

Itsejärjestävä piirrekartta SOM: joustava tilastomenetelmä Likert-asteikollisten kyselylomakeaineistojen ryhmittelyyn

ARI VOUTILAINEN

FT, dosentti, tutkija
Itä-Suomen yliopisto
Hoitotieteen laitos

TARJA KVIST

TtT, yliopistotutkija
Itä-Suomen yliopisto
Hoitotieteen laitos

RAIJA MÄNTYNEN

TtT, tutkija
Itä-Suomen yliopisto
Hoitotieteen laitos

KATRI VEHVILÄINEN-JULKUNEN

THT, Professori
Itä-Suomen yliopisto
Hoitotieteen laitos

TIIVISTELMÄ

Tutkimuksen tarkoituksena oli selvittää, miten hyvin itsejärjestävä piirrekartta (SOM) soveltuu epätäydellisen Likert-asteikollisen kyselylomakeaineiston ryhmittelyyn. SOM on fyysikko Teuvo Kohosen 1980-luvulla kehittämä tiedonlouhinnan apuväline. Tämä tutkimus tehtiin ryhmittelemällä neljän suomalaisen sairaalan hoitohenkilökunnalta At Safe -hankkeessa kehitetyllä Transformational Leadership -mittarilla vuonna 2008 kerätty aineisto. Jo aiemmin sama aineisto oli ryhmitelty eksploratiivisella faktorianalyysillä ja analysoitu epäparametrisellä varianssianalyysillä. SOM-verkko jakoi vastaajat (n = 1148) yleisesti tyytymättömiin, vain ylihoitajan toimintaan tyytyväisiin, vain osastonhoitajan toimintaan tyytyväisiin ja yleisesti tyytyväisiin. Vastajat olivat keskimäärin hyvin tyytyväisiä, vaikka muutamissa työyksiköissä suuri osa vastaajista oli tyytymättömiä. Tulosten perusteella SOM-verkko on varteenotettava vaihtoehto ryhmitellessä Likert-asteikollisia kyselylomakeaineistoja. Hoitotiede on aina etsinyt totuutta mahdollisimman pätevien ja luotettavien menetelmien avulla. Tänä päivänä, kun terveydenhuollon palvelujärjestelmää kehitetään entistä muutoskykyisemmäksi, luotettavuuden vaade edellyttää menetelmällisen osaamisen syventämistä ja laajen-

ABSTRACT

The self-organizing map: a flexible tool for clustering Likert scaled inquiry data sets

*Ari Voutilainen, PhD, Docent, Researcher
Tarja Kvist, PhD, University Researcher
Raija Mäntynen, PhD, Researcher
Katri Vehviläinen-Julkunen, PhD, Professor*

The aim of the study was to evaluate suitability of the self-organizing map (SOM) for clustering incomplete inquiry data set constructed from Likert scale questions. The SOM is originally developed for data mining in the 1980s by the physicist Teuvo Kohonen. Data used in the present study were collected with the Transformational Leadership Scale developed in the At Safe -project from nursing staff of four Finnish hospitals in 2008. The SOM divided the nursing staff (n = 1148) into four clusters, those who were generally unsatisfied, those who were unsatisfied with nursing director, those who were unsatisfied with general manager and those who were generally satisfied. In general, respondents were highly satisfied, although in a few working units most respondents were unsatisfied. The results of this study show that the SOM is a method of choice for clustering Likert scaled inquiry data sets. Nursing science has always used the most qualified and reliable methods available to search for truth. Nowadays, when health services are developed to be more flexible, the demand for reliability stresses both broadening and deepening of methodological know-how. Due to large complicated data sets

tamista. Suuret monimutkaiset aineistot tekevät vaateeseen vastaamisesta haasteellisen tehtävän. Avainsanat: henkilöstöjohtaminen, hoitohenkilökunta, kvantitatiivinen menetelmä, kyselytutkimukset, työtyytyväisyys

fulfilling the demand for reliability is a challenging task.

Key words: inquiry research, job satisfaction, leadership, nursing staff, quantitative method

Mitä tutkimusaiheesta jo tiedetään?

- Järjestysasteikollisista muuttujista rakentuvien suurten kyselylomakeaineistojen luotettava ja menetelmällisesti oikein tehty analysointi voi olla haasteellista.
- Kohosen (1990) itsejärjestyvä piirrekartta SOM ja muut neuroverkot ovat erinomaisia tiedonlouhinnan apuvälineitä.
- Neuroverkoja ei ole hoitotieteessä yleisesti käytetty.

Mitä uutta tietoa artikkeli tuo?

- Teknisesti SOM soveltuu hyvin myös hoitotieteellisten epätäydellisten Likert-asteikollisten kyselylomakeaineistojen ryhmittelyyn.
- SOM mahdollistaa täsmällisten ja moniulotteisten tulosten tuottamisen.

Mikä merkitys tutkimuksella on hoitotyölle, hoitotyön koulutukselle ja johtamiselle?

- Tutkimus tähdentää hoitotieteen tarvetta omaksua uusia menetelmiä.
- Neuroverkot ovat hyvä lisä hoitotieteelliseen menetelmävalikoimaan, jonka laajentaminen ja syventäminen parantaa mahdollisuuksia tehdä perusteltuja johtopäätöksiä tutkimustuloksista niin hoitotyön, koulutuksen kuin johtamisen kehittämiseksi.

Tutkimuksen lähtökohdat

Monet hoitotieteelliset tutkimuskysymykset edellyttävät suurten, useista muuttujista koostuvien aineistojen keräämistä. Tällaisten aineistojen luotettava analysointi voi olla vaikeaa, sillä esimerkiksi vinous (vrt. Altman & Bland 1995) ja puuttuvat havainnot eli kato (Little & Rubin 2002) ovat tavallisia aineistolähtöisiä tekijöitä, jotka rajoittavat käytettävissä olevien analyysimenetelmien lukumäärää ja tekevät analyysin tilastollisten vaatimusten saavuttamisesta työlästä. Aineisto on kuitenkin analysoitava, jotta tutkimus etenisi synteisiin ja uskottaviin johtopäätöksiin asti. Yksinkertaisimmillaan numeeristen aineistojen analyysi on periaatteellista toimintaa. Jos valittu tilastomenetelmä ei kyseisen aineiston analysointiin sovellu eli menetelmän ehdot eivät to-

