

# Monitasoanalyysit ja niiden soveltaminen sosiaalitieteissä<sup>1</sup>

**Noora Ellonen:** YTM, lehtori (ma.), Porin yksikkö, Tampereen yliopisto  
noora.ellonen@uta.fi

Janus vol. 14 (2) 2006, 127-138

Janus

## Tiivistelmä

Artikkelissa tarkastellaan monitasoanalyysien perusperiaatteita ja niihin liittyviä keskeisiä käsitteitä. Monitasoanalyysit esitellään artikkelissa hyvänä välineenä vastata sosiaalitieteiden peruskysymyseen yksilön ja sosiaalisen ympäristön välisestä suhteesta. Artikkelissa esitellään kaikki perusmonitasomallit: sekamallit, lineaarinen monitasomalli sekä logistinen monitasomalli. Monitasomallinusta lähestytään artikkelissa koulututkimuksen kautta. Koulumaailma tarjoaa yksinkertaisen monitasotutkimukselle sopivan hierarkkisen rakenteen sekä paljon hyviä esimerkkejä monitasoisista tutkimuksista. Artikkelissa pohditaan myös monitasoanalysoinnin keskeisiä rajoituksia sekä ongelmia.

## Abstract

Article deals with the basis of multilevel analysis. Multilevel modelling is seen as useful method for social science research to answer basic sociological questions about the relation between individuals and social environment. Article introduces the basic models of multilevel analysis: mixed models, linear multilevel models and logistic multilevel models. Also the main limitations and problems of multilevel analysis are discussed. In addition some practical advice and useful references are provided. Educational research is used as examples of multilevel modelling. It provides familiar and simple natural hierarchical structure: individual-level, class-level and school-level. It also provides a lot of examples of multilevel studies.

## Johdanto

**Sosiologisen ajattelun** yhtenä keskeisenä lähtökohtana on ajatus siitä, että yksilölliset ilmiöt ovat osittain selitettävissä ja ymmärrettävissä suhteessa sosiaaliseen ympäristöön (Valkonen 1970). Jo Durkheim (1895) kirjoitti aikanaan, että tutkimalla yksinomaan yksittäisiä ihmisiä

emme saa selville koko totuutta yksilöllisistä ilmiöistä. Durkheim (1897) esitti klassisia tuloksia muun muassa itsemurhien yhteydestä sosiaalisen ympäristön koheesion ja integraation asteeseen. Siitä lähtien yksilöllisiin ilmiöihin vaikuttavat yhteisölliset ominaisuudet ovat olleet

## ”Monitasoanalyysit...”

aktiivisen tutkimuksen kohteena. Kysymystä siitä, miksi ihmiset pärjäävät paremmin toisissa yhteisöissä kuin toisissa, on pohdittu niin teoreettisesti kuin empiirisestikin (Kawachi & Berkman 2000). Erityisesti sosiaalisen ympäristön merkitys terveyteen on saanut paljon huomiota erilaisten sosiaalisiin verkostoihin liittyvien teorioiden kautta (ks. Berkman & Glass 2000; Kawachi & Berkman 2000), mutta myös muut yksilölliset ilmiöt, kuten käyttäytyminen ja hyvinvointi, ovat olleet tutkimuksen kohteena (ks. esim. Sampson 2001). Myös viime aikoina virinnyt sosiaalisen pääoman keskustelu on tuonut oman lisänsä sosiaalisen ympäristön merkityksen tutkimiseen. Yhteisölliseen sosiaalisen pääoman on todettu vaikuttavan muun muassa terveyteen, käyttäytymiseen ja hyvinvointiin (ks. esim. Veenstra 2000; Goodwin & Armstrong-Esther 2004; Stolle & Hooghe 2002; Morrow 2000; Gordon & Jordan 1999; Ellonen 2005).

Käytännössä sosiaalisen ympäristön merkitys ihmisten toimintaan viittaa hierarkkiseen sosiaaliseen systeemiin. Ihmiset ovat osa mikro-tason ryhmiä, kuten perhettä, koulua, työpaikkaa tai asuinalueita sekä makro-tason ryhmiä, kuten kansallisuutta. Näillä kaikilla on vaikutuksensa ihmisten toimintaan. Toisin sanoen ihmiset ovat riippuvaisia sosiaalisesta ympäristöstään ja sen yksilöistä. Käytännössä tämä tulee esille muun muassa ryhmien samanlaistavana vaikutuksena: samaan ryhmään kuuluvat ihmiset ”altistuvat” samalle ryhmävaikutukselle, jonka seurauksena saman ryhmän jäsenet muistuttavat tietyiltä osin toisiaan. Esimerkiksi saman koulun oppilaat muistuttavat käyttäytymiseltään enemmän toisiaan kuin kahden eri koulun oppilaat ja suomalaiset muistuttavat enemmän toisiaan kuin suomalainen ja esimerkiksi intialainen. Tämän huomioon ottaminen on oleellinen lähtökohta kun tutkitaan sosiaalisen ympäristön merkitystä: yksilöllisiä ilmiöitä ei voida selittää yksinomaan yksilötason tekijöillä.

Empiirisesti tämä muodostaa merkittävän haasteen, sillä perinteiset analyysimenetelmät perustuvat olettamukseen siitä, että havainnot ovat toisistaan riippumattomia. Niihin ei myöskään samanaikaisesti voi sisällyttää muuttujia eri tasoilta. Esimerkiksi kun oppilaan koulumenestystä selitetään tavallisessa regressiomallissa oppilaan motivaatiolla, näyttäytyy motivaatio yksinomaan yksilötason ilmiönä ja ihmiset ympäristöstään irrallisina yksilöinä. Tosiasissa koulun sosiaalisella ympäristöllä on kuitenkin suuri merkitys yksilön motivaation muodostumiseen (ks. esim. Renchler 1992; Ryan 2001; Guay & Vallerand 1996), jolloin motivaatio ei olekaan yksinomaan yksilötason ilmiö, eivätkä ihmiset irrallisia sosiaalisesta ympäristöstään. Koulun ominaisuuksilla voi myös olla vaikutusta siihen, miten motivaatio on yhteydessä koulumenestykseen. Riippuen koulun muista ominaisuuksista toisissa kouluissa yksilöiden motivaatiolla voi olla suurempi merkitys koulumenestykseen kuin toisissa. Näiden vaikutusten huomioon ottaminen tavallisissa yksitasoisissa regressioanalyysissä on hankalaa, osittain jopa mahdotonta. Toisin sanoen yksitasoiset selitysmallit antavat arvokasta tietoa yksilöllisten ilmiöiden välisistä suhteista, mutta ovat usein puutteellisia sosiaalisen ympäristön vaikutusten osalta.

