

Vertailun vuoksi? Boolean algebra alueellisen vertailun välineenä

TOMMI INKINEN

Tietoyhteiskuntainstituutti, Tampereen yliopisto

Inkinen, Tommi (2003). Vertailun vuoksi? Boolean algebra alueellisen vertailun välineenä (For the sake of comparison? Using Boolean algebra in a regional analysis). Terra 115: 1, 19–29.



In this paper, the interpreted regional results obtained by using Qualitative Comparative Analysis (QCA) are evaluated and compared with results provided by the discriminant analysis, a statistical method. The focus is on evaluating a question included in a previous random sample survey. The principle of Occam's razor states that the simplest explanation should be the preferred one. The results suggest that the intuitive interpretations of both analyses provide similar outcomes, but QCA alone cannot provide the adequate understanding of a societal phenomenon. The logical power that the method has is clearly a strong addition to the explanation offered by the statistical method. The dichotomous approach, however, minimizes the human action into zeros and ones. These issues are never dichotomous in their essence. It seems that from a regional perspective QCA should be used in international comparisons as a supportive tool to confirm the interpretation results provided by variable-oriented statistical methods. In locally focused studies the method is too generalizing and variable-dependent. From a pedagogical perspective, the method should be introduced to undergraduate students. QCA is a logical and systematic tool and in methodological studies it can offer a valuable insight into alternative methods and methodological developments in geography and other social sciences.

Keywords: Boolean algebra, regional analysis, methodology, comparative research

Tommi Inkinen, Information Society Institute, University of Tampere, FIN-33014 Tampere, Finland. E-mail: <tommi.inkinen@uta.fi>

Tarkastelen artikkelissani Charles Raginin (1987, 1994a, 1994b, 2000) tunnetuksi tekemän kvalitatiivisen vertailevan analyysin (englanniksi Qualitative Comparative Analysis, QCA) käyttökelpoisuutta empiirisen aineiston alueellisessa vertailussa. Menetelmän antamien tulosten intuitiivista tulkintaa vertailen muuttujalähtöisen tilastotestin pohjalta tehtävien tulkintojen kanssa. Käyttäessään QCA-menetelmää tutkijan ensisijainen tehtävä on etsiä yhtäläisyyksiä ja eroja harvojen tapauksien kesken. QCA pyrkii kausaalisuuden löytämiseen ja systemaattiseen aineistoreduktioon, jolla tarkoitetaan oleellisten taustatekijöiden löytämistä selittävien muuttujien joukosta. Tyypillisesti QCA:ta käytetään vertailuissa, joissa tapauksia on muutamista aina noin viiteenkymmeneen. Raginin mukaan vertailevan tutkimuksen laajan kirjon tärkein yhteinen nimittäjä on erilaisuuden (diversiteetin) esiin tuominen, jolloin kulttuuriset ja historialliset erityispiirteet korostuvat tapausalueilla.

Vertailevan tutkimuksen lähtökohtana ovat useimmiten valtio- tai kansakuntatasoiset aggregaattialueet. Sosiologisessa vertailevassa tutkimuksessa maiden välisen erojen rakentumi-

nen nähdään pääasiallisesti historian kautta. Täten historiallinen selittäminen on elimellinen osa menetelmää. Selittävinä muuttujina voivat toimia esimerkiksi hallintojärjestelmän kehittymiseen liittyvät suuret muutokset, kuten vallankumoukset tai sodat. Tämä on kuitenkin vain yksi tapa tehdä vertailevaa tutkimusta (ks. Ragin 2000; Hellström 2001; Rantala 2001). Vertailevan tutkimuksen näkökulmasta valtiot historiallisina yksikköinä ovat kuitenkin helppoja tutkimuskohteita: Kehittyneissä teollisuusvaltioissa kansallisella tilastoinnilla on pitkät perinteet ja valtioiden poliittiset tapahtumat ovat usein yksikäsitteisesti todennettavissa historiankirjoituksen kautta (esim. vallankumoukset).