teudu, on käytettävä jotain muuta menetelmää. Kyselylomakeaineistolle on hyvä tehdä myös katoanalyysi ja tekstissä olisi tuotava selkeästi ilmi, 1) miten kadon ongelma on ratkaistu/ohitettu ja 2) miten aineiston epätäydellisyys mahdollisesti vääristää tuloksia. Analyysia ei siis voi aloittaa aineiston ”keskeltä” vaan on tehtävä analyysia ennen ja jälkeen varsinaisen tulosanalyysin. Epätäydellisten ja/tai muutoin hankalien aineistojen analysoinnista on julkaistu lukuisia kirjoja, esimerkiksi Statistical Analysis With Missing Data (Little & Rubin 2002), sekä luotettavia Internet-sivustoja, kuten Kvantitatiivisten menetelmien tietovaranto, KvantiMOTV (<http://www.fsd.uta.fi/menetelmaopetus/intro.html>). Kadon ongelma on tuotu esiin myös tunnetun hoitotieteen tilastollista analyysia käsittelevän kirjan Statistics and Data Analysis for Nursing Research uudemmassa painoksessa (Polit 2009). Puuttu-

vien havaintojen periaatteellista ongelmaa ei kuitenkaan postanalyttisesti voi koskaan täydellisesti ratkaista. Jos datamatriisin yksittäisen solun arvoa ei aineistoa kerättäessä ole saatu, säilyy se ikuisesti tuntemattomana. Jälkeenpäin pystyy vain tuottamaan enemmän tai vähemmän todennäköisiä vaihtoehtoja esimerkiksi interpoloimalla. Kyselytutkimuksen tapauksessa yksi matriisin ”solu” tarkoittaa yhden osallistujan vastausta yhteen väittämään/kysymykseen.

Strukturoiduissa kyselylomakkeissa käytetty Likert-asteikko tuottaa järjestysasteikollisia muuttujia. Tämä on ongelma, jos analyysimenetelmäksi valitaan menetelmä, joka olettaa muuttujien olevan välimatkat tai suhdeasteikollisia. Yleinen tapa on ryhmitellä järjestysasteikollisista muuttujista koostuvaa aineistoa suurimman uskottavuuden (engl. Maximum Likelihood) faktori-analyysillä. Suurimman uskottavuuden menetelmä olettaa riippumattomuutta, normaalijakaumaa sekä jakaumien samanlaisuutta ryhmien välillä. Jos oletukset eivät toteudu, faktorit rakentuvat eri tavoin eri ryhmissä, eivätkä tulokset ole enää luotettavia. Likert-asteikollisen aineiston kohdalla tällainen tulosten vääristymisen riski on erittäin suuri (Lubke & Muthén 2004). Likert-asteikolliset muuttujat on mahdollista muuntaa normaalisuusoletusta paremmin toteuttaviksi, mutta muuntaminen ei yksiselitteisesti poista edellä kuvattua ongelmaa (Wu 2007). Parempi ratkaisu on siis käyttää joko välimatka- tai suhdeasteikollisia muuttujia tuottavia mittareita tai analysoida järjestysasteikollisista muuttujista rakentuva aineisto sille oikeasti soveltuvalla menetelmällä, joka voi olla esimerkiksi tässä tutkimuksessa esiteltävä itsejärjestyvä piirrekartta.

Itsejärjestyvä piirrekartta SOM

Fyysikko Teuvo Kohosen 1980-luvulla kehittämä itsejärjestyvä piirrekartta SOM (engl. self-organizing map) (Kohonen 1990) on niin sanottu neuroverkko eli mallinnus aivojen toiminnasta. Aivojen tavoin SOM-verkko rakentuu neuroneiksi kutsutuista alkioista ja niiden välisistä yhteyksistä.

SOM-verkko on erinomainen apuväline tiedonlouhinnassa eli ryhmien etsimisessä suurista aineistoista. SOM-verkon laskentatulokset on mahdollista esittää ”sellaisenaan” visuaalisessa muodossa (Vesanto 1999) tai ne voidaan ryhmitellä erilaisten algoritmien avulla (Vesanto & Alhoniemi 2000). Neuroverkoista on kehitetty lukuisia yleispäteviä sovelluksia, joita ei kuitenkaan hoitotieteessä ole laajalti käytetty (huom. Ladstätter ym. 2010), vaikka perinteisempiin ryhmittely- ja analyysimenetelmiin, kuten faktori-, pääkomponentti- ja varianssi-analyysit, verrattuna neuroverkoilla on lukuisia etuja. SOM-verkko ei tarvitse ennakkotietoa muuttujien välisistä yhteyksistä, vaan sallii aineiston järjestyä itsenäisesti, mahdollisimman objektiivisesti. Tämä ei kuitenkaan tarkoita teoriattomuutta. Teoria on vahvasti mukana ennen varsinaista analyysivaihetta tutkimusongelmaa muodostettaessa ja varsinaisen analyysivaiheen jälkeen johtopäätöksiä esitettäessä. Kun SOM-verkkoa käytetään kyselylomakkeella kerätyn aineiston ryhmittelyssä, on teoria ohjannut myös itse tutkimusmenetelmän eli kyselylomakkeen rakentamista.

Teknisesti SOM-verkko kykenee käsittelemään epätäydellistä aineistoa ilman, että matriisin tyhjät solut eli puuttuvat tiedot täytetään esimerkiksi approksimoimalla niiden arvot – tai että johonkin kysymykseen vastaamatta jättäneet henkilöt poistetaan aineistosta. SOM-verkko ei järjestä aineistoa yksittäisten tietojen perusteella, vaan katsoo tutkimusyksikköä, joka kyselylomakeaineistossa yleensä tarkoittaa yhtä vastaajaa, kokonaisuutena ja sijoittaa toisiaan eniten vastaavat yksiköt mahdollisimman lähelle. Yksiköiden ei siis tarvitse olla yksittäisten muuttujien suhteen vahvasti korreloituneita. Jos aineistosta puuttuu paljon tietoja, voi rakentuneen SOM-kartan tulkitavuutta parantaa interpoloimalla tyhjien karttasolujen arvot niitä ympäröivien solujen arvojen avulla. Tämä ei kuitenkaan tapahdu ennen analyysiä vaan sen jälkeen, mikä tekee interpoloinnista analyttisessä mielessä hyväksyttävämmän. Keskeisimmät SOM-menetelmään liittyvät termit on selitetty taulukossa 1.

Taulukko 1. Keskeisimpien tekstissä esiintyvien matemaattisten termien selitykset.