Viime vuosina yleistyneet monitasoanalyysit ovat tuoneet yhden mahdollisuuden tämän ongelman ratkaisuun. Monitasoanalyysissä käytetyt monitasoiset analyysimallit huomioivat aineiston sisäisen korreloituneisuuden ja mahdollistavat eri tasoja kuvaavien muuttujien tuomisen samaan selitysmalliin. Toisin sanoen niiden avulla voidaan tarkastella sosiologian peruskysymyksiä yksilön ja sosiaalisen ympäristön välisestä suhteesta samassa analyysissä. Tässä artikkelissa esitellään monitasoanalyysin peruseräitä ja niihin liittyviä keskeisiä käsitteitä. Monitasoanalyysijä lähestytään artikkelissa koulututkimuksen kautta. Koulumaailma tarjoaa rakenteensa vuoksi yksinkertaisen ja kaikille tutun esimerkin monitasoisten aineisto-

## ”Monitasoanalyysit...”

jen tarkasteluun sekä antaa paljon hyviä tutkimusesimerkkejä. Monitasoanalyysijä on tehty paljon myös muilla sosiaalitieteiden aloilla, joista hyviä kotimaisia esimerkkejä ovat muun muassa huono-osaisuuden ja kuolleisuuden alueellisia eroja käsittelevät tutkimukset (Blomgren 2005; Valkonen & Kauppinen 2001), terveys- ja terveydenhuoltotutkimukset (Nordberg & Häkinen 1997; Elovainio ym. 2000; Rissanen ym. 2002; ks. myös Keskimäki ym. 2001) sekä nuorisotutkimukset (Karvonen ym. 2005; Karvonen ym. 2001; Ellonen 2005).

### Kontekstuaalianalyysistä monitasomalleihin

Monitasoisten analyysimenetelmien kenttä on vielä toistaiseksi suhteellisen sekava eikä terminologakaan ole mitenkään vakiintunutta. Käsitteitä kuten monitasotutkimus, (*multilevel research*), monitasoanalyysi (*multilevel analysis*), kontekstuaalianalyysi (*contextual analysis*), sekamalli (*mixed model*), satunnaiskertoimien malli (*random coefficient model, random effects model*) ja varianssikomponenttimalli (*variance component model*) käytetään sekaisin, jopa päällekkäin (Goldstein 2003; Diez-Roux 2002). Vaikka kaikki termit viittaavat analyyseihin, joissa otetaan huomioon yksilön ja ympäristön suhde, eivät kaikki ole täysin yhdenmukaisia.

Kontekstuaalianalyysillä viitataan yleisesti ottaen tutkimustapaan, jossa huomioidaan sosiaalisen ympäristön merkitys yksilöllisiä ilmiöitä selittäessä. Kontekstuaalianalyysi tarkoittaa yksinkertaisesti yhteisöllisten tekijöiden ja sosiaalisten ympäristöjen ominaisuuksien systemaattista käsittelyä survey-tutkimuksessa (Jokivuori & Hietala, tulossa). Kontekstuaalianalyysi-käsitettä käytetään usein analyyseistä, joissa yhteisötason vaikutus on otettu huomioon jollain muulla tavalla, kuin monitasoisia tilastomenetelmiä käyttäen. Hyvä esimerkki tästä löytyy kontekstuaalianalyysistä käsittelevästä luvusta kirjassa *Mo-*

*nimuuttujamenetelmien käyttö survey-aineistossa* (Jokivuori & Hietala, tulossa).

Tutkimuksista, jotka käyttävät erilaisia monitasoisia tilastomenetelmiä selvittämään yhteisön vaikutusta yksilöön, käytetään sen sijaan usein nimeä monitasoanalyysi tai monitasotutkimus. Monitasoanalyysit ovat usein kehittyneempiä kuin kontekstuaalianalyysit. Ne ottavat huomioon muun muassa eri tasojen välisen vaihtelun.

Käsitteet, jotka viittaavat analyysi-sanalla sijaan sanaan malli, kuten esimerkiksi sekamalli, satunnaiskertoimien malli ja varianssikomponenttimalli, viittaavat puolestaan monitasoisiin tilastomenetelmiin, joilla käytännössä toteutetaan monitasoanalyysijä. Monitasomalleilla voidaan käsitellä useita eritasoisia muuttujia samassa mallissa ja ne mahdollistavat aineiston sisäisen korreloituneisuuden ja eri tasojen välisen vaihtelun huomioimisen. Mallit poikkeavat hieman toinen toisistaan ja mallin valinta määräytyy aineiston rakenteen sekä tutkimuskysymyksen pohjalta. Käytännössä erot ovat suhteellisen pieniä, ja tutkimuksissa näkeekin usein käytettävien analyyseistä yleistä monitasomalli-nimitystä.

### Hierarkkiset aineistot

Monitasoanalyysin lähtökohta on tutkittavan ilmiön luonnollisesti hierarkkinen rakenne. Toisin sanoen tutkimuksessa tarkastellaan ilmiöitä, joita on luonnollista jäsentää monitasoisesti ja joissa on käsitteitä usealta eri tasolta. Tämä tulee ottaa huomioon jo aineistoa kerättäessä, ja siten monitasomalleja ei tule soveltaa aineistoihin, joiden keruuvaiheessa ei ole otettu huomioon hierarkkista rakennetta. Myös hierarkkisuuden luonnollisuus on oleellista. Toisin sanoen ryhmätaso on todellisuudessa olemassa oleva yhteisö, eikä sitä muodosteta yksinomaan tutkimusta varten. Kouluaineistot ovat tyypillisiä esimerkkejä aineistoista, jotka sisältävät luonnollisen monitasoisen rakenteen. Aineisto on kolmita-

## ”Monitasoanalyysit...”

soinen: alinna on yksilötaso, keskellä luokkataso ja ylinnä koulutaso. (Malin 2005.)

Hierarkkisessa aineistossa havainnot todennäköisesti korreloivat keskenään ainakin joidenkin ilmiöiden osalta. Henkilöiden jakautumisesta osajoukkoihin, kuten esimerkiksi koulussa luokkiin, seuraa homogenisoitumista ryhmän sisällä, jolloin ryhmän jäsenien välille muodostuu riippuvuussuhde. Korrelaatio syntyy näin ollen ryhmän kaikille jäsenille yhteisistä tekijöistä, jotka vaihtelevat ryhmittäin. (Malin 2005.) Saman luokan kaikki oppilaat esimerkiksi ”altistuvat” saman opettajan opetukselle.

Perinteisiä yksitasoisia tutkimusmenetelmiä käytettäessä aineiston hierarkkisuuutta pyritään usein huomioimaan disaggregoimalla ylemmän tason muuttujat alemmalle tasolle tai aggregoimalla alemman tason muuttujat ylemmälle tasolle. Toisin sanoen monistetaan ylemmän tason muuttujat jokaista ryhmän jäsentä kuvaavaksi muuttujaksi (disaggregointi) tai lasketaan tietystä muuttujasta ryhmäkeskiarvo ryhmämuuttujaksi summaamalla ja kokoamalla yksilötason muuttujia ryhmätason muuttujiksi (aggregointi). Molemmat tavat ovat kuitenkin ongelmallisia. Disaggregointi rikkoo muuttujien välistä riippumattomuusoletusta ja aggregoidun aineiston analyysi jättää huomiotta ryhmien sisäisen vaihtelun. (Malin 2005.) Molemmat johtavat helposti sellaiseen virheelliseen toimintatapaan, jossa aineisto analysoidaan yhdellä tasolla ja johtopäätökset tehdään koskemaan toista tasoa (Hox & Kreft 1994). Jos ylemmän tason havaintoja tulkitaan alemmalle tasolle, on kyseessä niin sanottu ekologinen harha (Robinson 1950), ja jos alemman tason havaintoja tulkitaan koskemaan ylempää tasoa, on kyseessä niin sanottu atomistinen harha (Malin 2005; Diez-Rouxin 1998).

