorimallit yhdistettiin kuuden faktorin eksploraatiiviseksi faktorimalliksi. Tarkoituksena oli tutkia, latautuvatko kunkin mittauspäivän negatiiviset mielialat omille faktoreilleen ja positiiviset mielialat omille faktoreilleen. Faktorointimenetelmäksi valittiin jälleen pääakselifaktorointi ja rotaatiomenetelmäksi Promax.

Nyt eksploraatiivisen faktorianalyysin suorittamisessa ilmeni ongelmia, sillä SPSS-ohjelma ei kyennyt suorittamaan faktorointia loppuun asti. Eräs syy eksploraatiivisten faktorianalyysien epäonnistumiseen saattoi olla tutkimusaineiston pieni koko ($N = 101$). Pienellä aineistolla korrelaatiomatriisin estimointi ei ole kovinkaan tarkkaa. Muita syitä analyysin epäonnistumiseen yritettiin löytää analysoimalla valittua mallia pienemmissä osissa. Ensin malli jaettiin kahtia negatiivisiin mielialoihin ja positiivisiin mielialoihin. Sen jälkeen kummankin mielialamallin faktorirakenteita tutkittiin eksploraatiivisen faktorianalyysin avulla. Negatiivisten mielialojen analyysi onnistui hyvin, mutta positiivisten mielialojen analyysissä ilmeni jälleen ongelmia faktorointi-

vaiheessa.

Epäonnistumisen syiden paikantamista jatkettiin tarkastelemalla vain positiivisia mielialoja kuvaavia muuttujia. Seuraavassa vaiheessa tutkittavaksi valittiin mittauskertapareja (ma-ke, ma-pe ja ke-pe). Kullekin mittauskertaparille suoritettiin jälleen eksploratiivinen faktorianalyysi parittaisten mittauskertojen faktorirakenteiden tutkimiseksi. Analyysi epäonnistui nyt mittauskertapareilla, joissa oli mukana keskiviikko eli pareilla ma-ke ja ke-pe. Viimeisessä vaiheessa tutkittiin ainoastaan keskiviikon mittauksen positiivisia mielialamuuttujia. Muuttujien tarkemmissa tutkimuksissa ei kuitenkaan löytynyt mitään selkeää syytä faktorointien epäonnistumiselle.

Osa tutkimusaineiston koehenkilöistä edusti samoja perheitä, joten heidän mielialoarvionsa olivat toisistaan riippuvia. Tämä riippuvuus saattoi myös olla syynä faktoroinnin yhteydessä ilmenneisiin ongelmiin. Riippuvuusrakennetta yritettiin poistaa valitsemalla aineistosta koehenkilöiksi vain äidit. Aineistossa oli saman verran äitejä kuin koko Vanhemmat, opettajat ja lapsen oppiminen -tutkimuksessa oli mukana lapsia. Myös tälle aineistolle suoritettussa eksploratiivisissa faktorianalyysissä ilmeni kuitenkin ongelmia faktorointivaiheessa, joten myöskään vanhempien mieliala-arvioiden riippuvuus ei ollut syynä faktoroinnin epäonnistumiseen.

Koska eksploratiivisten faktorianalyysien suorittamisessa ilmeni ongelmia, kun analyysissä oli mukana kolme mittauskertaa ja kultakin mittauskerralta kolme positiivista ja kolme negatiivista mielialaa eikä selkeää syytä tähän kyetty löytämään, päätettiin valittua mallia muuttaa. Alunperin mukaan analyysiin haluttiin riittävästi mittauskertoja, jotta saataisiin mahdollisimman hyvä kuva lasten mielialojen pysyvyydestä koko tutkimusviikon ajalta. Aineiston pienuuden takia mittauskertojen määrää jouduttiin kuitenkin rajoittamaan kolmeen. Samat mittauskerrat haluttiin sisällyttää myös uuteen malliin. Lisäksi eksploratiivisten faktorianalyysien perusteella havaittiin, että ainoastaan negatiivisten mielialojen malli toimi, kun mukana oli muuttujia kaikilta kolmelta mittauskerralta. Siten ainoa mahdollinen valinta uudeksi mielialamalliksi oli negatiivisten mielialojen malli, jossa on mukana kolme mittauspäivää ja kolme havaittua muuttujaa kultakin mittauspäivältä.

4.2 Konfirmatorinen faktorianalyysi

Koska kaikki havaitut muuttajat ovat järjestyasteikkollisia muuttujia, on konfirmatorisen faktorianalyysin perustana havaittujen muuttujien polykorinen korrelaatiomatriisi. Muuttujien väliset polykoriset korrelaatiot on esitetty taulukossa 3. Mittauskertojen sisällä mielialamuuttujien välillä on voimakasta positiivista korrelaatiota (.566–.806) ja mittauskertojen välillä korrelaatio ovat vähintään kohtalaista (.213–.615). Vain VIHake- ja HERMma -muuttujien välinen korrelaatio on erittäin heikkoa (.001), samoin muuttujien VIHApe ja VIHAma välinen korrelaatio (.145) sekä muuttujien VIHApe ja HERMma välinen korrelaatio (.084). Havaittujen muuttujien polykoristen korrelaatioiden perusteella negatiivisilla mielialoilla näyttäisi olevan voimakasta riippuvuutta sekä mittauskertojen sisällä että mittauskertojen välillä.