Artikkelini rakenne on kolmiosainen. Ensin selvitän QCA:n perusteet, mikä käytännössä tarkoittaa Boolean laskennan perusteita. Toiseksi analysoin QCA-menetelmän avulla aiemmin tekemääni kyselytutkimukseen (ks. Inkinen 2001) sisällynyttä kysymystä: ”Koetteko itsenne (a) kaupunkilaisiksi, (b) maaseudun asukkaaksi vai (c) edellisten välimuodoksi?” Kysymys on kolmiluokkainen, mutta analyysissä se on helposti muutettavissa kaupunki-

maaseutu-dikotomiaksi analyysityökalulaajennuksen avulla. Analysoinnissa pyrin löytämään selittäviä tekijöitä kokemukseen perustuvan mielipiteen taustalta ja tulkitsemaan tutkittavaa ilmiötä. Kolmanneksi arvioin QC-analyysistä saatuja tuloksia tilastollisen testin (erotteluanalyysin) antaman tulosten valossa. Täten vertailen artikkelissa kahden eri menetelmän tuottamia tulkintoja: kysyn, poikkeavatko tulokset toisistaan ja täydentävätkö menetelmien tulokset toisiaan. Päätelmissäni peilaan artikkelin tuloksia aiempiin empiirisiin tutkimuksiin, joissa QCA:ta on käytetty.

QCA-menetelmän perusteet

Boolean algebra, totuustaulut ja dikotomia

Kvalitatiivisten tekijöiden vertailumenetelmä pohjautuu Boolean algebraan. Sen kehitti George Boole 1800-luvun puolivälin tienoilla. Boolean algebraa on eniten käytetty virtapiirikentäkaavioiden yksinkertaistamiseen. Tämä on myös yhteys, josta se parhaiten tunnetaan. Yhteiskuntatutkimuksessa Boolean algebra tarjoaa loogisen vaihtoehdon systemaattiseen toimintaan. Esitettävät laskentaperiaatteet ovat yksinkertaisia, mutta niiden ymmärtäminen vaatii perustietämyksen Boolean laskennasta.

Vertailu aloitetaan jakamalla tutkimusalueet ryhmiin riippuen siitä, poikkeavatko alueet toisistaan tutkimusongelmaan nähden: jos ilmiö F esiintyy alueella A muttei alueella B , nämä kaksi aluetta luokitellaan eri ryhmiin. Seuraavaksi arvioidaan mielekkäät selittävät muuttujat, jotka todennäköisesti vaikuttavat ilmiön esiintymiseen. Muuttujien valinta voi perustua esimerkiksi aiempiin tutkimustuloksiin. Tehty muuttujaluokittelu muodostaa aineistomatriisin, jota kutsutaan *totuustauluksi*. Analyysin tavoitteena on löytää samankaltaisuuksia niiden tapausalueiden välillä, joilla tutkittava ilmiö esiintyy. Menetelmässä sovelletaan binaari-logiikkaa eroja ja yhtäläisyyksiä tuottavien tekijä-yhdistelmien löytämiseksi.

Boolean algebrassa on kaksi ilmiön olotilaa: tosi (*true*) ja epätosi (*false*). Edellisessä ilmiö esiintyy, jälkimmäisessä ilmiö ei esiinny tapausalueella. Käytettävät muuttujat ovat täten nominaaliasteikollisia. Numeeriset muuttujat muutetaan nominaalisiksi luokittelemalla. Tämä aiheuttaa jonkin verran tiedon häviämistä. Kvalitatiiviset ilmiöt ovat tosin usein jo valmiiksi nominaaliasteikollisia. Mikäli kategorisessa muuttujassa on enemmän kuin kaksi luokkaa, kysymys esitetään usean binaarimuuttujan avulla. Esimerkiksi viisi-luokkainen koulutusluokitus (akateeminen, ammattikorkeakoulu, opisto, ylioppilas, peruskoulu) esitetään viitenä omana muuttujanaan muodossa

”Onko henkilöllä akateeminen koulutus?”, johon vastataan ”kyllä” tai ”ei”. Kysymys siis toistetaan jokaisen koulutusasteen kohdalla. Aineistotaulut voivatkin kasvaa helposti hyvin suuriksi käytettäessä perinteisiä sosioekonomisia taustamuuttujaluokituksia. Lopullisessa totuustaulussa yhdistetään sellaiset tapaukset, joissa esiintyy samanlainen lopputila eli 0 (ilmiö ei esiinny) tai 1 (ilmiö esiintyy). Totuustauluun muodostuu niin monta loogista yhdistelmää (riviä) kuin selittävien muuttujien binaariarvojen kautta on mahdollista rakentaa. Esimerkiksi neljän selittävän muuttujan taulussa on kuusitoista riviä ($2^4 = 16$).

