



**TURUN  
YLIOPISTO**

MATEMATIIKKA- SEKÄ LUKEMIS- JA  
KIRJOITTAMISAHDISTUSKYSELYIDEN FAKTORIRAKENTEET JA  
MITTAUSINVARIANSSI

Santeri Holopainen

Pro gradu -tutkielma  
Huhtikuu 2023

Tarkastajat:  
FT Janne Kujala  
FT Joni Virta

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

Turun yliopiston laatu­järjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck-järjestelmällä

TURUN YLIOPISTO

Matematiikan ja tilastotieteen laitos

SANTERI HOLOPAINEN: Matematiikka- sekä lukemis- ja kirjoittamisahdistuskyselyiden faktorirakenteet ja mittausinvarianssi

Pro gradu -tutkielma, 46 s., 7 liites.

Tilastotiede

Huhtikuu 2023

---

Ahdistuneisuus on tunnetila, johon liittyy huoli jostain tapahtumasta. Matematiikka- ja lukemisahdistus ovat esimerkkejä tietynlaisista ahdistuneisuuden tyypeistä, joiden on todettu olevan yhteydessä erityisesti lapsen menestykseen akateemisessa ympäristössä tavalla tai toisella. Tällaisia ihmisen tunnetiloja ei voi havaita suoraan vaan ne ovat niin sanottuja latentteja muuttujia eli faktoreita, joita mitataan usein joillakin havaittavissa olevilla muuttujilla, kuten kyselypatterin kysymyksillä tai väittämillä. Näin tehtiin vuoden 2021 syksyllä, kun 3.–9.-luokkalaiset lapset vastasivat muun FUNA-tehtävistön ohella matematiikka- sekä lukemis- ja kirjoittamisahdistuskyselyihin.

Faktorianalyysi on eräs tilastollinen menetelmä, jolla faktorien ja havaittujen muuttujien muodostamaa rakennetta analysoidaan. Silloin kun tutkijalla on jo etukäteen ymmärrys kyselyn hypoteettisesta faktorirakenteesta, sitä tutkitaan konfirmatiivisella faktorianalyysillä. Faktorirakenteen tutkimiseen kuuluu usein myös mittausinvarianssin testaaminen, joka kertoo, mitataanko tutkittavaa rakennetta samalla tavalla ominaisuuksiltaan erilaisissa ihmisryhmissä. Mittausinvarianssia testataan ryhmiteltyllä konfirmatiivisella faktorianalyysillä, jossa vertaillaan sisäkkäisiä malleja, joista aina toinen on toista jollain tavalla rajoitetumpi tutkittavien ihmisryhmien suhteen.

Tässä tutkielmassa esitetään faktorianalyysin teoriaa ja menetelmiä testata sovitettua faktorimallin hyvyttä. Tämän lisäksi esitetään mittausinvarianssin määrittely ja tapa testata faktorirakenteen mittausinvarianssia. Esitetyt menetelmiä sovellettiin matematiikka- sekä lukemis- ja kirjoittamisahdistuskyselyillä kerättyihin aineistoihin. Analyysien tulokset osoittivat, että matematiikka-ahdistuskyselyn faktorirakenne oli sellainen kuin oletettiin ja se oli samanlainen tytöillä ja pojilla mutta ei eri luokka-asteiden lapsilla. Lukemis- ja kirjoittamisahdistuskyselyn faktorirakenne ei ollut sellainen kuin oletettiin, joten sen mittausinvarianssia ei voinut testata mielekkäästi.

Asiasanat: ahdistuneisuus, matematiikka-ahdistus, lukemis- ja kirjoittamisahdistus, faktorianalyysi, mittausinvarianssi.



# Sisällys

<b>1</b>	<b>Johdanto</b>	<b>1</b>
<b>2</b>	<b>Faktorianalyysi</b>	<b>3</b>
2.1	Rajoittamattomat lineaariset faktorimallit . . . . .	3
2.1.1	Mallin määrittely . . . . .	3
2.1.2	Parametrien identifiointi . . . . .	5
2.2	Rajoitetut lineaariset faktorimallit . . . . .	6
2.2.1	Mallin määrittely . . . . .	6
2.2.2	Parametrien identifiointi . . . . .	6
2.2.3	Estimointi . . . . .	8
2.2.4	Yksinkertainen esimerkki . . . . .	10
2.3	Ryhmitellyt rajoitetut faktorimallit . . . . .	11
2.3.1	Mallin määrittely . . . . .	11
2.3.2	Parametrien identifiointi . . . . .	12
2.3.3	Estimointi . . . . .	13
2.4	Faktorimallien sopivuuden testaaminen . . . . .	14
2.4.1	Absoluuttiset sopivuusindeksit . . . . .	15
2.4.2	Inkrementaaliset sopivuusindeksit . . . . .	17
2.4.3	Sopivuusindeksien estimaattien harhaisuus . . . . .	18
2.4.4	Ryhmiteltyjen faktorimallien sopivuuden testaaminen . . . . .	19
2.5	Faktorimallien uudelleenmäärittely . . . . .	20
<b>3</b>	<b>Mittausinvarianssi</b>	<b>22</b>
3.1	Määritelmä . . . . .	22
3.2	Faktorisen invarianssin testaaminen . . . . .	22
3.2.1	Testattavat mallit . . . . .	23
3.2.2	Testausmenettely . . . . .	24
3.2.3	Osittainen invarianssi . . . . .	25
<b>4</b>	<b>Ahdistuskyselyjen faktorirakenteet ja mittausinvarianssi</b>	<b>26</b>
4.1	Metodit . . . . .	28
4.1.1	Aineisto . . . . .	28
4.1.2	Kyselyt . . . . .	28
4.1.3	Proseduuri . . . . .	30
4.1.4	Analyysi . . . . .	30
4.2	Tulokset . . . . .	31
4.2.1	Kuvailevat tunnusluvut . . . . .	31
4.2.2	Kyselyiden faktorirakenteet . . . . .	32
4.2.3	Sopivuusindeksiestimaattorien tarkkuuden arviointi . . . . .	34
4.2.4	Kyselyiden mittausinvarianssitestaukset . . . . .	35
<b>5</b>	<b>Pohdintaa</b>	<b>38</b>
	<b>Viitteet</b>	<b>42</b>

<b>Liitteet</b>	<b>47</b>
A Faktorimalliesimerkin kovarianssimatriisin määrittely . . . . .	47
B Sovitusfunktion approksimatiivisen kvadraattisen muodon perustelu .	48
C MAQ- ja RWAQ-kyselyjen osiokohtaiset kuvailevat tunnusluvut . . .	49
D Esimerkkikuvat osiokohtaisista jakaumista . . . . .	52
E RWAQ-mallin rajoitettujen parametrien MI- ja EPC-arvojen tulkinta	53

# 1 Johdanto

Vuoden 2021 syksyllä useassa eri peruskoulussa Varsinais-Suomessa ja Uudellamaalla teetettiin FUNA-mittaukset [32] 3.–9.-luokkalaisille lapsille. Muun FUNA-tehtävistön lisäksi lapset vastasivat kolmeen ahdistusta käsittelevään kyselyyn, joista kahta käsitellään tässä tutkielmassa. Näistä kahdesta kyselystä ensimmäisessä lapset vastasivat kysymyksiin, jotka koskivat heidän kokemaansa ahdistusta matematiikkaan liittyvissä tilanteissa. Toisen kyselyn kysymykset koskivat lasten kokemaa ahdistusta lukemiseen ja kirjoittamiseen liittyvissä tilanteissa. Mutta miksi tällaiset kyselyt sitten teetettiin ja miten niitä tutkitaan? Ennen näihin kysymyksiin vastaamista pitää määritellä, mitä ahdistuneisuus edes ylipäätensä on.

Psykiatrian erikoislääkärin Matti Huttusen [18] mukaan ahdistuneisuudella tarkoitetaan ”pelonsekaista tai huolestunutta tunnetilaa, johon liittyy osin tietoinen tai tiedostamaton huoli jostain nykyhetkeen tai tulevaisuuteen liittyvästä tapahtumasta”. Ciporan et al. [11] mukaan matematiikka-ahdistus on oma erillinen ahdistuksen tyyppinsä, mutta se on määritelty aiheutta käsittelevässä kirjallisuudessa eri tavoin. Erään määritelmän mukaan matematiikka-ahdistus on tunnetila, joka voi vaihdella eri henkilöiden ja tilanteiden kesken; toisen määritelmän mukaan se on tietynlainen fobia. Piccolo et al. [34] määrittelivät lukemisahdistuksen myös tietynlaisena fobiana, joka riippuu tilanteesta. Tällaisia ihmisen tunnetiloja tai fobioita on vaikea mitata suoraan, joten erilaiset ahdistuneisuuden tyypit – kuten matematiikka-, lukemista tai kirjoittamisahdistus – voidaan käsittää niin sanottuina latentteina eli taustalla olevina muuttujina. Latentteja muuttujia voidaan mitata havaittavissa olevien muuttujien avulla, jotka voivat olla esimerkiksi kyselypatterin kysymyksiä kuten FUNA-mittausten ahdistuneisuuskyselyjen tapauksessa.

Latentteja muuttujia tutkitaan usein faktorianalyysillä ja niin tehdään myös tämän tutkielman empiirisessä osiossa. Luvussa 2 esitetään faktorianalyysin teoriaa lähtien liikkeelle rajoittamattomista faktorimalleista, joiden analysointi on luonteeltaan eksploratiivista. Tämän jälkeen siirrytään rajoitettujen faktorimallien teoriaan, joita varten analyysien tekijällä pitää olla etukäteen jokin teoreettinen näkemys tutkittavan mittarin/kyselyn faktorirakenteesta: hänellä täytyy olla käsitys (1) latenttien muuttujien lukumäärästä ja (2) siitä, mikä latentti muuttuja vaikuttaa mihinkin havaittuihin muuttujiin. Faktorianalyysiä rajoitetuilla faktorimalleilla kutsutaan konfirmatiiviseksi faktorianalyysiksi; sen avulla selvitetään, onko teorisoitu faktorirakenne validi. Ahdistuskyselyillä oli entuudestaan mietityt ja perustellut faktorirakenteet, joten niiden analysointiin käytettiin rajoitettuja faktorimalleja.

Faktorianalyysin olennainen osa on arvioida estimoidun faktorimallin sopivuutta aineistoon. Tätä varten on kehitetty useita tilastollisia menetelmiä, joita esitetään luvussa 2.4. Joskus sovitettu faktorimalli ei sovi aineistoon hyvin, jolloin teorisoitu faktorirakenne ei ole validi. Tällöin mallia halutaan usein modifioida, jolloin tarvitaan empiiristä tukea modifikaatioiden tekemiselle. Luvun 2 lopuksi esitetään eräs menetelmä, jota käytetään usein apuna, kun faktorimalleja modifioidaan. Luku 2 perustuu pääosin Bollenin [4] kirjaan *Structural Equations with Latent Variables* ja osittain Kaplanin [20] kirjaan *Structural Equation Modeling: Foundations and Extensions* sekä muutamaaan tieteelliseen artikkeliin.

Ahdistuskyselyjen tapaisten psykologisten mittarien tutkimisessa ollaan usein

kiinnostuneita faktorirakenteen lisäksi mittausinvarianssista. Se kertoo, pystytäänkö mittarilla mittaamaan samaa asiaa ominaisuuksiltaan erilaisissa ihmisryhmissä. Mittausinvarianssin tutkiminen on olennaista etenkin silloin, kun teorisoitu faktorirakenne on validi ja eri ihmisryhmiä halutaan vertailla mitattavien latenttien muuttujien suhteen. Ahdistuskyselyjen tapauksessa haluttiin tietää, olivatko teorisoitut faktorirakenteet invariantteja tyttöjen ja poikien sekä eri luokka-asteiden lasten suhteen. Luvussa 3 puhutaan mittausinvarianssista ja sen testaamisesta faktorimalleilla. Ensin esitetään mittausinvarianssin tilastollinen määritelmä, jonka jälkeen esitetään neljä mittausinvarianssin tutkimisessa testattavaa mallia. Lopuksi puhutaan lyhyesti osittaisesta invarianssista, jota testataan usein silloin, kun mittausinvarianssi ei jollain invarianssitasolla toteudu tutkittavalla faktorimallilla. Luku 3 perustuu pääosin Millsapin [30] teokseen *Statistical Approaches to Measurement Invariance* ja osittain muutamaan tieteelliseen artikkeliin.

Edellä vastattiin lähinnä kysymykseen 'Miten tällaisia ahdistuskyselyjä tutkitaan?'. Kysymykseen 'Miksi tällaiset ahdistuskyselyt teetettiin?' vastataan luvun 4 alussa, jossa ahdistuskyselyjä taustoitetaan kertomalla matematiikka-, lukemis- ja kirjoittamisahdistuskäsitteiden historiasta. Luvun 4 alussa puhutaan myös siitä, miten ja missä näitä ahdistuksen tyyppisiä on tutkittu aikaisemmin. Luvussa 4.1 esitetään empiirisen tutkimuksen menetelmät. Luvussa 4.2 esitetään tuloksia, jotka saatiin soveltamalla luvuissa 2 ja 3 esitettyjä menetelmiä ahdistuskyselyillä kerättyihin aineistoihin. Tutkielman lopuksi luvussa 5 pohditaan empiirisen osion tuloksia erityisesti ahdistuskyselyjen jatkotutkimuksen kannalta.



## 2 Faktorianalyysi

Erityisesti yhteiskuntatieteissä ja psykologiassa ollaan kiinnostuneita latenteista muuttujista eli faktoreista, joista hyviä esimerkkejä ovat ihmisten asenteet, tuntemukset ja motiivit. Sana 'latentti' viittaa siihen, että faktorit eivät ole suoraan havaittavissa, joten niitä yleensä mitataan ja tulkitaan joidenkin havaittavissa olevien muuttujien eli indikaattorien avulla. [17] Yleinen tapa hankkia mittaustietoa on teettää kyselypatteri, joka koostuu erilaisista tutkittavaan aiheeseen liittyvistä kysymyksistä tai väittämistä.

Yleisesti käytetty tilastollinen menetelmä latenttien muuttujien tutkimisessa on faktorianalyysi, jonka tavoitteena on tiivistää suurehkon määrän indikaattoreita sisältävä informaatio pienempään määrään faktoreita. Tässä on oletuksena siis se, että faktorit ovat jossain määrin aiheuttaneet havaitut muuttujat. Faktorianalyysissä ei kuitenkaan oleteta, että indikaattorit mittaavat faktoreita erehtymättömästi. [17] Faktorianalyysi on siis tietyssä mielessä datan reduktiomenetelmä. Faktorianalyysi eroaa yleisimmistä tilastollisista menetelmistä – kuten regressio- tai varianssianalyysistä – siinä, että yksittäisten havaintojen sijasta mallinnetaan kovariansseja, ja tavoitteena on minimoida otoskovarianssien ja mallin ennustamien kovarianssien välinen ero. [4]

Hyvä esimerkki faktorianalyysin käytöstä on FUNA-konsortion<sup>1</sup> kehittämä FUNA-DB-seulontatesti [32], joka on tarkoitettu perusopetusikäisten lasten matematiikan oppimisvaikeuksien tunnistamiseen. Testin kahden eri matematiikan taitotalueen (kahden faktorin) ja yhteensä kuuden eri tehtävän muodostama rakenne on perusteltu faktorianalyysillä Räsänen et al. [37] artikkelissa. Toinen hyvä esimerkki on Careyn et al. [9] mAMAS (modified Abbreviated Math Anxiety Scale) -mittarin faktorirakenteen tutkiminen. Tämän tutkielman empiirisessä osiossa tutkittavat MAQ- ja RWAQ-kyselyt perustuvat osittain mAMAS-mittariin.

### 2.1 Rajoittamattomat lineaariset faktorimallit

#### 2.1.1 Mallin määrittely

Olkoon  $q$  havaittujen muuttujien eli indikaattorien määrä ja  $k$  latenttien muuttujien eli faktorien määrä, joita indikaattorit mittaavat. Faktorimalleissa muuttujat esitetään usein poikkeamina keskiarvoistaan, ja täten eksploratiivisessa faktorianalyysissä (engl. exploratory factor analysis, EFA) käytettävä rajoittamaton lineaarinen faktorimalli on muotoa

$$\mathbf{x} = \mathbf{\Lambda}\xi + \delta, \tag{1}$$

jossa  $\mathbf{x}$  on havaituista muuttujista koostuva  $q \times 1$  -vektori,  $\mathbf{\Lambda}$  on faktorilatauksista koostuva  $q \times k$  -matriisi,  $\xi$  on faktoreista koostuva  $k \times 1$  -vektori ja  $\delta = \mathbf{e} + \mathbf{s}$  on puhtaasta mittavirheestä  $\mathbf{e}$  sekä spesifisestä virheestä  $\mathbf{s}$  koostuva  $q \times 1$  -vektori. [20]

Spesifinen virhe  $\mathbf{s}$  johtuu muuttujien spesifisestä valinnasta malliin. Tämä nähdään tutkimalla niin sanottuja 'aitoja pistemääriä'  $\mathbf{t}$ , kun havaittujen muuttujien  $\mathbf{x}$

<sup>1</sup>Ks. FUNA-konsortion jäsenet: [<https://sites.google.com/funa.nmi.fi/eng/members>].

ajatellaan olevan niin sanottuja 'havaittuja pistemääriä'. Aidot pistemäärät määritellään kaavalla  $\mathbf{t} = \mathbf{x} - \mathbf{e}$ , jossa  $\mathbf{x}$  ja  $\mathbf{e}$  on määritelty kuten edellä. Tällöin yhtälön (1) faktorimalli voidaan kirjoittaa aidoille pistemäärille yhtälöllä  $\mathbf{t} = \mathbf{\Lambda}\xi + \mathbf{s}$ , jossa  $\mathbf{\Lambda}$  ja  $\xi$  on määritelty kuten edellä ja  $\mathbf{s}$  sisältää yhtälössä (1) tehdystä muuttujien tietystä valinnasta johtuvat spesifiset virheet. Täten  $\mathbf{x} - \mathbf{e} = \mathbf{\Lambda}\xi + \mathbf{s}$ , eli  $\mathbf{x} = \mathbf{\Lambda}\xi + \mathbf{e} + \mathbf{s}$ , jossa  $\mathbf{e} + \mathbf{s} = \delta$ . [20]

Yhtälön (1) faktorimallin matriisiesitys voidaan kirjoittaa auki, jolloin saadaan

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1k} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{q1} & \lambda_{q2} & \cdots & \lambda_{qk} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_q \end{bmatrix}, \quad (2)$$

tai vastaavasti

$$\begin{cases} x_1 = \lambda_{11}\xi_1 + \lambda_{12}\xi_2 + \cdots + \lambda_{1k}\xi_k + \delta_1, \\ x_2 = \lambda_{21}\xi_1 + \lambda_{22}\xi_2 + \cdots + \lambda_{2k}\xi_k + \delta_2, \\ \vdots \\ x_q = \lambda_{q1}\xi_1 + \lambda_{q2}\xi_2 + \cdots + \lambda_{qk}\xi_k + \delta_q, \end{cases}$$

jossa faktorilataukset eli kertoimet  $\lambda_{ij}$ ,  $i \in \{1, \dots, q\}$ ,  $j \in \{1, \dots, k\}$ , voidaan tulkita havaitun muuttujan arvon odotetun muutoksen suuruutena, kun latentin muuttujan arvo kasvaa yhden yksikön verran. Ne ovat niin sanottuja suoria vaikutuksia faktoreista havaittuihin muuttujiin. [4]

Faktorianalyysin fundamentaalinen kovarianssirakennetta koskeva hypoteesi on

$$\mathbf{\Sigma} = \mathbf{\Sigma}(\eta), \quad (3)$$

jossa

$$\mathbf{\Sigma} = \begin{bmatrix} \text{Var}(x_1) & & & & \\ \text{Cov}(x_2, x_1) & \text{Var}(x_2) & & & \\ \vdots & \vdots & \ddots & & \\ \text{Cov}(x_{q-1}, x_1) & \text{Cov}(x_{q-1}, x_2) & \cdots & \text{Var}(x_{q-1}) & \\ \text{Cov}(x_q, x_1) & \text{Cov}(x_q, x_2) & \cdots & \text{Cov}(x_q, x_{q-1}) & \text{Var}(x_q) \end{bmatrix}$$

on havaittujen muuttujien  $q \times q$  -populaatio-kovarianssimatriisi ja  $\eta$  on mallin  $t \times 1$  -parametrivektori<sup>2</sup>. Yhtälö (3) tarkoittaa, että populaatio-kovarianssimatriisin jokainen alkio on funktio mallin yhdestä tai useammasta rakenteellisesta parametrusta. [4] Linearisille faktorimalleille pätee lisäksi oletukset  $E(\delta) = \mathbf{0}$ ,  $\text{Cov}(\xi, \delta) = \mathbf{0}$  ja  $\text{Cov}(\delta_i, \delta_j) = 0$ , kaikilla  $i \neq j$ ,  $i, j \in \{1, \dots, q\}$ . Koska havaitut muuttujat  $\mathbf{x}$  ovat poikkeamia keskiarvoistaan, niin  $E[\mathbf{x}] = \mathbf{0}$  ja niiden populaatio-kovarianssimatriisi voidaan yhtälön (3) mukaan esittää muodossa

$$\begin{aligned} \mathbf{\Sigma} &= E[\mathbf{xx}'] = E[(\mathbf{\Lambda}\xi + \delta)(\mathbf{\Lambda}\xi + \delta)'] = \mathbf{\Lambda}E[\xi\xi']\mathbf{\Lambda}' + E[\delta\delta'] \\ &= \mathbf{\Lambda}\Phi\mathbf{\Lambda}' + \Theta = \mathbf{\Sigma}(\eta). \end{aligned} \quad (4)$$

<sup>2</sup>Tässä tutkielmassa symmetriset matriisit esitetään niin, että vain diagonaali ja sen alapuoliset alkioit ovat näkyvissä ja diagonaalin yläpuoli jätetään tyhjäksi.

Yhtälössä (4)  $\Lambda$  on määritelty kuten yhtälöissä (1) ja (2), ja

$$\Phi = \begin{bmatrix} \phi_{11} & & & \\ \phi_{21} & \phi_{22} & & \\ \vdots & \vdots & \ddots & \\ \phi_{k1} & \phi_{k2} & \cdots & \phi_{kk} \end{bmatrix}$$

on faktorien  $k \times k$  -kovarianssimatriisi ja  $\Theta = \text{diag}(\theta_1, \dots, \theta_q)$  on virheiden variansseista koostuva  $q \times q$  -diagonaalimatriisi [20].

### 2.1.2 Parametrien identifointi

Parametrien määrittelemättömyys on yleisten faktorimallien ominainen ongelma. Tämä nähdään, kun määritellään ei-singulaarinen  $k \times k$  -matriisi  $\mathbf{T}$ . Tällöin, jos  $\Lambda^* = \Lambda\mathbf{T}$ ,  $\xi = \mathbf{T}\xi^*$  ( $\xi^* = \mathbf{T}^{-1}\xi$ ) ja  $\Theta^* = \Theta$ , niin

$$E[\xi^*\xi^{*'}] = E[\mathbf{T}^{-1}\xi\xi'\mathbf{T}^{-1'}] = \mathbf{T}^{-1}\Phi\mathbf{T}^{-1'} = \Phi^*$$

ja

$$\Sigma^* = \Lambda^*\Phi^*\Lambda^{*'} + \Theta^* = \Lambda\mathbf{T}\mathbf{T}^{-1}\Phi\mathbf{T}^{-1'}\mathbf{T}'\Lambda' + \Theta = \Lambda\Phi\Lambda' + \Theta = \Sigma.$$

Nähdään, että faktorirakenteen transformoitu kovarianssimatriisi  $\Sigma^*$  on ekvivalentti alkuperäisen kovarianssimatriisin  $\Sigma$  kanssa ja kyseessä oleva faktorimalli ei siis ole identifioituva. Ongelma voidaan ratkaista asettamalla  $k^2$  rajoitusta matriiseihin  $\Lambda$  ja  $\Phi$ . Eräs tapa on aloittaa asettamalla rajoitus  $\Phi = \mathbf{I}_k$ , jolloin  $\Phi^* = \mathbf{T}^{-1}\Phi\mathbf{T}^{-1'} = \mathbf{T}^{-1}\mathbf{T}^{-1'} = \mathbf{I}_k$ . Tällöin matriisi  $\mathbf{T}$  ja faktorit ovat ortogonaalisia ja puhutaan faktorirakenteen rotaationaalista määrittelemättömyydestä. Jos  $\Phi$  ei ole diagonaalinen, niin faktorien sanotaan olevan vinoja (engl. oblique). [2]

Kun asetetaan rajoitus  $\Phi = \mathbf{I}_k$  ja faktorien lukumäärä on  $k = 1$ , niin  $\Phi$  on skalaari ja kaikki parametrit ovat yksiselitteisesti identifioituvia. Kun  $k \geq 2$ , niin tuntemattomia parametreja on  $k^2 \geq 4$ , joista  $k(k+1)/2 \geq 3$  kappaletta poistuu asettamalla rajoitus  $\Phi = \mathbf{I}_k$ . Jäljelle jäävät  $k^2 - k(k+1)/2 = k(k-1)/2 \geq 1$  tuntematonta parametria poistetaan asettamalla rajoituksia faktorilatausten matriisiin  $\Lambda$ . Tapa jolla nämä rajoitukset asetetaan riippuu estimointimenetelmästä. [20]

Rajoituksen  $\Phi = \mathbf{I}_k$  asettamisen jälkeen mallia voidaan rotatoida. Olkoon  $\mathbf{T}$  ortogonaalinen  $k \times k$  -transformaatiomatriisi. Kun  $k = 2$ , transformaatiomatriisi on muotoa

$$\mathbf{T} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix},$$

jossa parametri  $\theta$  vastaa rotatointikulman suuruutta. Kun  $\theta$ :n arvo on valittu, jäljelle jäänyt yksi tuntematon parametri on saatu poistettua ja malli on identifioituva. Yleisessä tapauksessa  $k \geq 2$  ja tällöin täytyy valita  $k(k-1)/2$  rotatointikulmaa. Faktoriratkaisun rotatointi tehdään usein juuri yleisten faktorimallien tapauksessa, jotta saadaan yksinkertainen esitys faktoreille. [20]