Termi	Selitys
Algoritmi	Yksityiskohtainen kuvaus ongelmanratkaisutehtävän suorittamisesta, esimerkiksi ohje siitä, miten luku jaetaan toisella jakokulman avulla, on yksinkertainen algoritmi.
Batch-algoritmi	SOM-klusteroinnin ohje, menetelmän eksakti matemaattinen kuvaus
Interpolointi	Tuntemattomien arvojen laskeminen tunnettujen avulla
Jakautumisindeksi J	Tätä tutkimusta varten muotoiltu versio alun perin monimuotoisuuden laskemiseen kehitetystä Shannon-Wiener -indeksistä
Klusteri	Joukko toisiaan suuresti muistuttavia kohteita, joissakin yhteyksissä synonyyminä käytetään termiä ryväs/rypä.
Kvantittumis- ja topografinen virhe	SOM-kartta ei ole koskaan täydellinen esitys. SOM-kartan hyvyttä voi arvioida esimerkiksi kvantittumis- ja topografisen virheen perusteella. Kvantittumisvirhe kertoo, kuinka kaukana tutkimusyksiköt (aineistomatriisin yksi rivi edustaa yhtä tutkimusyksikköä) keskimäärin ovat niille parhaiten sopivista soluista. Topografinen virhe puolestaan ilmoittaa niiden tutkimusyksiköiden osuuden, joille toiseksi parhaiten sopiva solu ei sijaitse niille parhaiten sopivan solun vieressä. Jos topografinen virhe on suuri, muodostuneisiin klustereihin on syytä suhtautua varauksellisesti. SOM ei siis pysty sijoittamaan kaikkia tutkimusyksiköitä ideaaleille paikoilleen, vaan raportoitava SOM-kartta on aina kompromissi.
Matriisi	Tilastomatematiikassa yleisesti käytetty suorakulmainen riveistä ja sarakkeista rakentuva taulukko
Neuroverkko	Aivojen toiminnan matemaattinen mallinnus
SOM	Itsejärjestyvä piirrekartta (engl. self-organizing map), fyysikko Teuvo Kohosen kehittämä neuroverkko, jota voidaan hyödyntää tiedonlouhinnassa.
SOM-solu	SOM-menetelmä rakentaa aineistosta visuaalisen esityksen eli kartan , joka koostuu soluiksi nimitetyistä osista. SOM sijoittaa aineiston näihin soluihin batch-algoritmin määräämällä tavalla. SOM-solun muoto on valittavissa. Yksi yleisimmin käytetyistä muodoista on kuusikulmio . Muoto vaikuttaa solun viereisten solujen lukumäärään: kuusikulmiolla kuusi ja esimerkiksi neliöllä kahdeksan.
SOM Toolbox	Teknillisessä korkeakoulussa kehitetty MATLAB-nimisellä numeerisen matematiikan ohjelmistolla käytettävä matemaattinen työkalu, jolla voi tehdä SOM-klusteroiteja.
Tiedonlouhinta	Joukko tilastollisia menetelmiä, joilla pyritään järjestämään suuria aineistoja helpommin ymmärrettävään muotoon ja löytämään tutkittavan ilmiön kannalta oleellimmat seikat.

Tutkimuksen tarkoitus

Tutkimuksen tarkoituksena oli selvittää, miten hyvin SOM-verkko soveltuu laajan kyselyaineiston tilastolliseen ryhmittelyyn ja esitellä SOM vaihtoehtona perinteisemmille ryhmittelymenetelmille, kuten pääkomponentti- ja faktorianalyysit. Käytettävä aineisto, jota jo Kvist ym. (2012) analysoivat omassa julkaisussaan, on valittu menetelmällisten tarpeiden perusteella. Tulokset- ja Pohdinta-osioissa painotetaan menetelmällisiä kysymyksiä aineistolähtöisten sisällöllisten kysymysten sijaan.

Aineisto ja menetelmät

Aineiston keruu

Tutkimusaineisto kerättiin neljän suomalaisen sairaalan (yksi yliopistollinen sairaala ja kolme keskussairaala) hoitohenkilökunnalta At Safe -hankkeessa kehitetyllä Transformational Leadership -mittarilla, joka sisältää 54 viisiportaista Likert-asteikollista väittämää (1 = *täysin erimieltä*, 5 = *täysin samaa mieltä*). Mittaria kehitettäessä tehtiin laaja kirjallisuuskatsaus vuonna 2007, lisäksi käytettiin asiantuntijapaneelia ja pilotti-

utkimuksen tuloksia vuodelta 2007 (Määttä ym. 2009, Eneh ym. 2012, Kvist ym. 2012). Marraskuussa 2008 sähköinen kysely lähetettiin 5466 hoitohenkilökuntaan kuuluvan sähköpostiosoitteeseen. Kahden viikon kulluttua lähetettiin muistutusviesti kyselyyn vastaamisesta. Kyselyyn vastasi 1151 henkilöä. Pohjois-Savon sairaanhoitopiirin eettiseltä toimikunnalta saatiin puoltava lausunto tutkimukselle (46/2007). Tutkimusluvut saatiin tutkimusorganisaation lupakäytäntöjen mukaisesti hallintoylihoitajilta ja henkilöstojohtajilta. Tästä vuoden 2008 aineistosta tehty eksploratiivinen faktorianalyysi tuotti viisi faktoria, jotka olivat: 1. Eettinen johtaminen (14 väittämää, esimerkiksi *Osastonhoitajani arvostaa jokaista työntekijää*); 2. Hoitotyön prosessin johtaminen (16 väittämää, *Osastonhoitajani ohjaa arvioimaan hoitotyön tulosta hoitoprosessissa*); 3. Palautteen antaminen ja palkitseminen (6 väittämää, *Osastonhoitajani antaa säännöllisesti palautetta työsuorituksista*); 4. Työssä kehittymisen tukeminen (7 väittämää, *Osastonhoitajani mahdollistaa yksilöllisen kehittymisen kaikille työyksikössän*) ja 5. Ylihoitajan johtamistoiminta (11 väittämää, *Ylihoitajani toimintaan voi luottaa*). Faktoroinnin perusteella muodostettiin keskiarvosummamuuttujat, joiden arvot vaihtelivat välillä 1 (huonoin) – 5 (paras). Mittarin reliabiliteetti oli hyvä, Cronbachin α -kertoimet olivat välillä 0.909–0.997 (ks. Burns & Grove 2005). Aineiston kerääminen ja vastaajien taustamuuttujat on selitetty yksityiskohtaisesti julkaisussa Kvist ym. (2012). Samoista sairaaloista kerättiin aineisto vastaavalla tavalla vuonna 2010 (Eneh ym. 2012). Tämä aineisto analysoitiin ristiintaulukoimalla sekä laskemalla frekvenssejä, prosenttiosuuksia ja tunnuslukuja.

Ennen SOM-ryhmittelyä aineistosta poistettiin 12 vastaajaa, jotka eivät olleet vastanneet yhteenkään osastonhoitajaa tai yhteenkään ylihoitajaa koskevaan kysymykseen. Täten lopullisen aineiston kooksi muodostui 1148 vastaajaa, 21 prosenttia kyselyn vastaanottaneista. Puuttuvat vastaukset jakaantuivat tasaisesti väittämien kesken ja niiden lukumäärä oli hyvin alhainen ($n=296$), alle viisi promillea mahdollisten

vastausten kokonaismäärästä ($n=61992$). Aineistomatriisissa puuttuvat havainnot nimettiin ”tuntemattomiksi” epälukukoodilla NaN (vrt. Wang 2003).