QCA:n ongelmat liittyvät pääasiallisesti muuttujien muuntamiseen ja niiden tyyppittelyyn. Alueellisia ominaisuuksia täytyy redusoida muuttujia arvottaen. Esimerkiksi kysymys ”Onko alueella runsaasti kulttuuritarjontaa?” esitetään muodossa ”kyllä” tai ”ei”. Linjaus ”runsaaseen” ja ”vähäiseen” on kuin häilyvä veteen piirretty viiva. Selittävän muuttujan muotoilu on tutkijan harkinnan varassa. Toisaalta on tarpeen muistaa, että yleistäminen esimerkiksi suorista jakaumista voi pahimmillaan olla yhtä mielipiteenvaraista. Kategorisointi on aina arvosidonnainen prosessi, joka on erityisen tärkeässä asemassa yleistäviä menetelmiä käytettäessä.

Boolean laskutoimitukset ovat logiikkaa – eivät aritmetiikkaa. Lopputuloksella F on vain arvo 0 tai 1. Lopputilaehtojen eli selittävien muuttujien olemassaolo esitetään ”ISOINA” tai ”pieninä” kirjaimina. Isolla kirjoitetut muuttujat merkitsevät kausaaliehdon läsnäoloa (tosi) ja pienellä kirjoitetut ehdon puuttumista (epätosi). Ehtoja käsitellään loogisten operaattoreiden eli yhteenlaskun ja kertolaskun avulla. Boolean yhteenlaskussa + merkki symboloi loogista TAI (englanniksi OR) -operaattoria. Esimerkiksi $F = A + B$ tarkoittaa sitä, että jos A tai B ilmenevät toisistaan riippumatta eli saavat aineistossa arvon 1, F tapahtuu. A ja B voivat esiintyä myös yhtäaikaisesti. Kertolaskussa yhtälön $F = A \cdot B$ tulkinta on, että lopputulos F (0 tai 1) on riippuvainen yhtäaikaisesta A :n läsnäolosta sekä b :n ja c :n puuttumisesta. Boolean kertolaskun tulkinta on sama kuin looginen JA (englanniksi AND) -operaattori. Boolean laskennan eksaktit matemaattiset perusteet löytyvät esimerkiksi *Matematiikan käsikirjasta* (Thompson 1994).

Totuustaulujen minimointi

Kausaalisuuden näkökulmasta tärkeimpiin selittäviin tekijöihin päästään käsiksi minimoimalla. Boolean analyysi perustuu kombinatoriseen logiikkaan ja tällöin selittävän ilmiön esiintymisellä (1) ja puuttumisella (0) on sama merkitys laskentaprosessin

Taulukko 1. Esimerkkiyhtälön minimoidut päätekijät (sovellettu lähteestä Ragin 1987: 96).

Table 1. Minimized prime implicants in the example equation (modified from Ragin 1987: 96).

	Abc	abC	ABc	AbC
Ac	X		X	
Ab	X			
bC		X		X

kulussa. Yhtälöt voivatkin rakentua monimutkaisiksi ja laajoiksi, jolloin tarvitaan keino minimoida kausaalisten selittävien muuttujien joukko siten, että jäljelle jäävät tärkeimmät selittävät tekijät. Ragin (1987: 93) esittää seuraavan säännön:

Jos kaksi Boolean lausetta eroavat toisistaan vain yhden kausaalisen ehdon kohdalla ja kuitenkin tuottavat saman lopputuloksen, voidaan erottava ehto poistaa ja näin tuottaa yksinkertaisempi kombinoitu lause.