Esitetään vielä lyhyesti eräs yleisesti käytetty rotatointimenetelmä, joka olettaa faktorien olevan ortogonaalisia (rajoitus  $\Phi = \mathbf{I}_k$ ). Varimax-rotatointi (tai -kriteeri)

maksimoi neliöityjen faktorilatausten ja vastaavien sarakekeskiarvojen erotusten neliöiden summan

$$V = \sum_{j=1}^k \left[ \sum_{i=1}^q \left( \lambda_{ij}^{*2} - d_j/q \right)^2 \right],$$

jossa  $\lambda_{ij}^*$  on matriisiin  $\mathbf{\Lambda}^* = \mathbf{\Lambda}\mathbf{T}$   $ij$ -alkio ja  $d_j = \sum_{i=1}^q \lambda_{ij}^{*2}$ . Maksimoimalla  $V$  saadaan ratkaistua matriisiin  $\mathbf{T}$  sisältämät rotatointikulmat varimax-kriteerin mukaisesti. [20]

## 2.2 Rajoitetut lineaariset faktorimallit

Edellä esitetyillä rajoittamattomilla faktorimalleilla (eli EFA:lla) on kolme keskeistä ongelmaa, jotka rajoittavat menetelmän tehokkuutta latenttien muuttujien analysoinnissa. Ensinnäkin, EFA:ssa ei sallita faktorilatausten asettamista nolliksi, vaan kukin faktori vaikuttaa jokaiseen havaittuun muuttujaan. Toiseksi, EFA ei salli korreloituneita virhetermejä, joka tietyissä tilanteissa voisi olla validi oletus. Kolmas ongelma EFA:ssa on faktorien lukumäärän määrittäminen, joka tehdään usein pääakselimenetelmällä. Rajoitetuilla faktorimalleilla, joita käytetään konfirmatiivisessa faktorianalyysissä (engl. confirmatory factor analysis, CFA), ei ole kyseisiä ongelmia. [4]

### 2.2.1 Mallin määrittely

Rajoitettujen lineaaristen faktorimallien määrittely on melkein täysin sama kuin yleisten eli rajoittamattomien faktorimallien määrittely yhtälöissä (1) ja (4). Ero on se, että CFA:ssa ei välttämättä tehdä oletusta  $\text{Cov}(\delta_i, \delta_j) = 0$ , kaikilla  $i \neq j$ ,  $i, j \in \{1, \dots, q\}$ , vaan voi olla, että  $\text{Cov}(\delta_i, \delta_j) \neq 0$ , joillakin  $i \neq j$ ,  $i, j \in \{1, \dots, q\}$ . Rajoitetut lineaariset faktorimallit ovat siis muotoa

$$\mathbf{x} = \mathbf{\Lambda}\xi + \delta, \quad \text{E}(\delta) = \mathbf{0}, \quad \text{Cov}(\xi, \delta) = \mathbf{0}, \quad (5)$$

ja havaittujen muuttujien populaatio-kovarianssimatriisi on muotoa

$$\mathbf{\Sigma} = \mathbf{\Sigma}(\eta) = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}' + \mathbf{\Theta}, \quad (6)$$

jossa on edelleen näkyvissä matriisin  $\mathbf{\Sigma}$  alkioiden riippuvuus matriisin  $\mathbf{\Sigma}(\eta) = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}' + \mathbf{\Theta}$  alkioista. Tämä tarkoittaa, että kun faktorimalli on spesifioitu, havaittujen muuttujien varianssit ja kovarianssit voidaan kirjoittaa mallin rakenteellisten parametrien funktioina. [4]

### 2.2.2 Parametrien identifointi

Rajoitetuissa faktorimalleissa menettely parametrien identifoinnissa on hieman erilaista kuin rajoittamattomissa faktorimalleissa. Ensinnäkin rajoitetuissa faktorimalleissa asetetaan yleensä suurempi määrä rajoituksia. Ne asetetaan faktorilatausten matriisiin  $\mathbf{\Lambda}$  niin, että jäljelle jäävä faktorirakenne vastaa jotain ennalta määriteltyä yksinkertaista hypoteettista rakennetta. Toiseksi, rajoitetuissa faktorimalleissa

rotatointi ei ole mahdollista, sillä tällöin rajoitteiden paikat vaihtuisivat eikä hypoteettinen faktorirakenne toteutuisi. [20]

Tavoitteena on siis löytää yksikäsitteinen ratkaisu matriisien  $\mathbf{\Lambda}$ ,  $\mathbf{\Phi}$  ja  $\mathbf{\Theta}$  rakenteellisille parametreille  $\eta$ . Olkoot  $\eta_1$  ja  $\eta_2$  joitakin tiettyjä  $\eta$ :n arvoja. Tällöin parametrit  $\eta$  ovat identifioituja, jos  $\mathbf{\Sigma}(\eta_1) \neq \mathbf{\Sigma}(\eta_2)$ , kun  $\eta_1 \neq \eta_2$ , eli  $\mathbf{\Sigma}(\eta_1) = \mathbf{\Sigma}(\eta_2)$ , jos ja vain jos  $\eta_1 = \eta_2$ . Toisin sanoen rajoitetun faktorimallin parametrivektori  $\eta$  on identifioituva, jos sen alkiot ovat yksikäsitteisesti määritellyjä funktioita matriisin  $\mathbf{\Sigma}$  alkiosta. Tämän toteutumiseksi osaan parametreista täytyy asettaa rajoitteita aivan kuten rajoittamattomissa faktorimalleissa. [4]

Rajoitetuissa faktorimalleissa parametrirajoitteiden lisäksi täytyy määrittää latenttien muuttujien mitta. Yleisimpiä tapoja tehdä tämä on kaksi, joista ensimmäisessä faktorien varianssit standardoidaan ykkösiksi<sup>3</sup>. Toinen vaihtoehto on asettaa jokaisen faktorin skaala yhtäsuureksi yhden sen indikaattorin skaalan kanssa. Usein käytetty menetelmä on asettaa ensimmäinen faktorilataus ykköseksi kussakin faktorissa. [20] Seuraavaksi esitetään  $t$ -sääntö (engl.  $t$ -rule), jossa oletetaan, että latenttien muuttujien mitta on määritetty. Säännön avulla pystytään määrittämään, onko faktorimalli identifioituva vai ei.

Rajoitetut faktorimallit määriteltiin yhtälöillä (5) ja (6). Yhtälön (6) mukaan  $\mathbf{\Sigma}(\eta) = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}' + \mathbf{\Theta}$ , jossa matriisissa  $\mathbf{\Lambda}$  on  $qk$  alkiota, matriisissa  $\mathbf{\Phi}$   $k(k+1)/2$  ei-redundanttia parametria, matriisissa  $\mathbf{\Theta}$   $q(q+1)/2$  yksikäsitteistä parametria ja matriisissa  $\mathbf{\Sigma}(\eta)$  (vektorissa  $\eta$ ) on siis yhteensä  $qk + k(k+1)/2 + q(q+1)/2$  parametria. Koska vektorissa  $\mathbf{x}$  on  $q$  havaittua muuttujaa, niin populaatio-kovarianssimatriisissa  $\mathbf{\Sigma}$  on  $q(q+1)/2$  yksikäsitteistä alkiota. Täten, jos matriisin  $\mathbf{\Sigma}(\eta)$  kaikki parametrit ovat tuntemattomia, niin jäljelle jää  $qk + k(k+1)/2$  ei-identifiotuvaa parametria. Välttämätön mutta ei riittävä ehto mallin identifioitumiselle on  $t$ -sääntö, eli

$$t \leq \frac{q(q+1)}{2}, \quad (7)$$

jossa  $t$  on vapaiden (tuntemattomien ja ei-rajoitettujen) parametrien määrä vektorissa  $\eta$  ja  $q(q+1)/2$  on yksikäsitteisten alkioiden määrä havaittujen muuttujien  $\mathbf{x}$  kovarianssimatriisissa  $\mathbf{\Sigma}$ . [4]

Yhtälön (7)  $t$ -sääntö ei ole riittävä ehto faktorimallin identifioitumiselle, joten on syytä esittää jokin tällainen ehto tämän tutkielman empiirisen osion faktorimallien identifioitumiselle. Seuraavat neljä ehtoa liittyvät niin sanottuun kahden indikaattorin sääntöön<sup>4</sup> ja ovat yhdessä riittävä ehto useamman kuin yhden faktorin mallien identifioitumiselle: (1) matriisin  $\mathbf{\Lambda}$  jokaisella rivillä on yksi ja vain yksi nollasta poikkeava arvo, (2) matriisin  $\mathbf{\Lambda}$  jokaisella sarakkeella on vähintään kaksi nollasta poikkeavaa arvoa, (3) matriisin  $\mathbf{\Phi}$  jokaisella rivillä on vähintään yksi diagonaalien ulkopuolinen nollasta poikkeava arvo, ja (4) matriisi  $\mathbf{\Theta}$  on diagonaalinen. [4] Ehdot (1) ja (2) liittyvät faktorilatauksiin ja tarkoittavat yhdessä sitä, että kuhunkin indikaattorin vaikuttaa vain yksi faktori, ja että kullakin faktorilla on vähintään kaksi suoraa vaikutusta joihinkin indikaattoreihin. Ehto (3) puolestaan tarkoittaa sitä, että kukin faktori korreloi vähintään yhden muun faktorin kanssa. Lopuksi ehto (4) vastaa ai-

<sup>3</sup>Näin faktorien mitta määritettiin rajoittamattomassa mallissa.

<sup>4</sup>Kahden indikaattorin sääntö (engl. two-indicator rule) esitetään kirjan [4] luvussa 7 sivuilla 244-246.

emmin mainittua rajoittamattomien faktorimallien oletusta  $\text{Cov}(\delta_i, \delta_j) = 0$ , kaikilla  $i \neq j$ ,  $i, j \in \{1, \dots, q\}$ . Se siis tarkoittaa, että virhetermit eivät korreloi keskenään.

### 2.2.3 Estimointi

Kuten aikaisemmin mainittiin, faktorianalyysissä mallinnetaan kovariansseja, ja tavoitteena on minimoida otoskovarianssien ja mallin ennustamien kovarianssien välinen ero. Oleellisesti tärkeä osa estimointia faktorianalyysissä on siis havaittujen muuttujien kovarianssimatriisin  $\Sigma$  yhteys mallin rakenteellisiin parametreihin  $\eta$ . Jos faktorimalli on oikea ja populaatioparametrit ovat tunnettuja, niin pätee  $\Sigma = \Sigma(\eta) = \Lambda\Phi\Lambda' + \Theta$ . Käytännössä populaatioparametrit eivät kuitenkaan ole tunnettuja, joten tuntemattomat parametrit estimoidaan  $\mathbf{x}$ :n otoskovarianssimatriisin  $\mathbf{S}$  avulla, joka on muotoa

$$\mathbf{S} = \begin{bmatrix} \text{Var}(x_1) & & & & & \\ \text{Cov}(x_2, x_1) & \text{Var}(x_2) & & & & \\ \vdots & \vdots & \ddots & & & \\ \text{Cov}(x_{q-1}, x_1) & \text{Cov}(x_{q-1}, x_2) & \cdots & \text{Var}(x_{q-1}) & & \\ \text{Cov}(x_q, x_1) & \text{Cov}(x_q, x_2) & \cdots & \text{Cov}(x_q, x_{q-1}) & \text{Var}(x_q) & \end{bmatrix}.$$

Matriisin  $\mathbf{S}$  varianssit ja kovarianssit on siis laskettu havaitusta aineistosta. Faktorimallin estimoinnissa matriisit  $\mathbf{S}$  ja  $\Sigma(\eta)$  sijoitetaan johonkin soveltuvaan sovitusfunktioon  $F$ , joka minimoidaan ja saadaan estimaatti  $\Sigma(\hat{\eta})$ . [4] Tämän tutkielman empiirisessä osiossa faktorimallien estimoinnissa käytetään suurimman uskottavuuden funktiota, joten se esitetään faktorimallien kontekstissa seuraavaksi.

Oletetaan, että havaittujen muuttujien vektori  $\mathbf{x} = [x_1, \dots, x_q]'$  on multinormaalisesti jakautunut odotusarvovektorilla  $\mathbf{0}_q = [0, \dots, 0]'$  ja kovarianssimatriisilla  $\Sigma$  ja siitä saadaan  $n$  riippumatonta havaintoa  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . Havaintojen yhteistiheysfunktio on muotoa

$$f(\mathbf{x}_1, \dots, \mathbf{x}_n; \Sigma) = (2\pi)^{-nq/2} |\Sigma|^{-n/2} \exp \left[ -\frac{1}{2} \sum_{i=1}^n \mathbf{x}_i' \Sigma^{-1} \mathbf{x}_i \right].$$

Kovarianssimatriisin  $\Sigma$  paikalle voidaan sijoittaa  $\Sigma(\eta)$ , jolloin rakenteellisten parametrien  $\eta$  log-uskottavuusfunktio on

$$l(\eta) = -\frac{nq}{2} \log 2\pi - \frac{n}{2} \log |\Sigma(\eta)| - \frac{1}{2} \sum_{i=1}^n \mathbf{x}_i' \Sigma^{-1}(\eta) \mathbf{x}_i.$$

Sen viimeinen termi voidaan uudelleenkirjoittaa muotoon

$$\begin{aligned} -\frac{1}{2} \sum_{i=1}^n \mathbf{x}_i' \Sigma^{-1}(\eta) \mathbf{x}_i &= -\frac{1}{2} \sum_{i=1}^n \text{tr} [\mathbf{x}_i' \Sigma^{-1}(\eta) \mathbf{x}_i] \\ &= -\frac{n}{2} \sum_{i=1}^n \text{tr} [n^{-1} \mathbf{x}_i \mathbf{x}_i' \Sigma^{-1}(\eta)] \\ &= -\frac{n}{2} \text{tr} [\mathbf{S}^* \Sigma^{-1}(\eta)], \end{aligned} \tag{8}$$

jossa  $\mathbf{S}^* = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i'$  on suurimman uskottavuuden estimaattori havaittujen muut-  
tujen kovarianssimatriisille  $\Sigma$ . Kun uudelleenkirjoitettu termi (8) sijoitetaan log-  
uskottavuusfunktioon  $l(\eta)$ , saadaan

$$\begin{aligned} l(\eta) &= -\frac{nq}{2} \log 2\pi - \frac{n}{2} \log |\Sigma(\eta)| - \frac{n}{2} \text{tr} [\mathbf{S}^* \Sigma^{-1}(\eta)] \\ &= \text{vakio} - \frac{n}{2} \{ \log |\Sigma(\eta)| + \text{tr} [\mathbf{S}^* \Sigma^{-1}(\eta)] \}, \end{aligned} \quad (9)$$

joka maksimoimalla saadaan suurimman uskottavuuden estimaatti mallin rakenteel-  
lisille parametreille  $\eta$ . [4]

Oletetaan sitten, että otoskovarianssimatriisi  $\mathbf{S}$  ja matriisi  $\Sigma(\eta)$  ovat positiivisesti  
definiittejä, jolloin ne ovat kääntyviä. Tällöin lineaaristen faktorimallien suurimman  
uskottavuuden estimoinnissa sovituskuntiona voidaan käyttää myös funktiota

$$F = \log |\Sigma(\eta)| + \text{tr} [\mathbf{S} \Sigma^{-1}(\eta)] - \log |\mathbf{S}| - q, \quad (10)$$

joka eroaa yhtälön (9) log-uskottavuusfunktioista  $l(\eta)$  muutamalla tavalla. Ensini-  
näkin yhtälössä (9) vakio-termi on  $-\frac{nq}{2} \log 2\pi$ , mutta yhtälössä (10) vakio-termi on  
 $-\log |\mathbf{S}| - q$ . Vakio-termit eivät kuitenkaan vaikuta rakenteellisten parametrien es-  
timaattorin  $\hat{\eta}$  valintaan. Toiseksi yhtälössä (10) ei ole kerrointa  $-n/2$  termien  
 $\log |\Sigma(\eta)|$  ja  $\text{tr} [\mathbf{S} \Sigma^{-1}(\eta)]$  edessä, mikä tarkoittaa sitä, että sovituskuntio  $F$  mi-  
nimoidaan maksimoinnin sijasta, kun faktorimallia estimoidaan. Viimeiseksi yhtä-  
lössä (10) käytetään kovarianssimatriisin harhaisen suurimman uskottavuuden es-  
timaattorin  $\mathbf{S}^*$  sijasta harhatonta otoskovarianssimatriisia  $\mathbf{S}$ . Kun otoskoko  $n$  on  
suuri, niin  $\mathbf{S}^*$  ja  $\mathbf{S}$  ovat asympotoottisesti yhtäsuuria, sillä  $\mathbf{S}^* = \frac{n-1}{n} \mathbf{S}$ . Näistä eroista  
huolimatta funktiot  $l(\eta)$  ja  $F$  johtavat likimain samoihin estimaattoreihin  $\hat{\eta}$ . [4]

Koska  $F$  ja  $l(\eta)$  johtavat likimain samoihin estimaattoreihin  $\hat{\eta}$ , sillä ei ole vä-  
liä, kumpaa käyttää sovituskuntiona. Usein kuitenkin käytetään juuri funktiota  
 $F$ , koska sillä on seuraavat kätevät ominaisuudet: (1)  $F(\mathbf{S}, \Sigma(\eta))$  on skalaari, (2)  
 $F(\mathbf{S}, \Sigma(\eta)) \geq 0$ , (3)  $F(\mathbf{S}, \Sigma(\eta)) = 0$ , jos ja vain jos  $\Sigma(\eta) = \mathbf{S}$ , ja (4)  $F(\mathbf{S}, \Sigma(\eta))$  on  
jatkuva pisteissä  $\mathbf{S}$  ja  $\Sigma(\eta)$ . Kun funktiota  $F$  käytetään faktorimallin sovittamisessa  
aineistoon, eksplisiittisiä ratkaisuja ei aina löydetä, sillä  $F$  on usein monimutkai-  
nen epälineaarinen funktio mallin rakenteellisista parametreista  $\eta$ . Tällöin joudu-  
taan käyttämään iteratiivisia numeerisia menetelmiä  $F$ :n minimoivan parametries-  
timaatin  $\hat{\eta}$  ratkaisemiseksi. Kirjan [4] liitteessä 4C kuvaillaan ja havainnollistetaan  
Newtonin-Raphsonin menetelmää numeerisen ratkaisun löytämiseksi faktorimallin  
estimoinnissa.

Kun  $F$  on minimoitu ja on saatu joko eksplisiittinen tai numeerinen paramet-  
riestimaatti  $\hat{\eta}$ , niin sijoittamalla se funktioon  $F$  saadaan

$$F(\mathbf{S}, \Sigma(\hat{\eta})) = \log |\Sigma(\hat{\eta})| + \text{tr} [\mathbf{S} \Sigma^{-1}(\hat{\eta})] - \log |\mathbf{S}| - q.$$

Selkeästi, jos  $F(\mathbf{S}, \Sigma(\hat{\eta})) = 0$ , niin faktorimalli sopii aineistoon täydellisesti, sillä  
kun  $\Sigma(\hat{\eta}) = \mathbf{S}$ , niin

$$F(\mathbf{S}, \Sigma(\hat{\eta})) = \log |\mathbf{S}| + \text{tr}(\mathbf{I}_q) - \log |\mathbf{S}| - q = q - q = 0.$$

Lisäksi, koska  $\hat{\eta}$  on luonteeltaan suurimman uskottavuuden estimaattori, sillä on  
useita tärkeitä asympotoottisia ominaisuuksia: (1)  $\hat{\eta}$  on asympotoottisesti harhaton

estimaattori, (2) se on tarkentuva, (3) se on asymptoottisesti tehokas, eli millään muulla tarkentuvalla estimaattorilla ei ole pienempää asymptoottista varianssia, ja (4) se on asymptoottisesti normaalisesti jakautunut. Muita sovitusfunktion  $F$  hyviä ominaisuuksia – muutamia poikkeuksia lukuunottamatta – ovat sen 'skaalainvarianttisuus' (engl. scale invariance) ja 'skaalavapaus' (engl. scale freeness), joiden seurauksena sovitusfunktion arvot eivät riipu siitä, analysoidaanko korrelaatio- tai kovarianssimatriiseja, eivätkä siitä käytetäänkö alkuperäistä vai skaalattua aineistoa. [4], [39]

## 2.2.4 Yksinkertainen esimerkki

Tutkitaan esimerkkitilannetta konfirmatiivisesta faktorianalyysistä, jossa indikaattorien määrä on  $q = 6$  ja faktorien määrä  $k = 2$ . Lisäksi oletetaan, että faktoria  $\xi_1$  mittaavat indikaattorit  $x_1, x_2$  ja  $x_3$ , ja faktoria  $\xi_2$  indikaattorit  $x_4, x_5$  ja  $x_6$ . Tällöin asetetaan  $\lambda_{12} = \lambda_{22} = \lambda_{32} = \lambda_{41} = \lambda_{51} = \lambda_{61} = 0$  ja yhtälön (5) yleinen rajoittamaton lineaarinen faktorimalli saadaan rajoitetun faktorimallin muotoon

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{bmatrix},$$

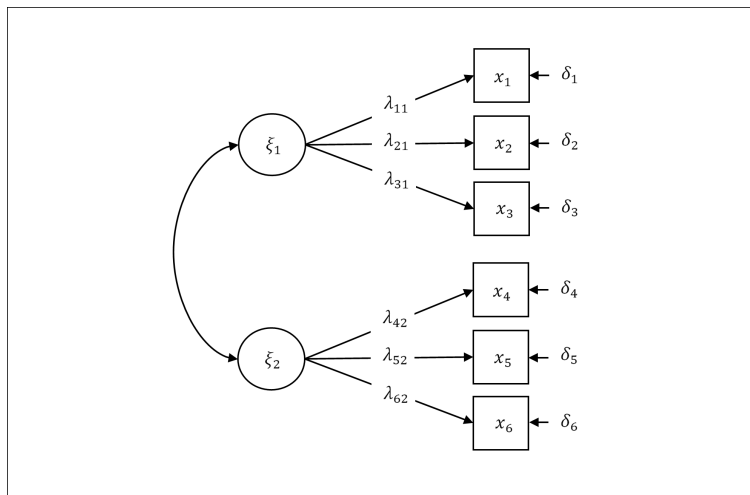
jossa  $E(\delta_j) = 0$ ,  $\text{Cov}(\xi_i, \delta_j) = 0$ ,  $\forall i \in \{1, 2\}$ ,  $\forall j \in \{1, \dots, 6\}$ . Oletetaan lisäksi, että  $\text{Cov}(\delta_i, \delta_j) = 0$ , kaikilla  $i \neq j$ ,  $i, j \in \{1, \dots, 6\}$ .

Havaittujen muuttujien  $x_1, \dots, x_6$  populaatio-kovarianssimatriisi  $\Sigma$  määritellään kuten yhtälössä (6), eli  $\Sigma = \Sigma(\eta) = \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}' + \mathbf{\Theta}$ . Matriisien  $\Sigma$  ja  $\Sigma(\eta)$  tarkemmat määrittelyt näkyvät liitteessä A, jonka perusteella esimerkiksi  $\text{Var}(x_1) = \lambda_{11}^2\phi_{11} + \theta_1$ . Tämä osoittaa, että  $x_1$ :n varianssi vaihtelee, jos  $\lambda_{11}$ ,  $\phi_{11}$  tai  $\theta_1$  saavat eri arvoja. Matriisin  $\Sigma$  muutkin alkiot voitaisiin kirjoittaa samaan tapaan matriisiin  $\Sigma(\eta)$  vastaavien alkioiden avulla. Matriisissa  $\Sigma$  on tässä tapauksessa  $6(6 + 1)/2 = 21$  ei-redundanttia alkioita, jotka voidaan estimoida laskemalla otoskovarianssimatriisi  $\mathbf{S}$ . Täten  $t$ -säännön mukaan estimoitavien parametrien määrän pitää olla  $t \leq 21$ . Tämä toteutuu, sillä nyt matriisissa  $\mathbf{\Lambda}$  on yhteensä 6 parametria, matriisissa  $\mathbf{\Phi}$   $2(2 + 1)/2 = 3$  ei-redundanttia parametria ja matriisissa  $\mathbf{\Theta}$  6 parametria, jolloin matriisissa  $\Sigma(\eta)$  on yhteensä  $t = 6 + 3 + 6 = 15 \leq 21$  tuntematonta parametria. Myös luvussa 2.2.2 esitetty riittävä ehto mallin identifioitumiselle toteutuu.

Kuvassa 1 esitetään esimerkkimallin polkukuvio<sup>5</sup>, jossa faktorit  $\xi_1$  ja  $\xi_2$  näkyvät ympyröiden sisällä ja havaitut muuttujat  $x_1, \dots, x_6$  suorakulmioiden sisällä. Faktorilataukset  $\mathbf{\Lambda}$  näkyvät nuolina faktoreista havaittuihin muuttujiin ja mallin virheet nuolina mallin ulkopuolelta indikaattoreihin. Lisäksi kuvassa näkyy faktorien välinen kovarianssi kaksipäisenä nuolena faktorien välillä. Mikäli olisi niin, että  $\text{Cov}(\delta_i, \delta_j) \neq 0$ , joillakin  $i \neq j$ ,  $i, j \in \{1, \dots, 6\}$ , niin kyseiset nolasta poikkeavat kovarianssit esitettäisiin kaksipäisinä nuolina virheiden välillä.

<sup>5</sup>Polkukuvio (engl. path diagram) on kuvallinen esitys samanaikaisten yhtälöiden ryhmästä (engl. system of simultaneous equations) [4].





Kuva 1: Esimerkki kahden faktorin ja kuuden havaitun muuttujan mallista.

## 2.3 Ryhmiteltyt rajoitetut faktorimallit

### 2.3.1 Mallin määrittely

Rajoitetun lineaarisen faktorimallin määritelmää on vielä syytä laajentaa tilanteeseen, jossa on useita alapopulaatioita, sillä mittausinvarianssin<sup>6</sup> testaaminen faktorimalleilla vaatii tämän. Tällöin kyseessä on konfirmatiivinen moniryhmäfaktorianalyysi (engl. multi-group confirmatory factor analysis, MG-CFA). [49] Meredith [29] esitti, että keskistettyä mallia käyttäen voidaan testata vain niin sanottua heikkoa invarianssia ja vahvan invarianssin testaaminen vaatii ei-keskistetyn mallin käyttämistä faktorianalyyseissä. Otetaan tämä huomioon ja määritellään siis ei-keskistetty ja ryhmitelty rajoitettu lineaarinen faktorimalli yhtälöllä

$$\mathbf{x} = \tau_g + \mathbf{\Lambda}_g \xi + \delta, \quad g = 1, \dots, G,$$

jossa  $\tau_g$  on latenteista vakiokertoimista  $\tau_{1g}, \dots, \tau_{qg}$  koostuva  $q \times 1$ -vektori alapopulaatiolle  $\Pi_g$ ,  $\mathbf{\Lambda}_g$  on faktorilatauksista koostuva  $q \times k$ -matriisi alapopulaatiolle  $\Pi_g$  ja vektorit  $\mathbf{x}$ ,  $\xi$  sekä  $\delta^7$  on määritelty kuten yhtälössä (1) [30].