Aineiston ryhmittely ja analyysi

SOM-ryhmittelyä varten aineisto järjestettiin 1148×54 matriisiksi, jossa yksi rivi vastasi yhtä vastaajaa ja yksi sarake yhtä kyselylomakkeen Likert-asteikollista väittämää. Matriisiin ei siis sisällytetty vastaajien taustamuuttujia. SOM-verkon perusidea on, että verkko oppii aineistosta ryhmittelyn edetessä. Ryhmittelyn lopputulos on sitä objektiivisempi mitä vähemmän ennako-oletuksia verkolle kerrotaan. Vastaajien taustamuuttujat ovat ennako-oletuksia vastaajia mahdollisesti luokittelevista tekijöistä. Mahdollisimman objektiivisesti muodostetut klusterit ovat aidompia kuin ennakkotiedon perusteella muodostetut. Jos näiivi SOM-verkko pelkkien vastausten perusteella muodostaa klusterin, joka kerää yhteen tietyn taustamuuttujan omaavat vastaajat, on tulos objektiivisesti totta. Jos klusteri olisi muodostettu kyseinen taustamuuttuja tietäen, olisi jälkepäin mahdotonta sanoa, kuinka paljon tämä ennakkotietä ohjasi klusterin rakentumista.

Ihanteellisin SOM-kartan koko valitaan niin kutsuttujen kvantittumis- (engl. quantization error) ja topografisten virheiden (engl. topographic error) perusteella. SOM-verkon tapauksessa kvantittumisvirhe ilmoittaa, kuinka kaukana tutkimusyksiköt rakentuneessa SOM-kartassa keskimäärin ovat niille parhaiten sopivista soluista. Topografinen virhe puolestaan kertoo niiden tutkimusyksiköiden osuuden, joille toiseksi parhaiten sopiva solu ei sijaitse kartassa niille parhaiten sopivan solun vieressä. Yleensä kvantittumisvirhe on sitä pienempi, mitä suurempi SOM-kartta on, mutta topografinen virhe ei samalla tavoin ole riippuvainen kartan koosta.

SOM-ryhmittely tehtiin MATLAB® R2012a -ohjelmalla käyttäen SOM Toolbox -työkalua ja siihen sisältyvää batch-algoritmia (Vesanto ym. 2000). Kartan solun muodoksi valittiin yleisesti käytetty kuusikulmio (Ve-

santo ym. 2000). SOM-kartasta muodostettujen klustereiden validiteetti arvioitiin k-means -algoritmillä ja klustereiden lukumäärä valittiin SS-indeksin (engl. simple structure index) perusteella (Borcard ym. 2011). SS-indeksi lasketaan virhesummien perusteella ja sen voikin mieltää eräänlaiseksi virhemitaksi; pieni SS-indeksi tarkoittaa vähäistä virhettä ja suuri SS-indeksi suurta virhettä. K-means -ryhmittely tehtiin ja SS-indeksi laskettiin R 2.11.1 -ohjelmalla käyttäen vegan-nimistä komentopakettia (Borcard ym. 2011).

SOM-tuloksia tarkasteltiin edelleen etsimällä eroja muodostuneiden neljän klusterin (näistä tarkemmin Tulokset-osiossa) välillä. Klustereille laskettiin työyksiköiden jakautumista osoittava indeksi J kaavalla:

$$J = -\sum_i p_i \log_{10} p_i / \log_{10} Y$$

missä p_i on yhden yksikön vastaajien osuus kaikkien vastaajien lukumäärästä tietyssä klusterissa ja Y on klusterin sisältämien työyksiköiden lukumäärä. Indeksillä on tätä tutkimusta varten muokattu sovellus alun perin monimuotoisuuden laskemiseen kehitettyä Shannon-Wiener -indeksistä (Spellerberg & Fedor 2003). Teoreettisesti indeksin arvo 1 tarkoittaa, että jokainen klusterin vastaaja työskentelee eri työyksikössä ja indeksin arvo 0, että jokainen vastaaja työskentelee saman sairaalan samassa työyksikössä. Eli mitä suurempi indeksi, sitä tasaisemmin klusterissa on vastaajia eri työyksiköistä.

Yksisuuntaisella varianssianalyysillä (ANOVA) tutkittiin, ovatko vastaajien taustamuuttujien (työkokemus nykyisessä työyksikössä ja työkokemus yhteensä) keskiarvot samat eri klustereissa? Vastaavasti Kruskallin-Wallisin yksisuuntaisella varianssianalyysillä tutkittiin, onko vastaajien mediaani-ikä sama eri klustereissa? Iän testaamiseen ei käytetty parametrissa varianssianalyysiä, koska muuttujan muunnoksista huolimatta virhevariانسien ryhmien välinen yhtäsuuruusehto ei toteutunut (Levenen testi, $p < 0.05$). Varianssianalyysit tehtiin ohjelmalla IBM® SPSS® 19.

Tulokset

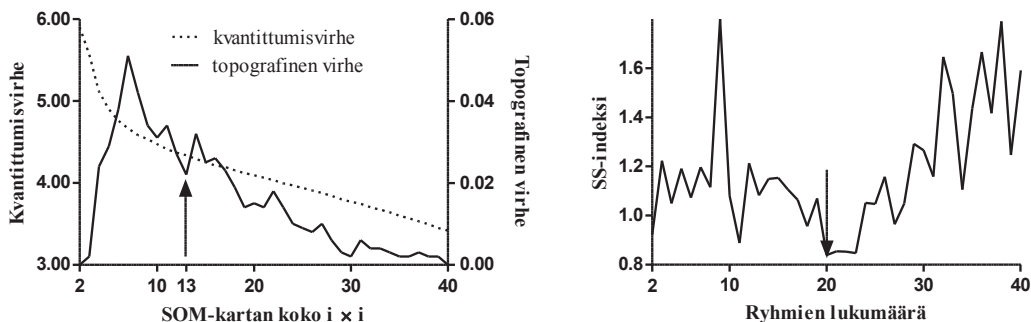
SOM-kartan valinta ja aineiston klusterointi

Ihanteellisimmaksi SOM-kartan kooksi valittiin 13×13 solua. Tämän kartan kvantittumisvirhe oli 4.337 ja topografinen virhe 0.022 (Kuvio 1). Suurempien ($> 17 \times 17$ solua) karttojen virheet olisivat olleet pienempiä (Kuvio 1), mutta nämä kartat sisältäisivät tyhjiä SOM-soluja, eli soluja, jotka eivät parhaalla mahdollisella tavalla edustaneet yhtäkään vastaajista. Suurempien karttojen käyttö olisi siten vaatinut tyhjien solujen arvojen interpolointia viereisten solujen arvojen perusteella. K-means -algoritmi (1000 iteraatiota) muodosti kartasta 20 ryhmää SS-indeksin ollessa tällöin 0.840 (Kuvio 1 ja 2).