Boolean minimoinnista voidaan esittää seuraava esimerkki. Ajatellaan hypoteettinen Boolean yhtälö $F = Abc + ABc + abC + AbC$. Looginen minimointi jatkuu kunnes yksinkertaisin yhtälö saavutetaan. Käytännössä minimointi tehdään yleensä tietokoneella mutta esimerkistämme havaitaan helposti toisensa kumoavat ehdot. Esimerkiksi Abc ja ABc aiheuttavat ehdon b häviämisen sen läsnäolon ja puuttumisen yhtäaikaisuuden vuoksi.

Lähtöyhtälön yksinkertaistettu muoto on $F = Ac + Ab + bC$. Saadun tuloksen intuitiivinen tulkinta on: tila F toteutuu, kun A tapahtuu ja C ei tapahdu tai A tapahtuu ja B ei tapahdu tai B ei tapahdu ja C tapahtuu. Edelliset ehdot voivat toteutua toistensa kanssa samanaikaisesti. Minimoinnissa täytyy muistaa, että vain yksi kausaalinen tekijä saa vaihdella lauseitten välillä. Asiaa voidaan edelleen kehittää käyttämällä *implikaation* käsitettä. Tämä tarkoittaa, että jos Boolean termi on jäsenenä sen alajoukkoon kuuluvassa Boolean termisä, se implikoi sitä. Esimerkiksi A käsittää kaikki Abc :n vaihtoehdot eli A implikoi Abc :tä. Implikaatio on tärkeä väline minimointioperaatioissa. Esimerkkiämme voidaan kehittää edelleen muotoon $F = Ac + bC$. Tämä käy ilmi taulukosta 1, jossa esitetään alkuperäistermien ja minimoitujen termien (*prime implicants*) välinen suhde. Komponentti Ab esiintyy molempien päätekijöiden kaut-

ta, minkä vuoksi lisäredusointi mahdollistuu esitettyyn muotoon.

Minimointioperaatioiden kautta on mahdollista löytää loogisesti yksinkertaisin tekijäjoukko ilmiö F :n taustalla. Tämä on positiivinen lähtökohta kausaaliteetin etsinnässä. Vaihtoehtoisesti voidaan etsiä ratkaisut negatiivisille tuloksille, jolloin ilmiö F ei esiinny. Tällöin sovelletaan *de Morganin lakia*, joka tarjoaa helpon tavan muuttaa asian tarkastelu päinvastaiseksi ja minimointia ei tarvitse enää toistaa alusta alkaen. Käyttäkäämme aikaisempaa esimerkkiä, jossa lopputulokseksi muodostui $F = Ac + bC$. Esiintyvät ilmiöt muutetaan nyt poissaoleviksi ($A \rightarrow a$), eli ne käännetään päinvastaisiksi. Seuraavaksi looginen JA-operaattori muutetaan loogiseksi TAI-operaattoriksi ja päinvastoin. Käyttämällä näitä kahta toimenpidettä saadaan $F = (a + C)(B + c) = aB + ac + CB$.

Saatu yhtälö on minimoitu ratkaisu ilmiön F esiintymättömyydelle. De Morganin laki tarjoaa nopean ja helppokäyttöisen tavan siirtyä vastakohtaiseen tarkasteluun, mikä antaa lisäsyvyyttä saatujen tulosten ymmärtämiseen. De Morganin lain avulla tuotettuja Boolean yhtälöitä on usein tarpeellista yksinkertaistaa helpommin ymmärrettävään muotoon. Yhtälöiden supistaminen ei poikkea tavanomaisesta algebrasta.

Ajatelkaamme uutta esimerkkilauseetta, joka on muotoa $F = AS + AF + AH$. Saatu yhtälö voidaan käsitellä muotoon, joka osoittaa, että A on *välttämätön* ehto kausaalisuudelle, jolloin $F = A(S + F + H)$. Lauseiden käsittelyn avulla saadaan helposti esille välttämättömät ehdot kausaalisuudelle. Lisäksi se osoittaa keskenään tasaveroiset (ekvivalentit) kausaalisuusehdot eli esimerkissämme S , F ja H .