### Sisäkorrelaatio

Hierarkkisessa aineistossa siis oletetaan, että samaan ryhmään kuuluvat havainnot, esimerkiksi saman luokan oppilaat, muistuttavat jossain

määrin toisiaan. Käytännössä ryhmän sisäinen korreloituneisuus kuvaa havaintojen tilastollista riippuvuutta tietyn ryhmän sisällä tietyn ilmiön suhteen. Tätä kutsutaan sisäkorrelaatioksi (intraclass correlation). Ajatellaan esimerkiksi tilannetta, jossa selitetään kouluviihtyvyyttä koulumenetyksellä. Toisissa kouluissa oppilaan viihtyvyys on hyvällä tasolla riippumatta koulumenestyksestä ja toisissa kouluissa viihtyvyys paranee koulumenestyksen myötä. Tällöin tilanteessa on negatiivinen sisäkorrelaatio ja koulujen väliset erot vähenevät koulumenestyksen kasvaessa. Ryhmän sisäinen korreloituneisuus onkin aina hyvä tarkistaa. Sisäkorrelaatio lasketaan jakamalla ryhmien välinen varianssi muuttujan kokonaisvarianssilla, eli ryhmien välisen ja ryhmän sisäisen varianssin summalla. (Goldstein 2003.)

$$\rho = \frac{\sigma_u^2}{(\sigma_u^2 + \sigma^2)}$$

Sisäkorrelaation itseisarvo vaihtelee arvojen 0 ja 1 välillä. Mikäli sisäkorrelaatiota ei ole eli sen arvo on 0, havaintoyksiköiden välillä ei ole keskinäistä riippuvuutta tutkitun muuttujan suhteen. Mikäli arvo on 1, havaintoyksiköiden välinen riippuvuus on täydellinen mitatun muuttujan suhteen. Tällöin kaikki ryhmän jäsenet ovat täysin toistensa kaltaisia kyseisen muuttujan suhteen. Kun sisäkorrelaatio on suurempi kuin nolla mutta alle yksi, ovat samaan ryhmään kuuluvat havaintoyksiköt selitettävän asian suhteen keskenään homogeenisempia kuin koko aineiston havaintoyksiköt. Tällöin oletus havaintoyksiköiden välisestä riippumattomuudesta ei täyty. Mikäli tätä riippuvuutta ei oteta huomioon tulokset poikkeavat oikeasta sitä enemmän, mitä enemmän sisäkorrelaatio poikkeaa nolasta. (Malin 2005.)

Sisäkorrelaatiota laskettaessa tulee muistaa, ettei se ole sama asia kuin perinteinen korrelaatiokerroin. Kun perinteisellä korrelaatiolla tarkas-

## ”Monitasoanalyysit...”

tellaan kahden muuttujan välistä riippuvuutta, tarkastellaan sisäkorrelaatiolla yhtä muuttujaa ja havaintoyksiköiden samankaltaisuutta tuon muuttujan suhteen. Sisäkorrelaatio kertoo suoraan myös sen, kuinka paljon ryhmäkeskiarvojen ero selittää muuttujan kokonaisvariانسista. Jos sisäkorrelaatio on esimerkiksi 0,20, on 20 % muuttujan kokonaisvariانسista selitettävissä ryhmien välisillä keskiarvoeroilla. (Malin 2005.)

### Yhteisötason muuttujat

Yhteisötason muuttujat jaetaan pääasiassa aitoihin ryhmätason muuttujiin sekä yksilötason muuttujista laskettuihin aggregaatteihin (Malin 2005). Luokkatasolla opettajaa kuvaavat muuttujat, kuten opettajan sukupuoli tai kokemus opettajan työstä, ovat aitoja ryhmätason muuttujia. Niiden arvo on täsmällisesti sama kaikilla luokan oppilailla, koskeehan kyseinen ominaisuus heitä kaikkia. Koulutasolla vastaavasti koulun koko on aito ryhmätason muuttuja. Aitojen ryhmätason muuttujien vaikutusta kutsutaan usein kontekstivaikutukseksi (Macintyre & Ellaway 2000).

Yhteisötasolle aggregoiduissa muuttujissa kaikilla ryhmään kuuluvilla on myös sama arvo, mutta niiden arvot on laskettu yksilötason muuttujista (Malin 2005). Esimerkiksi koulutasolla oppilaiden kokemaa sosiaalista tukea kuvaavista muuttujista lasketaan koko koulua kuvaava keskiarvo, joka kuvaa siis koulun ominaisuutta ja koskee siten koulun jokaista oppilasta. Näiden muuttujien vaikutuksia kutsutaan usein kompositionaali-vaikutukseksi (Macintyre & Ellaway 2000). Kun mallissa käytetään aggregoituja muuttujia, on yksilötason vaikutus pidettävä pääasiassa aina mukana mallissa sen vaikutuksen vakioimiseksi (Malin 2005).

### Monitasoiset analyysimallit

Monitasoiset analyysimallit sisältävät sekä yksilö- että ryhmätason vaikutuksia. Esimerkiksi oppilaan koulumenestykseen vaikuttaa sekä yksilötason ominaisuus, kuten motivaatio, että ryhmätason ominaisuus, kuten koulun sosiaalinen ilmapiiri. Nämä vaikutukset ovat huomioitavissa monitasoanalyysissä kahdella eri tapaa. Ensinnäkin koulutason kerroin voidaan estimoida yhteisesti kaikille aineiston kouluille, jolloin kerroin on kaikille kouluille sama. Tätä kutsutaan kiinteiden vaikutusten (*fixed effects*) analyysiksi. Estimoitujen kertoimien voidaan antaa myös vaihdella koulukohtaisesti, jolloin kerroin riippuu kyseisen koulun virhetermistä ja vaihtelee näin kouluittain. Tätä kutsutaan satunnaisvaikutusten (*random effects*) analyysiksi. (Diez-Roux 2002.) Esimerkiksi tutkittaessa koulun ilmapiirin vaikutusta koulumenestykseen, voidaan ilmapiiri-muuttujan antaa vaihdella koulujen välillä. Tällöin saadaan eri selitysmalleja eri koulujen välillä.