Määritellään vektoreille  $\xi$  ja  $\delta$  populaatiokohtaiset momentit

$$E_g[\xi] = \kappa_g, \quad \text{Cov}_g(\xi) = \mathbf{\Phi}_g, \quad E_g[\delta] = \mathbf{0}, \quad \text{Cov}_g(\delta) = \mathbf{\Theta}_g, \quad g = 1, \dots, G. \quad (11)$$

Lisäksi oletetaan, että  $\text{Cov}_g(\xi, \delta) = \mathbf{0}$ , jolloin havaittujen muuttujien odotusarvo- ja kovarianssirakenne ehdollistettuna latenttien muuttujien suhteen populaatiossa  $\Pi_g$  voidaan esittää momenttien (11) avulla yhtälöillä

$$E_g[\mathbf{x} | \xi] = E_g[\tau_g + \mathbf{\Lambda}_g \xi + \delta | \xi] = \tau_g + \mathbf{\Lambda}_g E_g[\xi | \xi] + E_g[\delta | \xi] = \tau_g + \mathbf{\Lambda}_g \xi, \quad (12)$$

<sup>6</sup>Luku 3 käsittelee mittausinvarianssia.

<sup>7</sup>Noudattaen Millsapin [30] merkintöjä, satunnaismuuttujia  $\mathbf{x}$ ,  $\xi$  ja  $\delta$  ei indeksoida alapopulaatioittain. Kaikki satunnaismuuttujat on kuitenkin määritelty kaikille alapopulaatioille ja ne saattavat noudattaa jakaumia, jotka riippuvat näistä populaatioista. Kyseiset vektorit eivät siis välttämättä ole yhtä suuria kaikilla  $g = 1, \dots, G$ .

ja

$$\begin{aligned}
\text{Cov}_g(\mathbf{x} \mid \xi) &= \mathbb{E}_g[\mathbf{x}\mathbf{x}' \mid \xi] - \mathbb{E}_g[\mathbf{x} \mid \xi]\mathbb{E}_g[\mathbf{x} \mid \xi]' \\
&= \mathbb{E}_g[(\tau_g + \mathbf{\Lambda}_g\xi + \delta)(\tau_g + \mathbf{\Lambda}_g\xi + \delta)' \mid \xi] - (\tau_g + \mathbf{\Lambda}_g\xi)(\tau_g + \mathbf{\Lambda}_g\xi)' \\
&= \mathbb{E}_g[(\tau_g + \mathbf{\Lambda}_g\xi)(\tau_g + \mathbf{\Lambda}_g\xi)' + (\tau_g + \mathbf{\Lambda}_g\xi)\delta' + \delta(\tau_g + \mathbf{\Lambda}_g\xi)' + \delta\delta' \mid \xi] \\
&\quad - (\tau_g + \mathbf{\Lambda}_g\xi)(\tau_g + \mathbf{\Lambda}_g\xi)' \\
&= \mathbb{E}_g[\delta\delta' \mid \xi] \\
&= \mathbf{\Theta}_g.
\end{aligned} \tag{13}$$

Yhtälöjen (12) ja (13) ehdolliset odotusarvo- ja kovarianssirakenteet ovat hyödyllisiä etenkin mittausinvarianssin määrittelyssä luvussa 3. Kun odotusarvo- ja kovarianssirakennetta ei ehdollisteta latenttien muuttujien suhteen populaatiossa  $\Pi_g$ , saadaan

$$\mu_g = \mathbb{E}_g[\mathbf{x}] = \mathbb{E}_g[\tau_g + \mathbf{\Lambda}_g\xi + \delta] = \tau_g + \mathbf{\Lambda}_g\mathbb{E}_g[\xi] + \mathbb{E}_g[\delta] = \tau_g + \mathbf{\Lambda}_g\kappa_g, \tag{14}$$

ja

$$\begin{aligned}
\mathbf{\Sigma}_g &= \text{Cov}_g(\mathbf{x}) = \mathbb{E}_g[\mathbf{x}\mathbf{x}'] - \mathbb{E}_g[\mathbf{x}]\mathbb{E}_g[\mathbf{x}]' \\
&= \mathbb{E}_g[(\tau_g + \mathbf{\Lambda}_g\xi + \delta)(\tau_g + \mathbf{\Lambda}_g\xi + \delta)'] - (\tau_g + \mathbf{\Lambda}_g\mathbb{E}_g[\xi])(\tau_g + \mathbf{\Lambda}_g\mathbb{E}_g[\xi])' \\
&= \mathbb{E}_g[(\tau_g + \mathbf{\Lambda}_g\xi)(\tau_g + \mathbf{\Lambda}_g\xi)' + (\tau_g + \mathbf{\Lambda}_g\xi)\delta' + \delta(\tau_g + \mathbf{\Lambda}_g\xi)' + \delta\delta'] \\
&\quad - (\tau_g + \mathbf{\Lambda}_g\mathbb{E}_g[\xi])(\tau_g + \mathbf{\Lambda}_g\mathbb{E}_g[\xi])' \\
&= \tau_g\tau_g' + \tau_g\mathbb{E}_g[\xi]'\mathbf{\Lambda}_g' + \mathbf{\Lambda}_g\mathbb{E}_g[\xi]\tau_g' + \mathbf{\Lambda}_g\mathbb{E}_g[\xi\xi']\mathbf{\Lambda}_g' + \mathbb{E}_g[\delta\delta'] \\
&\quad - \tau_g\tau_g' - \tau_g\mathbb{E}_g[\xi]'\mathbf{\Lambda}_g' - \mathbf{\Lambda}_g\mathbb{E}_g[\xi]\tau_g' - \mathbf{\Lambda}_g\mathbb{E}_g[\xi]\mathbb{E}_g[\xi]'\mathbf{\Lambda}_g' \\
&= \mathbf{\Lambda}_g(\mathbb{E}_g[\xi\xi'] - \mathbb{E}_g[\xi]\mathbb{E}_g[\xi]')\mathbf{\Lambda}_g' + \mathbb{E}_g[\delta\delta'] \\
&= \mathbf{\Lambda}_g\mathbf{\Phi}_g\mathbf{\Lambda}_g' + \mathbf{\Theta}_g,
\end{aligned} \tag{15}$$

joka selkeästi vastaa ei-ryhmitellyn rajoitetun faktorimallin kovarianssirakennetta koskevaa hypoteesia lukuunottamatta indeksointia alapopulaatioiden suhteen. [30]

### 2.3.2 Parametrien identifiointi

On selvää, että ryhmiteltyjä ja ei-keskistettyjä rajoitettuja lineaarisia faktorimalleja koskee sama identifiointiongelma kuin tavallisia rajoitettuja faktorimallejakin. Kun kyseessä on erityisesti ryhmitelty faktorimalli, mallin identifioituminen on monimutkaisempaa, sillä luvussa 2.2.2 esitelyjen parametrirajoitteiden lisäksi myös populaatioita koskevia rajoitteita voidaan asettaa malliin. Tästä seuraa parametrirajoitteiden 'kaksoisrooli' ryhmiteltyjen mallien tapauksessa. Toisaalta niiden ensisijainen tarkoitus on mahdollistaa vapaiden parametrien yksikäsitteinen estimointi, mutta toisaalta ne ovat myös niin sanottuja invarianssirajoitteita faktorimallien mittausinvarianssitestauksissa. [30]

Luvussa 2.2.2 esitelty kahden indikaattorin identifioituvuusääntö pätee myös, kun parametrirajoitteita asetetaan ryhmitellyn ja ei-keskistetyn faktorimallin kovarianssirakenteeseen, jolloin lisäksi vaaditaan, että säännön neljä ehtoa pätevät kussakin  $G$ :ssä eri alapopulaatiossa, ja yleensä halutaan löytää faktorilatausmatriisiin

$\Lambda_g$  asetetut 'nollalataukset' identtisesti kaikille populaatioille  $\Pi_1, \dots, \Pi_G$ . Tällöin puhutaan konfiguraalisesta invarianssirajoitteesta, joka on osa faktorimallin mittaussinvarianssitestauksia. [30]

Kovarianssirakenteeseen asetettujen rajoitteiden lisäksi erityisesti ei-keskistetyn faktorimallin identifioituminen vaatii parametrirajoitteiden asettamista mallin odotusarvorakenteeseen, mikä nähdään, kun määritellään uudet parametrit  $\kappa_g^* = \kappa_g + \mathbf{c}$  ja  $\tau_g^* = \tau_g - \Lambda_g \mathbf{c}$ . Tällöin

$$E_g[\mathbf{x}] = \tau_g + \Lambda_g \kappa_g = \tau_g^* + \Lambda_g \mathbf{c} + \Lambda_g (\kappa_g^* - \mathbf{c}) = \tau_g^* + \Lambda_g \kappa_g^*,$$

ja nähdään, että vakiotermparametri  $\tau_g$  ja latentti odotusarvoparametri  $\kappa_g$  eivät ole identifioituvia, vaikka  $\Lambda_g$  olisi tunnettu. Odotusarvorakenne vaatii vähintään  $k$  rajoitetta kussakin populaatiossa  $\Pi_g$ , jotta se olisi identifioituva. Esimerkiksi (1)  $\kappa_g = \mathbf{0}$  kaikilla  $g$ , (2)  $\tau_{jg} = 0$  joillakin  $k$ :lla  $j \in \{1, \dots, q\}$  ja kaikilla  $g$ , ja (3)  $\tau_g = \tau$  kaikilla  $g$  ja  $\kappa_g = 0$  jollakin valitulla  $g$ , ovat riittäviä vaihtoehtoisia tapoja saada odotusarvorakenne identifioituvaksi. Selkeästi kolmannella tavalla asetetaan enemmän kuin  $k$  rajoitetta, mutta se on erityisesti hyödyllinen, kun halutaan tutkia faktorimallin osittaista invarianssia<sup>8</sup>. [30]

### 2.3.3 Estimointi

Ryhmitellyn rajoitetun lineaarisen faktorimallin estimointi tapahtuu laajentamalla ei-ryhmiteltyä tapausta. Oletetaan siis ensin, että saadaan  $n$  havaintoa, jotka jakautuvat ominaisuuksiltaan erilaisiin alapopulaatioihin  $\Pi_1, \dots, \Pi_G$  otosko'illa  $n_1, \dots, n_G$ . Tällöin  $\mathbf{x}_{ig}$ ,  $i = 1, \dots, n_g$ ,  $g = 1, \dots, G$ , on havaintujen muuttujien  $q \times 1$ -vektori havainnolle  $i$  populaatiossa  $\Pi_g$ . Vektorien  $\mathbf{x}_{ig}$  oletetaan olevan riippumattomasti multinormaalisesti jakautuneita sekä populaatioiden sisällä että niiden välillä. Lisäksi koska nyt on kyseessä ei-keskistetty faktorimalli, niin mallin odotusarvorakenne täytyy ottaa huomioon. Oletetaan siis, että vektorilla  $\mathbf{x}_{ig}$  on  $N_q(\mu_g, \Sigma_g)$ -jakauma populaatiossa  $\Pi_g$ . Jakaumamerkinnässä alaindeksi  $q$  viittaa dimensioiden lukumäärään ja odotusarvovektori  $\mu_g = E_g[\mathbf{x}]$  sekä kovarianssimatriisi  $\Sigma_g = \text{Cov}_g(\mathbf{x})$  on määritelty kuten yhtälöissä (14) ja (15). Olkoot lisäksi

$$\bar{\mathbf{x}}_g = \frac{1}{n_g} \sum_{i=1}^{n_g} \mathbf{x}_{ig},$$

ja

$$\mathbf{S}_g = \frac{1}{n_g} \sum_{i=1}^{n_g} (\mathbf{x}_{ig} - \bar{\mathbf{x}}_g)(\mathbf{x}_{ig} - \bar{\mathbf{x}}_g)'$$

estimaattorit odotusarvolle  $\mu_g$  ja kovarianssimatriisille  $\Sigma_g$ . Yhtälön (10) suurimman uskottavuuden sovituskäytäntöä voidaan muokata ryhmitellylle ja ei-keskistetylle rajoitetulle lineaariselle faktorimallille sopivaksi, jolloin saadaan

$$F = \sum_{g=1}^G \frac{n_g}{n} F_g, \quad (16)$$

<sup>8</sup>Osittaisesta invarianssista puhutaan lyhyesti luvussa 3.2.3.

jossa

$$F_g = (\bar{\mathbf{x}}_g - \mu_g)' \boldsymbol{\Sigma}_g^{-1} (\bar{\mathbf{x}}_g - \mu_g) + \log \frac{|\boldsymbol{\Sigma}_g|}{|\mathbf{S}_g|} + \text{tr} [\mathbf{S}_g \boldsymbol{\Sigma}_g^{-1}] - q.$$

Kun oletetaan, että estimoitavaan faktorimalliin on asetettu tarvittavat parametrirajoitteet, joilla on saatu haluttu identifioituva malli, niin minimoimalla sovituskunktio (16) eksplisiittisesti tai numeerisesti saadaan estimaatti sekä odotusarvo- että kovarianssirakenteen vapaille parametreille. Mikäli estimoitavassa faktorimallissa ei ole invarianssirajoitteita tai muita populaatioiden välisiä parametrirajoitteita, niin sovituskunktio voidaan minimoida erikseen kussakin populaatiossa. Jos tällaisia parametrirajoitteita on, niin ne täytyy oleellisesti ottaa huomioon. [30] Invarianssirajoitteista puhutaan lisää luvussa 3.

## 2.4 Faktorimallien sopivuuden testaaminen

Kun faktorimalli on sovitettu aineistoon ja on saatu jokin eksplisiittinen tai numeerinen estimaatti sen rakenteellisille parametreille, mallin sopivuutta aineistoon on syytä tutkia tarkemmin. Mallin sopivuuden hyvyttä voidaan arvioida useilla erilaisilla sopivuusindekseillä, jotka määrittävät havaitun aineiston ja mallin implikoiman teoreettisen aineiston välisen suhteen. Faktorimallien sopivuutta aineistoon voi arvioida kahdella tavalla; lokaalisti ja globaalisti. Mallin lokaalia sopivuutta testataan esimerkiksi tarkistamalla, että yksittäisten parametrien estimoidut arvot ovat matemaattisten rajojen sisällä ja ne eroavat tilastollisesti merkitsevästi nolasta. [3] Lokaalin sopivuuden tutkimisessa on erityisen tärkeää tarkistaa, että estimoidut parametriestimaatit ovat 'suuruudeltaan merkityksellisiä', ja esimerkiksi faktorilatausestimaattien suuruus kertoo, kuinka vahvasti indikaattorit mittaavat niihin latautuvia faktoreita. Faktorilatausten riittävä suuruus riippuu tilanteesta. [6] Globaalia sopivuutta arvioidaan tutkimalla voidaanko useilla mittauksilla nykyisiä parametrirajoitteita käyttämällä saada otoskovarianssimatriisia hyvin jäljittelevä matriisi, jonka arvot eivät poikkea siitä liian paljon. Tässä olennaisena komponenttina toimii aiemmin määritellyt sovituskunktiot (10) ja (16). [3]

Globaalit sopivuusindeksit voidaan jakaa absoluuttisiin (engl. absolute), inkrementaalisiin (engl. incremental) ja säästäväisiin (engl. parsimonious) sopivuusindekseihin. Absoluuttisten sopivuusindeksien avulla voidaan määrittää, kuinka hyvin sovitettu malli sopii aineistoon verraten mallittomaan tilanteeseen. Khiin neliön testi ( $\chi^2$ -testi) ja RMSEA (Root Mean Square Error of Approximation) -indeksi ovat absoluuttisia sopivuusindeksejä, joita käytetään tämän tutkielman empiirisessä osiossa. Inkrementaaliset sopivuusindeksit vertaavat sovitetun mallin  $\chi^2$ -testisuureen arvoa niin sanotun 'lähtökohtamallin' (engl. baseline model)  $\chi^2$ -testisuureen arvoon. Tässä tutkielmassa käytetään inkrementaalisia Tuckerin–Lewisin indeksiä (engl. Tucker–Lewis Index, TLI) ja vertailevaa sopivuusindeksiä (engl. Comparative Fit Index, CFI) mallien arvioinnissa. [15] Säästäväisiä sopivuusindeksejä ei tässä tutkielmassa esitetä eikä käytetä. Määritellään tutkielman empiirisessä osiossa käytettävät sopivuusindeksit ensin tavallisille keskistetyille ja ei-ryhmitellyille faktorimalleille luvuissa 2.4.1 ja 2.4.2, minkä jälkeen luvussa 2.4.4 niiden määritelmät laajennetaan ei-keskistettyjen ja ryhmiteltyjen faktorimallien tilanteeseen.

### 2.4.1 Absoluuttiset sopivuusindeksit

Suurimman uskottavuuden estimaattorin käyttäminen faktorimallin sovittamisessa mahdollistaa mallin 'kokonaissopivuuden' (engl. overall fit) arvioimisen  $\chi^2$ -testillä. Mikäli jakaumaoletukset pitävät paikkansa,  $\chi^2$ -testillä voidaan testata pitääkö faktorianalyysin fundamentaalinen hypoteesi paikkansa; se siis testaa nollahypoteesia, että populaatio-kovarianssimatriisi  $\Sigma$  on yhtä suuri kuin mallin rakenteellisten parametrien kovarianssimatriisi  $\Sigma(\eta)$ , eli  $\Sigma - \Sigma(\eta) = \mathbf{0}$ . Vastahypoteesi on tällöin luonnollisesti  $\Sigma - \Sigma(\eta) \neq \mathbf{0}$ . Osoittautuu, että jos nollahypoteesi on voimassa, niin  $(n-1)F(\mathbf{S}, \Sigma(\hat{\eta}))$  noudattaa asympotoottisesti  $\chi^2$ -jakaumaa  $q(q+1)/2 - t$  vapausasteella, jossa  $q(q+1)/2$  on ei-redundanttien alkioiden määrä otoskovarianssimatriisissa  $\mathbf{S}$  ja  $t$  on vapaiden parametrien määrä vektorissa  $\eta$ . Pienet  $(n-1)F(\mathbf{S}, \Sigma(\hat{\eta}))$ -arvot suhteessa vapausasteisiin viittaavat siihen, että  $\Sigma$  ja  $\Sigma(\eta)$  eivät eroa toisistaan merkittävästi. Toisin sanoen, jos  $P((n-1)F(\mathbf{S}, \Sigma(\hat{\eta})) \leq \chi^2(v)) \geq \alpha$ , jossa  $v = q(q+1)/2 - t$  ja  $\alpha$  on ennalta määritelty merkitsevyystaso (esimerkiksi  $\alpha = 0,05$  tai  $\alpha = 0,001$ ), niin nollahypoteesia ei voida hylätä. [39]

Perustellaan, miksi  $(n-1)F(\mathbf{S}, \Sigma(\hat{\eta}))$  noudattaa asympotoottisesti  $\chi^2(v)$ -jakaumaa, jossa  $v = q(q+1)/2 - t$ . Kun nollahypoteesi  $\Sigma = \Sigma(\eta)$  on voimassa, niin tutkittavan mallin log-uskottavuusfunktion maksimoi suurimman uskottavuuden estimaatti  $\hat{\eta}$ , jolloin

$$l(\Sigma(\hat{\eta})) = l_0 \propto -\frac{n-1}{2} \{ \log |\Sigma(\hat{\eta})| + \text{tr} [\Sigma(\hat{\eta})^{-1}\mathbf{S}] \}.$$

Kun vastahypoteesiksi valitaan  $\Sigma = \mathbf{S}$ , niin sitä vastaava log-uskottavuusfunktion maksimi on pisteessä  $\mathbf{S}$ , jolloin

$$l(\mathbf{S}) = l_1 \propto -\frac{n-1}{2} \{ \log |\mathbf{S}| + \text{tr}(\mathbf{S}^{-1}\mathbf{S}) \} = -\frac{n-1}{2} \{ \log |\mathbf{S}| + q \}.$$

Funktioiden  $l_0$  ja  $l_1$  avulla saadaan muodostettua uskottavuusosamäärän testisuure

$$\begin{aligned} -2(l_0 - l_1) &= (n-1) \{ \log |\Sigma(\hat{\eta})| + \text{tr} [\Sigma(\hat{\eta})^{-1}\mathbf{S}] \} - (n-1) \{ \log |\mathbf{S}| + q \} \\ &= (n-1) \{ \log |\Sigma(\hat{\eta})| + \text{tr} [\Sigma(\hat{\eta})^{-1}\mathbf{S}] - \log |\mathbf{S}| - q \} \\ &= (n-1)F(\mathbf{S}, \Sigma(\hat{\eta})), \end{aligned}$$

joka noudattaa asympotoottisesti  $\chi^2$ -jakaumaa  $q(q+1)/2 - t$  vapausasteella, kun nollahypoteesi on voimassa. [4]

Khiin neliön testillä on heikkouksia, joista yksi liittyy otoskoon [3], [4], [20], [39]. Toisaalta testi olettaa, että otoskoko on riittävän suuri, mutta toisaalta suurilla otosko'oilla hyvilläkkin malleilla testi on usein tilastollisesti merkitsevä. Toinen  $\chi^2$ -testin heikkous on testisuureen arvon riippuvuus mallin kompleksisuudesta. Testisuureen arvo laskee, kun malliin lisätään parametreja, sillä tällöin vapausasteiden määrä laskee. Esimerkiksi saturoidussa mallissa on estimoitavia parametreja yhtä paljon kuin matriisissa  $\mathbf{S}$  on ei-redundantteja alkiota ja siitä laskettu  $\chi^2$ -testisuureen arvo on 0. [39] Khiin neliön testin heikkouksien takia on kehitetty monia vaihtoehtoisia sopivuusindeksejä, joista RMSEA-indeksi esitetään absoluuttisena sopivuusindeksinä seuraavaksi.

Steiger [44] määritteli RMSEA-indeksin populaatioarvon kaavalla

$$\epsilon = \sqrt{\frac{F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\hat{\eta}))}{v}}, \quad (17)$$

jossa  $F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\hat{\eta}))$  on sovitusfunktion arvo, kun mallin sovittamisessa on käytetty populaatio-kovarianssimatriisia  $\boldsymbol{\Sigma}$  otoskovarianssimatriisin  $\mathbf{S}$  sijasta, ja  $v = q(q + 1)/2 - t$  on vapausasteiden lukumäärä. Alunperin Steiger esitti, että arvoa  $F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\hat{\eta}))$  voisi käyttää sellaisenaan kovarianssirakenteen sopivuuden huonouden luonnollisena mittana, mutta kyseinen mitta ei ollut täydellinen. Lisäämällä jakajaksi vapausasteiden lukumäärän  $v$  mitta kykenee ottamaan huomioon mallin kompleksisuuden vaikutuksen. Lopuksi, kun saadusta osamäärästä otetaan vielä neliöjuuri, tuloksena saadaan indeksi, joka on luonteeltaan samanlainen kuin varianssianalyysissä käytettävä ryhmävaikutuksen suuruutta mittaava RMSSE (Root Mean Square Standardized Effect) -indeksi.

RMSEA-indeksin nimen 'Root Mean Square Error of Approximation' merkitys on nyt ilmeinen. Sana 'root' viittaa selkeästi neliöjuureen. Sanat 'mean' ja 'square' viittaavat siihen, että  $F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\hat{\eta}))$  voidaan Taylorin sarjaa hyödyntäen approksimoida tarkasti kvadraattisessa muodossa  $\text{tr}[\boldsymbol{\Sigma}(\hat{\eta})^{-1}(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\hat{\eta}))]^2/2$  [7]. Tämä tulos perustellaan liitteessä B. Nimen loppuosa 'error of approximation' viittaa approksimointivirheeseen, joka saadaan, kun malli sovitetaan populaatio-kovarianssimatriisiin  $\boldsymbol{\Sigma}$ . RMSEA-indeksi mittaa siis mallin approksimatiivista sopivuutta populaatiossa. [8]

Oletetaan, ettei ole olemassa sellaista parametrivektoria  $\eta$ , jolla nollahypoteesi  $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\eta)$  pitäisi paikkansa. Tällöin, jos sovitusfunktio on oikein määritelty, otoskoko ei ole liian pieni eikä  $F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\hat{\eta}))$  ole liian suuri, niin  $(n - 1)F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\eta}))$  noudattaa approksimatiivisesti epäkeskistä  $\chi^2$ -jakaumaa  $v$ :llä vapausasteella ja epäkeskisyysparametrilla  $(n - 1)F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\hat{\eta}))$ . Tästä seuraa, että muuttujan  $F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\eta}))$  odotusarvo voidaan approksimoida kaavalla  $E[F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\eta}))] \approx F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\hat{\eta})) + (n - 1)^{-1}v$  ja  $F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\hat{\eta}))$ :lle saadaan piste-estimaatti  $F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\eta})) - (n - 1)^{-1}v$ . Tämä voi kuitenkin saada negatiivisia arvoja, joten valitaan estimaatiksi  $\max\{F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\eta})) - (n - 1)^{-1}v, 0\}$ . Kun vielä hyödynnetään kaavaa (17), RMSEA-indeksille saadaan piste-estimaatti

$$\hat{\epsilon} = \sqrt{\max\left\{\frac{F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\eta}))}{v} - \frac{1}{n - 1}, 0\right\}} \quad (18)$$

ja 90 %:n luottamusväli

$$(\hat{\epsilon}_{0,05}, \hat{\epsilon}_{0,95}) = \left( \sqrt{\frac{\hat{\lambda}_{0,05}}{(n - 1)v}}, \sqrt{\frac{\hat{\lambda}_{0,95}}{(n - 1)v}} \right),$$

jossa  $\hat{\epsilon}_c$  on RMSEA-indeksin estimoitu raja  $100 \times c$ -persentiilissä ja  $\hat{\lambda}_c$  on epäkeskisen khiin neliön jakauman epäkeskisyysparametrin estimoitu raja  $100 \times c$ -persentiilissä. Epäkeskisyysparametrin rajat voidaan estimoida numeerisella approksimoinnilla. [8]

Browne ja Cudeck [8] luettelivat heidän käytännön kokemukseensa perustuvia RMSEA-indeksin kynnsarvoja, joita heidän mielestään voisi käyttää tutkittavan

mallin sopivuuden hyvyyden arvioinnissa. Malleilla, joilla  $\hat{\epsilon} \leq 0,05$ ; on vapausasteisiin nähden ”läheinen” sopivuus aineistoon. Sen sijaan, kun  $0,05 < \hat{\epsilon} \leq 0,08$ ; mallin sopivuus on ”hyväksyttävä”. Arvot  $\hat{\epsilon} > 0,10$  viittaavat siihen, että malli ei sovi aineistoon lainkaan. Browne ja Cudeck mainitsivat kuitenkin, että näitä kynnsarjoja ei voi pitää ”erehtymättöminä” tai ”oikeina”.