Taulukko 3: Havaittujen muuttujien polykorinen korrelaatiomatriisi.

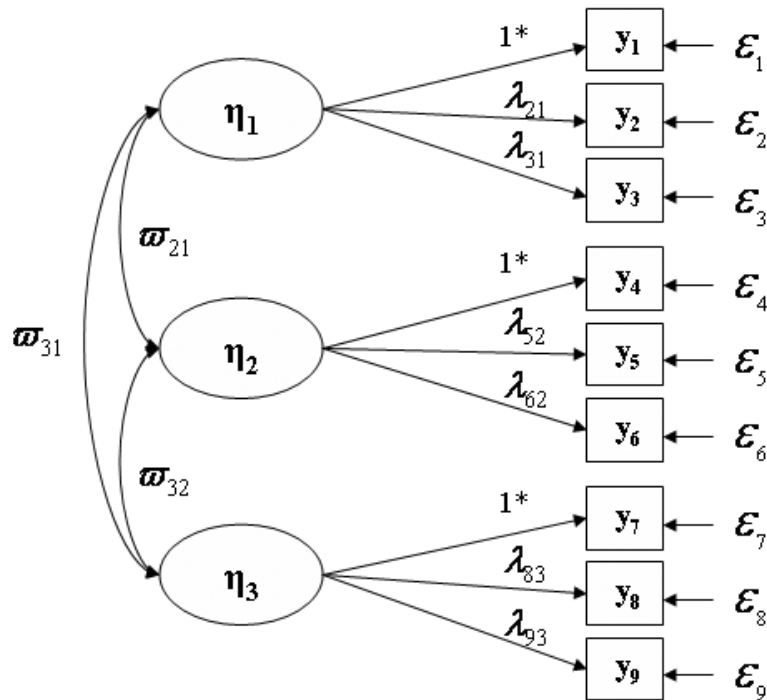
	1	2	3	4	5	6	7	8	9
1 VIHAma	1.00								
2 HERMma	.806	1.00							
3 TURHma	.638	.806	1.00						
4 VIHake	.213	.001	.144	1.00					
5 HERMke	.421	.434	.382	.653	1.00				
6 TURHke	.224	.408	.422	.677	.710	1.00			
7 VIHApe	.145	.084	.168	.236	.318	.325	1.00		
8 HERMpe	.358	.270	.249	.360	.574	.454	.610	1.00	
9 TURHpe	.392	.512	.439	.251	.615	.469	.566	.664	1.00

4.2.1 Kolmen mittauskerran pitkittäisfaktorimalli

Konfirmatorisen pitkittäisfaktorimallin rakentaminen aloitettiin yksinkertaisesta mallista, jota modifioitiin saatujen tulosten perusteella. Lähtökohtana oli kolme mittauspäivää käsittävä mittausmalli (kuva 4), jossa kunkin mittauspäivän kolme mielialamuuttujaa muodostavat oman faktorinsa. Faktoreiden skaalojen yksikäsitteisyyden vuoksi jokaiselta faktorilta on kiinnitetty ensimmäinen lataus ykköseksi. Faktoreiden

korrelointi sallittiin, koska pitkittäisasetelmassa eri mittauskertojen muuttujat korreloivat keskenään. Kaikki jäännöskovarianssit oletettiin nolliksi. Malli vastaa luvussa 2 esimerkissä 2.3 spesifioitua hypoteettista mallia.

Mallin estimoiduista parametreista vain osa oli tilastollisesti merkitseviä. Lisäksi osa estimoiduista jäännösvariansseista oli negatiivisia. Mallia ei voitu pitää tilastollisesti kelvollisena, joten sitä oli modifioitava. Mallin modifikaatioindeksien arvoista useat ylittivät yleisesti kriteerinä käytetyn rajan $MI > 8$. Muuttujien VIHake ja HERMma ($MI = 22.10$) sekä muuttujien TURHke ja VIHama ($MI = 21.31$) jäännösten välisten kovarianssien modifikaatioindeksit olivat suurimmat, joten kyseiset jäännöskovarianssit θ_{42} ja θ_{61} vapautettiin estimoitavaksi. Myös yksi kiinnitetyistä latausparametreista, muuttuja TURHma faktorille NEGke, tulisi modifikaatioindeksin perusteella vapauttaa ($MI = 10.60$), mutta tästä luovuttiin, koska kyseiselle vapautukselle ei ollut sisällöllistä tulkintaa.



Kuva 4: Hypoteettinen kolmen mittauskerran konfirmatorinen pitkittäisfaktorimalli.