Vastaavasti yksilötason kertoimet ovat estimoitavissa kahdella tapaa. Selittävän muuttujan kerroin voidaan pitää kiinteänä, jolloin se on kaikille havainnoille sama. Vaihtoehtoisesti kertoimen voidaan antaa vaihdella ylemmän tason yksiköiden välillä, jolloin muuttuja saa kussakin koulussa oman kertoimensa. Tällöin yksilötason satunnaiset termit siis estimoidaan koulujen sisällä. Tutkimuksen kysymyksenasettelusta riippuu, käytetäänkö kiinteiden vai satunnaisvaikutusten mallia. Mallissa voi olla periaatteessa rajoittamaton määrä selittäviä tekijöitä, joiden kertoimet voivat olla kumpia tahansa. Kannattaa kuitenkin huomioida, että satunnaisvaikutusten käyttäminen hankaloittaa tulosten tulkintaa ja useiden satunnaisvaikutusten käyttäminen saattaa tehdä tulosten tulkinnasta jopa mahdotonta.

Monitasoisen aineiston analyysimallit jaetaan usein kahteen ryhmään; sekamallit ja monitasomallit. Molemmat ovat yleisten lineaaristen mallien (regressiomallien ja varianssianalyysimalli-

## ”Monitasoanalyysit...”

en) laajennuksia. Sekamallit ovat perinteisempi tapa tarkastella erilaisia ympäristön vaikutuksia, kun taas monitasomallit ovat verraten uusia menetelmiä. Sekamallien teoria perustuu lähinnä normaalijakaumaan ja se soveltuu parhaiten kaksitasoisten aineistojen analysointiin. Monitasomallit ovat puolestaan monipuolisimpia useaa eri tasoa<sup>2</sup> käsitteleviä malleja. Niiden avulla voidaan selvittää myös erilaisia yhdysvaikutuksia, ja aineiston hierarkkinen rakenne on mallinnuksessa suhteellisen vapaa. Monitasomalleilla voidaan tarkastella myös useanlaisia vastemuuttujia<sup>3</sup>: lineaaristen monitasomallien avulla voidaan tutkia jatkuvia ja normaalisti jakautuneita vastemuuttujia ja yleistetyllä lineaarisella monitasomallilla dikotomisia tai luokka-asteisia muuttujia. (Hox 2002; Fitzmaurice ym. 2004.)

Käytännössä monitasomallit ovat sosiaalitieteissä selvästi sekamalleja yleisempiä. Käytännön esimerkkejä kouluaineistolla tehdyistä lineaarisista monitasoanalyysistä löytyy muun muassa artikkeleista *Multilevel modelling in repeated measures of the quality of Finnish school life* (Mallin & Linnakylä 2001) sekä *Evaluation of well-being in schools – A multilevel analysis of general subjective well-being* (Konu, Lintonen & Autio 2002). Artikkelissa *Paikallisyhteisöjen sosiaalisen pääoman merkitys nuorten masentuneisuudessa ja rikekäyttäytymisessä* (Ellonen 2005) on puolestaan sovellettu kaksitasoista sekamallia.

### Sekamallit ja lineaariset monitasomallit

Normaalisti jakautuneiden vastemuuttujien tarkasteluun soveltuvat sekamallit ja lineaariset monitasomallit toimivat kutakuinkin samalla periaatteella. Ne voidaan kuvata tavallisen yksitasoisen regressioanalyysin laajennuksena. Kun selitetään esimerkiksi oppilaiden koulumenestystä heidän motivaatiollaan, kirjoitetaan yksitasoinen regressiomalli muotoon  $Y_i = \beta_0 + \beta_1 X_i + e_i$  jossa  $Y_i$  viittaa selitettävään muuttujaan (koulumenestys),  $X_i$  selittävään muuttujaan (motivaatio),  $\beta_0$  ja  $\beta_1$  mallissa estimoitaviin regressiokertoimiin ( $\beta_1$  on selittävän tekijän kerroin ja  $\beta_0$  mallin

vakiotermin) sekä  $e_i$  virhetermiin. Kun mallista halutaan tehdä kaksitasoinen eli ottaa koulutaso mukaan tarkasteluun, tulee malliin sekä yksilötason virhetermi  $e_i$  että koulutason virhetermi  $u_j$ , joka kertoo koulujen poikkeaman vakiotermitä. Malli kirjoitetaan tällöin muotoon  $Y_{ij} = \beta_0 + \beta_1 X_{ij} + u_j + e_{ij}$ , jossa kiinteä osa  $\beta_0 + \beta_1 X_{ij}$  määrittää regressiokertoimet ja satunnaisosa  $u_j + e_{ij}$  virhetermit. Mallia kutsutaan kiinteiden kertoimien malliksi.

Kiinteiden kertoimien mallissa kaikilla kouluilla on siis yhteinen koulutason virhetermi. Kun mallista tehdään satunnaiskertoimien malli, saa jokainen koulu oman virheterminsä  $u_{ij}$ , joka kertoo koulukohtaisen eron vakiotermiin  $\beta_0$ . Tällöin oletetaan, että  $\beta_{ij} = \beta_1 + u_{ij}$ . Malli kirjoitetaan tällöin muotoon  $Y_{ij} = \beta_0 + \beta_1 X_{ij} + u_{ij} X_{ij} + u_{0j} + e_{ij}$ . Mallissa kiinteä osa  $\beta_0 + \beta_1 X_{ij}$  on ennallaan, mutta satunnaisosaan on nyt lisätty termi  $u_{ij} X_{ij}$ . Tätä termiä kutsutaan satunnaiskerroimeksi ja se riippuu virhetermistä. (Hox 2002; Goldstein 2003.)

Seuraavassa tarkastellaan erilaisia malleja graafisesti<sup>4</sup>. Otetaan esimerkiksi tilanne, jossa oppilaan koulumenestystä selitetään yksilön motivaatiolla. Kuviossa 1 on tavallinen yksitasoinen regressiomalli  $Y_i = \beta_0 + \beta_1 X_i + e_i$ , jossa  $Y_i$  viittaa koulumenestykseen,  $\beta_1 X_i$  motivaatioon,  $\beta_0$  vakioon ja  $e_i$  virhetermiin. Malli on kaikille oppilaille yhteinen eikä sisällä aineiston koulurakennetta. Regressiosuora kertoo, että mitä korkeampi motivaatio oppilaalla on, sitä paremmin hän pärjää koulussa.

Kun malliin otetaan koulutason vaikutus mukaan, tulee malliin myös koulutason virhetermi. Annetaan koulutason vaikutuksen vaihdella satunnaisesti, jolloin jokainen koulu saa oman virheterminsä ja tällöin myös oman regressiosuoransa (kuviokuva 2). Kaikkien koulujen regressiosuorat ovat kuitenkin samansuuntaisia, sillä selittävän muuttujan regressiokerroin pidetään mallissa kiinteänä. Mallin avulla voidaankin todeta kunkin

## ”Monitasoanalyysit...”

koulun osalta koulutason vaikutuksen suuruus läksyihin kulutetun ajan ja koulumenestyksen suhteessa sekä tarkastella koulujen välisiä eroja ja yhtäläisyyksiä. Mallia kutsutaan usein englanninkielisellä termillä *random intercept model*, jota suoraan suomennettuna voisi kutsua satunnaisvakiomalliksi. Mallista kuulee käytettävän myös satunnaistermien malli -nimitystä.