Hu ja Bentler [19] arvioivat simulaatiotutkimuksella RMSEA-indeksillä sekä muilla sopivuuksindekseillä heidän tutkimustansa ennen<sup>9</sup> tavanomaisesti käytettyjen kynnsarjojen soveltuvuutta ja esittivät vaihtoehtoisia arvoja, joilla he saivat pienempiä tyyppin II virheitä. Hu ja Bentler käyttivät kahta eri tyyppistä rajoitettua faktorimallia (yksinkertainen ja monimutkainen malli) generoimaan 200 otosta kullakin ennalta määritellyllä otoskoollla (150, 250, 500, 1000, 2500 ja 5000) seitsemässä eri olosuhteessa. Nämä seitsemän eri olosuhdetta viittasivat muuttujien ominaisuuksiin kuten niiden normaalisuuteen ja riippumattomuuteen. Otosten generoinnissa käytetyt faktorimallit sisälsivät kumpikin 15 havaittua muuttujaa ja kolme latenttia muuttujaa. Mallit erosivat niiden faktorilatausmatriisien suhteen; monimutkaisessa mallissa yhteen havaittuun muuttujaan saattoi latautua useampi kuin yksi faktori, mutta yksinkertaisessa mallissa yhteen havaittuun muuttujaan latautui vain yksi faktori.

Kun Hu ja Bentler [19] arvioivat tavanomaisesti käytettyjä kynnsarjoja eri sopivuuksindekseillä, he käyttivät yhtä oikein ja kahta väärin määriteltyä mallia, joista kunkin he sovittivat  $7 \times 6 \times 200 = 8400$  (olosuhteiden lukumäärä  $\times$  otoskokojen lukumäärä  $\times$  replikaatioiden lukumäärä) erilliseen otokseen. Tämän he toteuttivat kummallekin generoinnissa käytettävälle mallille ja kullekin testattavalle mallille erikseen. Kaikissa tapauksissa he laskivat eri sopivuuksindeksien arvot ja arvioivat sitten tavanomaisia kynnsarjoja käyttämällä saatavia tyyppin I ja II virheitä. Ensin Hu ja Bentler vertailivat tavanomaisia ja vaihtoehtoisia kynnsarjoja keskenään laskemalla kussakin tapauksessa, kuinka suuri osa malleista hylättiin, kun otoskoko oli  $\leq 250$ , tasan 500 tai  $\geq 1000$ . Tämän jälkeen he vertailivat tyyppin I ja II virheiden keskimääräisiä summia yhden ja kahden indeksin ”esitysstrategioilla”. Yhden indeksin esitysstrategiassa mallin hylkäys perustui vain yhteen sopivuuksindeksiin, joka oli jokin muu kuin SRMR (Standardized Root Mean squared Residual) -indeksi [39], ja kahden indeksin strategiassa hylkäys perustui SRMR-indeksiin sekä johonkin muuhun sopivuuksindeksiin (esimerkiksi RMSEA). RMSEA-indeksin tapauksessa he päätyivät tulokseen, että kynnsarvo lähellä lukua 0,06 johtaa pienimpiin tyyppin II virheisiin hyväksyttävillä tyyppin I virheillä.

#### 2.4.2 Inkrementaaliset sopivuuksindeksit

Inkrementaaliset sopivuuksindeksit vertaavat tutkittavan mallin  $\chi^2$ -testisuureen arvoa niin sanotun ’lähtökohtamallin’  $\chi^2$ -testisuureen arvoon. Toisin sanoen niiden avulla voidaan tutkia, onko tutkittava malli parannus verrattuna lähtökohtamalliin, jonka  $\chi^2$ -testisuureen arvo viittaa usein huonoon mallin sopivuuteen. Useimmiten lähtökohtamallina käytetään niin sanottua ’riippumattomuusmallia’ (engl. independence model), jossa kaikki virhevarianssit asetetaan nolliksi, kaikki faktorilataukset asetetaan ykkösiksi ja kaikki muuttujat ovat riippumattomia, jolloin vain muuttujien

<sup>9</sup>Hun ja Bentlerin [19] tutkimus julkaistiin vuonna 1999.

varianssit täytyy estimoida. Toinen usein käytetty lähtökohtamalli on niin sanottu 'nollamalli' (engl. null model), jossa kaikki parametrit asetetaan nolliksi. [39]

Inkrementaalisten TLI- ja CFI-indeksien otosestimaatit määritellään kaavoilla

$$\text{TLI} = \frac{\chi_i^2/v_i - \chi_t^2/v_t}{\chi_i^2/v_i - 1}, \quad (19)$$

$$\text{CFI} = 1 - \frac{\max\{\chi_t^2 - v_t, 0\}}{\max\{\chi_t^2 - v_t, \chi_i^2 - v_i, 0\}}, \quad (20)$$

missä  $\chi_i^2$  on lähtökohtamallin eli tässä tapauksessa riippumattomuusmallin  $\chi^2$ -testisuureen arvo,  $\chi_t^2$  on tutkittavan mallin  $\chi^2$ -testisuureen arvo ja  $v_i$  ja  $v_t$  ovat vastaavat vapausasteet. TLI-indeksi on muunneltu versio normitetusta sopivuusindeksistä (engl. Normed Fit Index, NFI)<sup>10</sup>, joka on riippuvainen otoskoosta, ja TLI-indeksiä kutsutaankin myös ei-normitetuksi sopivuusindeksiksi (engl. Nonnormed Fit Index, NNFI). Sen sijaan CFI-indeksi on muunneltu versio suhteellisesta epäkeskisyysindeksistä (engl. Relative Noncentrality Index, RNI) [28]. Yleisesti ottaen sekä TLI-että CFI-indeksin arvojoukko on suljettu väli  $[0, 1]$ , mutta koska TLI-indeksi ei ole normitettu, niin se voi saada arvoja myös kyseisen välin ulkopuolelta. Molempien sopivuusindeksien suuremmat arvot viittaavat parempaan mallin sopivuuteen. [39] Hun ja Bentlerin [19] simulaatiotutkimuksen perusteella<sup>11</sup> TLI- ja CFI-indeksien estimaattien arvot yli luvun 0,95 viittaavat mallin hyväksyttävään sopivuuteen. Schermelleh ja Moosbrugger [39] esittivät, että koska riippumattomuusmallin  $\chi^2$ -testisuureen arvo on melkein aina suuri, niin CFI:n ja TLI:n estimaatit ovat usein lähellä ykköstä. Tämän takia he ehdottivat sopivuudeltaan hyvän mallin CFI- ja TLI-estimaattien alarajaksi lukua 0,97.

### 2.4.3 Sopivuusindeksien estimaattien harhaisuus

Puhutaan vielä hieman edellä esiteltyjen RMSEA-, CFI- ja TLI-indeksien estimaattien harhaisuudesta. Kuten aikaisemmin mainittiin, RMSEA, CFI ja TLI esiteltiin ja niitä käytetään tutkielman empiirisessä osiossa siksi, että  $\chi^2$ -testisuureen estimaatin arvo riippuu suuresti otoskoosta ja mallin estimoitavien parametrien lukumäärästä. RMSEA, CFI ja TLI ovat rutiininomaisesti käytettyjä sopivuusindeksejä [41], minkä takia juuri ne valittiin  $\chi^2$ -testin tueksi vaihtoehtoisiksi tavoiksi arvioida faktorimallien sopivuutta.

RMSEA-, CFI- ja TLI-indeksien todellisia populaatioarvoja estimoidaan otoksesta lasketuilla estimaateilla yhtälöissä (18), (19) ja (20), minkä takia ne voivat olla harhaisia. Shi et al. [41] mainitsivat, että mallin koko eli malliparametrien lukumäärä saattaa vaikuttaa RMSEA-, CFI- ja TLI-indeksien estimaattien arvoihin, sillä ne ovat riippuvaisia  $\chi^2$ -testisuureesta. Shi kumppaneineen toteuttivat Monte Carlo -simulaation, jonka avulla he pyrkivät tutkimaan mallin koon vaikutusta juuri kyseisiin sopivuusindekseihin sekä oikein että väärin määritellyillä malleilla. He nuotoilivat yhteensä  $3 \times 5 \times 3 \times 2 = 90$  erilaista vaihtoehtoista tilannetta (kolme erilaista mallin määrittelyä, viisi erilaista mallikokoa, kolme erilaista otoskokoa

<sup>10</sup>NFI-indeksi määritellään artikkelissa [39] sivulla 40

<sup>11</sup>Ks. luku 2.4.1, jossa tätä tutkimusta referoitiin.



ja kaksi erilaista faktorilatausten tasoa), joista kutakin he simuloivat 1000 kertaa generoimalla havainnot multinormaalijakaumasta.

Shin et al. [41] simulaatiotutkimuksen tulosten perusteella RMSEA-indeksin populaatioarvot pienenevät, kun mallin koko kasvoi, riippumatta siitä, miten malli oli määritelty väärin. Sen sijaan CFI- ja TLI-indeksien populaatioarvot riippuivat mallin määrittelyvirheen tyypistä, kun mallin koko kasvoi: CFI- ja TLI-populaatioarvot pienenevät, kun mallin faktorien lukumäärä oli määritelty väärin; CFI- ja TLI-populaatioarvot kasvoivat, kun mallin määrittelyssä ei otettu huomioon korreloituneita residuaaleja. Shi kumppaneineen mainitsivat erityisesti sen, että suuren RMSEA-arvon tulkinnassa kannattaa olla varovainen silloin, kun mallin koko on pieni ja faktorilataukset ovat suuria, koska siinä tapauksessa jopa RMSEA-populaatioarvo oli suurempi kuin yleisesti käytetty kynnyisarvo (0,06). Myös Kenny et al. [21] tutkivat RMSEA-indeksin ja mallin koon välistä yhteyttä simulaatiotutkimuksella. He suosittelevat, että RMSEA-indeksin estimaattia ei kannata laskea ollenkaan, kun mallin vapausasteiden lukumäärä on pieni, eikä varsinkaan silloin, kun otoskoko on pieni vapausasteiden lukumäärän lisäksi.

Shi et al. [41] havaitsivat, että pienellä otoskolla sekä RMSEA-, CFI- että TLI-indeksin estimaatit osoittivat keskimäärin huonommasta mallin sopivuudesta kuin niitä vastaavat populaatioarvot. Kun mallin koko kasvoi, sopivuusindeksien keskimäärien estimaattien ja vastaavien populaatioarvojen erot kasvoivat. Nämä erot kasvoivat varsinkin silloin, kun faktorilataukset olivat matalia. Tulokset indikoivat erityisesti, että CFI- ja TLI-indeksi saattavat vaatia suuren otoskoon ( $n \geq 1000$ ), jotta niiden estimaatteja voidaan tulkita luotettavasti, kun malli on suuri ja faktorilataukset ovat matalia. RMSEA-indeksin keskimääräiset otosestimatit olivat suurempia kuin yleisesti käytetty kynnyisarvo (0,06) vain silloin, kun mallin koko oli pieni ja faktorilataukset olivat suuria.

Edellä tehtyjen havaintojen perusteella RMSEA-, CFI- ja TLI-indeksi eivät selvästikään ole täydellisiä tapoja arvioida faktorimallien hyvyyttä. Sopivuusindeksien estimaattien tulkinnan apuna tavanomaisesti käytettyjä kynnyisarvoja on täten syytä pitää vain suuntaa antavia suosituksina, kuten Marsh et al. [27] totesivat myös. Tämän takia mallin hyvyyden arvioinnissa kannattaa ottaa huomioon myös otoskoko ja parametrien lukumäärä. Kun otoskoko on pieni, RMSEA-, CFI- ja TLI-indeksien tulkinnassa kannattaa olla erityisen varovainen. Kun otoskoko on suuri, sopivuusindeksejä voi tulkita jo melko luotettavasti, mutta tällöin malliparametrien suuri lukumäärä on silti saattanut vaikuttaa sopivuusindeksien estimaatteihin.

#### 2.4.4 Ryhmiteltyjen faktorimallien sopivuuden testaaminen

Mittausinvarianssin testaaminen vaatii tutkittavan ryhmitellyn faktorimallin sopivuuden arvioinnin. Luvuissa 2.4.1 ja 2.4.2 esitellyt sopivuusindeksit voidaan laajentaa ryhmiteltyyn ja ei-keskistettyyn tilanteeseen. Luvussa 2.4.1 esiteltiin  $\chi^2$ -testi ja RMSEA-indeksi, jotka ovat absoluuttisia sopivuusindeksejä. Olkoot  $\Sigma_{0g}$  ja  $\mu_{0g}$  mallin implikoimat kovarianssimatriisi ja odotusarvovektori populaatiossa  $\Pi_g$ . Khiin-neliön testin nollahypoteesi on nyt muotoa  $\Sigma_g = \Sigma_{0g}$ ,  $\mu_g = \mu_{0g}$  kaikilla  $g = 1, \dots, G$ . Kun  $\hat{F}$  on sovitustfunktion (16) minimi, niin  $\chi^2$ -testisuure  $(n-1)\hat{F}$  noudattaa asymp-totottisesti  $\chi^2$ -jakaumaa  $q(q+1)/2 + q - t$  vapausasteella, jossa  $q(q+1)/2$  on otoskova-

rianssimatriisiin  $\mathbf{S}_g$  ei-redundanttien alkioden lukumäärä,  $q$  on otoskeskiarvovektorin  $\bar{\mathbf{x}}_g$  alkioden lukumäärä ja  $t$  on vapaiden parametrien määrä mallin odotusarvo- ja kovarianssirakenteissa. RMSEA-indeksi on nyt muotoa  $\hat{\epsilon}_g = \hat{\epsilon}\sqrt{G}$ , jossa  $\hat{\epsilon}$  on määritetty kuten yhtälössä (18) ottaen tietenkin huomioon sovituskäytännön muodon (16) ja vapausasteiden lukumäärän  $q(q+1)/2 + q - t$  ryhmiteltyssä ja ei-keskistetyssä tilanteessa. [30]

Luvussa 2.4.2 esiteltiin TLI- ja CFI-indeksit, jotka ovat inkrementaalisia sopivuusindeksejä. Ne määritellään nyt samalla tavalla kuin yhtälöissä (19) ja (20) ottaen huomioon  $\chi^2$ -testisuureen muodon ja vapausasteiden lukumäärän ryhmiteltyjen ja ei-keskistettyjen faktorimallien tapauksessa. TLI- ja CFI-indeksien perusteella tulkittu faktorimallin sopivuus riippuu vahvasti lähtökohtamallin valinnasta, jossa täytyy nyt ottaa huomioon kovarianssirakenteen lisäksi odotusarvorakenne. Mitään selvää parasta valintaa odotusarvorakenteen lähtökohtamallille tässä tapauksessa ei ole. Yleensä lähtökohtamalli on kuitenkin vähintään yhtä rajoitettu kuin tutkittava malli. Ryhmiteltyjen faktorimallien tapauksessa  $\chi^2$ -testiä tulkitaan samalla tavalla ja RMSEA- sekä TLI- ja CFI-indekseille pätevät samat kynnysarvot kuin ei-ryhmiteltyjen faktorimallien tapauksessa, kun niitä käytetään arvioimaan mallin sopivuutta. [30]

## 2.5 Faktorimallien uudelleenmäärittely

Toisinaan ensimmäinen tutkittava faktorimalli ei sovi aineistoon hyvin, jolloin se pyritään usein uudelleenmäärittämään. Mahdollisesti parhaat ideat mallin modifioimiseen tulevat tutkijan omasta teoreettisesta ymmärryksestä tutkimuksen aiheesta. Kuitenkin yleensä tutkija on jo ennen ensimmäisen malliversion määrittelyä tehnyt laadullista tutkimusta ja käyttänyt kaiken kerryttämänsä tiedon rakentaakseen parhaan mahdollisen teoreettisen mallin. Tällöin tutkijan teoreettisesta ymmärryksestä ei välttämättä ole sen enempää hyötyä ja hän voi turvautua erilaisiin tilastollisiin menetelmiin mallin uudelleenmäärittelyssä. [4] Faktorimallien modifiointi ei kuitenkaan ole tämän tutkielman tutkimusaiheena, joten seuraavaksi esitellään vain pintapuolisesti eräs tällainen faktorimallin uudelleenmäärittelyn apuna käytettävä tilastollinen menetelmä, jotta lukija ymmärtää luvussa 5 esitetyjä pohdintoja koskien luvussa 4 tutkittavien ahdistuskyselyjen mahdollisia jatkotutkimuksia.

Lagrangen kertoimen testi (LM-testi) vertaa enemmän rajoitetun mallin sopivuutta vähemmän rajoitetun mallin sopivuuteen. Vain enemmän rajoitettu malli täytyy estimoida, ja tässä tapauksessa kyseinen malli on tutkittava teoreettinen faktorimalli. Sen sijaan vähemmän rajoitettu malli on tässä tapauksessa alkuperäinen faktorimalli, josta yksi tai useampi rajoitettu rakenteellinen parametri on vapautettu estimoitavaksi. Faktorianalyysin kontekstissa LM-testiä kutsutaan muunnosindeksiksi (engl. modification index, MI). Sen yksiulotteinen versio määritellään kaavalla

$$MI_i = \left[ \frac{\delta l(\eta)}{\delta \eta_i} \right]^2 [\mathbf{I}^{-1}(\hat{\eta}_r)]_{ii}, \quad (21)$$

jossa  $l(\eta)$  on vähemmän rajoitetun mallin log-uskottavuusfunktio,  $\eta_i$  on vapautettu rakenteellinen parametri,  $\hat{\eta}_r$  on rajoitetumman mallin eli tutkittavan faktorimallin rakenteellisten parametrien vektorin estimaatti, ja  $\mathbf{I}^{-1}(\hat{\eta}_r)$  on matriisin

$I^{-1}(\eta) = \{-E[\delta^2 l(\eta)/\delta\eta\delta\eta']\}^{-1}$  arvo pisteessä  $\hat{\eta}_r$  ja  $[I^{-1}(\hat{\eta}_r)]_{ii}$  on sen  $i$ :nnes diagonaalialkio. Yhtälön (21) tuloksena saadaan  $\chi^2$ -estimaatti yhdellä vapausasteella, jonka avulla voidaan arvioida kuinka paljon mallin sopivuus paranee, kun parametri  $\eta_i$  vapautetaan estimoitavaksi. Muunnosindeksejä käyttäen tutkija voi etsiä parametrin, jonka vapauttamisella saadaan suurin MI-indeksin arvo eli suurin  $\chi^2$ -estimaatin pieneneminen, ja näin saadaan päivitetty malli, joka asetetaan yhtälössä (21) rajoitetun mallin paikalle. Tätä prosessia toistetaan, kunnes saavutetaan sopivuudeltaan riittävä malli. [4]

LM-testillä on kuitenkin lukuisia ongelmia. Ensimmäinen ongelma on se, että järjestys jossa lukittuja parametreja vapautetaan saattaa vaikuttaa jäljelle jääneiden lukittujen parametrien LM-testien tilastolliseen merkitsevyyteen. Toiseksi, jos estimointi ja mallin modifiointi tapahtuu samalla aineistolla, niin tällöin LM-testiin liittyvät todennäköisyydet eivät luultavasti ole tarkkoja. Kolmanneksi, LM-testillä ei pystytä kertomaan mitään siitä, minkä suuruinen on vapautettavan parametrin estimaatti. Neljänneksi, jos tutkittava malli ei ole oikea, niin tällöin LM-testi on harhainen versio LM-testistä, joka tehtäisiin oikealle mallille. Viides ongelma on se, että LM-testi on hyödyllisimmillään silloin, kun virheet mallin määrittelyssä ovat pieniä. Viimeiseksi, LM-testien indikoimilla vapautetuilla parametreilla ei välttämättä ole selkeää tulkintaa. Erityisesti virheiden väliset vapautetut korrelaatioparametrit saattavat parantaa mallin sopivuutta, mutta ne eivät välttämättä tuo tutkijalle mitään uutta tietoa, mikäli hän ei pysty selittämään parametrin vapauttamista millään muulla tavalla kuin empiirisen todistusaineiston avulla. [4]

LM-testin kolmannen ongelman ratkaisemiseksi voidaan tutkia rajoitettujen parametrien estimoituja muutoksia (engl. expected parameter change, EPC) muunnosindeksien lisäksi [4]. Rajoitetun parametrin EPC-arvo voidaan tulkita estimaattina sille, kuinka paljon kyseisen parametrin rajoitettu arvo eroaa sen oikeasta arvosta. EPC määritellään kaavalla

$$EPC_i = \eta_i - \eta_0 = \frac{MI_i}{\delta l(\eta)/\delta\eta_i},$$

jossa  $\eta_i$  on vapautettu parametri,  $\eta_0$  on vapautetun parametrin rajoitettu arvo (usein  $\eta_0 = 0$ ) ja muunnosindeksi  $MI_i$  on määritelty kuten yhtälössä (21). Kun rajoitettujen parametrien MI- ja EPC-arvoja tutkitaan samanaikaisesti, vastaan saattaa tulla neljä erilaista tapausta. Mikäli rajoitetun parametrin MI- sekä EPC-arvo ovat suuria, on perusteltua vapauttaa kyseinen parametri estimoitavaksi. Sen sijaan, jos rajoitetun parametrin MI- sekä EPC-arvo ovat pieniä, on selvää, ettei kyseisen parametrin vapauttaminen ole mielekäästä. Kaksi jäljelle jäänyttä tapausta eivät ole yhtä selkeitä kuin kaksi ensimmäistä: rajoitetun parametrin MI-arvo on suuri, mutta EPC-arvo on pieni, tai päinvastaisesti rajoitetun parametrin MI-arvo on pieni, mutta EPC-arvo on suuri. [38] Huomattavaa on siis, että MI- ja EPC-arvojen tutkiminen samanaikaisesti ei siltikään tee mallien uudelleenmäärittelystä suoraviivaista.

### 3 Mittausinvarianssi

Psykologisten testien tutkimisessa ollaan usein kiinnostuneita siitä, mittaako testi tutkittavaa asiaa samalla tavalla ominaisuuksiltaan erilaisissa ihmisryhmissä eli alapopulaatioissa (esim. naiset ja miehet). Tällöin tutkitaan ovatko testin tulokset näissä erilaisissa tutkittavissa alapopulaatioissa vertailtavissa samalla mitta-asteikolla. Toisin sanoen tällöin testataan toteutuuko mittausinvarianssi (engl. measurement invariance) tutkittavien alapopulaatioiden yli kyseisen testin tapauksessa. Mikäli mittausinvarianssi toteutuu, alapopulaatioiden eroja tutkittavissa piirteisissä voidaan vertailla mielekkäästi. Sen sijaan, jos mittausinvarianssi ei toteudu, havaitut erot alapopulaatioiden välillä voivat olla virheellisiä ja harhaanjohtavia. [48]

#### 3.1 Määritelmä

Oletetaan, että on teetetty testi ja ollaan saatu  $n:n$  kokoinen otos, joka koostuu  $G$ :stä ominaisuuksiltaan erilaisesta alapopulaatiosta  $\Pi_1, \dots, \Pi_G$  otosko'oilla  $n_1, \dots, n_G$ . Olkoon  $R_i$  satunnaismuuttuja, jonka realisaatio  $r_i \in \{1, \dots, G\}$  kertoo mihin alapopulaatioon  $i$ :nnes havainto kuuluu. Olkoot sitten  $\mathbf{t}_i$  vektori, joka sisältää testissä tutkittavan ilmiön tuntemattomat ja ei-havaittavissa olevat 'aidot pistemäärät', ja  $\mathbf{y}_i$  testissä havaittujen pistemäärien vektori, jolla mitataan vektoria  $\mathbf{t}_i$ . Tällöin  $\mathbf{T}_i$  ja  $\mathbf{Y}_i$  ovat vektoreita  $\mathbf{t}_i$  ja  $\mathbf{y}_i$  vastaavat satunnaisvektorit  $i$ :nnelle havainnolle. Satunnaismuuttujan  $\mathbf{Y}_i$  sanotaan olevan mittausinvariantti suhteessa satunnaismuuttujan  $R_i$  realisaatioon, jos

$$P(\mathbf{Y}_i = \mathbf{y}_i \mid \mathbf{T}_i = \mathbf{t}_i, R_i = r_i) = P(\mathbf{Y}_i = \mathbf{y}_i \mid \mathbf{T}_i = \mathbf{t}_i), i \in \{1, \dots, n\}. \quad [49] \quad (22)$$

Mittausinvarianssi toteutuu siis jos ja vain jos havaitun pistemäärän todennäköisyys ehdolla aito pistemäärä ja ryhmäjäsensyys on yhtä suuri kuin havaitun pistemäärän todennäköisyys ehdolla vain aito pistemäärä. Faktoriansalyysin kontekstissa havaittujen pistemäärien vektori vastaa indikaattorivektoria  $\mathbf{x}_i$  ja aitojen pistemäärien vektori vastaa faktoripisteiden vektoria  $\xi_i$ . Tällöin mittausinvarianssia kutsutaan myös faktoriseksi invarianssiksi (engl. factorial invariance) ja tarkoittaa, että samaa latenttia muuttujaa mitataan samalla mitalla eri alaryhmissä, jolloin eri alaryhmien faktoripisteet ovat vertailtavissa. [49]

#### 3.2 Faktorisen invarianssin testaaminen

Oletetaan, että tutkittava rajoitettu faktorimalli on estimoitu ja sen sopivuuden aineistoon on arvioitu olevan riittävä. Seuraavaksi halutaan tutkia, onko eri alapopulaatioiden välillä tilastollisesti merkitseviä eroja faktoripisteissä. Tätä ennen täytyy varmistua siitä, että malli on faktorisesti invariantti tutkittavien alapopulaatioiden suhteen, jotta populaatioiden eroja faktoripisteissä voidaan vertailla luotettavasti. Faktorisen invarianssin testaaminen koostuu sisäkkäisten mallien peräkkäisistä hypoteesitestauksista lähtien liikkeelle vähiten rajoitetusta mallista [49]. Meredith [29] esitti, että käytettäisiin neljää sisäkkäistä mallia.