Saatua mallia ei voitu pitää riittävänä kuvaamaan aineistoa, sillä χ^2 -testisuureen arvoksi saatiin $\chi^2(22) = 443.89$ ($p = 0.000$) ja esimerkiksi $RMSEA = 0.438$. Myöskään muuttujien HERMma ja TURHpe jäännösvarianssit θ_{22} ja θ_{99} eivät olleet tilas-

tollisesti merkitseviä (t -arvot vastaavassa järjestyksessä 0.75 ja 1.43). Lisäksi mallin modifikaatioindekseissä oli edelleen jäännöskovariansseja, joiden modifikaatioindeksit ylittivät rajan $MI > 8$.

Koska edellä saatu mallia ei ollut riittävä, jatkettiin mallin modifointia jäännöskovariansseja vapauttamalla. Modifiointeja suoritettiin kaikkiaan 10 kertaa. Lopulta löydettiin malli, joka sopi aineistoon. Tähän malliin vapautettiin estimoitavaksi muuttujien TURHma, TURHke ja TURHpe jäännösten väliset autokovarianssit θ_{63} , θ_{93} ja θ_{96} sekä muuttujien HERMke ja HERMpe jäännösten välinen autokovarianssi θ_{85} . Jotta malli saatiin sopimaan aineistoon, jouduttiin lisäksi vapauttamaan jäännöskovarianssit θ_{42} , θ_{61} , θ_{72} , θ_{81} , θ_{91} , θ_{92} ja θ_{95} .

Saatus mallia voidaan pitää riittävyysindeksien mielessä hyvänä, sillä koko mallia koskevat riittävyysindeksien arvot läpäisevät mallin riittävyuden rajat: $\chi^2(13) = 16.42$, ($p = 0.227$), $RMSEA = 0.051$, $CFI \approx 1.00$, $NFI = 0.94$ ja $SRMR = 0.096$. Mallin estimoidut latausparametrit, jäännökset keskivirheineen sekä havaittujen muuttujien reliabiliteetit ja faktoreiden korrelaatiot on esitetty taulukossa 4. Kaikki muut mallin parametrit ovat tilastollisesti merkitseviä, paitsi muuttujan HERMma jäännösvarianssi θ_{22} ($t = 0.52$). Muuttujien reliabiliteetit $\hat{\mathbf{R}}_y^2$ ovat hyviä. Suurimman reliabiliteetin sai muuttuja HERMma, jonka reliabiliteetti on 0.97. Vähiten mallia selittää muuttuja VIHape, jonka reliabiliteetti on 0.51. Tämäkin reliabiliteettikerroin on kuitenkin niin korkea, ettei muuttujaa kannattanut poistaa mallista.

Faktoreiden väliset korrelaatiot ovat melko voimakkaita. Peräkkäisten faktoreiden NEGma ja NEGke välinen korrelaatio on 0.53 ja faktoreiden NEGke ja NEGpe välinen korrelaatio 0.57. Ensimmäisen ja kolmannen mittauksen faktoreiden NEGma ja NEGpe välinen korrelaatio on hieman heikompi (0.26) kuin peräkkäisten mittausten faktoreiden. Koska peräkkäisten mittausten faktoreiden väliset korrelaatiot ovat melko voimakkaita, voidaan sanoa, että lasten negatiivisten mielialojen mittaaminen on ollut kohtalaisen stabiilia peräkkäisten mittauspäivien välillä. Sen sijaan ensimmäisen ja kolmannen mittauspäivän välillä lasten negatiivisten tunteiden mittaaminen ei ole ollut kovinkaan stabiilia.

Taulukko 4: Mallin 1 parametrien estimaatit keskivirheineen, havaittujen muuttujien reliabiliteetit sekä faktoreiden väliset korrelaatiot. Tähdellä (*) merkityt arvot on kiinnitetty.

y	$\hat{\Lambda}_{(s.e.)}$			$\hat{\theta}_{(s.e.)}$	\hat{R}_y^2
	NEGma	NEGke	NEGpe		
VIHama	1.00*	0*	0*	.35(.05)	0.62
HERMma	1.16(.11)	0*	0*	.02(.04)	0.97
TURHma	0.94(.10)	0*	0*	.36(.06)	0.58
VIHAke	0*	1.00*	0*	.33(.06)	0.64
HERMke	0*	1.08(.12)	0*	.29(.05)	0.70
TURHke	0*	1.14(.10)	0*	.24(.05)	0.76
VIHApe	0*	0*	1.00*	.48(.08)	0.51
HERMpe	0*	0*	1.21(.17)	.25(.08)	0.74
TURHpe	0*	0*	1.06(.15)	.39(.08)	0.59

Estimoitujen faktoreiden korrelaatiot: $\omega_{21} = 0.53$, $\omega_{31} = 0.26$ ja $\omega_{32} = 0.57$

Estimoidut jäännöskovarianssit (autokovarianssit lihavoitu):