Kuviossa 3 koulutason vaikutuksen lisäksi myös selittävän muuttujan kertoimesta tehdään satunnainen vaikutus eli sen annetaan vaihdella kouluittain. Virhetermejä on tällöin kolme: oppilaan, selittävän muuttujan kertoimen koulukohtainen sekä koulutason vaikutuksen virhetermi. Mallissa kullakin koululla on edelleen oma regressiosuoransa, mutta ne eivät välttämättä ole enää samansuuntaisia: selittävät muuttujat saavat erisuuruisia kertoimia eri kouluissa. Mallin avulla voidaan muun muassa selvittää ja verrata koulukohtaisia tietoja läksyihin kulutetun ajan ja koulumenestyksen suhteesta. Mallia kutsutaan usein englanninkielisellä termillä *random coefficient model*, joka voidaan suomentaa satunnaiskertoimien malliksi.

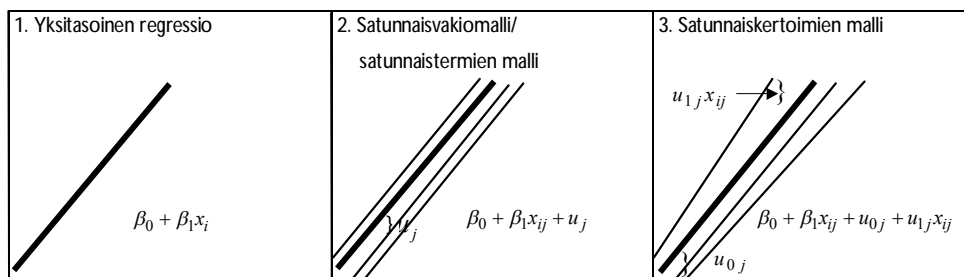
Mikäli koulun merkitystä koulumenestyksen ja motivaation yhteyteen halutaan tarkastella vain jossain tietyssä koulussa, voidaan sitä tarkastella myös koulukohtaisella yksilötason regressiomallilla. Malli tulee kuitenkin muodostaa jokaisen tutkittavan koulun kohdalta erikseen. Mikäli ha-

lutaan yleistää tuloksia koskemaan koulujen välistä vaihtelua, tulee mallista tehdä kaksitasoinen. Laajojen aineistojen ollessa kyseessä monitasomallit ovat myös lukijaystävällisempiä, sillä koko analyysi saadaan sovitettua samaan malliin.

### Logistinen monitasomalli

Lineaaristen monitasomallien soveltuessa yksinomaan jatkuvien ja normaalijakautuneiden vastemuuttujien analysointiin, voidaan yleistetyillä lineaarisilla monitasomalleilla (*generalized linear multilevel models*) analysoida myös muunlaisia vastemuuttujia (Fitzmaurice ym. 2004). Esimerkiksi logistisella monitasomallilla, jota pidetään yleistettyjen lineaaristen monitasomallien erikoistapauksena, voidaan analysoida dikotomisia vastemuuttujia. Dikotomisella muuttujalla tarkoitetaan muuttujaa, jonka arvot vaihtelevat nollan ja yhden välillä: ominaisuus joko on olemassa kyseisellä havaintoyksiköllä (1) tai sitä ei ole olemassa (0). Esimerkiksi oppilas joko on motivoitunut tai ei ole motivoitunut.

Logistisessa monitasomallissa alemman tason varianssi on yleensä kiinnitetty ( $\sigma^2 = 1$ ), ja kiinnostus on ylempien tasojen kertoimissa. Tällöin ei myöskään sisäkorrelaatiota voida laskea. Kiinnostuksen kohteena voivat esimerkiksi olla koulukohtaiset erot koulumenestyksessä ja siihen vaikuttavat yksilö- ja yhteisötason tekijät. Analyysi alkaa tällöin ylempien tason yksiköiden välisten erojen laskemisesta tutkittavan asian, tässä



Kuvio 1. Yksi- ja monitasoisten analyysimallien regressiosuorat

## ”Monitasoanalyysit...”

esimerkissä koulumenestyksen, suhteen. Tämän niin sanotun tyhjän mallin jälkeen malliin lisätään erilaisia yksilö- ja yhteisötason selittäviä muuttujia seuraten samalla koulujen välisen vaihtelun muutosta. Tällöin saadaan selville, paljonko käytetyt koulutason muuttujat selittävät koulujen välisiä eroja koulumenestyksessä, kun mallissa olevat yksilölliset tekijät on vakioitu pois. Alemman tason varianssi voidaan tietyissä tapauksissa myös vapauttaa, jolloin puhutaan niin sanotusta extra-binomiallisesta vaihtelusta. Sen käyttö on kuitenkin vähäisempää ja toteutus sekä tulkinta monimutkaisempaa<sup>5</sup>. Artikkelista *Factors behind low reading literacy achievement* (Linnakylä, Malin & Taube 2004) löytyy kotimainen esimerkki tavallisesta kaksitasoisesta logistisesta monitasomallista.

### Ongelmista ja rajoituksista

Monitasoanalyysiin liittyy suhteellisen paljon ongelmia ja rajoituksia (ks. esim. Diez-Roux 1998; Pickett & Pearl 2000; Maas & Hox 2004). Lähtökohtaisen rajoituksen monitasomallinnukselle muodostavat aineiston ominaisuudet. Ensinnäkin aineiston luonnollinen monitasoisuus tulee osoittaa niin teoreettisesti kuin empiiristäkin. Toisin sanoen hierarkkisuus tulee ottaa huomioon jo aineiston keruuvaiheessa ja sen olemassaolo tulee osoittaa esimerkiksi sisäkorrelaation tai niin sanotun tyhjän mallin avulla. Toiseksi aineistossa tulee olla riittävästi havaintoja. Erityisesti ylempien tasojen osajoukkoja, esimerkiksi kouluja, tulee olla riittävän paljon, jotta niitä kuvaavia parametreja on mahdollista estimoida. Sen sijaan alemman tason havaintojen suuri määrä ei ole yhtä välttämätöntä. Toisin sanoen vaikka jossain koulussa olisi huomattavasti vähemmän oppilaita kuin toisissa, se ei estä monitasomallinnuksen tekoa, mikäli kouluja on kokonaisuudessaan riittävästi. (Keskimäki ym. 2001.) Virheiden mahdollisuus kuitenkin kasvaa, jos alemman tason havainnot ovat kovin vähäisiä. Esimerkiksi kouluaineistossa reilusti alle kah-

denkymmenen oppilaan koulut muodostavat jo riskin luotettavan mallin muodostamiseen. Snijders & Bosker (2002) ovat esittäneet teoksessaan hyviä argumentteja otoskoon suuruuden arvioimisesta.