### 3.2.1 Testattavat mallit

Invarianssitestausten vähiten rajoitettu malli on konfiguraalisen invarianssin (engl. configural invariance) malli, joka mainittiin jo luvussa 2.3.2. Sen nollahypoteesi on muotoa

$$H_{01}: \Sigma_g = \Lambda_{gc} \Phi_g \Lambda'_{gc} + \Theta_g, \mu_g = \tau_g + \Lambda_{gc} \kappa_g, g = 1, \dots, G, \quad (23)$$

jossa alaindeksi  $c$  viittaa johonkin tiettyyn rakenteeseen faktorilatausten matriiseissa  $\Lambda_g$ . Hypoteesi (23) tarkoittaa siis, että kussakin populaatiossa  $\Pi_1, \dots, \Pi_G$  on sama määrä faktoreita ja kyseiset faktorit liittyvät samoihin muuttujien osajoukkoihin. Konfiguraalisen invarianssin malli ei vielä aseta mitään ylimääräisiä rajoitteita mallin odotusarvorakenteeseen. [30]

Oletetaan, että nollahypoteesia (23) ei voida hylätä, jolloin seuraavaksi testattava malli on metrisen invarianssin (engl. metric invariance) malli. Sen nollahypoteesi on muotoa

$$H_{02}: \Sigma_g = \Lambda \Phi_g \Lambda' + \Theta_g, \mu_g = \tau_g + \Lambda \kappa_g, g = 1, \dots, G, \quad (24)$$

joka asettaa faktorilatausmatriisit täysin invarianteiksi populaatioiden välillä, mutta ei vielä mitään ylimääräisiä rajoitteita odotusarvorakenteeseen. Metristä invarianssia kutsutaan myös heikoksi faktoriseksi invarianssiksi. [30]

Oletetaan sitten, että myöskään nollahypoteesia (24) ei voida hylätä, jolloin seuraavaksi testattava malli on skalaarisen invarianssin (engl. scalar invariance) malli. Sen nollahypoteesi on muotoa

$$H_{03}: \Sigma_g = \Lambda \Phi_g \Lambda' + \Theta_g, \mu_g = \tau + \Lambda \kappa_g, g = 1, \dots, G, \quad (25)$$

jossa asetetaan faktorilatausmatriisien lisäksi odotusarvorakenteen vakiotermit invarianteiksi populaatioiden välillä. Nollahypoteesi (25) olettaa, että populaatioiden väliset erot havaittujen muuttujien keskiarvoissa  $\mu_g$  johtuvat pelkästään eroista faktorikeskiarvoissa  $\kappa_g$ . Skalaarista invarianssia kutsutaan myös vahvaksi faktoriseksi invarianssiksi. [30]

Jos nollahypoteesia (25) ei voida hylätä, niin seuraavaksi ja viimeiseksi testattava malli on tiukan faktorisen invarianssin (engl. strict factorial invariance) malli. Sen nollahypoteesi on muotoa

$$H_{04}: \Sigma_g = \Lambda \Phi_g \Lambda' + \Theta, \mu_g = \tau + \Lambda \kappa_g, g = 1, \dots, G, \quad (26)$$

jossa asetetaan faktorilatausten ja vakiotermin lisäksi virhetermien kovarianssimatriisit  $\Theta_g$  invarianteiksi populaatioiden välillä. Mikäli tiukka invarianssi toteutuu, populaatioiden väliset erot havaittujen muuttujien keskiarvoissa ja kovarianssirakenteissa johtuvat pelkästään eroista faktorien jakaumissa. [30] Meredith [29] väitti, että tiukan invarianssin toteutuminen vaaditaan, jotta faktoripisteitä voidaan vertailla luotettavasti populaatioiden välillä. Wu et al. [49] mainitsivat, että tästä näkemyksestä on ollut eriäviä mielipiteitä, mutta hallitseva tutkimuskäytäntö on ollut, että skalaarisen invarianssin toteutuminen on riittävä todiste mittausinvarianssille. Schoot et al. [40] esittivät, että vaikka tiukka invarianssi ei toteudu, niin ryhmien faktoripisteitä voidaan silti vertailla, mutta tällöin ryhmien välillä on erisuuruiset

mittausvirheet, mikä on syytä huomioida päätelmien tekemisessä. Tämän tutkielman empiirisen osion invarianssitestauksissa pysähdytään skalaariseen invarianssiin ja sen toteutumista pidetään riittävänä todistamaan faktorimallin täysi mittausinvarianssi.

Mainittakoon vielä, että malleissa (23), (24), (25) ja (26) ei aseteta invarianssirajoitteita faktorien kovarianssimatriiseihin  $\Phi_g$  eikä niiden odotusarvovektoreihin  $\kappa_g$ , sillä ne eivät ole osa faktorista invarianssia. Tämä nähdään selkeästi, kun tutkitaan tarkemmin mittausinvarianssin määritelmää (22). Kun havaittujen muuttujien jakauma ehdollistetaan faktoripisteiden suhteen, niin yhtälöiden (12) ja (13) perusteella havaittujen muuttujien ehdolliset odotusarvo- ja kovarianssirakenteet ovat  $E_g[\mathbf{X} | \xi] = \tau_g + \mathbf{\Lambda}_g \xi$  ja  $\text{Cov}_g(\mathbf{X} | \xi) = \mathbf{\Theta}_g$ . Huomataan, että parametrit jotka määrittävät faktoripisteiden jakauman, eivät ole enää relevantteja, ja täten populaatioiden väliset erot matriiseissa  $\Phi_g$  tai vektoreissa  $\kappa_g$  eivät vaikuta mittausinvarianssin toteutumiseen. [30]

### 3.2.2 Testausmenettely

Vielä täytyy määritellä, miten malleja (23), (24) ja (25) testataan ja miten voidaan päätellä, että mittausinvarianssi toteutuu. Koska konfiguraalisen invarianssin malli (23) on vähiten rajoitettu malli, sitä ei tarvitse verrata enemmän rajoitettuihin malleihin. Konfiguraalinen invarianssi todetaan sovittamalla tutkittava faktorimalli erikseen kussakin alapopulaatiossa ja sitten testaamalla, että kunkin populaatiokohtaisen mallin sopivuus on riittävä luvussa 2.4 esiteltyjen sopivuusindeksien perusteella. Vaihtoehtoisesti voidaan sovittaa ryhmitelty faktorimalli ja tutkia sen sopivuutta aineistoon. [40] Mikäli konfiguraalinen invarianssi ei toteudu, eri populaatioissa mitattiin erilaisia rakenteita. Täten konfiguraalisen invarianssin toteutuminen vaaditaan, jotta faktorisen invarianssin testauksia on mielekästä jatkaa. [49]

Metrisen ja skalaarisen invarianssin testaaminen vaatii sisäkkäisten mallien (23), (24) ja (25) vertailemisen; kun testataan metristä invarianssia, vertaillaan konfiguraalista (23) ja metristä (24) mallia keskenään, ja kun testataan skalaarista invarianssia, vertaillaan metristä (24) ja skalaarista (25) mallia keskenään. Eräs tapa vertailla näitä malleja toisiinsa on muotoilla niiden  $\chi^2$ -testisuureiden väliset erotukset  $\Delta\chi^2$ . Nämä erotukset noudattavat  $\chi^2$ -jakaumia vapausasteilla, jotka ovat yhtä suuria kuin vertailtavien mallien vapausasteiden erotukset. Jos khiin neliön erotuksen testi ei ole tilastollisesti merkitsevä, voidaan päätellä, että rajoitetumman mallin sopivuus aineistoon ei ole huomattavasti huonompi kuin vähemmän rajoitetun mallin, ja tätä pidettäisiin todisteena sille, että tutkittava invarianssin taso toteutuu. Khiin neliön erotuksen testiä ei kuitenkaan ole syytä pitää ainoana aputekijänä päätöksenteossa, sillä se on riippuvainen otoskoosta ja mallin kompleksisuudesta. [49]

Khiin neliön erotuksen testin sijasta voi tutkia vertailtavien mallien sopivuusindeksien erotuksia. Tämän tutkielman empiirisessä osiossa käytetään erotuksia  $\Delta\text{CFI}$  ja  $\Delta\text{RMSEA}$ , joissa aina vähemmän rajoitetun mallin sopivuusindeksin arvosta vähennetään enemmän rajoitetun mallin sopivuusindeksin arvo. [49] Kun testataan metristä tai skalaarista invarianssia, seuraavia kynnyksarvoja suositellaan käyttämään päätöksenteon apuna: kun koko populaation otoskoko on  $n \leq 300$ , kynnyksarvot  $\Delta\text{CFI} \leq -0,005$  ja  $\Delta\text{RMSEA} \geq 0,010$  yhdessä osoittavat, että tutkittava invarianssin taso ei toteudu; kun koko populaation otoskoko on  $n > 300$ , kynny-

sarvot  $\Delta\text{CFI} \leq -0,010$  ja  $\Delta\text{RMSEA} \geq 0,015$  yhdessä osoittavat, että tutkittava invarianssin taso ei toteudu. [10]

### 3.2.3 Osittainen invarianssi

Lopuksi on vielä syytä puhua osittaisesta invarianssista (engl. partial invariance), vaikkei se varsinaisesti kuulu tämän tutkielman aiheisiin. Osittainen invarianssi esitellään tässä lyhyesti siksi, että lukija ymmärtää, mitä sillä tarkoitetaan, kun luvun 5 pohdinnoissa puhutaan tutkielman empiirisessä osiossa tutkittavien ahdistuskyselyiden jatkotutkimuksesta, joka mahdollisesti sisältää kyselyiden osittaisen invarianssin tutkimista. Osittaisen invarianssin tutkiminen on olennaista silloin, kun metrinen, skalaarinen tai tiukka invarianssi tai mikään niistä ei toteudu.

Faktorimalli on osittain invariantti silloin, kun koko parametrivektori tai -matriisi ei ole invariantti vaan vain jokin sen alkioiden osajoukko on invariantti alapopulaatioiden suhteen. Tällöin invarianssimalleissa (24), (25) ja (26) invarianssirajoitteita asetetaan vain joihinkin odotusarvovektorin  $\tau_g$ , faktorilatausmatriisiin  $\Lambda_g$  ja virheiden kovarianssimatriisiin  $\Theta_g$  alkioiden osajoukkoihin. Osittaisen invarianssin tutkimisen avulla voidaan siis löytää ne parametrit, joilla täysi mittausinvarianssi ei toteudu. Esimerkiksi metrisen invarianssin mallissa on mahdollista valita jokin faktorien osajoukko ja asettaa metrinen invarianssirajoite vain valittua osajoukkoa vastaaviin sarakkeisiin faktorilatausten matriisissa  $\Lambda_g$ . [30] Tällöin – luvun 3.2.2 testausmenettelyn mukaisesti – mikäli konfiguraalisen ja metrisen mallin erot CFI- ja RMSEA-sopivuusindekseissä eivät ole suuria, voidaan päätellä, että osittainen metrinen invarianssi toteutuu. Samalla tavalla voi menetellä skalaarisen ja tiukan faktorisen invarianssin testaamisessa.

## 4 Ahdistuskyseilyjen faktorirakenteet ja mittausinvarianssi

Seuraavaksi tutkitaan matematiikka- (MAQ) sekä lukemis- ja kirjoittamisahdistuskyseilyjen (RWAQ) faktorirakenteita ja mittausinvarianssia. Idea on peräisin Turun yliopiston Oppimisanalytiikan tutkimusinstituutin [33] kansainvälisestä yhteistyöhankkeesta (Pekka Räsänen, varajohtaja, Oppimisanalytiikan tutkimusinstituutti, Turun yliopisto; Annemarie Fritz ja Lars Orbach, Duisburg-Essenin yliopisto, Saksa; Vitor Haase, Federal University of Minas Gerais (UFMG), Belo Horizonte, Brasilia; Ann Dowker, Oxfordin yliopisto, Englanti). Matematiikka-ahdistusta on tutkittu melko laajalti aikaisemminkin. On todettu, että matematiikka-ahdistuksella ja heikolla menestyksellä matemaattisia saavutuksia mittaavissa testeissä on yhteys toisiinsa. Henkilöille, joilla on matematiikka-ahdistusta, on tyypillistä karttaa matematiikkaa kokonaan. [14] Lukemisahdistusta on myös tutkittu laajasti aikaisemminkin, mutta lähinnä vain toisen kielen omaksumisen yhteydessä, ja lukemisahdistuksen tutkiminen oman äidinkielen oppimisessa on ollut vähäistä [34]. Lisäksi vaikka lapsen aikainen pätevyys lukemisessa on tärkeä osa akateemista menestystä, niin aikaisen lukemisahdistuksen ja lukemismenestyksen välistä yhteyttä ei ole tutkittu systemaattisesti [36]. Kirjoittamisahdistusta on tutkittu muun muassa Turkissa, jossa Yaman [50] kehitti mittarin peruskouluikäisten turkkilaisten lasten kirjoittamisahdistuksen mittaamiseen.

Cipora et al. [11] pohdiskelivat, mitä matematiikka-ahdistuksesta tiedetään jo aiemman tutkimuksen perusteella. Ensimmäisenä asiana Cipora kumppaneineen nostivat esille sen, että matematiikka-ahdistuksen tiedetään todellakin olevan olemassa omana rakenteenaan muista ahdistuksen tyypeistä erottuvana. Muita ahdistuksen tyyppisiä ovat esimerkiksi tila-ahdistus (engl. state anxiety), piirreahdistus (engl. trait anxiety) ja koeahdistus (engl. test anxiety), joiden kanssa matematiikka-ahdistus korreloi positiivisesti. He mainitsivat myös sen, että matematiikka-ahdistusta voidaan mitata reliabelisti mittareilla, joilla tutkittavat henkilöt kuvailevat itse omaa ahdistustaan (esim. kyselyt, joihin vastataan tietyllä asteikolla). Lisäksi tiedetään, että matematiikka-ahdistus ei ole yksiulotteinen rakenne, vaan se jakautuu useampaan kuin yhteen matematiikka-ahdistuksen tyyppiin. Näiden lisäksi Cipora kumppaneineen mainitsivat useita muita kvalitatiiviseen ja kvantitatiiviseen tutkimukseen pohjautuvia tietämyksiä matematiikka-ahdistuksesta.

Cipora et al. [11] esittivät näkemyksiään myös siitä, mitä matematiikka-ahdistuksesta ei vielä tiedetä. He listasivat kymmenen kysymystä, joihin matematiikka-ahdistus-tutkijoiden olisi heidän mielestään syytä vastata. He esimerkiksi esittivät kysymyksen ”What is the empirical support for the validity of the MA as a construct?”<sup>12</sup>, johon tässä luvussa pyritään ainakin osittain vastaamaan. Heidän muita esittämiään tärkeitä kysymyksiä olivat esimerkiksi ”What is the prevalence of MA?” ja ”When and how does MA begin?”. Viimeistä kysymystä he jatkoivat kysymyksillä ”To what extent can it be seen in the early school years and even preschool years?” ja ”Could very early intervention prevent it from developing?”, jotka viitta-

<sup>12</sup>Tässä termi *MA* viittaa matematiikka-ahdistukseen.



vaat siihen, voiko lapsen matematiikka-ahdistuksen kehityksen jollain tavalla estää varhaisessa iässä. Selvästikin matematiikka-ahdistuksen tutkimisessa on vielä paljon tehtävää.

Suomessa Sorvo et al. [42] tutkivat matematiikka-ahdistuksen yhteyttä aritmeettisiin perustaitoihin alakouluikäisillä lapsilla. Heidän ahdistuskyselynsä koostui kahdesta osa-alueesta, joiden faktorirakennetta Sorvo kumppaneineen tutkivat konfirmatiivisella faktorianalyysillä: ahdistus epäonnistumisesta matematiikassa ja ahdistus matematiikkaan liittyvissä tilanteissa. Tulosten perusteella nämä kaksi eri osaluuetta olivat eroteltavissa mutta korreloituneita ja niistä jälkimmäinen oli enemmän yhteydessä aritmeettiseen sujuvuuteen kuin ensimmäinen. Sorvo et al. [43] jatkoivat työtään pitkittäisellä tutkimuksella, jossa he yhden kouluvuoden aikana tutkivat kehitystä alakouluikäisten lasten matematiikka-ahdistuksessa ja sen yhteydessä aritmeettiseen osaamiseen. Tulosten perusteella keskimääräinen ahdistuksen taso matematiikkaan liittyvissä tilanteissa laski toisen, kolmannen ja neljännen luokka-asteen lapsilla mutta pysyi samana viidennen luokka-asteen lapsilla. Sen sijaan keskimääräinen ahdistuksen taso epäonnistumisessa matematiikassa pysyi samana toisen luokka-asteen lapsilla mutta laski muiden luokka-asteiden lapsilla. Lisäksi tulokset osoittivat, että alun alhaiset aritmeettiset taidot ennustivat myöhempää korkeaa ahdistuneisuutta matematiikassa.

Tuomola [45] tutki Pro gradu -tutkielmassaan lukemisahdistusta ja sen yhteyttä matematiikka-ahdistukseen, lukusujuvuuteen, sukupuoleen ja luokka-asteeseen. Sekä lukemisahdistusta että matematiikka-ahdistusta mitattiin kolmella väittämällä. Matematiikka-ahdistuksen väittämät olivat Sorvon et al. [42] käyttämät väittämät, jotka koskivat ahdistusta matematiikkaan liittyvissä tilanteissa. Lukemisahdistuksen väittämät vastasivat matematiikka-ahdistuksen väittämiä, mutta ne oli muokattu vastaamaan lukemiseen liittyviä tilanteita. Väittämien pienen määrän takia mitään ylempää latenttien muuttujien rakennetta ei tutkielmassa analysoitu.

Seuraavaksi tutkittavat kyselyt (MAQ ja RWAQ) on luotu Careyn et al. [9] mAMAS-mittarin pohjalta. Carey kumppaneineen tutkivat matematiikka-ahdistusta englantilaisilla lapsilla ja sen yhteyttä koeahdistukseen ja yleiseen ahdistuneisuuteen. He pitivät tärkeänä kehittää luotettava mittari, jolla voi mitata matematiikka-ahdistusta sekä erittäin nuorilla että vanhemmilla lapsilla. Heidän mielestään aiemmin tutkituilla lapsille suunnatuilla ahdistusmittareilla on ollut ongelmana niiden ikärajoitteisuus ja validiteettia tukevien tilastollisten tunnusluku-  
jen puute. Niinpä Carey kumppaneineen kehittivät lapsia ja nuoria varten modifioidun version aikuisille suunnatusta Abbreviated Math Anxiety Scale (AMAS) -kyselystä [16] ja osoittivat konfirmatiivisella faktorianalyysillä, että sillä voidaan mitata matematiikka-ahdistusta englantilaisilla lapsilla ja nuorilla luotettavasti. He eivät kuitenkaan tutkineet mittarin mittausinvarianssia ollenkaan, minkä takia he ehdottivatkin, että mittarin jatkotutkimuksissa voisi arvioida eroaako sen faktorirakenne esimerkiksi sukupuolten suhteen. Tämän takia MAQ- ja RWAQ-kyselyjen faktorirakenteita ei analysoida pelkästään konfirmatiivisella faktorianalyysillä yhden populaation suhteen vaan myös niiden mittausinvarianssia tutkitaan sukupuolten sekä luokka-asteiden suhteen.

## 4.1 Metodit

### 4.1.1 Aineisto

Syksyn 2021 FUNA-mittauksiin [32] osallistui kahtena eri päivänä yhteensä 10573 peruskouluikäistä lasta useassa eri paikkakunnassa Varsinais-Suomessa ja Uudellamaalla. Ensimmäisenä päivänä osallistujat suorittivat matematiikkaan liittyviä testejä ja vastasivat lukemis- ja kirjoittamisahdistuskyselyyn (RWAQ). Toisena päivänä osallistujat suorittivat lukemiseen liittyviä testejä ja vastasivat matematiikkaahdistuskyselyyn (MAQ). Testien ja ahdistuskyselyjen päiväkohtaiset aiheet (matematiikka ja lukeminen) olivat ristikkäiset, jotta testien tekeminen ei vaikuttaisi lasten vastauksiin kyselyjen kysymyksiin. Tämän takia kyselyillä kerätyt otokset olivat kooltaan hyvin erilaisia (ks. seuraava kappale).

Molempina mittauspäivinä osallistujat vastasivat taustakyselyihin ennen varsinaisia testejä. Osa lapsista vastasi taustakyselyihin epäsystemaattisesti ja kyseiset vastaajat poistettiin aineistosta. Tämän jälkeen aineistoon jäi jäljelle 10035 lasta, joista MAQ-kyselyyn vastasi yhteensä 7410 lasta ja RWAQ-kyselyyn yhteensä 9222 lasta. Analyysieihin jätettiin vain ne nuoret, joille kyselyt teetettiin suomen kielellä ja jotka eivät olleet yhdysluokkalaisia<sup>13</sup>. Täten lopulliset otokoot olivat 6362 (MAQ) ja 7837 (RWAQ). Aineistoissa ei ollut yhtään puuttuvia arvoja.

Vastaajat koostuivat seitsemän eri ikäluokan lapsista peruskoulutasolla. MAQ-kyselyyn vastanneista lapsista alakoulun kolmannella vuosiluokalla olevia oli yhteensä 849, neljännellä 988, viidennellä 1062 ja kuudennella 992. Yläkoulun 7.-luokkalaisia oli yhteensä 875, 8.-luokkalaisia 841 ja 9.-luokkalaisia 755. RWAQ-kyselyn vastaavat luvut olivat 970 (3. luokka), 1180 (4. luokka), 1281 (5. luokka), 1278 (6. luokka), 1092 (7. luokka), 998 (8. luokka) ja 1038 (9. luokka). MAQ-kyselyn vastaajista 3252 oli tyttöjä ja 3110 oli poikia. RWAQ-kyselyn vastaavat luvut olivat 3975 tyttöä ja 3862 poikaa.

### 4.1.2 Kyselyt

Matematiikka- (MAQ) sekä lukemis- ja kirjoittamisahdistusta (RWAQ) mitattiin kyselyillä, joiden ensimmäiset versiot luotiin Careyn et al. [9] mAMAS-mittaria mukailleen. Careyn ja kumppaneiden mittari koostui yhdeksästä osiosta (kysymyksestä), joihin lapset vastasivat viiden pisteen Likert-asteikolla osoittaen kuinka ahdistuneelta heistä tuntuisi joissakin tietyissä matematiikkaan liittyvissä tilanteissa (1 = matala ahdistus, low anxiety; 5 = korkea ahdistus, high anxiety). Carey kumppaneineen validoivat mAMAS-mittarille kahden faktorin rakenteen, joka koostui oppimistilanne- ja koetilannefaktorista.

MAQ- ja RWAQ-kyselyjen ensimmäisiä versioita testattiin keväällä 2021. MAQ-kyselyyn oli otettu mAMAS-mittarista [9] suoraan kaikki sen yhdeksän osiota, jotka käännettiin suomen kielelle. MAQ-kyselyn huomattava ero mAMAS-mittariin oli kaksi uutta sosiaalisiin tilanteisiin liittyvää kysymystä. Mielenkiintona oli selvittää muodostivatko nämä kaksi kysymystä matematiikan oppimis- ja koeahdistuslottuvuuksista erottuvan sosiaalisen matematiikka-ahdistuneisuuden ulottuvuuden. RWAQ-kysely oli suora kopio MAQ-kyselystä, mutta sen kysymykset oli muokat-

<sup>13</sup>Yhdysluokissa on lapsia eri ikäluokista.

tu vastaamaan lukemiseen tai kirjoittamiseen liittyviä tilanteita. Lapset vastasivat kysymyksiin viiden pisteen Likert-asteikolla, jonka vastausvaihtoehdot olivat seuraavanlaiset: (0) innostunut, (1) rauhallinen, (2) hiukan jännittynyt tai ahdistunut, (3) jännittynyt tai ahdistunut ja (4) erittäin jännittynyt tai ahdistunut.

Kevään 2021 analyysien tulokset osoittivat, että edellä mainitut kaksi uutta kysymystä muodostivat muista ulottuvuuksista erottuvan sosiaaliseen ahdistuneisuuteen liittyvän ulottuvuuden sekä MAQ- että RWAQ-kyselyn tapauksessa. Lisäksi tulokset osoittivat, että lukemis- ja kirjoittamisahdistuneisuuden kolmas ulottuvuus oppimisahdistuksen sijasta saattaakin olla lukemisen ja kirjoittamisen työmäärään liittyvä ahdistuneisuus. Näiden tulosten perusteella rakennettiin uudet versiot MAQ- ja RWAQ-kyselyistä, joita käytettiin syksyn 2021 FUNA-mittauksissa. MAQ-kysely jakautui kolmeen alakyselyyn, joilla haluttiin mitata matematiikkaahdistuksen eri ulottuvuuksia: osiot MAL1-MAL4 muodostivat oppimistilanteisiin liittyvän kyselyn (Learning), osiot MAE1-MAE4 koe-/arviointitilanteisiin liittyvän kyselyn (Evaluation) ja osiot MAS1-MAS4 sosiaalisiin tilanteisiin liittyvän kyselyn (Social). Myös RWAQ jakautui kolmeen alakyselyyn, mutta oppimistilanteisiin liittyvän alakyselyn tilalla oli edellä mainittuun työmääräahdistuneisuuteen liittyvä alakysely (Workload). Lisäksi RWAQ:n alakyselyiden sisällä osiot jakautuivat vielä kahteen lukemiseen (r-alkuiset osiot) ja kahteen kirjoittamiseen (w-alkuiset osiot) liittyvään tilanteeseen. Kummassakin kyselyssä oli siis 12 kysymystä, joihin lapset vastasivat samalla asteikolla kuin kyselyiden ensimmäisissä versioissa. Kysymykset näkyvät taulukoissa 1 ja 2.

Taulukko 1: MAQ-kyselyn osiot alakyselyittäin.

<i>Alakysely</i>	<i>Osio</i>	<i>Kysymys: Miltä sinusta tuntuu, kun...</i>
Learning	MAL1	sinun pitää tehdä sivu laskutehtäviä valmiiksi?
	MAL2	katselit opettajan ratkaisevan matematiikan tehtävää taululla?
	MAL3	kuuntelet opettajan puhuvan pitkään matematiikasta?
	MAL4	matematiikassa aloitetaan uusi aihe?
Evaluation	MAE1	ajattelet seuraavana päivänä pidettävää matematiikan koetta?
	MAE2	teet matematiikan koetta?
	MAE3	saat tietää tunnin alussa, että teillä on yllätyksenä matematiikan kokeet?
	MAE4	harjoittelet matematiikan koetta varten?
Social	MAS1	opettaja kysyy sinulta vastausta matematiikan tehtävään?
	MAS2	sinun pitää näyttää luokalle, miten ratkaisit matematiikan tehtävän?
	MAS3	opettaja katsoo, miten lasket laskutehtävää tunnilla?
	MAS4	joku kysyy sinulta vastausta laskutehtävään?

Taulukko 2: RWAQ-kyselyn osiot alakyselyittäin.