$\hat{\theta}_{42} = -0.22(.04)$, $\hat{\theta}_{61} = -0.15(.04)$, $\hat{\theta}_{63} = \mathbf{0.08(.04)}$, $\hat{\theta}_{72} = -0.11(.03)$,
 $\hat{\theta}_{81} = 0.13(.03)$, $\hat{\theta}_{85} = \mathbf{0.14(.04)}$, $\hat{\theta}_{91} = 0.15(.06)$, $\hat{\theta}_{92} = 0.22(.07)$,
 $\hat{\theta}_{93} = \mathbf{0.17(.07)}$, $\hat{\theta}_{95} = 0.26(.05)$ ja $\hat{\theta}_{96} = \mathbf{0.11(.05)}$

4.2.2 Kolmen mittauskerran stationaarinen pitkittäisfaktorimalli

Seuraavaksi tutkittiin negatiivisten mielialojen pitkittäisfaktorimallin latausrakenteen stationaarisuutta. Jos pitkittäisfaktorimallin latausrakenne ei ole stationaarinen, aiheutuu mittauskertojen välisissä vertailuissa ongelmia. Latausten stationaarisuuden tutkimiseksi eri mittauskertojen latausparametrit estimoitiin yhtä suurina seuraavasti:

$$\lambda_{21} = \lambda_{52} = \lambda_{83}, \lambda_{31} = \lambda_{62} = \lambda_{93}.$$

Kunkin faktorin ensimmäinen lataus on jälleen kiinnitetty ykköseksi. Jotta stationaa-

risuusmalli olisi vertailukelpoinen edellä käsitellyn pitkittäisfaktorimallin kanssa, vapautettiin stationaarisuusmalliin estimoitavaksi samat jäännösten väliset kovarianssit kuin pitkittäisfaktorimallissakin.

Stationaarisuusmallin voimassaoloa testattiin χ^2 -peräkkäistestin avulla, sillä stationaarisuusmalli on sisäkkäinen edellä käsitellyn pitkittäisfaktorimallin kanssa. Stationaarisuusmallissa oli käytössä enemmän vapausasteita kuin pitkittäisfaktorimallissa, koska mittauskertojen latausparametrit estimoitiin yhtäsuurina. Pitkittäisfaktorimallille χ^2 -testisuureen arvoksi saatiin $\chi^2(13) = 16.42$. Stationaarisuusmallille vastaava arvo oli $\chi^2(17) = 20.35$. Tällöin χ^2 -peräkkäistestin arvoksi saadaan

$$D = 20.35 - 16.42 = 3.93$$

ja vapausasteiden määräksi

$$df = 17 - 13 = 4.$$

Valitsemalla peräkkäistestin merkitsevyytasoksi $\alpha = 0.10$ on χ^2 -jakauman vertailuarvo $\chi_{0.10}^2(4) = 7.779$. Koska vertailuarvo on suurempi kuin peräkkäistestin arvo, saadaan peräkkäistestin p -arvoksi $p > 0.10$. Siten negatiivisten mielialojen pitkittäisfaktorimallin latausrakenne on stationaarinen.

Stationaarisuusmallin latausparametrien estimaatit ja keskivirheet on esitetty taulukossa 5. Stationaarisuusmallin lataukset ovat samaa luokkaa kuin mallin 1 lataukset. Jäännösvarianssien ja jäännöskovarianssien estimaatit keskivirheineen sekä muuttujien reliabiliteetit ovat hyvin lähellä mallin 1 vastaavia arvoja, samoin faktoreiden korrelaatiot. Erot ovat suurimmillaan sadasosia.

Taulukko 5: Mallin 2 latausparametrien estimaatit ja keskivirheet. Tähdellä (*) merkityt arvot on kiinnitetty.

y	$\hat{\Lambda}$ (s.e.)		
	NEGma	NEGke	NEGpe
VIHama	1.00*	0*	0*
HERMma	1.13*(.07)	0*	0*
TURHma	1.06*(.07)	0*	0*
VIHAke	0*	1.00*	0*
HERMke	0*	1.13*(.07)	0*
TURHke	0*	1.06*(.07)	0*
VIHApe	0*	0*	1.00*
HERMpe	0*	0*	1.13*(.07)
TURHpe	0*	0*	1.06*(.07)

4.2.3 Kolmen faktorin rakenneyhtälömalli

Koska negatiivisten mielialojen pitkittäisaineiston mittauskertoja vastaavilla faktoreilla on olemassa järjestys, sovittiin aineistoon myös faktoreiden rakenneyhtälömallia. Mallissa on mukana faktoreiden suorat yhteydet peräkkäisten mittausten välillä. Jotta rakenneyhtälömalli olisi täysin vertailukelpoinen aiempien mallien kanssa, myös tähän malliin vapautettiin estimoitavaksi samat jäännösten väliset kovarianssit kuin edellisissäkin malleissa. Lisäksi mallin lataukset estimoitiin yhtä suurina samoin kuin mallissa 2.