On myös tarkkaan pohdittava, mitä monitasoanalyysillä halutaan selvittää ja minkä tyyppinen mallinnus siihen sopii, sillä suurin monitasomallinnuksen rajoite (toisaalta myös mahdollisuus) on teoreettisten oletusten sisällyttäminen malliin. Monitasoanalyysit ovat suhteellisen helppo rakentaa, mutta niiden tulkinta on huomattavasti vaikeampaa. Tilastollisesti merkitsevät selitykset voivat olla sisällöllisesti täysin merkityksettömiä. Malli tulee rakentaa niin, että se on helposti tulkittavissa. Mikäli tulkinnasta herää epäily, tulee malli kyseenalaistaa. Tulosten tulkintaa helpottaa tulosten graafinen tarkastelu, joka on suositeltavaa myös analyysin eri vaiheissa.

Mahdollinen rajoittava tekijä on myös se, että toistaiseksi esimerkiksi SPSS ei tarjoa mahdollisuutta hyvään monitasomallinnukseen. Lisäpalikan kautta tietyt kaksitasoiset lineaariset sekamallit ovat mahdollisia, mutta monimutkaisempaan monitasomallinnukseen vaaditaan muita ohjelmia. Käyttökelpoisia monitasomallinnuksiin sopivia tilasto-ohjelmia on useita, joista suosituimpia lienevät MLwiN (Rasbash ym. 2000) sekä HLM (Raudenbush ym. 2000). Muita soveltuvia tilasto-ohjelmia ovat muun muassa SAS ja R, jotka soveltuvat vain kaksitasoisten analyysien tekoon<sup>6</sup>.

Keskeisimpinä ongelmina voidaan mainita riski virhepäätelmistä. Diez-Rouxin (1998) erottaa neljä eri monitasoanalyysiin liittyvää virhepäätelmän mahdollisuutta: ekologinen, atomistinen, psykologistinen ja sosiologistinen. Ekologinen virhepäätelmän tarkoittaa sitä, että ylempien tason havainnoista tehdään johtopäätöksiä alemman tason ilmiöihin. Atomistinen virhepäätelmä on tämän vastakohta, eli alemman tason



## ”Monitasoanalyysit...”

havainnoista johdetaan päätelmiä ylemmälle tasolle. Psykologistinen virhepäätelmä puolestaan viittaa siihen, että analyysissä ei oteta huomioon jotain olennaista ryhmätason muuttujaa, jonka seurauksena johtopäätökset ovat puutteelliset. Esimerkiksi tietty käyttäytyminen voi olla tilanne- tai tapasidonnaista, jolloin sen selittäminen yksinomaan esimerkiksi koulun rakenteellisilla ominaisuuksilla johtaa virhepäätelmiin. Vastavasti virheellisiin päätelmiin päädytään, mikäli jotain olennaista yksilötason tekijää ei huomioida mallissa. Esimerkiksi jos kiinnostuksen kohteena on oppilastoimikuntaan osallistumisen aktiivisuus, ei sitä voida selittää yksinomaan koulun aktiivisuuteen innostavalla ilmapiirillä, vaan yksilöllinen kiinnostus vaikuttamiseen on otettava myös huomioon. Ilman sitä päädyttään sosiologistiseen virhepäätelmään.

Ongelman voivat muodostaa myös niin sanotut sekoittavat tekijät. Näillä tarkoitetaan tekijöitä, jotka vaikuttavat siihen, miten selitettävä muuttuja vaikuttaa vastemuuttujaan tai ominaisuuksia, joiden perusteella tutkittavat ryhmät ovat eri asemassa suhteessa vastemuuttujaan. Tyypillisiä sekoittavia tekijöitä ovat ikä ja sukupuoli. Esimerkiksi vertailtaessa eri koulujen sosiaalisen ympäristön merkitystä oppilaiden masennustasoon, tulee ikä ja sukupuoli huomioida mallissa, sillä esimerkiksi tytöillä on huomattavasti suurempi riski masennusoireisiin kuin pojilla. Hierarkkisissa aineistoissa näiden sekoittavien tekijöiden vaikutukset ovat huomattavasti monimutkaisempia kuin yksitasoisissa malleissa, sillä ne voivat esiintyä paitsi samalla tasolla, myös eri tasojen välillä (Keskimäki ym. 2001). Ongelma kärjistyy juuri sosiaali- ja ihmistieteille tyypillisissä kysymyksenasetteluissa. Usein voidaan kysyä, selittyykö tutkittavan muuttujan vaihtelu todella selittävällä muuttujalla, vai tuntemattomaksi jääneillä sekoittavilla tekijöillä (ks. esim. Rinne ym. 2004).

Myös multikollineaarisuus on vaikeampi ongelma monitasoanalyysissä kuin yksilötason

analyysissä. Hankalaksi siitä tekee sen, että multikollineaarisuutta voi esiintyä sekä samalla tasolla että eri tasojen välillä. (Keskimäki ym. 2001.) Multikollineaarisuudella tarkoitetaan tilannetta, jossa selittävät muuttujat korreloivat erityisen voimakkaasti keskenään. Esimerkiksi tietyissä tilanteissa motivaatio ja lahjakkuus voivat korreloida voimakkaasti keskenään. Tällöin molempien muuttujien käyttäminen saman vastemuuttujan selittämiseen voi johtaa multikollineaarisuuteen. Tämä on tyypillinen ongelma sosiaalitieteissä; useat sosiaalitieteille tyypilliset käsitteet ovat vaikeasti muodostettavissa yksiselitteisiksi havaintokäsitteiksi.

## Lopuksi

Tässä artikkelissa on esitelty monitasoanalyysin peruseriaatteita ja niihin liittyviä keskeisiä käsitteitä. Monitasomalleja lähestyttiin artikkelissa koulututkimuksen kautta, sillä se tarjoaa rakenteensa vuoksi yksinkertaisen, kaikille tutun esimerkin monitasoisten aineistojen tarkasteluun sekä paljon hyviä tutkimusesimerkkejä. Koulututkimusten lisäksi monitasomallinnusta on tehty muun muassa hyvinvointitutkimuksen, terveystutkimuksen ja kaupunkitutkimuksen kentällä.

Monitasomallit on tarkoitettu hierarkkisten aineistojen analysointiin. Niiden avulla voidaan analysoida eritasoisia muuttujia samassa analyysimallissa ja ottaa huomioon aineiston sisäinen korreloituneisuus sekä muuttujien välinen vaihtelu eri tasoilla. Toisin sanoen niiden avulla päästään kiinni sosiologian peruskysymykseen sosiaalisen ympäristön merkityksestä yksilön käyttäytymiseen. Tästä huolimatta monitasomallit ovat suhteellisen vähän käytettyjä kotimaisessa sosiaalitieteellisessä tutkimuksessa. Osittain tämä johtunee siitä, että monitasoanalyysin kenttä on vielä suhteellisen selkiytymätön ja terminologialtaan sekava. Myös analyysien toteutus vaatii erilaisten, sosiaalitieteilijöiden peruskoulu-

## ”Monitasoanalyysit...”

tukseen kuulumattomien ohjelmien opettelua. Monitasoanalyysillä tehdyt tutkimustulokset puhuvat kuitenkin niiden käyttökelpoisuuden puolesta. Monitasoanalyysin avulla on saatu arvokkaita tuloksia yksilön, yhteisön ja sosiaalisen ympäristön välisistä suhteista.