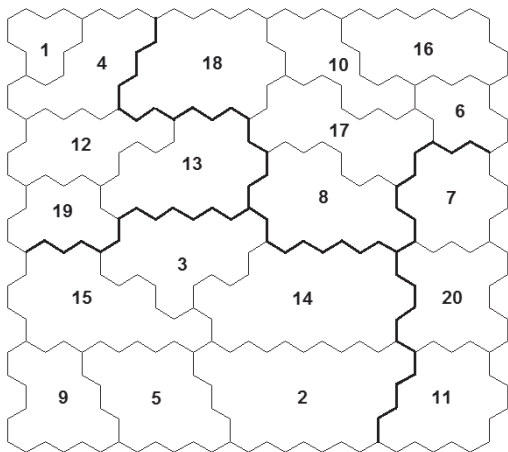
SOM-kartta rakentui siten, että kartan oikeaan yläkulmaan sijoittuneissa ryhmissä 6, 8, 10, 16, 17 ja 18 (klusteri A) vastaajat olivat keskimääräistä tyytymättömmämpiä sekä osaston- että ylihoitajan toimintaan. Vasempaan yläkulmaan sijoittuneissa ryhmissä 1, 4, 12, 13, 18 ja 19 (klusteri B) vastaajat olivat tyytymättömiä osastonhoitajan toimintaan, mutta tyytyväisiä ylihoitajan toimintaan. Oikeaan alakulmaan sijoittuneissa ryhmissä 7, 11 ja 20 (klusteri C) vastaajat olivat tyytyväisiä osastonhoitajan toimintaan, mutta tyytymättömiä ylihoitajan toimintaan. Vasempaan alakulmaan sijoittuneissa ryhmissä 2, 3, 5, 9, 14 ja 15 (klusteri D) vastaajat olivat tyytyväisiä sekä osaston- että ylihoitajan toimintaan. SOM-karttaa esittävässä Kuviossa 2 nämä neljä aluetta on erotettu toisistaan tummennetulla rajaviivalla. Jakautumisindeksi J oli klustereissa A ja B 0.90 ja klustereissa C ja D 0.96.

Analyysit klustereiden sisällä ja välillä

Vastaajien taustamuuttujien keskiarvot (työkokemus nykyisessä työpaikassa ja työkokemus yhteensä) tai mediaani (ikä) eivät eronneet klustereiden välillä: nykyinen työkokemus (ANOVA, $p = 0.375$), työkokemus yhteensä (ANOVA, $p = 0.880$) ja ikä (Kruskal-Wallis, $p = 0.076$). Taustamuuttujien tun-



Kuvio 1. SOM-kartan ihanteellisen koon valinta kvantittumis- ja topografinen virheen perusteella (vasemmalla) ja k-means -algoritmin muodostamien ryhmien lukumäärän valinta SS-indeksin perusteella (oikealla). Valitut vaihtoehdot on osoitettu kuvioissa nuolella.



Kuvio 2. SOM-kartan klusterointi k-means -algoritmilla 20 ryhmään. Tummennetulla rajaviivalla on erotettu toisistaan neljä vastaajien keskimääräisen tyytyväisyyden perusteella muodostettua klusteria (ks. teksti ja Taulukko 1 yksityiskoktaisempaa selitystä varten).

nucluvut on esitetty Taulukossa 2. Samassa taulukossa on lisäksi ilmoitettu määräaikaisten jaksotyöläisten, määräaikaisten päivätyöläisten, vakinaisten jaksotyöläisten ja vakinaisten päivätyöläisten prosenttiosuudet klustereittain. Käyttäen arcsin-neliöjuurimuunnosta (Zar 1999) ja k-means -algoritmin muodostamia ryhmiä tilastoyksiköinä olisi ollut mahdollista testata, ovatko edellä mainitut prosenttiosuudet samat eri kluste-

reissa. Tätä ei kuitenkaan tehty, koska k-means -ryhmien kokonaislukumäärä oli sängen pieni (20) eivätkä ne jakautuneet tasaisesti eri klustereiden kesken. Huomionarvoisinta prosenttiosuuksissa on vakinaisten jaksotyöläisten suurehko osuus klusterissa A (56 %), jossa siis vastaajat olivat keskimääräistä tyytymättömämpiä sekä osaston- että ylihoitajan toimintaan (Taulukko 2). Muissa klustereissa vakinaisia jaksotyöläisiä oli 48 tai 49 % (Taulukko 2).

Kun aineiston jakaa neljään edellä mainittuun klusteriin ja piirtää vastausten keskiarvoista kuvaajat on helppo havaita seuraavaa: osastonhoitajan toimintaan liittyvien väittämien osalta klusterit A ja B sekä klusterit C ja D muodostivat lähes identtiset parit (Kuvio 3). Ylihoitajan toimintaan liittyvien väittämien osalta parit vaihtuivat siten, että klusterit A ja C sekä klusterit B ja D muistuttivat paljolti toisiaan. Toinen huomionarvoinen seikka on väittämien numero 15: *Osastonhoitajani antaa säännöllisesti palautetta työsuorituksista* ja numero 16: *Osastonhoitajani palkitsee työssä kehittymisestä* saamat alhaiset arviot klusterista riippumatta. Edelleen huomio kiinnittyy ensimmäiseen ylihoitajan toimintaan koskevaan väittämään: *Ylihoitajani on vaikutusvaltaisen johtaja tulosityksikössään*. Tämä väittäjä jakaa klusterit pareihin osastonhoitajan toimintaa koskevien väittämien tavoin (A ja B sekä C ja D), ja näin ollen siis toisin kuin muut ylihoitajan toimintaa koskevat väittä-

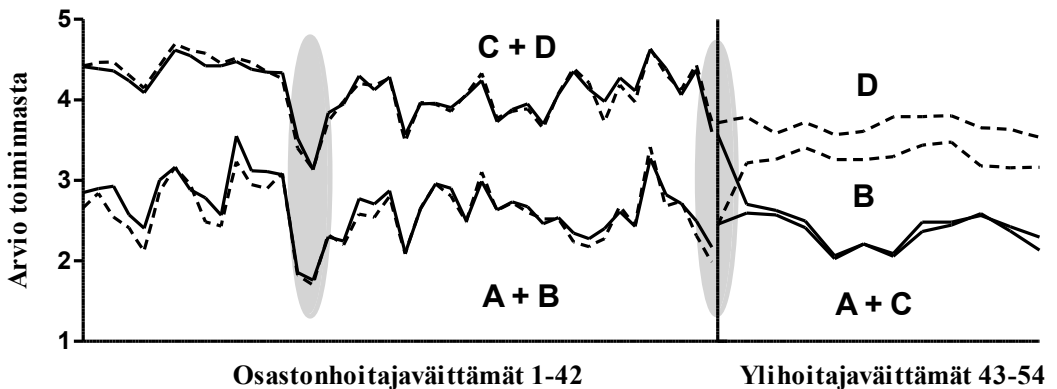
Taulukko 2. K-means -ryhmien ja vastaajien jakautuminen neljään klusteriin (A, B, C ja D). Kustakin muuttujasta on ilmoitettu prosenttiosuus tai keskiarvo (ka) ja keskibajonta (kb). Osaston- ja ylihoitajan toiminnan arvioista on ilmoitettu klusterikobtainen mediaani (md) ja keskiarvo.