Rakenneyhtälömallin χ^2 -testisuureen arvoksi saatiin $\chi^2(18) = 20.19$ ($p = 0.322$), joka on lähes sama kuin stationaarisuusmallin testisuureen arvo $\chi^2(17) = 20.35$ ($p = 0.256$). Vapausasteita rakenneyhtälömallin tapauksessa on yksi enemmän, sillä faktoreiden suora yhteys maanantain ja perjantain mittausten välillä jätettiin nyt estimoimatta. χ^2 -testin perusteella rakenneyhtälömalli olisi näistä kahdesta mallista parempi. Riittävyysindeksien perusteella rakenneyhtälömalli sopii aineistoon hyvin: CFI ≈ 1.00 , NFI = 0.92, RMSEA = 0.035 ja SRMR = 0.11. Verrattaessa koko mallia koskevia riittävyysindeksejä mallin 2 riittävyysindekseihin voidaan todeta, että rakenneyhtälömalli on indeksistä riippuen joko yhtä hyvä tai parempi malli kuin malli 2. χ^2 -testin ja riittävyysindeksien perusteella näyttäisi, että yksinkertaisempi

rakenneyhtälömalli olisi riittävä mallinnettaessa lasten negatiivisten mielialoja tässä aineistossa.

Latausparametrien estimaatit keskivirheineen, jäännösvarianssit ja -kovarianssit keskivirheineen sekä havaittujen muuttujien reliabiliteetit ovat hyvin lähellä mallin 1 vastaavia parametrien estimaatteja. Erot ovat suurimmillaankin sadasosia. Rakenneyhtälömallin β -rakenneparametrien estimaatit ovat tilastollisesti merkitseviä. Maanantain ja keskiviikon mittausten välisen β_{21} -parametrin estimaatin arvo on 0.54 (s.e. = .11) ja keskiviikon ja perjantain mittausten välisen β_{32} -parametrin estimaatin arvo on 0.53 (s.e. = .09). Faktoreiden selitysasteet ovat faktorille NEGke 0.27 ja faktorille NEGpe 0.31. Rakenneyhtälömallin faktoreiden väliset korrelaatiokertoimet eli pysyvyykertoimet kertovat negatiivisten mielialojen pysyvyydestä. Negatiiviset mielialat näyttäisivät olevan kohtalaisen pysyvyviä sekä maanantain ja keskiviikon mittausten välillä ($R_{\eta_2} = 0.52$) että keskiviikon ja perjantain mittausten välillä ($R_{\eta_3} = 0.55$).

4.2.4 Rakenneyhtälömallin rakenneparametrien aikainvarianttisuus

Koska aiemmin mallin 2 kohdalla todettiin, että negatiivisten mielialojen pitkittäisfaktorimalli on stationaarinen, voidaan rakenneyhtälömallin erikoistapauksena tutkia β -rakenneparametrien aikainvarianttisuutta. Tällöin rakenneparametrit estimoidaan yhtäsuurina:

$$\beta_{21} = \beta_{32}.$$

Mallin lataukset estimoidaan samoin kuin mallissa 2 ja 3. Jotta rakenneparametrien aikainvarianttisuusmalli olisi vertailukelpoinen aiempien mallien kanssa, myös tähän malliin vapautettiin estimoitavaksi samat jäännöskovarianssit kuin edellä.

Aikainvarianttisuusmallin voimassa oloa testattiin χ^2 -peräkkäistestillä. Vertailumallina oli nyt negatiivisten mielialojen rakenneyhtälömalli. Aikainvarianttisuusmallin χ^2 -testisuureen arvoksi saatiin $\chi^2(19) = 20.19$ ($p = 0.383$). Tässä mallissa on enemmän vapausasteita kuin rakenneyhtälömallissa, sillä rakenneparametrit estimoitiin yhtäsuurina. Rakenneyhtälömallille χ^2 -testisuureen arvo oli $\chi^2(18) = 20.19$, joten peräkkäistestin arvoksi saadaan nyt

$$D = 20.19 - 20.19 = 0.$$

Peräkkäistestin vapausasteiden lukumäärä on

$$df = 19 - 18 = 1.$$

Kun merkitsevyystasoksi valitaan $\alpha = 0.10$, saadaan χ^2 -jakauman vertailuarvoksi $\chi_{0.10}^2(1) = 2.706$. Peräkkäistestin arvo on pienempi kuin vertailuarvo, joten lasten negatiivisten mielialojen rakenneyhtälömallin rakenneparametrit ovat aikainvariantteja ($p > 0.10$).