Yhtälöiden luokittelua käytetään myös selkeyttämään ehtoja. Voidaan ajatella hieman monimutkaisempi kausaalimalli $F = ABC + aBC + aDR + Y$. Siinä tekijä A ilmenee molemmissa olotiloissa (0 ja 1). Hedelmällinen lähestyminen on tarkastella F :n ilmenemistä A :n kautta. Tällöin yhtälö voidaan järjestää uudelleen muotoon, josta havaitaan F :n toteutumisen asiayhteydet joko A :n läsnä- tai poissaolon tapauksissa. Yhtälöön on otettu mukaan niin sanottu ulkoinen tekijä Y , joka esiintyy molemmissa tilanteissa. Edellinen on esimerkki Boolean ja normaalilaskennan välisistä eroista. Esimerkin lopullinen yhtälö on muotoa $F = a(BC + DR + Y) + A(BC + Y)$. A :n ilmeneminen tai puuttuminen tutkittavan ilmiön selittämisessä riippuu täten muiden selittävien tekijöiden muodoista.

Empiirinen evaluointi

Tutkimusasetelmasta

QCA:n kehittämisen taustalla on ollut tavoite löytää välimuoto kvantitatiiviselle ja kvalitatiiviselle lähestymistavalle. Menetelmän käytettävyyden arvioinnissa on tällöin keskityttävä tarkastelemaan käytettävien muuttujien ominaisuuksia kriittisesti. Tämän artikkelin tutkimusasetelmaa havainnollistaa empiirinen esimerkki, jossa sekä QCA:n että tilastollisen erotteluanalyysin tuottamat tulokset kuvautuvat tutkijan tekemien valintojen (käytettävien selittävien tekijöiden) kautta.

Kaikessa tutkimuksessa tavoitteena tulisi olla olemassa olevan maailman ymmärtäminen. Kvantitatiivisessa tutkimusotteessa kausaalisuuden etsintä on perinteisesti edustanut objektiivisen tieteen ihannetta. Ihmisiä ja heidän yhteiskuntaan tutkivissa tieteissä syy–seuraus-suhteen osoittaminen on kuitenkin ongelmallista. Hyväksyttävissä kausaliiteetin ongelmallisuus sekä luokittelevien menetelmien aineistoa jäsentävä luonne niin QCA kuin tilastoanalyysikin palautuvat periaatteessa samanhenkisiksi työkaluiksi. Käyttämäni esimerkin tavoitteena on valottaa tiedon tuottamisen prosessia sekä peilata saatuja tuloksia muihin tutkimuksiin, joissa QCA:ta ja tilastoanalyysiä on käytetty rinnakkaisina menetelminä.

Aineisto ja arvioitava kysymys

Käyttämäni aineiston keräsin laajempaa kvantitatiivista tutkimusta varten Varsinais-Suomen kunnista helmikuussa 2001. Tutkimuskunnat ovat Aura, Kaarina, Korppoo, Marttila, Nauvo, Raisio, Rymättylä, Tarvasjoki ja Turku. Aineistosta irrotin yhden kysymyksen tämän artikkelin empiiriseksi esimerkiksi. Kysymys on: ”Koetteko itsenne (a) kaupunkilaiseksi, (b) maaseudun asukkaaksi vai (c) edellisten välimuodoksi?” Sen erilliseen tarkasteluun on kolme syytä. Ensinnäkin konkreettinen kysymys havainnollistaa tuloksia paremmin kuin yleisen muodon tarkastelu. Toiseksi kysymys on maantieteellisesti mielenkiintoinen, koska se selvittää ihmisten kokemusta heidän asuinalueidentiteetistään. Muuttujaa on luontevaa käsitellä vastakkainasetteluna (kaupunki–maaseutu). Käytettävässä QCA-ohjelmistossa on mahdollista käsitellä myös ”sumeaa” välimuokkaa. Kolmanneksi muuttuja mittaa kvalitatiivista mielipidettä ja kokemuspohjaista tekijää, jollaista varten QCA alun perin kehitettiin.