Monitasomallit ovat hyödyllinen väline myös nykyhetken sosiaalipolitiikan ja sosiaalityön tutkimuksen polttavissa kysymyksissä. Kysymykset alueellisista hyvinvointieroista, alueiden ja erilaisten toimintamallien terveys- ja hyvinvointivaikutuksista sekä kansalaisosallistumisesta ovat monitasoanalyysien ominta kenttää. Vastaavasti monitasoanalyysiä voidaan käyttää kansainvälisissä vertailuissa esimerkiksi selvittämään hyvinvointivaltiomallien tai poliittisten järjestelmien yhteyttä tiettyihin kansallisiin ja kansainvälisiin ilmiöihin. Monitasomallit ovat jo tuoneet uutta näkökulmaa muun muassa sosiaalisen pääoman keskusteluun.

Monitasotutkimuksen haasteina tulee puolestaan ottaa huomioon tutkimuskentän selkeyttäminen niin terminologian kuin menetelmien käytönkin osalta. Myös tutkittavat asiat ja tulokset vaativat usein selventämistä. Yhteisötason vaikutus käsitteenä jää tutkimuksissa helposti liian epäselväksi, ellei sitä konkretisoida suhteessa tutkimusasetelmaan. Suurin haaste monitasomallinnuksessa onkin teoreettisten oletusten sijoittaminen tilastolliseen mallin. Vaikeus näkyy myös tulosten tulkinnassa; ne jäävät usein analyysin antiin nähden suhteellisen kapeaksi ja yksioikoiseksi. Tulosten tulkinnan kautta haasteeksi muodostuu myös tutkimustulosten sovellettavuus. Miten arvokkaita tutkimustuloksia yksilön ja yhteisön välisestä suhteesta voidaan soveltaa käytännön toimien parantamiseen? Vielä toistaiseksi tähän vastaaminen jää useissa tutkimuksissa lapsen kenkiin.

## Viitteet

<sup>1</sup> Artikkelit on kirjoitettu osana Suomen Akatemian rahoittamaa Tampereen yliopiston Porin yksikön tutkimusprojektia Koulu ja paikallisyhteisö nuoren tukiverkostona (207367). Väitöskirjani ohjaajien lisäksi haluan kiittää Antero Malinia Jyväskylän yliopistosta, Tapio Nummea Tampereen yliopistosta, Ville Autiota Puolustusvoimista sekä Juha Kääriäistä Poliisiammattikorkeakoulusta asiantuntevista kommentteista artikkelin käsikirjoitusvaiheessa.

<sup>2</sup> Realistisesti monitasomalleissa kannattaa käyttää maksimissaan kolmea tasoa tulosten tulkittavuuden turvaamiseksi.

<sup>3</sup> Nykyään tosin myös yleistettyjen lineaaristen sekamallien käyttö tarjoaa jo huomattavasti laajemmat mahdollisuudet erilaisten vastemuuttujien jakaumien mallintamiseen.

<sup>4</sup> Kuviot ovat mahdollisimman yksinkertaistettuja, jotta perusidea tulee selkeästi esille.

<sup>5</sup> Lisää extra-binomiaalisesta mallista on luettavissa Snijders & Boskerin (2002) teoksesta *Multilevel analysis. An introduction to basic and advanced multilevel modeling*.

<sup>6</sup> Kaikkien ohjelmien kotisivuilta on löydettävissä lisätietoja ohjelmista ja niiden käytöstä (ks. URL osoitteet lähdeluettelosta). MLwiN ja HLM ohjelmien kotisivuilta on saatavilla ohjelmista ilmaiset demot ja R-ohjelma on kokonaisuudessaan ladattavissa ilmaiseksi ohjelman kotisivuilta.

## Kirjallisuus

Bergman, Lisa & Glass, Thomas (2000) *Social Integration, Social Networks, Social Support, and Health*. Teoksessa Lisa Bergman & Ichiro Kawachi (toim.) *Social Epidemiology*. Oxford: University Press, 137-173.

Bergman, Lisa & Kawachi, Ichiro (toim.) (2000) *Social Epidemiology*. Oxford: University Press.

Blomgren, Jenni (2005) *Huono-osaisuus Suomen kaupunkiseutukunnissa: alue-erot ja sosiaalisen ympäristön vaikutukset 1990-luvulla*. Helsingin yliopiston Sosiologian laitoksen tutkimuksia nro 246. Helsinki: Yliopistopaino.

Diez-Roux, Ana (2002) *A Glossary for Multilevel Analysis*. *Journal of Epidemiology and Community of Health* 56 (8), 588-594.

Diez-Roux, Ana (1998) *Bringing Context Back into Epidemiology: Variables and Fallacies in Multilevel Analysis*. *American Journal of Public Health* 88 (2), 216-222.

Durkheim, Emil (1897, 1997) *Suicide: A Study in Sociology*. New York: Free Press.

Durkheim, Emil (1895) *The Rules of Sociological Method*. New York: Free Press.