Aikainvarianttisuusmallin latausparametrien estimaatit keskivirheineen, jäännösvarianssien ja -kovarianssien estimaatit keskivirheineen sekä havaittujen muuttujien reliabiliteetit ovat jälleen hyvin lähellä edellä käsiteltyjen mallien vastaavia arvoja. Mallin rakenneparametrien estimaatit ovat tilastollisesti merkitseviä. Estimaattien arvot ovat samat, $\beta_{21} = \beta_{32} = 0.53$ (s.e. = 0.07). Faktoreiden selitysasteet $R_{\eta_t}^2$ ovat faktorille NEGke 0.26 ja faktorille NEGpe 0.31., joten NEGke -faktorin selitysaste laski sdasosalla verrattuna rakenneyhtälömallin keskiviikon faktorin selitysasteeseen. NEGpe -faktorin selitysasteet ovat kummassakin mallissa samat. Aikainvarianttisuusmallin pysyvyyskerroimet ovat lähes samoja kuin rakenneyhtälömallillakin: maanantain ja keskiviikon mittausten välinen kerroin on $R_{\eta_2} = 0.51$ ja keskiviikon ja perjantain mittausten välinen kerroin on $R_{\eta_3} = 0.56$. Aikainvarianttisuusmallin R_{η_2} -kerroin laski ja R_{η_3} -kerroin nousi sadasosalla verrattuna rakenneyhtälömallin vastaaviin kerroimiin. Erot ovat kuitenkin niin pieniä, että negatiivisten mielialojen mittaamisen voidaan sanoa olevan edelleen kohtalaisen pysyvää.

5 Yhteenveto

Tutkielman alkuosassa esiteltiin konfirmatorisen pitkittäisfaktorimallin teoriaa keskittyen erityisesti mallinrakentamisen vaiheisiin. Pitkittäisfaktorimallin rakentamista käsiteltiin kahdessa osassa. Ensin tarkasteltiin konfirmatorisen faktorimallin rakentamista yksittäisellä mittauskerralla, jonka jälkeen tarkasteltiin, kuinka yksittäisten mittauskertojen faktorimalleista voitiin rakentaa mittausten välisten kovarianssimatriisien avulla konfirmatorinen pitkittäisfaktorimalli. Konfirmatoristen pitkittäisfaktorimallien erikoistapauksina teoriaosassa käsiteltiin stationaarisia ja invariantteja pitkittäisfaktorimalleja sekä pitkittäisfaktorimallien rakentamista ryhmävertailutilanteissa. Teoriaosan lopuksi tarkasteltiin pitkittäisfaktorimallien rakenneyhtälömalleja.

Tutkielman soveltavassa osassa aineistona käytettiin Vanhemmat, opettajat ja lapsen oppiminen -tutkimuksen pitkittäisaineistoa. Aineistosta valittiin tätä tutkielmaa varten kolmelta eri viikonpäivältä kultakin kolme negatiivista mielialaa kuvaava muuttujaa ja kolme positiivista mielialaa kuvaavaa muuttujaa. Mielialamuuttujat olivat vanhempien ensimmäisellä luokalla olevan lapsensa mielialoista tekemien arviointien pistemääriä. Tavoitteena oli tutkia positiivisten ja negatiivisten mielialamuuttujien faktorirakenteita sekä faktorirakenteiden stationaarisuutta ja mielialamuuttujien mittaamisen pysyvyyttä.

Sopivien mielialoja mittaavien muuttujien ja faktorirakenteiden löytämiseksi käytettiin alustavana analyysimenetelmänä eksploratiivista faktorianalyysia. Eksploratiivisen faktorianalyysin avulla aineistoon sovitettiin kolmen mittauskerran mallia, jossa kultakin mittauskerralta oli mukana kaksi faktoria. Mallin sovituksessa ilmeni kuitenkin ongelmia, sillä faktorointi ei onnistunut. Ongelmakohdaksi paljastui keskimmäisen mittauksen positiivisia mielialoja kuvaavat muuttujat. Selkeää syytä faktoroinnin epäonnistumiselle ei kyseisten muuttujien tarkastelussa kuitenkaan löytynyt.

Eksploratiivisen faktorianalyysin suorittamisessa ilmenneiden faktorointiongelmien takia lopullisessa mallinnuksessa päätettiin mallintaa ainoastaan lasten negatiivisia mielialoja. Negatiivisten mielialamuuttujien faktorirakenteita testattiin konfirmatorista faktorianalyysia käyttäen. Aineistoon sovitettiin ensin konfirmatorinen pitkittäisfaktorimalli ja sen jälkeen testattiin pitkittäisfaktorimallin stationaarisuutta. Seuraavaksi faktoreiden mittauskertojen väliset kovarianssit kuvattiin rakenneyhtälöiden avulla. Lopuksi testattiin rakenneparametrien aikainvarianttisuutta. Näistä faktori-

malleista parhaiten ensimmäisellä luokalla olevien lasten negatiivisia mielialoja kuvaavaksi malliksi osoittautui rakenneyhtälömalli, jonka rakenneparametrit olivat aikainvariantteja. Lasten negatiiviset mielialat osoittautuivat kohtalaisen pysyviksi peräkkäisten mittauspäivien välillä, mutta ensimmäisen ja kolmannen mittauspäivän välillä ei mielialoilla ollut juurikaan pysyvyyttä.