QCA:ssa selittävinä muuttujina toimii neljä Tilastokeskuksen kunta-aineistosta (SVT 1999) muovattua muuttujaa sekä yksi kyselylomakkeesta saatu muuttuja. Kyselylomakkeesta sisällyttiin analyysiin kysymyksen ”Koetteko etäisyydet

kotikunnassanne pitkiksi?” Muuttuja perustuu kategorisen järjestysasteikollisen muuttujan uudelleenluokitteluun. Alkuperäisessä muodossa kysymys oli neliluokkainen, jonka dikotomisoin tätä analyysiä varten. QCA:n vertailukohtana toimivassa tilastoanalyysissä käytän laajempaa kvantitatiivisten muuttujien joukkoa, johon sisältyy kuusitoista alueellista tekijää. Tilastolliset testit tein henkilökohtaisten muuttujien sekä alueellisten aggregaattimuuttujien suhteen.

QCA:ssa ja tilastoanalyysissä päädyin käyttämään eri selittäviä muuttujia. Ratkaisuun on monta perustetta. Ensinnäkin alueelliset taustamuuttujat, jotka käytännöllisesti katsoen aina kuvaavat prosenttiosuuksia (esim. teollisuustyöntekijöiden osuus alueen kaikista työntekijöistä) muuttuvat dikotomiaa tehdessä helposti samoiksi. QCA:han on valittu tässä sellaiset muuttujat (kunnan työttömyysaste ja asukkaiden keski-ikä), joilla riittävät erot ovat. Toiseksi tilastollisessa analyysissä käytettävät muuttujat eivät ole kvalitatiivisia. QCA:ssa on käytettävä intuitiivisia muuttujia kuten ”Onko kunta maaseutumainen?” tai ”Onko kunnassa paljon palveluita?” Näiden asioiden rajaamisessa ei ole objektiivisia mittareita ja luokitukset pohjautuvat stereotyyppeihin. Kolmanneksi käyttämäni tilastomenetelmä on *askeltava* eli muuttujajoukkoa rajaava. Mikäli analyysissä käytettäisiin samoja taustamuuttujia, QCA:n minimointialgoritmi käsittelisi koko taustamuuttujajoukon samantapaisena. QCA:ssa analysoitavat muuttujat on valittu siten, että niiden välillä on riittävästi eroja. Menetelmälähtöisyys onkin eräs QCA:n perusongelmista, joka on otettava huomioon analyysiä tehtäessä.

Analysoitavat totuustaulut käsittelen neljänä kokonaisuutena (taulukko 2). Ensimmäiseksi sovellan Boolean menetelmää siten, että luon totuustaulun täydelliseksi (taulukossa 2 taulu A) ja suoritan tiukat yleistävät linjaukset kuntien kohdalla sen mukaan, kuinka enemmistö on vastannut. Toisin sanoen: jos 45 prosenttia vastaajista kokee itsensä kaupunkilaiseksi ja 51 prosenttia maaseudun asukkaaksi, tapaus luokitellaan maaseutua korostavaan ryhmään. Ensimmäisessä tapauksessa välimuotoluokan vastaukset jätetään pois ja muuttujaa käsitellään vain kaksiluokkaisena.

Toisesta taulusta (2B) eteenpäin ”sumea” välimuokkaa otetaan mukaan ja merkitään kysymysmerkillä (?) ne tapaukset, joissa välimuotoluokka on saanut eniten vastauksia. QCA-ohjelmistossa on mahdollista käsitellä *don't care* -luokkia, jotka soveltuvat niille tapauksille, joita ei luontevasti voida jakaa suuntaan tai toiseen. Kolmannessa taulussa (2C) tällaisten luokkien määrää lisätään intuitiivisen tulkinnan mukaan. Esimerkik-

Taulukko 2. Neljä totuustaulua aluekokemuksen selittämisen lähtökohtana.

Table 2. Four truth tables as a starting point in the analysis of perceived rural–urban affiliations.