	A	B	C	D
K-means -ryhmien lukumäärä	6	5	3	6
Vastaajien lukumäärä	292	222	153	481
Määräaikaisia jaksotyöläisiä (%)	16.3	14.0	19.9	16.5
Määräaikaisia päivätyöläisiä (%)	4.2	2.3	4.0	5.1
Vakinaisia jaksotyöläisiä (%)	56.2	48.4	49.0	49.1
Vakinaisia päivätyöläisiä (%)	23.3	35.3	27.2	29.3
Vastaajien ikä (vuotta, ka ± kh)	41 ± 10	43 ± 9	42 ± 10	43 ± 10
Nykyinen työkokemus (vuotta, ka ± kh)	10 ± 8	9 ± 8	9 ± 8	9 ± 9
Työkokemus yhteensä (vuotta, ka ± kh)	17 ± 10	17 ± 10	17 ± 10	17 ± 11
Osastonhoitajan toiminta [md (ka)]	3 (2.7)	2 (2.6)	4 (4.1)	4 (4.1)
Ylihoitajan toiminta [md (ka)]	3 (2.4)	3 (3.3)	3 (2.4)	4 (3.7)

mät (Kuvio 3). Nämä erityisen mielenkiintoiset aineiston kohdat on korostettu harmaalla värillä kuviossa 3.

Osastonhoitajan toimintaa koskevien väittämien osalta vastaajat olivat sangen yksimielisiä siitä, mikä on hyvää ja mikä huonoa (Kuvio 3). Tästä osoituksena klustereiden A, B, C ja D väittämäkohtaiset vastausmediaanit korreloivat positiivisesti (Spearmanin järjestyskorrelaatiokerroin $\rho \geq 0.32$ ja $p \leq 0.035$ kaikissa pareissa A × B, A × C,

A × D, B × C, B × C ja C × D, n=42). Vastausten yleisen tason mukaan vastaajat sen sijaan jakoutuivat kahteen ryhmään. Klustereihin A ja B ryhmittyneet vastaajat käyttivät osastonhoitajan toimintaa arvioidessaan pääsääntöisesti Likert-asteikon kielteisiä vaihtoehtoja 1 = *täysin eri mieltä* sekä 2 = *osittain eri mieltä* ja klustereihin C ja D ryhmittyneet vastaajat myönteisiä vaihtoehtoja 4 = *osittain samaa mieltä* sekä 5 = *täysin samaa mieltä*.



Kuvio 3. Väittämäkohtaiset vastauskeskiarvot klustereittain (A, B, C ja D). Harmaalla on korostettu väittämät, jotka saivat subteellisesti kielteisimpiä arvioita kaikissa klustereissa (vasemmalla) sekä siirtyminen osastonhoitajan toimintaa koskevista väittämistä ylihoitajan toimintaa koskeviin väittämiin (oikealla).

Pohdinta

Tulosten tarkastelu

SOM-ryhmittelyn avulla löydettiin kolme keskeistä havaintoa: 1) vastaajien joukko ($n = 1148$) jakautui sekä osaston- että ylihoitajan toimintaan tyytymättömiin, vain osastonhoitajan toimintaan tyytymättömiin, vain ylihoitajan toimintaan tyytymättömiin ja sekä osaston- että ylihoitajan toimintaan tyytyväisiin, 2) vastaajien tyytyväisyys yksiköittäin jakautui tasaisemmin klustereiden sisällä kuin tyytymättömyys tarkoittaen, että aineistossa oli työyksiköitä, joissa suuri osa vastaajista oli tyytymättömiä ja 3) vastaajien taustamuuttujien keskiarvot (työkokemus nykyisessä työpaikassa, työkokemus yhteensä) tai mediaani (ikä) olivat samat eri klustereissa. Yksittäisistä väittämistä kannattaa huomioida ainakin numero 15: *Osastonhoitajani antaa säännöllisesti palautetta työsuorituksista* ja numero 16: *Osastonhoitajani palkitsee työssä kehittämisestä*, jotka molemmat, mutta erityisesti numero 16, saivat suhteellisesti kielteisimpiä arvioita kaikissa klustereissa. Tärkeimmäksi kehittämiskohteeksi hoitotyön johtajien toiminnassa Kvist ym. (2012) mainitsivat palautteen antamisen. Kehittämistarpeen palautteen antamisessa saattoi todeta myös SOM-ryhmittelyn perusteella. SOM-tulokset olivat siis samansuuntaisia kuin eksploratiivisellä faktorianalyysillä ja Kruskallin-Wallisin yksisuuntaisella varianssianalyysillä saadut tulokset (Kvist ym. 2012).

Jakautumisindeksi J oli kaikissa klustereissa varsin korkea osoittaen yksiköiden jakautuneen tasaisesti klustereiden sisällä. Indeksien perusteella voi päätellä, että tyytyväisyys osastonhoitajan toimintaan oli jakautunut tasaisemmin kuin tyytymättömyys. Toisin sanoen, aineistossa oli työyksiköitä, joissa sangan moni vastaajista oli tyytymättömän osastonhoitajan toimintaan. Osa havaitusta hoitohenkilökunnan tyytymättömyydestä selittyy näin ollen enemmän osastonhoitaja- kuin vastaajalähtöisillä tekijöillä. Hoitohenkilökunnan yleistä tyytyväisyyttä osastonhoitajansa toimintaa lienee helpointa lisätä kiinnittämällä erityistä huomiota

työyksiköihin, joissa tyytymättömyys osastonhoitajan toimintaan on yleinen ilmiö. Tyytymättömiä löytyy toki kaikkialta, mutta keskimääräinen tyytyväisyys on työyksiköiden välillä tasaisemmin jakautunut kuin tyytymättömyys. Sinällään jo varsin korkea yleistä tyytyväisyyttä on työyksikötasoisilla toimilla vaikea enää lisätä, mutta puuttamalla yksiköihin, joissa on paljon tyytymättömiä hoitajia, on mahdollista vähentää paikallista tyytymättömyyttä. Yleistä voidaan todeta, että työyksiköiden erot keskimääräisessä tyytyväisyydessä osaston- ja ylihoitajan toimintaan eivät johdu tyytyväisyyden vaan tyytymättömyyden vaihtelusta. Tyytymättömyydellä näyttää olevan suurempi taipumus kasautua kuin tyytyväisyydellä.