Jatkossa olisi tärkeää kyetä löytämään syyt, jotka aiheuttivat ongelmia positiivisten tunteiden analysoinnissa. Tutkielmassa näitä syitä yritettiin kartoittaa eri menetelmillä. Ainoaksi varteenotettavaksi syyksi faktorointien epäonnistumiselle jäi tutkimusaineiston pieni koko. Vanhemmat, opettajat ja lapsen oppiminen -tutkimuksen aineisto tulee kuitenkin kasvamaan, kun mukaan aineistoon saadaan myös lukuvuosi-
na 2007–2008 ja 2008–2009 tutkimukseen osallistuvat vanhemmat. Siten myös tämä syy saadaan eliminoitua.

Kiinnostavaa olisi myös toistaa tässä tutkielmassa tehdyt analyysit tuleville Vanhemmat, opettajat ja lapsen oppiminen -tutkimukseen osallistuville vanhemmille sekä verrata saatuja tuloksia keskenään. Uusille aineistoille olisi lisäksi mielenkiintoista tehdä ryhmävertailuja vanhempien ja opettajien välillä.

Lähteet

- Aunola, K. & Nurmi, J.-E. (2005). Motivaation ja koulutaitojen kehitys arjen oppimistilanteissa. *Tutkimussuunnitelma*. Jyväskylän yliopisto.
- Aunola, K. & Nurmi, J.-E. (2006). VALO: Vanhemmat, opettajat ja lapsen oppiminen. *EriKa. Erityisopetuksen tutkimus- ja menetelmätieto*, 3(2006), 5–8.
- Bentler, P.M. & Bonett, D.G. (1980). Significance Tests and Goodness of Fit in the Analysis of Covariance Structures. *Psychological Bulletin*, 88, 588–606.
- Bollen, K.A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons.
- Browne, M.W. & Cudeck, R. (1993). Alternative Ways of Assessing Model Fit. Teoksessa K.A. Bollen & J.S. Long (toim.). *Testing Structural Equation Models*, 230–258. Newbury Park, CA: Sage Publications.
- Leskinen, E. (1987). *Faktorianalyysi. Konfirmatoristen faktorimallien teoria ja rakentaminen*. Jyväskylän yliopiston tilastotieteen laitoksen julkaisuja 1/1987.
- Leskinen, E. (1995). Pitkittäisaineistojen faktorimalleista. *Suomen Tilastoseuran vuosikirja 1994*, 117–129.
- Leskinen, E. (2005). Pitkittäisaineistojen analyysistä latenttien muuttujien malleilla. *Psykologia*, 5–6, 476–488.
- Tanaka, J.S. (1993). Multifaceted Conceptions of Fit in Structural Equation Models. Teoksessa K.A. Bollen & J.S. Long (toim.). *Testing Structural Equation Models*, 10–39. Newbury Park, CA: Sage Publications.
- Tisak, J. & Meredith, W. (1990). Longitudinal Factor Analysis. Teoksessa A. von Eye (toim.). *Statistical Methods in Longitudinal Research*, 125–149. Boston, MA: Academic Press.

Watson, D., Clark, L.A. & Tellegen, A. (1988). Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070.

Liitteet

Liite A: Vanhempien päiväkirjan kysymyssarja 8.2.

8.2. Arvioi lapsen mielialaa tämän päivän aikana asteikolla 1–5 (1 = ei ollenkaan, 2 = ei juurikaan, 3 = jonkin verran, 4 = aika paljon, 5 = erittäin paljon):

Lapseni vaikutti tänään:

1. vihaiselta/ärtyneeltä	1	2	3	4	5
2. hermostuneelta/levottomalta	1	2	3	4	5
3. syyllisen oloiselta	1	2	3	4	5
4. iloiselta/onnelliselta/innostuneelta	1	2	3	4	5
5. turhautuneelta	1	2	3	4	5
6. ylpeältä	1	2	3	4	5
7. surulliselta/alakuloiselta	1	2	3	4	5
8. ahdistuneelta	1	2	3	4	5
9. kiitolliselta	1	2	3	4	5
10. nolostuneelta/häpeilevältä	1	2	3	4	5
11. väsyneeltä	1	2	3	4	5
12. joltakin muulta miltä?	1	2	3	4	5

Liite B: LISREL-komentojonot

Malli 1: Kolmen faktorin pitkittäismalli.

Pitkittäismalli

DA NI=9 NO=101 MA=PM

LA

VIHama HERMma TURHma VIHake HERMke TURHke VIHape HERMpe TURHpe

PM FI=Negat.pm

MO NY=9 NE=3 LY=FU,FI PS=FU,FI TE=FU,FI

LE

NEGATma NEGATke NEGATpe/

VA 1 LY 1 1 LY 4 2 LY 7 3

FR LY 2 1 LY 3 1

FR LY 5 2 LY 6 2

FR LY 8 3 LY 9 3

FR PS 1 1 PS 2 2 PS 3 3

FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 TE 7 7 TE 8 8 TE 9 9

ST 1 LY 2 1 LY 3 1 LY 5 2 LY 6 2 LY 8 3 LY 9 3

ST 0.6 PS 1 1 PS 2 2 PS 3 3

ST 0.3 TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6

ST 0.3 TE 7 7 TE 8 8 TE 9 9

FR PS 2 1 PS 3 1 PS 3 2

FR TE 4 2 TE 6 1 TE 6 3 TE 7 2 TE 8 1 TE 8 5 TE 9 1 TE 9 2 TE 9 3

FR TE 9 5 TE 9 6

PATH DIAGRAM

OUT SS RS FS MI ME=GL IT=1000 NS

Malli 2: Stationaarinen kolmen faktorin pitkittäismallin.