<i>Alakysely</i>	<i>Osio</i>	<i>Kysymys: Miltä sinusta tuntuu, kun...</i>
Workload	RAW1	sinun pitää koulussa lukea monta sivua?
	RAW2	sinulle annetaan seuraavaksi päiväksi pitkiä lukemistehtäviä kotiläksyksi?
	WAW1	sinun pitää kirjoittaa koulussa pitkä kirjoitelma tai essee?
	WAW2	sinulla on kotiläksyksi kirjoitettava monta pitkää vastausta kysymyksiin?
Evaluation	RAE1	kuulet, että kaikille tehdään lukutaitojen testi koulussa?
	RAE2	opettaja arvioi, osaatko lukea vaikeita sanoja oikein?
	WAE1	kuuntelet opettajan selittävän oikeinkirjoitussäännöistä?
	WAE2	opettaja on palauttamassa ainekirjoituksen kokeita takaisin?
Social	RAS1	opettaja pyytää sinua lukemaan ääneen koko luokalle?
	RAS2	sinun pitää lukea kirjasta ääneen kokonainen sivu osana ryhmätyötä?
	WAS1	sinun pitää kirjoittaa ryhmätyön tulokset paperille niiden esittämistä varten?
	WAS2	sinun pitäisi mennä luokan eteen kirjoittamaan lause taululle?

#### 4.1.3 Proseduuri

Kyselyt teetettiin osana syksyn 2021 FUNA-mittauksia [32], joissa lapset tekivät ahdistuskyselyjen lisäksi muita FUNA-tehtävistöön liittyviä testejä. Mittaukset toteutettiin hankeyhteistyössä Turun yliopiston Oppimisanalytiikan tutkimusinstituutin [33], Konsti-hankkeen [22] ja KouluKunnossa-hankkeen [23] kesken. Lapset osallistuivat testien tekemiseen anonymisti tavallisen koulupäivän aikana opettajien valvonnan alla. FUNA-testit toteutettiin sähköisesti Turun yliopiston tulevaisuuden teknologioiden laitoksella kehitetyssä ViLLE-oppimisjärjestelmässä, jota käytetään laajasti eri koulutusasteilla. Sen on todettu parantavan oppimista [24], [25], ja muun muassa eri osatekijöiden vaikutuksia on selvitetty yliopistotason tutkimuksissa [26], [47]. ViLLE-järjestelmä sai arvostetun UNESCO-palkinnon vuonna 2020 [46]. Opettajat jakoivat testeihin osallistujille käyttäjätunnukset, joilla he pääsivät kirjautumaan ViLLE:een internet-selaimessa omilla tietokoneillaan tai tableteillaan.

#### 4.1.4 Analyysi

MAQ- ja RWAQ-kyselyjen analysointi koostui niiden faktorirakenteiden tutkimisesta ja mittausinvarianssin testaamisesta Mplus-ohjelman versiolla 8.6 [31]. Kyselyiden faktorirakenteita tutkittiin konfirmatiivisella faktorianalyysillä. Kummallekin kyselylle oletettiin alakyselyjaottelun mukaiset kolmen faktorin rakenteet, joissa faktorien sallittiin korreloida keskenään. Latenttien muuttujien mitat määritettiin asettamalla jokaisen faktorin ensimmäinen lataus ykköseksi kummankin kyselyn tapauksessa. Kyselyiden mittausinvarianssia testattiin sukupuolen ja luokka-asteen suhteen kahdessa vaiheessa. Ensin testattiin konfiguraalista invarianssia tutkimalla, olivatko ryhmäkohtaisten faktorimallien sopivuudet aineistoon riittäviä. Tämän

jälkeen testattiin metristä ja skalaarista invarianssia toteuttamalla konfirmatiivinen moniryhmäfaktorianalyysi sukupuolten sekä luokka-asteiden suhteen ja sitten vertailemalla konfiguraalisen, metrisen sekä skalaarisen mallin eroja RMSEA- ja CFI-sopivuuksindekseissä. Kaikki mallit estimoitiin käyttäen suurimman uskottavuuden menetelmää olettamalla muuttujien olevan normaalisti jakautuneita.

Estimoitujen faktorimallien sopivuuden arviointi perustui RMSEA-, TLI- ja CFI-indeksien estimaatteihin. Sopivuuksindekseille valittiin etukäteen kynnsarvoihin perustuvat säännöt, joiden avulla määritettiin tutkittavan mallin sopivuuden taso. RMSEA-indeksin tapauksessa kynnsarvot olivat  $\hat{\epsilon} < 0,06$  (hyvä sopivuus);  $0,06 \leq \hat{\epsilon} < 0,08$  (hyväksyttävä sopivuus);  $0,08 \leq \hat{\epsilon} < 1$  (huono sopivuus) ja  $\hat{\epsilon} \geq 1$  (käyttökelvoton malli). TLI- ja CFI-indeksien tapauksessa estimaattien arvot alle luvun 0,95 indikoivat huonosta sopivuudesta aineistoon; arvot välillä  $[0,95; 0,97]$  hyväksyttävästä sopivuudesta ja arvot, jotka olivat vähintään 0,97; indikoivat hyvästä sopivuudesta. Nämä kynnsarvot perustuivat Brownen ja Cudeckin [8], Hun ja Bentlerin [19] sekä Schermellehin ja Moosbruggerin [39] tutkimuksiin, ja kuten luvussa 2.4.3 jo mainittiin, ne ovat vain suuntaa antavia suosituksia.

Sopivuuksindeksien estimaattorien tarkkuuden arvioinnissa hyödynnettiin epäparametrinen saapasremmimenetelmää. Saapasremmi (engl. bootstrap) on uudelleenotantamenetelmä, joka voidaan toteuttaa seuraavanlaisella Monte Carlo -algoritmilla:

1. arvo satunnaisesti palauttaen alkuperäisestä otoksesta tarpeeksi suuri määrä ( $B$ ) alkuperäisen otoksen suuruisia saapasremmiotoksia,
2. laske kustakin saapasremmiotoksesta mielenkiinnon kohteena olevan tunnusluvun (kuten sopivuuksindeksin) estimaatti, ja
3. arvioi tunnusluvun estimaattorin tilastollisia ominaisuuksia hyödyntäen kohdassa 2 laskettua  $B$ :tä estimaattia. [12]

Tässä tapauksessa algoritmin kohdassa 3 kullekin sopivuuksindeksille laskettiin saapasremmiestimaatti, -keskivirhe ja -harha sekä 90 %:n saapasremmiluottamusväli  $BC_a$  (bias corrected and accelerated) -menetelmällä [12]. Saapasremmiotosten lukumääräksi valittiin  $B = 10000$ , joka on kymmenkertainen Monte Carlo -saapasremmiotosten ”karkeaan minimiin” verrattuna, kun estimaattorin tarkkuutta arvioidaan luottamusvälin avulla [12]. Koska RMSEA-indeksin tapauksessa oli mahdollista laskea havaittuun otokseen perustuva parametrinen luottamusväli, niin erityisen mielenkiintoista oli verrata tätä havaittua luottamusväliä ja saapasremmiluottamusväliä keskenään. Saapasremmimenetelmä toteutettiin R 4.0.5 -ohjelmalla [35]. Saapasremmiestimaatit laskettiin hyödyntäen R-pakettia MplusAutomation [13].

## 4.2 Tulokset

### 4.2.1 Kuvailevat tunnusluvut

Koko MAQ-kyselyn keskimääräinen kokonaispistemäärä oli 16,27 (keskihajonta = 8,38). Sen sijaan koko RWAQ-kyselyn keskimääräinen kokonaispistemäärä oli 16,43 (keskihajonta = 7,86). MAQ-alakyselyjen keskimääräiset kokonaispistemäärät olivat Learning: 4,18 (keskihajonta = 2,57), Evaluation: 6,38 (keskihajonta = 3,55)

Taulukko 3: MAQ- ja RWAQ-kyselyjen kuvailevat tunnusluvut sukupuolittain sekä luokka-asteittain.

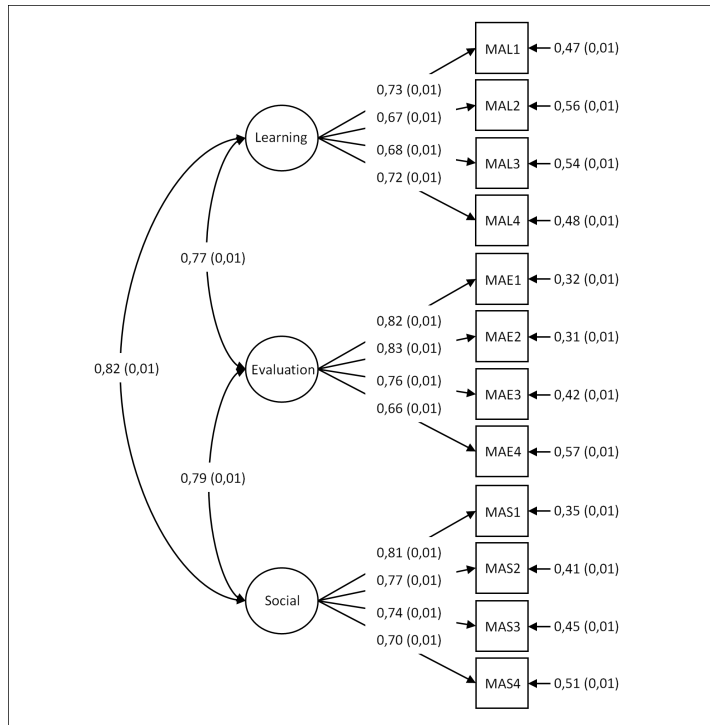
		MAQ								
		Koko kysely		Learning		Evaluation		Social		
		<i>n</i>	<i>M</i>	<i>Sd</i>	<i>M</i>	<i>Sd</i>	<i>M</i>	<i>Sd</i>	<i>M</i>	<i>Sd</i>
Sukupuoli										
	Tyttö	3252	18,19	8,44	4,38	2,60	7,37	3,58	6,44	3,49
	Poika	3110	14,27	7,82	3,97	2,54	5,34	3,21	4,96	3,03
Luokka-aste										
	3lk	849	13,84	8,30	3,57	2,65	5,16	3,64	5,11	3,33
	4lk	988	14,56	7,82	3,70	2,47	5,67	3,48	5,20	3,17
	5lk	1062	14,83	7,96	3,78	2,40	5,83	3,35	5,22	3,22
	6lk	992	15,88	7,08	3,99	2,04	6,33	3,15	5,55	2,99
	7lk	875	17,85	8,29	4,53	2,50	7,05	3,43	6,27	3,48
	8lk	841	18,75	8,38	4,77	2,60	7,47	3,48	6,51	3,48
	9lk	755	19,19	9,45	5,22	3,01	7,50	3,70	6,47	3,58
		RWAQ								
		Koko kysely		Workload		Evaluation		Social		
		<i>n</i>	<i>M</i>	<i>Sd</i>	<i>M</i>	<i>Sd</i>	<i>M</i>	<i>Sd</i>	<i>M</i>	<i>Sd</i>
Sukupuoli										
	Tyttö	3975	17,16	7,98	5,63	3,05	5,33	2,77	6,19	3,64
	Poika	3862	15,68	7,66	5,64	3,07	4,65	2,65	5,39	3,12
Luokka-aste										
	3lk	970	15,00	8,09	5,11	3,24	4,49	2,84	5,39	3,45
	4lk	1180	15,43	7,57	5,34	3,09	4,66	2,70	5,42	3,39
	5lk	1281	15,03	7,44	5,22	2,89	4,63	2,60	5,17	3,23
	6lk	1278	15,91	7,05	5,47	2,77	4,90	2,50	5,54	3,20
	7lk	1092	17,95	8,29	6,04	3,13	5,42	2,83	6,48	3,55
	8lk	998	18,13	7,65	6,17	2,99	5,46	2,65	6,49	3,39
	9lk	1038	18,05	8,28	6,23	3,16	5,51	2,85	6,30	3,50

*Huomioita.* *n* = otoskoko; *M* = keskiarvo; *Sd* = keskihajonta.

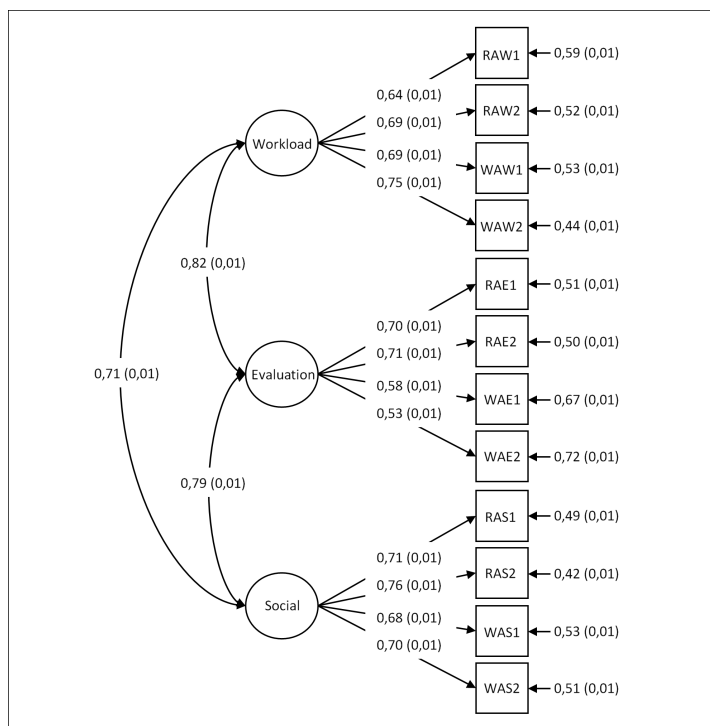
ja Social: 5,72 (keskihajonta = 3,36). RWAQ-alakyselyjen vastaavat luvut olivat Workload: 5,64 (keskihajonta = 3,06), Evaluation: 5,00 (keskihajonta = 2,73) ja Social: 5,80 (keskihajonta = 3,42). Taulukossa 3 näkyvät kyselyjen kuvailevat tunnusluvut sukupuolittain sekä luokka-asteittain. Osiokohtaiset kuvailevat tunnusluvut ovat liitteessä C.

#### 4.2.2 Kyselyiden faktorirakenteet

Kuvissa 2 ja 3 esitetään estimoitujen mallien polkukuviot. Faktorilatausten estimaatti on standardisoitu sekä faktorien että havaittujen muuttujien keskihajontojen suhteen kaavalla  $\hat{\lambda}_{std} = \hat{\lambda} \times \sqrt{\hat{\phi}} / \sqrt{\text{Var}(x)}$ , jossa  $\hat{\lambda}$  on ei-standardisoitu faktorilatausestimaatti,  $\hat{\phi}$  on estimaatti faktorilatausta  $\lambda$  vastaavan faktorin varianssille ja  $\text{Var}(x)$  on faktorilatausta  $\lambda$  vastaavan havaitun muuttujan  $x$  otosvariassi [4]. Standardisoinnin takia kuvissa näkyvät faktorien väliset estimoidut korrelaatiot kovarianssien sijasta. Estimaattien keskivirheet ovat niiden vieressä sulkumerkkien sisällä.



Kuva 2: MAQ-kyselyn konfirmatiivinen faktorianalyysi: estimoidun faktorimallin polkukuvio.



Kuva 3: RWAQ-kyselyn konfirmatiivinen faktorianalyysi: estimoidun faktorimallin polkukuvio.

MAQ-mallin kaikki parametriestimaatit erosivat tilastollisesti merkitsevästi nol-  
lasta 0,1 %:n merkitsevyystasolla. Kuten oli odotettavissa, suuren otoskoon ta-  
kia  $\chi^2$ -testin perusteella malli erosi tilastollisesti merkitsevästi ideaalista mallista:  
 $\chi^2(51) = 1827,79$ ;  $p < 0,001$ . Sen sijaan RMSEA-, CFI- ja TLI-sopivuusindeksit  
osoittivat mallin sopivuuden olevan riittävä: RMSEA-indeksin estimaatti oli 0,074  
(90 %:n luottamusväli oli 0,071-0,077), CFI = 0,953 ja TLI = 0,940.

Myös RWAQ-mallin kaikki parametriestimaatit erosivat tilastollisesti merkitse-  
västi nolasta 0,1 %:n merkitsevyystasolla. Tässäkin tapauksessa suuren otoskoon  
takia  $\chi^2$ -testin perusteella malli erosi tilastollisesti merkitsevästi ideaalista mallista  
( $\chi^2(51) = 1992,66$ ;  $p < 0,001$ ). RMSEA-estimaatti 0,070 (90 %:n luottamusväli oli  
0,067-0,072) osoitti mallin sopivuuden olevan riittävä. Sen sijaan CFI = 0,943 ja  
TLI = 0,927 indikoivat, että malli ei sopinut aineistoon hyvin.

### 4.2.3 Sopivuusindeksiestimaattorien tarkkuuden arviointi

Sopivuusindeksien 10000 saapasremmiotokseen perustuneet saapasremmiestimaatit,  
-luottamusvälit (90 %), -keskivirheet ja -harhat ovat esillä taulukossa 4. Sekä MAQ-  
että RWAQ-mallin tapauksessa sopivuusindeksien saapasremmiestimaatit olivat lä-  
hellä havaitusta otoksesta laskettuja estimaatteja; saapasremmimenetelmällä laske-  
tut harhat olivat itseisarvoltaan enintään 0,003 ja lisäksi saapasremmikeskivirheet  
olivat pieniä (enintään 0,004).  $BC_a$ -menetelmällä lasketut 90 %:n saapasremmiluo-  
ttamusvälit sisälsivät havaitusta otoksesta lasketut estimaatit molemmilla malleil-  
la kaikkien sopivuusindeksien tapauksessa. Mikäli mallien hyvyyden arviointi olisi  
perustunut pelkästään sopivuusindeksien saapasremmiestimaatteihin, niin päättely  
olisi ollut samanlaista kuin havaitusta aineistosta laskettuihin estimaatteihin perus-  
tunut päättely luvussa 4.2.2.

Erityisen mielenkiintoista oli verrata RMSEA-indeksin havaitusta aineistosta las-  
kettuja 90 %:n luottamusvälejä ja  $BC_a$ -menetelmään perustuvia 90 %:n saapasrem-  
miluottamusvälejä keskenään. Osoittautui, että saapasremmiluottamusvälit olivat  
vain hieman leveämpiä. MAQ-mallin tapauksessa luottamusvälien alarajat olivat  
yhtä suuret kahden desimaalin tarkkuudella ja ylärajojen erotus oli itseisarvoltaan  
vain 0,002. RWAQ-mallin tapauksessa alarajojen erotus oli itseisarvoltaan 0,001 ja  
ylärajojen erotus 0,003.

RMSEA-indeksin tapauksessa molemmilla malleilla sekä havaittuun aineistoon  
perustuvan luottamusvälin että saapasremmiluottamusvälin yläraja oli alle 0,08; jo-  
ka on sopivuudeltaan hyväksyttävän mallin RMSEA-arvon yläraja. Tämä havainto  
tuki luvussa 4.2.2 tehtyjä RMSEA-indeksiin perustuvia päätelmiä mallien hyvyy-  
destä. Myöskään CFI- ja TLI-indeksien saapasremmiluottamusvälit eivät juurikaan  
muuttaneet CFI:n ja TLI:n perusteella tehtäviä päätelmiä mallien hyvyydestä. Huo-  
mionarvoista tosin oli MAQ-mallin tapauksessa CFI:n saapasremmiluottamusvälin  
alarajan arvo, joka oli alle luvun 0,95. Tämän perusteella ei voitu olla täysin var-  
moja, että CFI-populaatioarvo oli oikeasti vähintään 0,95; joka on sopivuudeltaan  
hyväksyttävän mallin CFI-arvon alaraja.



Taulukko 4: MAQ- ja RWAQ-mallien sopivuusindeksien saapasremmiestimaatit, -luottamusvälit (90 %), -keskivirheet ja -harhat.

MAQ						
Sopivuus- indeksi	Estimaatti		90 % <i>lv</i>		<i>se<sub>BS</sub></i>	<i>b<sub>BS</sub></i>
	<i>est<sub>hav</sub></i>	<i>est<sub>BS</sub></i>	<i>lv<sub>hav</sub></i>	<i>lv<sub>BS</sub></i>		
RMSEA	0,074	0,076	0,071-0,077	0,071-0,079	0,003	0,002
CFI	0,953	0,951	-	0,947-0,957	0,003	-0,002
TLI	0,940	0,937	-	0,929-0,943	0,004	-0,003

RWAQ						
Sopivuus- indeksi	Estimaatti		90 % <i>lv</i>		<i>se<sub>BS</sub></i>	<i>b<sub>BS</sub></i>
	<i>est<sub>hav</sub></i>	<i>est<sub>BS</sub></i>	<i>lv<sub>hav</sub></i>	<i>lv<sub>BS</sub></i>		
RMSEA	0,070	0,071	0,067-0,072	0,068-0,075	0,002	0,001
CFI	0,943	0,941	-	0,936-0,947	0,003	-0,002
TLI	0,927	0,924	-	0,916-0,930	0,004	-0,003

*Huomioita.* *Estimaatti* = havaitusta aineistosta laskettu estimaatti (*est<sub>hav</sub>*) tai saapasremmiestimaatti (*est<sub>BS</sub>*); 90 % *lv* = havaitusta aineistosta laskettu 90 %:n luottamusväli (*lv<sub>hav</sub>*) tai  $BC_a$ -menetelmällä laskettu 90 %:n saapasremmiluottamusväli (*lv<sub>BS</sub>*); *se<sub>BS</sub>* = saapasremmikeskivirhe; *b<sub>BS</sub>* = saapasremmiharha.

#### 4.2.4 Kyselyiden mittausinvarianssitestaukset

Ryhmäkohtaisten MAQ- ja RWAQ-mallien sopivuusindeksit näkyvät taulukoissa 5 ja 6. MAQ-mallien sopivuusindeksien perusteella konfiguraalinen invarianssi toteutui ainakin sukupuolten suhteen. Luokka-asteiden suhteen konfiguraalinen invarianssi ei toteutunut, sillä erityisesti luokka-asteiden 7-9 mallien sopivuusindeksit indikoivat niiden huonosta sopivuudesta aineistoon (RMSEA-estimaatit olivat yli 0,1 ja molemmat inkrementaaliset sopivuusindeksit olivat alle 0,95). Näiden tarkastelujen perusteella eri luokka-asteiden oppilaat käsittivät MAQ-kysymykset eri tavalla, jolloin eri luokka-asteilla mitattiin erilaisia faktorirakenteita.

RWAQ-mallien sopivuusindeksien perusteella konfiguraalinen invarianssi ei toteutunut ainakaan luokka-asteiden suhteen (jälleen luokka-asteiden 7-9 mallien sopivuudet aineistoon olivat huonoja), ja sen toteutuminen sukupuolten suhteen oli myös kyseenalaista, sillä tyttöjen mallin molemmat inkrementaaliset sopivuusindeksit olivat alle 0,95. Näiden tarkastelujen perusteella sekä tytöt ja pojat että eri luokka-asteiden oppilaat käsittivät RWAQ-kysymykset eri tavalla.

Konfiguraalisten invarianssitestausten perusteella invarianssitestauksia olisi ollut mielekästä jatkaa vain MAQ-mallin tapauksessa ja vain sukupuolten suhteen. Harjoituksen vuoksi metristä ja skalaarista invarianssia testattiin kuitenkin MAQ-mallin tapauksessa myös luokka-asteiden suhteen ja RWAQ-mallin tapauksessa sukupuolten sekä luokka-asteiden suhteen. Pitää vain muistaa, että täysi mittausinvarianssi ei voinut toteutua näissä tapauksissa vaikka metrinen ja skalaarinen invarianssi olisivatkin toteutuneet.

Koska otoskoko oli yli 300, niin luvussa 3.2.2 esitellyn testausmenettelyn mukaisesti sisäkkäisten invarianssimallien sopivuusindeksivertailuissa kynnyksarvot  $\Delta\text{RMSEA} < 0,015$  ja  $\Delta\text{CFI} > -0,010$  yhdessä osoittivat metrisen ja/tai skalaarisen invarianssin toteutumisesta. Metrinen ja skalaarinen invarianssi toteutuivat

MAQ-mallille sekä sukupuolten että luokka-asteiden suhteen:  $\Delta RMSEA \leq 0,004$  ja  $\Delta CFI \geq -0,004$ , kun vertailtiin konfiguraalisia ja metrisiä malleja keskenään (metrinen invarianssi), ja  $\Delta RMSEA \leq 0,001$  ja  $\Delta CFI \geq -0,007$ , kun vertailtiin metrisiä ja skalaarisia malleja keskenään (skalaarinen invarianssi). RWAQ-mallin tapauksessa metrinen invarianssi toteutui sekä sukupuolten että luokka-asteiden suhteen:  $\Delta RMSEA \leq 0,004$  ja  $\Delta CFI \geq -0,003$ , kun vertailtiin konfiguraalisia ja metrisiä malleja keskenään. Skalaarisen invarianssin toteutuminen oli kyseenalaista sekä sukupuolten että luokka-asteiden suhteen, sillä  $\Delta CFI \leq -0,010$ , kun vertailtiin metrisiä ja skalaarisia malleja keskenään.

Invarianssitestausten perusteella MAQ-kysely oli täysin mittausinvariantti sukupuolten suhteen ja olisi ollut myös luokka-asteiden suhteen, jos konfiguraalinen invarianssi olisi toteutunut luokka-asteiden yli. RWAQ-kysely ei ollut täysin invariantti sukupuolten eikä luokka-asteiden suhteen, mikä nähtiin jo konfiguraalisen invarianssin testaamisessa. Vain MAQ-kyselyn faktoripisteitä eri sukupuolilla voitaisiin siis vertailla luotettavasti, mutta se ei kuulu tämän tutkielman aiheisiin. Kumpaan kyselyä on syytä tutkia jatkossa vielä tarkemmin, jotta saadaan koottua reliabeilit ja validit ahdistuskyselykokonaisuudet, joilla voidaan mitata eri ikäisten lasten matematiikka- ja lukemis- sekä kirjoittamisahdistuneisuutta ja vertailla luotettavasti sukupuolten lisäksi eri luokka-asteiden ja eri kieliryhmien välisiä eroja ahdistuneisuuden tasossa. Kyselyjen jatkotutkimuksesta puhutaan lisää luvussa 5.