SOM-tulosten perusteella pohdittiin myös käytettyä aineiston hankintamenetelmää eli Likert-asteikollista strukturoitua kyselylomaketta. Likert-asteikko ei ole objektiivinen mittari, vaan vastausten sijoittaminen asteikolle on sekä yksilö- että kulttuurisidonnaista (Lee ym. 2002, Johnson ym. 2005, Möttöus ym. 2012). Yleisesti erotetaan kaksi toisistaan poikkeavaa vastaajatyyppeä: asteikon ääripäitä suosiva ja vastauksiaan lieventävä (Johnson ym. 2005). Tässä tutkimuksessa osastonhoitajan toimintaa koskevat väittämät erottivat toisistaan ”perustyytyväiset” (klusterit C ja D) ja ”perustyytymättömät” vastaajat (klusterit A ja B). Aineiston perusteella on mahdotonta todeta, johtuuko perustyytyväisyys ja -tytymättömyys vastaajan persoonallisuudesta vai tutkimuksen kohteesta (osastonhoitajan toiminta). Ensimmäistä vaihtoehtoa tukee Tulokset-osiossa mainittu väittämäkohtaisten vastausmediaanien korreloiminen kaikkien klustereiden kesken ja erityisesti perustyytyväisten ($C \times D$, Spearmanin $\rho = 0.70$) ja perustyytymättömien kesken ($A \times B$, Spearmanin $\rho = 0.75$). Jälkimmäistä vaihtoehtoa puolestaan tukee klustereiden järjestyminen eri tavoin osastonhoitajan ja ylihoitajan toimintaa koskevien väittämien perusteella. Osastonhoitajan toimintaa arvioidessaan pääsääntöisesti myönteisiä vastausvaihtoehtoja valinneet (klusteri C) siirtyivät käyttämään pääsääntöisesti kielteisiä vaihtoehtoja ylihoitajan toimintaa arvioides-

saan. Klusteriin B ryhmittyneiden vastaajien kohdalla tapahtui päinvastoin. Näin ollen klusteriin C ryhmittyneitä ei koko aineiston perusteella voi nimittää ”kaikkean myönteisesti suhtautuviksi” eikä klusteriin B ryhmittyneitä ”kaikkean kielteisesti suhtautuviksi”. Koska klusteriin D ryhmittyneet vastaajatkaan eivät olleet ylihoitajan toimintaa arvioidessaan yhtä selkeän myönteisiä kuin osastonhoitajan toimintaa arvioidessaan, jäljelle jää klusteri A, johon ryhmittyneet vastaajat säilyttivät keskimäärin kielteisen mielipiteensä läpi koko kyselylomakkeen.

Tutkimuksen luotettavuus

Myös SOM-ryhmittelyyn liittyy osittain subjektiivisia eli tutkijasta riippuvia työvaiheita, esimerkiksi ihanteellisimman SOM-kartan koon valinta. Subjektiivisuutta voi vähentää perustelemalla valintansa avoimesti ja yksityiskohtaisesti tai käyttämällä laskennallisia valintaperusteita (Vesanto & Alhoniemi 2000). Laskennallisesta hyvä esimerkki on k-means -algoritmin (Borcard ym. 2011) hyödyntäminen klustereiden määrää päätettäessä. Vaihtoehtona olisi klusterimäärän valinta esimerkiksi visuaalisin perustein (Vesanto 1999), jolloin useammat lopputulokset tulisivat mahdollisiksi. Kaikkien tilastollisten menetelmien luotettavuuden kannalta tärkeintä on, että 1) menetelmällä käsitellään vain sille soveltuvia aineistoja eli ymmärretään, mitä saa ja mitä ei saa tehdä, 2) menetelmää käytetään teknisesti oikein ja 3) tulokset osataan tulkita.

Keskiarvojen laskeminen järjestysasteikollisista muuttujista ei ole sallittua. Tällaisia muuttujia ovat esimerkiksi Likert-asteikolliset muuttujat ja luokitellut taustamuuttujat. Tässäkin tutkimuksessa on menetely sääntöä vastaan, kun SOM-verkolla tuotettu aineisto on jaettu klustereihin ”keskimääräisten” tyytyväisyyksien perusteella (Kuvio 2). Tunnuslukujen laskeminen ei kuitenkaan ole analyysiä eikä keskiarvoja testattu tilastollisesti. Niitä käytettiin jaettaessa aineisto sekä osaston- että ylihoitajan toimintaa koskeviin väittämiin annettujen vastausten perusteella kahteen yhtä suureen jouk-

koon (Kuvio 2) sekä konkretisoitaessa klustereiden välisiä eroja ja samankaltaisuuksia (Kuvio 3). Vaikka usean Likert-asteikollisen väittämän vastausten yhdistämistä summuuttujiksi ja niistä laskettujen keskiarvojen käyttämistä jatkuvina muuttujina pidetään usein eräänlaisena tilastollisena rajatapauksena (vrt. Jakobsson 2004), ei epäjatkuvan ja jatkuvan muuttujan välillä olevaa perustavaa eroa saa koskaan mieltää merkityksettömäksi. Järjestysasteikollisista muuttujista rakentuneen aineiston virheellinen analysointi ei valitettavasti ole hoitotieteellekään vieras ilmiö (Jakobsson 2004).

Johtopäätökset ja jatkotutkimusehdotus

Tutkimustulosten ja pohdinnan perusteella esitetään seuraavat johtopäätökset:

- 1) SOM-verkko on varteenotettava vaihtoehto ryhmiteltäessä Likert-asteikollista kyselylomakeaineistoa.
- 2) Muilla tieteenaloilla kehitettyjen analyysimenetelmien soveltaminen ja menetelmien kehittäminen edelleen on hoitotieteellisen tutkimuksen yksi tärkeimmistä tulevaisuuden haasteista.
- 3) Koska hoitotieteen rooli niin yhteiskunnallisessa kuin tieteellisessä päätöksenteossa on erittäin merkittävä, on alan tutkijoiden aina parhaalla mahdollisella tavalla vastattava tutkimustulostensa oikeellisuudesta. Tämä ei onnistu ilman avointa suhtautumista uusiin mahdollisuuksiin, mikä usein tarkoittaa tutkimus- ja tilastomenetelmien monipuolista soveltamista.