## ”Monitasoanalyysit...”

- Ellonen, Noora (2005) Paikallisyhteisöjen sosiaalisen pääoman merkitys nuorten masentuneisuudessa ja rikekäyttäytymisessä. *Nuorisotutkimus* 23 (2), 32-46.
- Elovainio, Marko & Kivimäki, Mika & Steen, Nick & Kalliomäki-Levanto, Tiina (2000) Organizational and Individual Factors Affecting Mental Health and Job Satisfaction: A Multilevel Analysis of Job Control and Personality. *Journal of Occupational Health Psychology* 5 (2), 369-277.
- Fitzmaurice, Garrett & Laird, Nan & Ware, James (2004) *Applied Longitudinal Analysis*. Hoboken: Wiley & Sons.
- Goldstein, Harvey (2003) *Multilevel Statistical Models*. Third Edition. London: Arnold. Vuoden 1995 painos luettavissa osoitteesta [www.mlwin.com/hgpersonal/multibook1995.zip](http://www.mlwin.com/hgpersonal/multibook1995.zip), luettu 15.6.2006.
- Goodwin, Mark & Armstrong-Esther, Dawn (2004) Children, Social Capital and Health: Increasing the Well-being of Young People in Rural Wales. *Children's Geographies* 2 (1), 49-63.
- Gordon, Jack & Jordan, Bill (1999) Social Capital and Welfare. *Children & Society* 13 (4), 242-256.
- Guay, Frederic & Vallerand, Robert (1996) Social Context, Student's Motivation, and Academic Achievement: Toward a Process Model. *Social Psychology of Education* 1 (3), 211-233.
- Hox, Joop (2002) *Multilevel Analysis. Techniques and Applications*. Mahwah (New Jersey): Lawrence Erlbaum. Vuoden 1995 painos luettavissa osoitteesta <http://www.fss.uu.nl/ms/jh/publist/amaboek.pdf>, luettu 15.6.2006.
- Hox, Joop & Kreft, Ita (1994) *Multilevel Analysis Methods. Sociological Methods & Research* 22 (3), 283-299.
- Jokivuori, Pertti & Hietala, Risto (tulossa) Monimuuttujamenetelmien käyttö survey -aineistossa. Käsikirjoitus.
- Karvonen, Sakari & Vikat, Andres & Rimpelä, Matti (2005) The Role of School Context in the Increase in Young People's Health Complaints in Finland. *Journal of Adolescence* 28 (1), 1-16.
- Karvonen, Sakari & Vikat, Andres & Rimpelä, Matti (2001) Monitasoanalyysi nuorten tupakoinnin seutukuntavaihtelun selittämisessä. *Sosiaalilääketieteellinen Aikakauslehti* 38 (4), 336-345.
- Kawachi, Ichiro & Bergman, Lisa (2000) Social Cohesion, Social Capital, and Health. Teoksessa Lisa Bergman & Ichiro Kawachi (toim.) *Social Epidemiology*. Oxford: University press, 174-190.
- Keskimäki, Ilmo & Karvonen, Sakari & Sund, Reijo & Leyland, Alastair (2001) Monitasomallien käyttö terveystutkimuksessa. *Sosiaalilääketieteellinen Aikakauslehti* 38 (4), 327-335.
- Konu, Anne & Lintonen, Tomi & Autio, Ville (2002) Evaluation of Well-being in Schools – A Multilevel Analysis of General Subjective Well-being, School Effectiveness and School Improvement 13 (2), 187-200.
- Linnakylä, Pirjo & Malin, Antero & Taube, Karin (2004) Factors behind Low Reading Literacy Achievement. *Scandinavian Journal of Educational Research* 48 (3), 231-249.
- Maas, Cora & Hox, Joop (2004) Robustness Issues in Multilevel Regression Analysis. *Statistica Neerlandica* 58 (2), 127-137.
- Macintyre, Sally & Ellaway, Anne (2000) *Ecological Approaches: Rediscovering the Role of the Physical and Social Environment*. Teoksessa Lisa Bergman & Ichiro Kawachi (toim.) *Social Epidemiology*. Oxford: University Press, 332-348.
- Malin, Antero (2005) Tutkimusaineiston rakenne ja kvantitatiiviset analyysimenetelmät. *Psykologia* 40 (5-6), 489-501.
- Malin, Antero & Linnakylä, Pirjo (2001) Multilevel Modelling in Repeated Measures of the Quality of Finnish School Life. *Scandinavian Journal of Educational Research* 45 (2), 145-166.
- Morrow, Virginia (2000) “Dirty Looks” and “Trampy Places” on Young People's Accounts of Community and Neighbourhood: Implications for Health Inequalities. *Critical Public Health* 10 (2), 141-152.
- Nordberg, Mai & Häkkinen, Unto (1997) Sairaalakäytön alueellinen vaihtelu – sairastavuuden, tarjonnan ja sosioekonomisten tekijöiden vaikutus. *Sosiaalilääketieteellinen Aikakauslehti* 34, 203-211.
- Pickett, Kate & Pearl, Michelle (2001) Multilevel Analysis of Neighbourhood Socioeconomic Context and Health Outcomes: A Critical Review. *Journal of Epidemiology of Community Health* 55, 111-122.
- Rasbash, Jon & Browne, William & Goldstein, Harvey & Yang, Min & Plewis, Ian & Healy, Michael & Woodhouse, Geoff & Draper, David & Langford, Ian & Lewis, Toby (2000) *A User's Guide to MLwiN. Version 2.1b*. Institute of Education. University of London. [http://www.mlwin.com/1\\_10/userman.pdf](http://www.mlwin.com/1_10/userman.pdf), luettu 12.6.2006.
- Raudenbush, Stephen & Bryk, Anthony & Cheong, Yuk & Congdon, Richard (2000) HLM 5. Hierarchical Linear and Nonlinear Modeling. *Scientific Software International*. <http://www.ssicentral.com/>, luettu 12.6.2006.
- Renchler, Ron (1992) Student Motivation, School Culture, and Academic Achievement. *What School Leaders Can Do. Trends & Issues*: <http://cepm.uoregon.edu/pdf/trends/motivation.pdf>, luettu 12.6.2006.
- Rinne, Risto & Kivirauma, Joel & Lehtinen, Erno (2004) Johdatus kasvatustieteisiin. Helsinki: WSOY.
- Rissanen, Pekka & Sund, Reijo & Nordback, Isto & Rousi, Timo & Idänpään-Heikkilä, Ulla (2002) Lonkkamurtuman hoidon vaikuttavuuden rekisteriperusteinen mittaaminen ja vertailu. *Stakesin aiheita* 21/2001. Helsinki: Stakes.
- Robinson, William (1950) Ecological Correlations and the Behaviour of Individuals. *American Sociological Review* 15, 351-357.
- Ryan, Allison (2001) The Peer Group as a Context for the Development of Young Adolescent Motivation and

## ”Monitasoanalyysi...”

---

Achievement. *Child Development* 72 (4), 1135-1150.

Sampson, Robert (2001) *The Embeddedness of Child and Adolescent Development: Community-Level Perspective on Urban Violence*. Teoksessa Joan McCord (toim.) *Violence and Childhood in the Inner City*. Cambridge: University Press, 31-77.

Snijders, Tom & Bosker, Roel (2002) *Multilevel Analysis. An introduction to Basic and Advanced Multilevel Modelling*. London: Sage Publications.

Stolle, Dietlind & Hooghe, Marc (2002) *The Roots of Social Capital. The Effects of Youth Experiences on Participation and Value Patterns in Adult Life*. Paper presented at the 98<sup>th</sup> annual meeting of the American Political Science Association Boston, August 29 – September 1, 2002.

Valkonen, Tapani & Kauppinen, Timo (2001) *Miesten kuolleisuuden alue-erot ja sosiaalinen segregaatio pääkaupunkiseudulla*. *Kvartti* (1), 7-21.

Valkonen, Tapani (1970) *Yhteisön vaikutukset yksilöön*

ja niiden kvantitatiivinen tutkiminen. Helsinki: Helsingin yliopisto.

Veenstra, Gerry (2000) *Social Capital, SES and Health: an Individual-level Analysis*. *Social Science & Medicine* 50 (5), 619-629.

### Lisätietoja tilasto-ohjelmista:

HLM: <http://www.ssicentral.com/hlm/index.html>

MLwiN: <http://www.mlwin.com/>

R: [www.r-project.org](http://www.r-project.org)

SAS: <http://www.psych.yorku.ca/lab/sas/>