Taulukko 5: Ryhmäkohtaisten MAQ-mallien sopivuusindeksit

	<i>n</i>	$\chi^2$ -testi			<i>RMSEA</i>		<i>CFI</i>	<i>TLI</i>
		$\chi^2$	<i>v</i>	<i>p-arvo</i>	$\hat{\epsilon}$	90 % <i>lv</i>		
Sukupuoli								
Tyttö	3252	888,51	51	<0,001	0,071	0,067-0,075	0,955	0,942
Poika	3110	946,97	51	<0,001	0,075	0,071-0,079	0,950	0,936
Luokka-aste								
3lk	849	173,12	51	<0,001	0,053	0,045-0,062	0,963	0,953
4lk	988	271,11	51	<0,001	0,066	0,058-0,074	0,948	0,933
5lk	1062	346,11	51	<0,001	0,074	0,067-0,081	0,953	0,940
6lk	992	266,67	51	<0,001	0,065	0,058-0,073	0,960	0,948
7lk	875	585,21	51	<0,001	0,109	0,102-0,117	0,916	0,891
8lk	841	550,95	51	<0,001	0,108	0,100-0,116	0,914	0,888
9lk	755	558,86	51	<0,001	0,115	0,106-0,124	0,927	0,906

*Huomioita.*  $\chi^2$  = testisuureen arvo; *v* = vapausasteet;  $\hat{\epsilon}$  = RMSEA-indeksin estimaatti; 90 % *lv* = RMSEA-estimaatin 90 %:n luottamusväli.

Taulukko 6: Ryhmäkohtaisten RWAQ-mallien sopivuusindeksit

	$n$	$\chi^2$ -testi			RMSEA		CFI	TLI
		$\chi^2$	$v$	$p$ -arvo	$\hat{\epsilon}$	90 % $lv$		
Sukupuoli								
Tyttö	3975	1100,81	51	<0,001	0,072	0,068-0,076	0,937	0,919
Poika	3862	844,13	51	<0,001	0,063	0,060-0,067	0,955	0,941
Luokka-aste								
3lk	970	223,82	51	<0,001	0,059	0,051-0,067	0,946	0,930
4lk	1180	258,25	51	<0,001	0,059	0,052-0,066	0,947	0,931
5lk	1281	391,04	51	<0,001	0,072	0,066-0,079	0,939	0,921
6lk	1278	430,41	51	<0,001	0,076	0,070-0,083	0,926	0,904
7lk	1092	412,29	51	<0,001	0,081	0,073-0,088	0,941	0,924
8lk	998	397,70	51	<0,001	0,083	0,075-0,090	0,927	0,906
9lk	1038	554,85	51	<0,001	0,098	0,090-0,105	0,917	0,893

*Huomioita.*  $\chi^2$  = testisuuren arvo;  $v$  = vapausasteet;  $\hat{\epsilon}$  = RMSEA-indeksin estimaatti; 90 %  $lv$  = RMSEA-estimaatin 90 %:n luottamusväli.

## 5 Pohdintaa

Edellä esiteltiin lineaaristen faktorimallien ja mittausinvarianssin teoriaa, ja sitten kyseisiä menetelmiä sovellettiin oikeaan empiiriseen aineistoon. Faktorimallien teoriassa lähdettiin liikkeelle rajoittamattomien faktorimallien määrittelystä ja parametrien identifioituvuudesta, josta oli luontevaa siirtyä rajoitettuihin faktorimalleihin. Rajoitettujen faktorimallien yhteydessä puhuttiin määrittelyn ja parametrien identifioituvuuden lisäksi mallin estimoinnista ja lopuksi esiteltiin yksinkertainen esimerkki rajoitetusta faktorimallista. Rajoitetuille faktorimalleille esiteltiin erikseen niiden ryhmitellyt versiot, joita tarvittiin faktorisen invarianssin testaamisessa. Faktorimallien teoriassa lopuksi puhuttiin mallien sopivuuden testaamisesta ja mallien uudelleenmäärittelystä. Mittausinvarianssin teoriassa esiteltiin mittausinvarianssin tilastollinen määritelmä, minkä jälkeen kerrottiin miten faktorista invarianssia testataan.

Tutkielman empiirinen osio koostui rajoitettujen lineaaristen faktorimallien soveltamisesta syksyn 2021 FUNA-mittauksissa kerättyihin aineistoihin ja faktorisen invarianssin testaamisesta sovelletuille malleille. Aineistot koostuivat peruskoulutason luokka-asteilla 3-9 olevista lapsista useista eri kouluista Varsinais-Suomessa ja Uudellamaalla. Lapset vastasivat muun FUNA-tehtävistön lisäksi matematiikka- sekä lukemis- ja kirjoittamisahdistuskyselyihin, joita analysoitiin tässä tutkielmassa. Kevään 2021 aikaisempien analyysien tulosten perusteella kyselyille rakennettiin omat teoreettiset faktorimallinsa, joista kummassakin oli kolme faktoria ja 12 indikaattoria. MAQ-mallin faktorirakenne koostui faktoreista, jotka mittasivat oppimistilanteisiin liittyvää ahdistusta (Learning), koetilanteisiin liittyvää ahdistusta (Evaluation) ja sosiaalisin tilanteisiin liittyvää ahdistusta (Social). RWAQ-mallin faktorirakenne koostui MAQ-mallin tapaan Evaluation- ja Social-faktoreista, mutta Learning-faktorin tilalla oli Workload-faktori, joka mittasi lukemisen ja kirjoittamisen työmäärään liittyvää ahdistusta.

Teoreettisia faktorimalleja analysoitiin ensin konfirmatiivisella faktorianalyysillä. Sitten testattiin faktorimallien mittausinvarianssia kahdessa vaiheessa: ensimmäisessä vaiheessa testattiin konfiguraalista invarianssia jakamalla aineistot tyttöjen ja poikien sekä eri luokka-asteiden omiin ala-aineistoihin, jonka jälkeen faktorimallit sovitettiin näihin ala-aineistoihin ja sovitettujen mallien sopivuutta aineistoissa arvioitiin; toisessa vaiheessa testattiin metristä ja skalaarista invarianssia asettamalla tiettyjä invarianssirajoitteita mallien parametreihin ja sitten tutkimalla sovitettujen invarianssimallien sopivuusindeksien eroja.

Jokainen malli estimoitiin suurimman uskottavuuden menetelmällä olettaen havaittujen muuttujien olevan normaalisti jakautuneita. Tämän oletuksen toteutuminen on vähintäänkin kyseenalaista ei-jatkuvien Likert-asteikollisten muuttujien tapauksessa. Lisäksi liitteen D esimerkkikuvat 4 ja 5 osioiden MAL1, MAE3, RAE1 ja WAE2 havaituista jakaumista havainnollistavat, miten MAQ- ja RWAQ-kyselyiden osiokohtaiset jakaumat olivat oikealle vinoja ja huipukkaita lukuunottamatta osion MAE3 jakaumaa, joka oli muodoltaan täysin erilainen kuin muut. Tässä tapauksessa normaalisuusoletus ei selvästikään toteutunut.

Yhtälössä (10) esitettiin suurimman uskottavuuden estimoinnissa käytettävä sovitusfunktio, jonka johtamisessa havaittujen muuttujien vektorin oletettiin olevan

multinormaalisti jakautunut. Bollen [4] esitti, miten havaittujen muuttujien jakauma vaikuttaa suurimman uskottavuuden estimaattorin ominaisuuksiin. Jos havaittujen muuttujien jakauma ei ole multinormaalijakauma, niin – riippuen jakauksen muodosta – estimaattori ei välttämättä ole asymptoottisesti tehokas ja sen asymptoottinen kovarianssimatriisi saattaa olla virheellinen. Lisäksi  $\chi^2$ -testisuureen  $(n - 1)F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\eta}))$  asymptoottinen jakauma ei välttämättä ole  $\chi^2$ -jakauma, jolloin  $\chi^2$ -testin tulokset ovat virheellisiä.

Koska RMSEA-, CFI- ja TLI-indeksien estimaatit yhtälöissä (18), (19) ja (20) riippuvat sovituskäytännön arvosta  $F(\mathbf{S}, \boldsymbol{\Sigma}(\hat{\eta}))$ , niin selvästi havaittujen muuttujien jakauma vaikuttaa myös niihin. Jos havaittujen muuttujien jakauma ei ole multinormaalijakauma, niin sopivuuksien estimaattorit eivät välttämättä mittaa haluttua asiaa. Luvussa 4.2 kuitenkin huomattiin, että havaittuun ei-normaaliseen aineistoon perustuvat RMSEA-, CFI- ja TLI-estimaatit olivat lähellä niitä vastaavia saapasremmiestimaatteja. Lisäksi RMSEA-indeksin havaittuun aineistoon perustuvat 90 %:n luottamusvälit olivat lähellä niitä vastaavia  $BC_a$ -menetelmällä laskettuja 90 %:n saapasremmiluottamusvälejä.

Vaikka normaalisuusoletus ei toteutunut, niin edellä tehtyjen huomioiden perusteella havaitusta aineistosta laskettuihin sopivuuksien estimaatteihin perustuva päättely mallien hyvyydestä oli luotettavaa. Konfirmatiivisen faktorianalyysin tulosten perusteella MAQ-mallin sopivuus aineistoon oli riittävä, mikä osoitti, että MAQ-kyselyn tapauksessa oli tarpeeksi vahvaa näyttöä rakennevaliditeetin (engl. construct validity) [1] toteutumiseksi. Täten kyselyn hypoteettista kolmen alakyselyn rakennetta ei voitu hylätä. Mainittakoon, että kun puhutaan MAQ-mallin sopivuudesta aineistoon, niin painotus on nyt sanassa 'riittävä', sillä RMSEA-estimaatin ja sen 90 %:n luottamusvälin perusteella mallin sopivuus aineistoon oli todellakin vain riittävä. Edelleen on kuitenkin syytä muistaa, että mallin sopivuuden hyvyydestä indikoivat sopivuuksien kynnyksarvot ovat vain suuntaa antavia suosituksia, ja huomioon pitää ottaa koko kokonaisuus. Mallin sopivuuden hyvyyttä puolsi CFI-indeksin korkea arvo. Tämän lisäksi kaikki estimoidut parametriestimaatit erosivat tilastollisesti merkitsevästi nolasta ja standardoidut faktorilataukset olivat korkeita (lähellä ykköstä). Tämä viittasi siihen, että kukin alakysely muodosti oman toimivan rakenteensa oikeasta joukosta kysymyksiä, mikä antoi lisää varmuutta sosiaalisen matematiikka-ahdistuneisuuden olemassaolosta omana erillisenä ulottuvuutenaan.

Myös RWAQ-mallin sopivuus oli riittävä RMSEA-estimaatin ja sen 90 %:n luottamusvälin perusteella. Tässä tapauksessa tosin sekä CFI- että TLI-indeksin arvo indikoivat mallin sopivuuden olevan riittämätön. Nämä tulokset viittasivat siihen, että RWAQ-kyselyn rakennevaliditeetti ei toteutunut. Erityisen mielenkiinnon kohteena olivat työmäärään ja sosiaalisiin tilanteisiin liittyvät lukemis- ja kirjoittamisahdistustyypit, joiden olemassaolosta omina erillisinä ulottuvuuksinaan ei voitu olla yhtä varmoja kuin sosiaalisen matematiikka-ahdistusulottuvuuden olemassaolosta. Niiden olemassaoloa puolsivat kuitenkin vastaavien faktorien korkeat faktorilataukset. RWAQ-mallia voisi yrittää modifioida luvussa 2.5 esiteltyä menetelmää hyödyntäen. Tämä tulee erityisesti olemaan mielenkiintoista RWAQ-kyselyn jatkotutkimuksessa, kun siitä yritetään luoda uusi paranneltu versio. Liitteessä E on laskettu RWAQ-mallin rajoitettujen parametrien MI- ja EPC-arvoja ja pohdittu, mitä ne kertovat RWAQ-kyselyn faktorirakenteesta.

Empiirisen osion mittausinvarianssitestausten ensimmäisessä vaiheessa todettiin, että toiseen vaiheeseen eli metrisen ja skalaarisen invarianssin testaamiseen oli mielekästä siirtyä vain MAQ-kyselyn tapauksessa ja vain silloin, kun aineisto ryhmiteltiin sukupuolittain, sillä konfiguraalinen invarianssi ei toteutunut MAQ-kyselylle luokka-asteiden suhteen ja RWAQ-kyselylle konfiguraalinen invarianssi ei toteutunut ollenkaan. Tästä huolimatta harjoituksen vuoksi täyttää mittausinvarianssia testattiin kummallakin kyselyllä ja kumpaakin ryhmäjaottelua käyttäen. Mittausinvarianssitestausten toisen vaiheen perusteella MAQ-kyselyn faktorirakenne oli täysin mittausinvariantti sukupuolten suhteen. Täten sukupuolten eroja MAQ-kyselyn faktoripisteissä olisi voinut vertailla luotettavasti ja siten tehdä päätelmiä mahdollisista eroista sukupuolten välillä matematiikka-ahdistuneisuuden tasoissa sen eri osa-alueilla (Learning, Evaluation, Social). Matematiikka-ahdistuneisuuden keskimääraisten tasojen on osoitettu konsistentisti olevan korkeampia tytöillä kuin pojilla [14]. Olisi mielenkiintoista tutkia, pätekö tämä tyttöjen ja poikien ero tässäkin tapauksessa.

Mittausinvarianssitestausten perusteella MAQ-kysely käsitettiin eri tavalla eri luokka-asteilla ja RWAQ-kysely käsitettiin eri tavalla sekä eri sukupuolilla että eri luokka-asteilla. Kumpaakin kyselyä on siis syytä tutkia tarkemmin, jotta pystytään kehittämään reliaabelit ja validit ahdistuskyselykokonaisuudet, jotka käsitetään samalla tavalla eri lapsiryhmissä (sukupuoli, ikä, kieli). On vielä epäselvää, mikä erityisesti RWAQ-kyselyssä aiheutti sen, että jopa eri sukupuolet käsittivät sen eri tavalla. Ovatko tyttöjen ja poikien lukemis- ja kirjoittamistottumukset jo peruskouluiässä niin erilaisia, että he kokevat lukemiseen ja/tai kirjoittamiseen liittyvää ahdistusta eri tavalla? Ainakin tämän tutkielman konfiguraalisen invarianssin testaamisen tulosten perusteella teorisoitu faktorirakenne toimi pojilla mutta tytöillä ei niinkään. Voinee siis olettaa, että juuri tytöt kokevat lukemis- ja/tai kirjoittamisahdistuksen eri tavalla kuin oletettiin. Voi myös olla niinkin, että lukemis- ja kirjoittamisahdistus muodostavat oikeasti omat erilliset kokonaisuutensa ja niitä ei voi mitata yhdessä, minkä takia saadaan huonosti aineistoon sopivia malleja. Tätäkin on syytä tutkia tarkemmin.

On myös vielä epäselvää, miksi eri luokka-asteiden lapset käsittivät kummatkin kyselyt eri tavalla. Konfiguraalisen invarianssin testaamisen tulosten perusteella sekä MAQ-malli että RWAQ-malli sopivat alakoulun luokka-asteiden (3-6) ala-aineistoihin vähintäänkin kohtuullisesti, mutta kaikilla yläkoulun luokka-asteilla (7-9) kumpikaan malli ei sopinut aineistoihin lainkaan. Erityisesti MAQ-kyselyn tapauksessa tämä oli hieman yllättävää, sillä Carey et al. [9] osoittivat, että heidän mAMAS-mittarinsa – johon MAQ-kysely osittain perustuu – mittasi samoja matematiikka-ahdistuneisuuden ulottuvuuksia Englannin ala-asteen (primary school) neljännen vuosiluokan (8-9-vuotiaat) ja yläasteen ja lukion (secondary school) seitsemännen ja kahdeksannen vuositasen (11-13-vuotiaat) lapsilla. Toisaalta, Suomen yläkoulutason luokka-asteiden 7-9 lapset ovat 12-16-vuotiaita ja Careyn et al. tutkimus ei kattanut 14-16-vuotiaita, joka vastaisi Englannin yläasteen ja lukion 10. ja 11. vuositasoja [5]. Lisäksi he yhdistivät seitsemännen ja kahdeksannen vuositasen lapset yhdeksi ikäryhmäksi, jolloin eri ikäryhmiä oli vain kaksi. Myös MAQ-kyselyn tapauksessa olisi voinut yhdistää esimerkiksi alakoulun ja yläkoulun lapset omiksi ikäryhmikseen. Lopuksi, ainakaan tämän tutkielman laatijalle ei ole

selkeää, miten matematiikan opetus eroaa suomalaisissa ja englantilaisissa kouluissa. Ovatko matematiikan opetustavat Suomessa ja Englannissa niin erilaisia ja etenevät eri tahdeissa, että tämän takia suomalaiset ja englantilaiset lapset kokevat matematiikka-ahdistusta eri tavalla, jolloin matematiikka-ahdistuksen mittaaminen suomalaisilla lapsilla vaatisi aivan uudenlaisen mittarin?

Mainittakoon vielä, että MAQ- ja RWAQ-kyselyjen jatkotutkimuksessa on mahdollista tutkia niiden osittaista invarianssia. Olisi mielenkiintoista löytää ne yksittäiset parametrit, jotka eivät ole invariantteja. Voi olla esimerkiksi niin, että erityisesti RWAQ-kyselyssä, joka ei ollut mittausinvariantti sukupuolten eikä luokka-asteiden suhteen, vain jokin tietty kysymysten osajoukko aiheuttaa ongelmia täyden mittausinvarianssin toteutumisessa. Tällöin RWAQ-kyselyn faktorirakenne olisi osittain invariantti ja tällainen havainto auttaisi kyselyn kehittämistä suuresti.

Kuten aikaisemmin mainittiin, erityisesti matematiikka-ahdistuksella ja matemaattisella osaamisella on todettu olevan yhteys toisiinsa. Lisäksi mainittiin, että lukemisahdistuksen ja lukemismenestyksen välistä yhteyttä ei ole vielä systemaattisesti tutkittu, vaikka lapsen aikainen pätevyys lukemisessa on tärkeä osa myöhempää akaatemista menestystä. Lienee siis perusteltua sanoa, että on äärimmäisen tärkeää kehittää luotettavat mittarit matematiikka- ja lukemisahdistuksen mittaamista varten lapsilla ja nuorilla. On vielä epäselvää, voidaanko lukemis- ja kirjoittamisahdistusta mitata yhtenäisenä kokonaisuutena vai pitääkö niitä mitata erikseen. Voi olla, että lukemis- ja kirjoittamisahdistus ovat hyvinkin vahvasti korreloituneita, mutta silti toisistaan erottuvat ahdistuneisuuden osa-alueet. Tämä tulee vaatimaan paljon jatkotutkimusta.

## Viitteet

- [1] American Psychological Association (2023) *APA Dictionary of Psychology: Construct Validity* [Online]. Saatavilla: <https://dictionary.apa.org/construct-validity> (Haettu: 16. helmikuuta 2023)
- [2] Anderson, T. W. (2003) *An Introduction to Multivariate Statistical Analysis*. 3. painos. Hoboken, NJ: John Wiley & Sons, Inc.
- [3] Babyak, M. A. & Green, S. B. (2010) 'Confirmatory Factor Analysis: An Introduction for Psychosomatic Medicine Researchers', *Psychosomatic Medicine*, 72(6), 587—597. <https://doi.org/10.1097/PSY.0b013e3181de3f8a>
- [4] Bollen, K. A. (1989) *Structural Equations with Latent Variables*. New York, NY: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118619179>
- [5] Bright World Education (2019) *British Education System* [Online]. Saatavilla: <https://www.brightworldguardianships.com/en/guardianship/british-education-system/> (Haettu: 16. helmikuuta 2023)
- [6] Brown, T. A. (2015) *Confirmatory Factor Analysis for Applied Research*. New York, NY: The Guilford Press.
- [7] Browne, M. W. (1973) 'Generalized Least Squares Estimators in the Analysis of Covariance Structures', *ETS Research Bulletin Series*, 1973(1), i—36. <https://doi.org/10.1002/j.2333-8504.1973.tb00197.x>
- [8] Browne, M. W. & Cudeck, R. (1992) 'Alternative Ways of Assessing Model Fit', *Sociological Methods & Research*, 21(2), 230—258. <https://doi.org/10.1177/0049124192021002005>
- [9] Carey, E., Hill, F., Devine, A. & Szűcs, D. (2017) 'The Modified Abbreviated Math Anxiety Scale: A Valid and Reliable Instrument for Use with Children', *Front. Psychol.*, 8(11). <https://doi.org/10.3389/fpsyg.2017.00011>
- [10] Chen, F. F. (2007) 'Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance', *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464—504. <https://doi.org/10.1080/10705510701301834>
- [11] Cipora, K., Santos, F. H., Kucian, K. & Dowker, A. (2022) 'Mathematics Anxiety—Where Are We and Where Shall We Go?', *Annals of the New York Academy of Sciences*, 1513(1), 10—20. <https://doi.org/10.1111/nyas.14770>
- [12] Efron, B. & Tibshirani, R. (1986) 'Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy', *Statistical Science*, 1(1), 54—75. <https://doi.org/10.1214/ss/1177013815>
- [13] Hallquist, M. N. & Wiley, J. F. (2018) 'MplusAutomation: An R Package for Facilitating Large-scale Latent Variable Analyses in Mplus', *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 621—638. <https://doi.org/10.1080/10705511.2017.1402334>



- [14] Hembree, R. (1990) 'The Nature, Effects, and Relief of Mathematics Anxiety', *Journal for Research in Mathematics Education*, 21(1), 33—46. <https://doi.org/10.2307/749455>
- [15] Hooper, D., Coughlan, J., & Mullen, M. (2008) 'Structural Equation Modelling: Guidelines for Determining Model Fit', *Electronic Journal of Business Research Methods*, 6(1), 53—60. <https://doi.org/10.21427/D7CF7R>
- [16] Hopko, D. R., Mahadevan, R., Bare, R. L. & Hunt, M. K. (2003) 'The Abbreviated Math Anxiety Scale (AMAS): Construction, Validity, and Reliability', *Assessment*, 10(2), 178—182. <https://doi.org/10.1177/1073191103010002008>
- [17] Hoyle, R. H. & Duvall, J. L. (2004) 'Determining the Number of Factors in Exploratory and Confirmatory Factor Analysis' teoksessa Kaplan, D. (toim.) *The SAGE Handbook of Quantitative Methodology for the Social Sciences*. Thousand Oaks, CA: SAGE Publications, Inc.; 302—317. <http://dx.doi.org/10.4135/9781412986311>
- [18] Huttunen, M. (2018) *Ahdistuneisuus* [Online]. Saatavilla: <https://www.terveyskirjasto.fi/dlk00188> (Haettu: 25. heinäkuuta 2022)
- [19] Hu, L. & Bentler, P. M. (1999) 'Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives', *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1—55. <https://doi.org/10.1080/10705519909540118>
- [20] Kaplan, D. (toim.) (2009) *Structural Equation Modeling: Foundations and Extensions*. 2. painos. Thousand Oaks, CA: SAGE Publications, Inc. <https://dx.doi.org/10.4135/9781452226576>
- [21] Kenny, D. A., Kaniskan, B. & McCoach D. B. (2015) 'The Performance of RMSEA in Models with Small Degrees of Freedom', *Sociological Methods & Research*, 44(3), 486—507. <https://doi.org/10.1177/0049124114543236>
- [22] Konsti-hanke (2021) *KONSTI* [Online]. Saatavilla: <https://sites.google.com/edu.uki.fi/konsti/etusivu> (Haettu: 14. heinäkuuta 2022)
- [23] KouluKunnossa-hanke (2020) *KOULUKUNNOSSA - SKOLSKICK* [Online]. Saatavilla: <https://www.koulukunnossa.fi/> (Haettu: 14. heinäkuuta 2022)
- [24] Kurvinen, E., Dagienė, V. & Laakso, M.-J. (2018) 'The Impact and Effectiveness of Technology Enhanced Mathematics Learning', *Constructionism*. <http://www.constructionism2018.fsf.vu.lt/proceedings>.
- [25] Laakso, M.-J., Kaila, E. & Rajala, T. (2018) 'ViLLE – Collaborative Education Tool: Designing and Utilizing an Exercise-based Learning Environment', *Educ. Inf. Technol.*, 23, 1655—1676. <https://doi.org/10.1007/s10639-017-9659-1>

- [26] Laakso, M.-J., Rajala, T., Kaila, E. & Salakoski, T. (2008) 'The Impact of Prior Experience in Using a Visualization Tool on Learning to Program. Proceedings of Cognition and Exploratory Learning in Digital Age (CELDA 2008)', *IEEE Technical Committee on Learning Technology and Japanese Society of Information and Systems in Education*. [A4]
- [27] Marsh, H. W., Hau, K.-T. & Wen, Z. (2004) 'In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings', *Structural Equation Modeling: A Multidisciplinary Journal*, 11(3), 320—341. [https://doi.org/10.1207/s15328007sem1103\\_2](https://doi.org/10.1207/s15328007sem1103_2)
- [28] McDonald, R. P. & Marsh, H. W. (1990) 'Choosing a Multivariate Model: Noncentrality and Goodness of Fit', *Psychological Bulletin*, 107(2), 247—255. <https://doi.org/10.1037/0033-2909.107.2.247>
- [29] Meredith, W. (1993) 'Measurement Invariance, Factor Analysis and Factorial Invariance', *Psychometrika*, 58, 525—543. <https://doi.org/10.1007/BF02294825>
- [30] Millsap, R. E. (2011) *Statistical Approaches to Measurement Invariance*. 1. painos. New York, NY: Routledge. <https://doi.org/10.4324/9780203821961>
- [31] Muthén, L. K. & Muthén, B. O. (1998-2017) *Mplus User's Guide*. 8. painos. Los Angeles, CA: Muthén & Muthén.
- [32] Oppimisanalytiikan tutkimusinstituutti (2022a) *FUNA – Functional Numeracy Assessment* [Online]. Saataavilla: <https://www.oppimisanalytiikka.fi/ville/funa/> (Haettu: 14. huhtikuuta 2022)
- [33] Oppimisanalytiikan tutkimusinstituutti (2022b) *Turun yliopisto – Oppimisanalytiikan tutkimusinstituutti* [Online]. Saataavilla: <https://www.oppimisanalytiikka.fi/> (Haettu: 14. heinäkuuta 2022)
- [34] Piccolo, L. R., Giacomoni, C. H., Julio-Costa, A., Oliveira, S., Zbornik, J., Haase, V. G. & Salles, J. F. (2017) 'Reading Anxiety in L1: Reviewing the Concept', *Early Childhood Education Journal*, 45, 537—543. <https://doi.org/10.1007/s10643-016-0822-x>
- [35] R Core Team (2021) 'R: A language and environment for statistical computing', *R Foundation for Statistical Computing*. Wien, Itävalta. <https://www.R-project.org/>
- [36] Ramirez, G., Fries, L., Gunderson, E., Schaeffer, M. W., Maloney, E. A., Beilock, S. L. & Levine, S. C. (2019) 'Reading Anxiety: An Early Affective Impediment to Children's Success in Reading', *Journal of Cognition and Development*, 20(1), 15—34. <https://doi.org/10.1080/15248372.2018.1526175>