Tämän tutkimuksen perusteella ei voi tehdä suoria johtopäätöksiä SOM-menetelmän hyvydestä. SOM kyllä soveltuu Likert-asteikollisen aineiston ryhmittelyyn, mutta vertaaminen muihin menetelmiin edellyttää aitoa tilastollista vertaamista. Tämä on tärkein jatkotutkimustehtävä.

Kadon vaikutusta erilaisilla klusterointimenetelmillä, mukaan luettuna SOM, tuotettuihin tuloksiin on tutkittava tulosten yleisen luotettavuuden parantamiseksi. Tähän tarvitaan suuria täydellisiä aineistoja,

joissa siis ei ole puuttuvia havaintoja. Kadon vaikutusta voi mallintaa poistamalla täydellisestä aineistosta satunnaisesti (tai systemaattisesti) havaintoja ja selvittämällä, miten nämä poistot tuloksiin vaikuttavat.

Kiitokset

Tutkimuksen rahoittivat Itä-Suomen yliopiston strateginen IRI hanke (At Safe) ja Sairaanhoidtajien koulutussäätiö. Käsikirjoituksen anonyymeista arvioijista toinen esit-

teli yksityiskohtaisen ja toteuttamiskelpoisen idean menetelmällisen vertailun toteuttamiseksi, joka on tämän tutkimuksen keskeisin jatkotutkimustehtävä. Hänelle osoitamme erityiskiitoksen.

VASTUUALUEET

Tutkimuksen suunnittelu: TK, KV-J, AV, aineistonkeruu: TK, RM, aineiston analysointi: AV, käsikirjoituksen kirjoittaminen: AV, käsikirjoituksen kommentointi: TK, KV-J

LÄHTEET

- Altman DG & Bland JM. 1995. The normal distribution. *BMJ* 310, 298.
- Borcard D, Gillet F & Legendre P. 2011. *Numerical Ecology with R*. Springer, New York City, New York.
- Burns N & Grove S. 2005. *The Practice of Nursing Research: Conduct, Critique, and Utilization*. Fifth Edition. Elsevier, St. Louis, Missouri.
- Eneh VO, Vehviläinen-Julkunen K & Kvist T. 2012. Nursing leadership practices as perceived by Finnish nursing staff: high ethics, less feedback and rewards. *Journal of Nursing Management* 20 (2), 159–169.
- Jakobsson U. 2004. Statistical presentation and analysis of ordinal data in nursing research. *Scandinavian Journal of Caring Sciences* 18 (4), 437–440.
- Johnson T, Kulesa P, Cho YI & Shavitt S. 2005. The relation between culture and response styles. *Journal of Cross-Cultural Psychology* 36 (2), 264–277.
- Kohonen T. 1990. The self-organizing map. *Proceedings of the IEEE* 78 (9), 1464–1480.
- Kvist T, Mäntynen R, Turunen H, Partanen P, Miettinen M, Wolf GA & Vehviläinen-Julkunen K. 2012. How magnetic are Finnish hospitals measured by transformational leadership and empirical quality outcomes? *Journal of Nursing Management* 21 (1), 152–164.
- Ladstätter F, Garrosa E, Badea C & Moreno B. 2010. Application of artificial neural networks to a study of nursing burnout. *Ergonomics* 53 (9), 1085–1096.
- Lee JW, Jones PS, Mineyama Y & Zhang XE. 2002. Cultural differences in responses to a Likert scale. *Research in Nursing & Health* 25 (4), 295–306.
- Little RJA & Rubin DB. 2002. *Statistical Analysis With Missing Data*. Second Edition. Wiley, Hoboken, New Jersey.
- Lubke GH & Muthén BO. 2004. Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural Equation Modeling* 11 (4), 514–534.
- Möttus R, Allik J, Realo A, Rossier J, Zecca G, Ah-Kion J, Amoussou-Yéyé D, Bäckström M, Barkauskiene R, Barry O, Bhowon U, Björklund F, Bochaver A, Bochaver K, de Bruin G, Cabrera HF, Chen SX, Church AT, Cissé DD, Dahourou D, Feng X, Guan Y, Hwang H-S, Idris F, Katigbak MS, Kuppens P, Kwiatkowska A, Laurinavicius A, Mastor KA, Matsumoto D, Riemann R, Schug J, Simpson B, Tseung-Wong CN & Johnson W. 2012. The effect of response style on self-reported conscientiousness across 20 countries. *Personality and Social Psychology Bulletin* 38 (11), 1423–1436.
- Määttä M, Kvist T & Kankkunen P. 2009. Osastonhoitaja – matkalla vetovoimaisen sairaalan hoitotyön johtajaksi. *Premissi* 5, 34–39.
- Polit DF. 2009. *Statistics and Data Analysis for Nursing Research*. Second Edition. Humanalysis, Inc., Saratoga Springs, New York.
- Spellerberg IF & Fedor PJ. 2003. A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the ‘Shannon-Wiener’ Index. *Global Ecology & Biogeography* 12 (3), 177–179.
- Vesanto J. 1999. SOM-based data visualization methods. *Intelligent Data Analysis* 3 (2), 111–126.
- Vesanto J & Alhoniemi E. 2000. Clustering of the self-organizing map. *IEEE Transactions on Neural Networks* 11 (3), 586–600.
- Vesanto J, Himberg J, Alhoniemi E & Parhankangas J. 2000. *SOM Toolbox for Matlab 5. Report A57*. SOM Toolbox Team, Helsinki University of Technology.
- Wang S. 2003. Application of self-organising maps for data mining with incomplete data sets. *Neural Computing & Applications* 12 (1), 42–48.
- Wu C-H. 2007. An empirical study on the transformation of Likert-scale data to numerical scores. *Applied Mathematical Sciences* 1 (58), 2851–2862.
- Zar JH. 1999. *Biostatistical analysis*. Prentis Hall, Upper Saddle River, New Jersey.

Ari Voutilainen, FT, dosentti, tutkija, Hoitotieteen laitos, Itä-Suomen yliopisto, PL 1627, 70211 Kuopio, sähköposti: ari.voutilainen@uef.fi

Tarja Kvist, TtT, yliopistotutkija, Hoitotieteen laitos, Itä-Suomen yliopisto, PL 1627, 70211 Kuopio, sähköposti: tarja.kvist@uef.fi

Raija Mäntynen, TtT, tutkija, Hoitotieteen laitos, Itä-Suomen yliopisto, PL 1627, 70211 Kuopio, sähköposti: raija.mantynen@uef.fi

Katri Vehviläinen-Julkunen, THT, professori, Hoitotieteen laitos, Itä-Suomen yliopisto, PL 1627, 70211 Kuopio, sähköposti: katri.vehvilainenjulkunen@uef.fi