Pitkittäismallin stationaarisuus

DA NI=9 NO=101 MA=PM

LA

VIHama HERMma TURHma VIHake HERMke TURHke VIHape HERMpe TURHpe

PM FI=Negat.pm

MO NY=9 NE=3 LY=FU,FI PS=FU,FI TE=FU,FI

LE

NEGATma NEGATke NEGATpe/

VA 1 LY 1 1 LY 4 2 LY 7 3

FR LY 2 1 LY 3 1

FR LY 5 2 LY 6 2

FR LY 8 3 LY 9 3

FR PS 1 1 PS 2 2 PS 3 3

FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 TE 7 7 TE 8 8 TE 9 9

ST 1 LY 2 1 LY 3 1 LY 5 2 LY 6 2 LY 8 3 LY 9 3

ST 0.6 PS 1 1 PS 2 2 PS 3 3

ST 0.3 TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6

ST 0.3 TE 7 7 TE 8 8 TE 9 9

FR PS 2 1 PS 3 1 PS 3 2

FR TE 4 2 TE 6 1 TE 6 3 TE 7 2 TE 8 1 TE 8 5 TE 9 1 TE 9 2 TE 9 3

FR TE 9 5 TE 9 6

EQ LY 2 1 LY 5 2 LY 8 3

EQ LY 3 1 LY 6 2 LY 9 3

PATH DIAGRAM

OUT SS RS FS MI ME=GL IT=1000 NS

Malli 3: Kolmen faktorin rakenneyhtälömalli.

Rakenneyhtälömalli

DA NI=9 NO=101 MA=PM

LA

VIHama HERMma TURHma VIHake HERMke TURHke VIHape HERMpe TURHpe

PM FI=Negat.pm

MO NY=9 NE=3 LY=FU,FI BE=FU,FI PS=FU,FI TE=FU,FI

LE

NEGATma NEGATke NEGATpe/

FR BE 2 1 BE 3 2 VA 1 LY 1 1 LY 4 2 LY 7 3

FR LY 2 1 LY 3 1

FR LY 5 2 LY 6 2

FR LY 8 3 LY 9 3

FR PS 1 1 PS 2 2 PS 3 3

FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 TE 7 7 TE 8 8 TE 9 9

ST 1 LY 2 1 LY 3 1 LY 5 2 LY 6 2 LY 8 3 LY 9 3

ST 0.6 PS 1 1 PS 2 2 PS 3 3

ST 0.3 TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6

ST 0.3 TE 7 7 TE 8 8 TE 9 9

FR TE 4 2 TE 6 1 TE 6 3 TE 7 2 TE 8 1 TE 8 5 TE 9 1 TE 9 2 TE 9 3

FR TE 9 5 TE 9 6

EQ LY 2 1 LY 5 2 LY 8 3

EQ LY 3 1 LY 6 2 LY 9 3

PATH DIAGRAM

OUT SS RS FS MI ME=GL IT=1000 NS

Malli 4: Rakenneyhtälömallin rakenneparametrien aikainvarianttisuus.

Rakenneparametrien yhtäsuuruus

DA NI=9 NO=101 MA=PM

LA

VIHama HERMma TURHma VIHake HERMke TURHke VIHape HERMpe TURHpe

PM FI=Negat.pm

MO NY=9 NE=3 LY=FU,FI BE=FU,FI PS=FU,FI TE=FU,FI

LE

NEGATma NEGATke NEGATpe/

FR BE 2 1 BE 3 2 VA 1 LY 1 1 LY 4 2 LY 7 3

FR LY 2 1 LY 3 1

FR LY 5 2 LY 6 2

FR LY 8 3 LY 9 3

FR PS 1 1 PS 2 2 PS 3 3

FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 TE 7 7 TE 8 8 TE 9 9

ST 1 LY 2 1 LY 3 1 LY 5 2 LY 6 2 LY 8 3 LY 9 3

ST 0.6 PS 1 1 PS 2 2 PS 3 3

ST 0.3 TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6

ST 0.3 TE 7 7 TE 8 8 TE 9 9

FR TE 4 2 TE 6 1 TE 6 3 TE 7 2 TE 8 1 TE 8 5 TE 9 1 TE 9 2 TE 9 3

FR TE 9 5 TE 9 6

EQ LY 2 1 LY 5 2 LY 8 3

EQ LY 3 1 LY 6 2 LY 9 3

EQ BE 2 1 BE 3 2

PATH DIAGRAM

OUT SS RS FS MI ME=GL IT=1000 NS