- [37] Räsänen, P., Aunio, P., Laine, A., Hakkarainen, A., Väisänen, E., Finell, J., Rajala, T., Laakso, M.-J. & Korhonen, J. (2021) 'Effects of Gender on Basic Numerical and Arithmetic Skills: Pilot Data from Third to Ninth Grade for a Large-scale Online Dyscalculia Screener', *Front. Educ.*, 6(683672). <https://doi.org/10.3389/feduc.2021.683672>
- [38] Saris, W. E., Satorra, A. & Sörbom, D. (1987) 'The Detection and Correction of Specification Errors in Structural Equation Models', *Sociological Methodology*, 17, 105—129. <https://doi.org/10.2307/271030>
- [39] Schermelleh-Engel, K. & Moosbrugger, H. (2003) 'Evaluating the fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures', *Methods of Psychological Research Online*, 8(8), 23—74. Saatavilla: <https://www.researchgate.net/publication/251060246>
- [40] Schoot, R. van de, Lugtig, P. & Hox, J. (2012) 'A Checklist for Testing Measurement Invariance', *European Journal of Developmental Psychology*, 9(4), 486—492. <https://doi.org/10.1080/17405629.2012.686740>
- [41] Shi, D., Lee, T. & Olivares-Maydeu, A. (2019) 'Understanding the Model Size Effect on SEM Fit Indices', *Educational and Psychological Measurement*, 79(2), 310—334. <https://doi.org/10.1177/0013164418783530>
- [42] Sorvo, R., Koponen, T., Viholainen, H., Aro, T., Räikkönen, E., Peura, P., Dowker, A. & Aro, M. (2017) 'Math Anxiety and Its Relationship with Basic Arithmetic Skills among Primary School Children', *British Journal of Educational Psychology*, 87(3), 309—327. <https://doi.org/10.1111/bjep.12151>
- [43] Sorvo, R., Koponen, T., Viholainen, H., Aro, T., Räikkönen, E., Peura, P., Tolvanen, A. & Aro, M. (2019) 'Development of Math Anxiety and Its Longitudinal Relationships with Arithmetic Achievement among Primary School Children', *Learning and Individual Differences*, 69, 173—181. <https://doi.org/10.1016/j.lindif.2018.12.005>
- [44] Steiger, J. H. (1990) 'Structural Model Evaluation and Modification: An Interval Estimation Approach', *Multivariate Behavioral Research*, 25(2), 173—180. [https://doi.org/10.1207/s15327906mbr2502\\_4](https://doi.org/10.1207/s15327906mbr2502_4)
- [45] Tuomola, E. (2020) *Lukemisahdistuksen yhteydet matematiikka-ahdistukseen, lukusujuvuuteen, sukupuoleen ja luokka-asteeseen* (Julkaisu nro 23948) [Pro Gradu -tutkielma, Jyväskylän Yliopisto]. <http://urn.fi/URN:NBN:fi:jyu-202005193337>
- [46] UNESCO (2022) *UNESCO Prize awarded to a collaborative learning platform ViLLE from Finland* [Online]. Saatavilla: <https://www.unesco.org/en/articles/unesco-prize-awarded-collaborative-learning-platform-ville-finland> (Haettu: 19. lokakuuta 2022)

- [47] Veerasamy, A. K., D'Souza, D., Lindén, R. & Laakso, M.-J. (2019) 'Relationship Between Perceived Problem-solving Skills and Academic Performance of Novice Learners in Introductory Programming Courses', *Journal of Computer Assisted Learning*, 35(2), 246—255. <https://doi.org/10.1111/jcal.12326>
- [48] Widaman, K. F. & Reise, S. P. (1997) 'Exploring the Measurement Invariance of Psychological Instruments: Applications in the Substance Use Domain' teoksessa Bryant, K. J., Windle, M. & West, S. G. (toim.) *The science of Prevention: Methodological Advances from Alcohol and Substance Abuse Research*. Washington, DC: American Psychological Association, 281—324. <https://doi.org/10.1037/10222-009>
- [49] Wu, A. D., Li, Z. & Zumbo, B. D. (2007) 'Decoding the Meaning of Factorial Invariance and Updating the Practice of Multi-group Confirmatory Factor Analysis: A Demonstration with TIMSS Data', *Practical Assessment, Research, and Evaluation*, 12(3). <https://doi.org/10.7275/mhqa-cd89>
- [50] Yaman, H. (2010) 'Writing Anxiety of Turkish Students: Scale Development and the Working Procedures in Terms of Various Variables', *International Online Journal of Educational Sciences*, 2(1), 267—289. ISSN: 1309-2707

# Liitteet

## A Faktorimalliesimerkin kovarianssimatriisin määrittely

Luvun 2.2 esimerkin rajoitetun faktorimallin havaittujen muuttujien populaatiovarianssi-kovarianssimatriisi on muotoa  $\Sigma = \Sigma(\eta) = \Lambda\Phi\Lambda' + \Theta$ , jossa

$$\Lambda = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{bmatrix},$$

$$\Phi = \begin{bmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{bmatrix}, \text{ ja}$$

$$\Theta = \begin{bmatrix} \theta_1 & & & & & \\ 0 & \theta_2 & & & & \\ 0 & 0 & \theta_3 & & & \\ 0 & 0 & 0 & \theta_4 & & \\ 0 & 0 & 0 & 0 & \theta_5 & \\ 0 & 0 & 0 & 0 & 0 & \theta_6 \end{bmatrix},$$

jolloin

$\Sigma(\eta)$

$$= \begin{bmatrix} \lambda_{11}^2\phi_{11} + \theta_1 & & & & & \\ \lambda_{11}\lambda_{21}\phi_{11} & \lambda_{21}^2\phi_{11} + \theta_2 & & & & \\ \lambda_{11}\lambda_{31}\phi_{11} & \lambda_{21}\lambda_{31}\phi_{11} & \lambda_{31}^2\phi_{11} + \theta_3 & & & \\ \lambda_{11}\lambda_{42}\phi_{21} & \lambda_{21}\lambda_{42}\phi_{21} & \lambda_{31}\lambda_{42}\phi_{21} & \lambda_{42}^2\phi_{22} + \theta_4 & & \\ \lambda_{11}\lambda_{52}\phi_{21} & \lambda_{21}\lambda_{52}\phi_{21} & \lambda_{31}\lambda_{52}\phi_{21} & \lambda_{42}\lambda_{52}\phi_{22} & \lambda_{52}^2\phi_{22} + \theta_5 & \\ \lambda_{11}\lambda_{62}\phi_{21} & \lambda_{21}\lambda_{62}\phi_{21} & \lambda_{31}\lambda_{62}\phi_{21} & \lambda_{42}\lambda_{62}\phi_{22} & \lambda_{52}\lambda_{62}\phi_{22} & \lambda_{62}^2\phi_{22} + \theta_6 \end{bmatrix}.$$

Koska

$$\Sigma = \begin{bmatrix} \text{Var}(x_1) & & & & & \\ \text{Cov}(x_2, x_1) & \text{Var}(x_2) & & & & \\ \text{Cov}(x_3, x_1) & \text{Cov}(x_3, x_2) & \text{Var}(x_3) & & & \\ \text{Cov}(x_4, x_1) & \text{Cov}(x_4, x_2) & \text{Cov}(x_4, x_3) & \text{Var}(x_4) & & \\ \text{Cov}(x_5, x_1) & \text{Cov}(x_5, x_2) & \text{Cov}(x_5, x_3) & \text{Cov}(x_5, x_4) & \text{Var}(x_5) & \\ \text{Cov}(x_6, x_1) & \text{Cov}(x_6, x_2) & \text{Cov}(x_6, x_3) & \text{Cov}(x_6, x_4) & \text{Cov}(x_6, x_5) & \text{Var}(x_6) \end{bmatrix},$$

ja  $\Sigma = \Sigma(\eta)$ , niin esimerkiksi  $\text{Var}(x_1) = \lambda_{11}^2\phi_{11} + \theta_1$ . Tämä osoittaa, että  $x_1$ :n varianssi vaihtelee, jos  $\lambda_{11}$ ,  $\phi_{11}$  tai  $\theta_1$  saavat eri arvoja. Matriisin  $\Sigma$  muutkin alkiot voitaisiin kirjoittaa samaan tapaan matriisin  $\Sigma(\eta)$  vastaavien alkioden avulla.

## B Sovitusfunktion approksimatiivisen kvadraattisen muodon perustelu

Perustellaan, miksi sovituskfunktion arvo  $F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\tilde{\eta}))$  voidaan approksimoida kvadraattisessa muodossa  $\text{tr}[\boldsymbol{\Sigma}(\tilde{\eta})^{-1}(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\tilde{\eta}))]^2/2$ . Sovitusfunktio voidaan ensin johtaa muotoon

$$\begin{aligned} F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\tilde{\eta})) &= \log |\boldsymbol{\Sigma}(\tilde{\eta})| + \text{tr} [\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}(\tilde{\eta})] - \log |\boldsymbol{\Sigma}| - q \\ &= \text{tr} [\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}(\tilde{\eta})] - \text{tr} [\mathbf{I}_q] - \log |\mathbf{I}_q - \mathbf{I}_q + \boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}(\tilde{\eta})| \\ &= \text{tr} [\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}(\tilde{\eta}) - \boldsymbol{\Sigma}(\tilde{\eta})\boldsymbol{\Sigma}^{-1}(\tilde{\eta})] - \log |\mathbf{I}_q - \boldsymbol{\Sigma}(\tilde{\eta})\boldsymbol{\Sigma}^{-1}(\tilde{\eta}) + \boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}(\tilde{\eta})| \\ &= \text{tr} [(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\tilde{\eta}))\boldsymbol{\Sigma}^{-1}(\tilde{\eta})] - \log |\mathbf{I}_q + (\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\tilde{\eta}))\boldsymbol{\Sigma}^{-1}(\tilde{\eta})|. \end{aligned}$$

Browne [7] esitti, että funktion jälkimmäinen termi voidaan esittää Taylorin sarjana

$$-\log |\mathbf{I}_q + (\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\tilde{\eta}))\boldsymbol{\Sigma}^{-1}(\tilde{\eta})| = \sum_{k=1}^{\infty} \frac{1}{k} \text{tr} [-(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\tilde{\eta}))\boldsymbol{\Sigma}^{-1}(\tilde{\eta})]^k,$$

jolloin

$$\begin{aligned} F(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}(\tilde{\eta})) &= \text{tr} [(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\tilde{\eta}))\boldsymbol{\Sigma}^{-1}(\tilde{\eta})] + \sum_{k=1}^{\infty} \frac{1}{k} \text{tr} [-(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\tilde{\eta}))\boldsymbol{\Sigma}^{-1}(\tilde{\eta})]^k \\ &= \sum_{k=2}^{\infty} \frac{1}{k} \text{tr} [-(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\tilde{\eta}))\boldsymbol{\Sigma}^{-1}(\tilde{\eta})]^k \\ &\approx \frac{1}{2} \text{tr} [\boldsymbol{\Sigma}(\tilde{\eta})^{-1}(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\tilde{\eta}))]^2. \end{aligned}$$

## C MAQ- ja RWAQ-kyselyjen osiokohtaiset kuvailevat tunnusluvut

Taulukoissa 7, 8, 9, 10, 11 ja 12 näkyvät MAQ- ja RWAQ-kyselyjen osiokohtaiset kuvailevat tunnusluvut. Huomattavaa on se, että MAQ-kysymyksiä vastaavat keskimääräiset pistemäärät ovat systemaattisesti korkeampia tytöillä kuin pojilla, ja nämä pistemäärät kasvavat myös systemaattisesti luokka-asteittain muutamia poikkeuksia lukuun ottamatta. RWAQ-kyselyn tapauksessa yhtä selkeää systemaattisuutta ei näy. Esimerkiksi RWAQ Workload -alakyselyn kysymysten keskimääräiset pistemäärät ovat yhtä suuria tytöillä ja pojilla. Evaluation- ja Social-alakyselyissä pistemäärät ovat suurimmassa osassa kysymyksistä suurempia tytöillä kuin pojilla. Yläkouluikäisten lasten keskimääräiset pistemäärät RWAQ-kysymyksissä ovat suurempia kuin alakouluikäisillä lapsilla. Sen sijaan kun pistemääriä verrataan yläkoulun luokka-asteiden kesken tai alakoulun luokka-asteiden kesken, erot ovat pieniä.

Taulukko 7: MAQ Learning -alakyselyn osiokohtaiset kuvailevat tunnusluvut.

	MAL1			MAL2			MAL3			MAL4		
	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>
Tyttö	1	1,18	0,88	1	0,98	0,77	1	1,21	0,81	1	1,00	0,88
Poika	1	1,05	0,86	1	0,95	0,74	1	1,08	0,78	1	0,89	0,82
3lk	1	0,92	0,93	1	0,86	0,89	1	1,03	0,90	1	0,76	0,95
4lk	1	0,98	0,87	1	0,86	0,76	1	1,09	0,84	1	0,76	0,86
5lk	1	0,98	0,80	1	0,88	0,74	1	1,08	0,72	1	0,83	0,82
6lk	1	1,04	0,73	1	0,92	0,58	1	1,11	0,67	1	0,92	0,71
7lk	1	1,22	0,85	1	1,03	0,68	1	1,22	0,76	1	1,07	0,81
8lk	1	1,34	0,88	1	1,07	0,74	1	1,25	0,82	1	1,12	0,79
9lk	1	1,40	0,95	1	1,22	0,84	1	1,33	0,84	1	1,27	0,88
Yht.	1	1,11	0,87	1	0,97	0,76	1	1,15	0,80	1	0,95	0,85

*Huomioita.* *Md* = mediaani; *M* = keskiarvo; *Sd* = keskihajonta.

Taulukko 8: MAQ Evaluation -alakyselyn osiokohtaiset kuvailevat tunnusluvut.

	MAE1			MAE2			MAE3			MAE4		
	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>
Tyttö	2	1,95	1,11	2	1,77	1,05	2	2,29	1,23	1	1,36	0,90
Poika	1	1,36	1,01	1	1,24	0,93	1	1,60	1,17	1	1,14	0,84
3lk	1	1,35	1,20	1	1,20	1,08	1	1,49	1,32	1	1,12	1,01
4lk	1	1,45	1,12	1	1,32	1,06	2	1,78	1,30	1	1,12	0,88
5lk	1	1,50	1,07	1	1,38	0,97	2	1,86	1,24	1	1,10	0,78
6lk	1	1,64	1,02	1	1,51	0,90	2	1,99	1,15	1	1,19	0,75
7lk	2	1,86	1,02	1	1,65	0,98	2	2,18	1,19	1	1,35	0,83
8lk	2	1,97	1,04	2	1,79	1,03	2	2,27	1,20	1	1,44	0,86
9lk	2	1,98	1,08	2	1,83	1,04	2	2,18	1,17	1	1,51	0,94
Yht.	2	1,66	1,10	1	1,51	1,03	2	1,95	1,25	1	1,25	0,88

*Huomioita.* *Md* = mediaani; *M* = keskiarvo; *Sd* = keskihajonta.

Taulukko 9: MAQ Social -alakyselyn osiokohtaiset kuvailevat tunnusluvut.

	MAS1			MAS2			MAS3			MAS4		
	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>
Tyttö	1	1,51	1,06	2	2,07	1,21	1	1,58	1,01	1	1,28	0,97
Poika	1	1,12	0,91	1	1,51	1,09	1	1,25	0,88	1	1,08	0,85
3lk	1	1,07	1,02	2	1,71	1,28	1	1,15	0,99	1	1,17	1,07
4lk	1	1,07	0,93	2	1,71	1,24	1	1,30	0,94	1	1,11	0,95
5lk	1	1,18	0,96	1	1,63	1,12	1	1,32	0,92	1	1,10	0,90
6lk	1	1,29	0,87	2	1,75	1,11	1	1,40	0,86	1	1,11	0,78
7lk	1	1,51	1,04	2	1,95	1,17	1	1,56	0,96	1	1,26	0,89
8lk	1	1,61	1,05	2	1,99	1,17	1	1,61	0,99	1	1,30	0,91
9lk	1	1,60	1,04	2	1,90	1,18	1	1,65	1,01	1	1,31	0,91
Yht.	1	1,32	1,01	2	1,79	1,19	1	1,42	0,96	1	1,19	0,92

*Huomioita.* *Md* = mediaani; *M* = keskiarvo; *Sd* = keskihajonta.

Taulukko 10: RWAQ Workload -alakyselyn osiokohtaiset kuvailevat tunnusluvut.

	RAW1			RAW2			WAW1			WAW2		
	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>
Tyttö	1	1,20	0,88	1	1,45	0,99	1	1,55	1,13	1	1,44	0,96
Poika	1	1,19	0,86	1	1,46	1,03	1	1,54	1,03	1	1,45	0,97
3lk	1	1,00	0,92	1	1,34	1,14	1	1,38	1,12	1	1,38	1,06
4lk	1	1,05	0,86	1	1,42	1,06	1	1,50	1,12	1	1,37	1,01
5lk	1	1,05	0,76	1	1,34	0,95	1	1,49	1,06	1	1,36	0,91
6lk	1	1,16	0,77	1	1,41	0,95	1	1,50	1,04	1	1,40	0,89
7lk	1	1,35	0,90	1	1,55	0,98	1	1,63	1,09	1	1,51	0,94
8lk	1	1,38	0,89	1	1,58	0,98	1	1,65	1,05	1	1,57	0,94
9lk	1	1,42	0,93	1	1,58	0,99	2	1,68	1,05	1	1,55	0,97
Yht.	1	1,19	0,87	1	1,46	1,01	1	1,54	1,08	1	1,44	0,96

*Huomioita.* *Md* = mediaani; *M* = keskiarvo; *Sd* = keskihajonta.

Taulukko 11: RWAQ Evaluation -alakyselyn osiokohtaiset kuvailevat tunnusluvut.

	RAE1			RAE2			WAE1			WAE2		
	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>
Tyttö	1	1,50	1,00	1	1,48	0,96	1	1,04	0,69	1	1,31	1,13
Poika	1	1,27	0,92	1	1,21	0,85	1	1,09	0,75	1	1,08	1,00
3lk	1	1,27	1,10	1	1,24	1,02	1	0,93	0,78	1	1,06	1,14
4lk	1	1,34	1,03	1	1,29	0,96	1	0,98	0,73	1	1,05	1,12
5lk	1	1,32	0,92	1	1,29	0,89	1	0,98	0,66	1	1,04	1,03
6lk	1	1,40	0,93	1	1,32	0,83	1	1,03	0,61	1	1,15	1,03
7lk	1	1,47	0,96	1	1,46	0,91	1	1,15	0,71	1	1,34	1,05
8lk	1	1,46	0,90	1	1,43	0,88	1	1,15	0,68	1	1,42	1,05
9lk	1	1,44	0,91	1	1,42	0,92	1	1,24	0,80	1	1,40	1,03
Yht.	1	1,38	0,97	1	1,35	0,92	1	1,06	0,72	1	1,20	1,08

*Huomioita.* *Md* = mediaani; *M* = keskiarvo; *Sd* = keskihajonta.

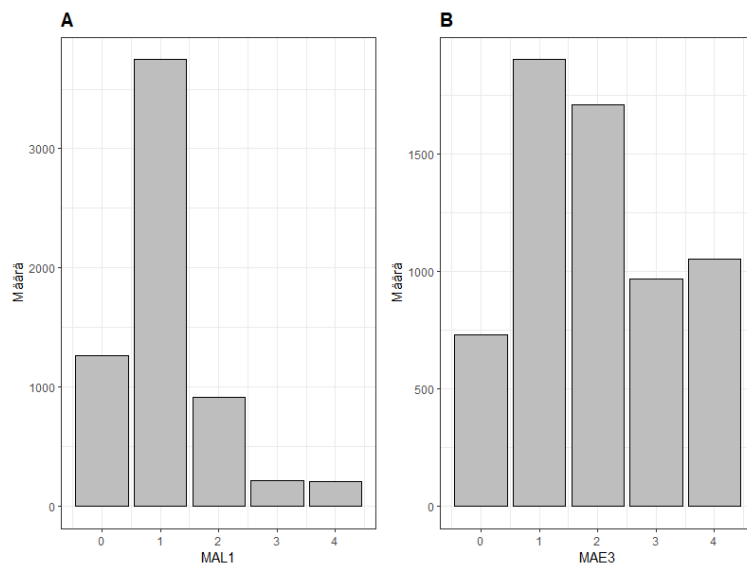


Taulukko 12: RWAQ Social -alakyselyn osiokohtaiset kuvailevat tunnusluvut.

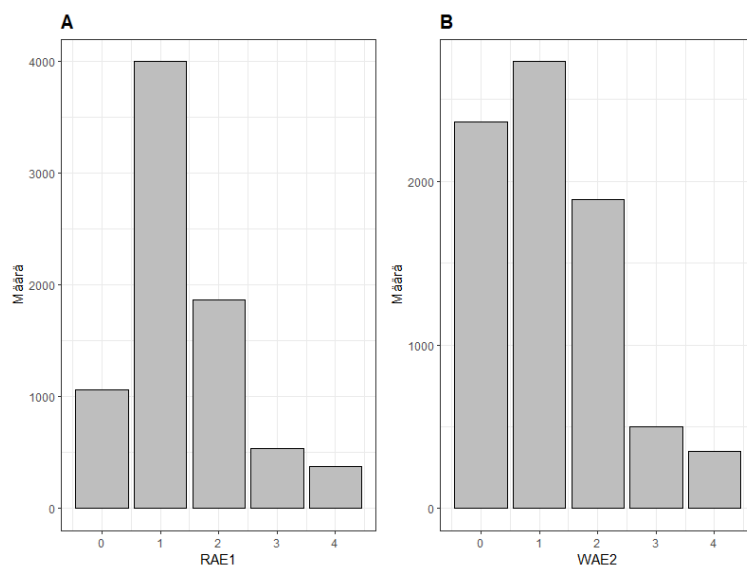
	RAS1			RAS2			WAS1			WAS2		
	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>	<i>Md</i>	<i>M</i>	<i>Sd</i>
Tyttö	2	1,89	1,19	1	1,46	1,16	1	1,30	0,96	1	1,54	1,21
Poika	1	1,54	1,08	1	1,23	0,96	1	1,28	0,89	1	1,34	1,06
3lk	1	1,67	1,28	1	1,07	1,04	1	1,28	1,09	1	1,37	1,29
4lk	2	1,64	1,17	1	1,25	1,11	1	1,21	0,97	1	1,33	1,18
5lk	1	1,50	1,06	1	1,17	1,02	1	1,24	0,93	1	1,26	1,07
6lk	1	1,59	1,07	1	1,32	1,01	1	1,23	0,82	1	1,40	1,08
7lk	2	1,92	1,14	1	1,56	1,08	1	1,39	0,91	1	1,61	1,14
8lk	2	1,93	1,12	1	1,57	1,05	1	1,36	0,85	1	1,63	1,10
9lk	2	1,85	1,16	1	1,52	1,07	1	1,36	0,88	1	1,56	1,08
Yht.	2	1,72	1,15	1	1,35	1,07	1	1,29	0,92	1	1,44	1,14

*Huomioita.* *Md* = mediaani; *M* = keskiarvo; *Sd* = keskihajonta.

## D Esimerkkikuvat osiokohtaisista jakaumista



Kuva 4: MAQ-kyselyn kaksi osiokohtaista jakaumaa. **A:** Osion MAL1 jakauma. **B:** Osion MAE3 jakauma. Kaikki muut MAQ:n osiokohtaiset jakaumat muistuttivat muodoltaan MAL1:n jakaumaa lukuunottamatta kuvassa **B** näkyvää MAE3:n jakaumaa.



Kuva 5: RWAQ-kyselyn kaksi osiokohtaista jakaumaa. **A:** Osion RAE1 jakauma. **B:** Osion WAE2 jakauma. Kaikki muut RWAQ:n osiokohtaiset jakaumat muistuttivat muodoltaan RAE1:n jakaumaa lukuunottamatta kuvassa **B** näkyvää WAE2:n jakaumaa.

## E RWAQ-mallin rajoitettujen parametrien MI- ja EPC-arvojen tulkinta

Taulukossa 13 näkyvät RWAQ-mallin lukitut parametrit, joiden muunnosindeksin arvo on yli 100, sekä kyseiset muunnosindeksien arvot ja vastaavat standardisoidut EPC-arvot. Taulukossa symboli  $\rightarrow$  viittaa faktorilataukseen ja symboli  $\updownarrow$  virheiden väliseen kovarianssiin. Taulukon perusteella suurimmat MI-arvot ovat rajoitetuilla faktorilatauksilla Workload-faktorista osioihin WAS1 ja WAE1 ja Evaluation-faktorista osioon WAS1 sekä osioita RAS1 ja RAS2 vastaavien virheiden välisellä kovarianssilla. Ottaen huomioon näiden parametrien EPC-arvot, perustelluinta lie nisi vapauttaa Workload-faktorin lataus WAE1-osioon estimoitavaksi.

Huomattavaa on se, että osio WAS1 voisi mahdollisesti liittyä sekä Workload-että Evaluation-faktoriin. Lisäksi niiden parametrien, jotka vastaavat osion WAS1 virhetermin korrelaatioita osioiden RAS1 ja RAS2 virhetermien kanssa, EPC-arvot ovat negatiivisia. Nämä havainnot viittaavat siihen, että osio WAS1 ei välttämättä muodosta Social-alakyselyn muiden osioiden kanssa mielekästä sosiaalisiin tilanteisiin liittyvää lukemis- ja kirjoittamisahdistusta mittaavaa kokonaisuutta. Toisaalta alkuperäisessä mallissa Social-faktorilataus osioon WAS1 oli suuri ja positiivinen, joten WAS1 ei ainakaan mittaa täysin päinvastaista asiaa kuin muut Social-osiot. Tämä osoittaa sen, että faktorimallien modifiointi MI- ja EPC-arvojen avulla ei ole suoraviivaista.

Taulukko 13: RWAQ-mallin lukitut parametrit, joilla  $MI > 100$ .

	MI	EPC <sub>std</sub>
Faktorilataus		
Workload $\rightarrow$ WAS1	395,93	0,37
Evaluation $\rightarrow$ WAS1	331,04	0,44
Workload $\rightarrow$ WAE1	286,90	0,50
Workload $\rightarrow$ RAE2	105,85	-0,32
Virheiden kovarianssi		
RAS1 $\updownarrow$ RAS2	284,43	0,29
RAS1 $\updownarrow$ WAS1	182,84	-0,20
RAS2 $\updownarrow$ WAS1	172,02	-0,21
WAW1 $\updownarrow$ WAS1	160,47	0,17
RAW1 $\updownarrow$ WAE1	151,88	0,16
RAW1 $\updownarrow$ WAW2	113,46	-0,17
WAW1 $\updownarrow$ WAW2	108,79	0,17