A			B		
Ehto	Tulos	Tapaukset	Ehto	Tulos	Tapaukset
Condition	Result	Cases	Condition	Result	Cases
K E M I T	A		K E M I T	A	
1 1 1 1 1	1	1, Turku	1 1 1 1 1	1	1, Turku
1 1 1 0 1	1	1, Raisio	1 1 1 0 1	?	1, Raisio
1 1 1 0 0	1	1, Kaarina	1 1 1 0 0	1	1, Kaarina
0 1 0 0 1	0	1, Aura	0 1 0 0 1	0	1, Aura
0 0 1 1 0	0	2, Marttila, Tarvasjoki	0 0 1 1 0	0	2, Marttila, Tarvasjoki
0 0 1 1 0	0	1, Rymättylä	0 0 1 1 0	?	1, Rymättylä
0 0 0 1 1	0	1, Korppoo	0 0 0 1 1	0	1, Korppoo
0 0 0 1 0	0	1, Nauvo	0 0 0 1 0	0	1, Nauvo
K = Kaupunkistatus (City status)			K = Kaupunkistatus (City status)		
E = Lyhyet etäisyydet (Short distances)			E = Lyhyet etäisyydet (Short distances)		
M = Muuttovoitto (Positive net migration)			M = Muuttovoitto (Positive net migration)		
I = Korkea keski-ikä, >39 v. (High average age, >39 yrs.)			I = Korkea keski-ikä, >39 v. (High average age, >39 yrs.)		
T = Korkea työttömyysaste, >12 % (High unemployment rate)			T = Korkea työttömyysaste, >12 % (High unemployment rate)		
A = Kokemus elämäntavasta (kaupunki–maaseutu)			A = Kokemus elämäntavasta (kaupunki–maaseutu)		
Perceived lifestyle (urban–rural)			Perceived lifestyle (urban–rural)		

C			D		
Ehto	Tulos	Tapaukset	Ehto	Tulos	Tapaukset
Condition	Result	Cases	Condition	Result	Cases
K E M I T	A		K E M I T	A	
1 1 1 1 1	1	1, Turku	1 1 1 1 1	1	1, Turku
1 1 1 0 1	?	1, Raisio	1 1 1 0 1	?	1, Raisio
1 1 1 0 0	?	1, Kaarina	1 1 1 0 0	?	1, Kaarina
0 1 0 0 1	?	1, Aura	0 1 0 0 1	?	1, Aura
0 0 1 1 0	0	2, Marttila, Tarvasjoki	0 0 1 1 0	?	2, Marttila, Tarvasjoki
0 0 1 1 0	?	1, Rymättylä	0 0 1 1 0	?	1, Rymättylä
0 0 0 1 1	0	1, Korppoo	0 0 0 1 1	?	1, Korppoo
0 0 0 1 0	0	1, Nauvo	0 0 0 1 0	0	1, Nauvo
K = Kaupunkistatus (City status)			K = Kaupunkistatus (City status)		
E = Lyhyet etäisyydet (Short distances)			E = Lyhyet etäisyydet (Short distances)		
M = Muuttovoitto (Positive net migration)			M = Muuttovoitto (Positive net migration)		
I = Korkea keski-ikä, >39 v. (High average age, >39 yrs.)			I = Korkea keski-ikä, >39 v. (High average age, >39 yrs.)		
T = Korkea työttömyysaste, >12 % (High unemployment rate)			T = Korkea työttömyysaste, >12 % (High unemployment rate)		
A = Kokemus elämäntavasta (kaupunki–maaseutu)			A = Kokemus elämäntavasta (kaupunki–maaseutu)		
Perceived lifestyle (urban–rural)			Perceived lifestyle (urban–rural)		

si jos 35 prosenttia sanoo itseään kaupunkilaiseksi, 33 prosenttia välimuodoksi ja 32 prosenttia on maaseudulta, tapauskunta luokitellaan *don't care*-ryhmään. Tämä lisää välimuototapausten määrää. Taulukon 2 viimeisessä taulussa (2D) käsitellään voimakkaimmat ääripäät, jotka aineistossa ovat Turku (kaupunki) ja Nauvo (maaseutu). Turkulaisista vastaajista kaupunkilaiseksi itsensä koki 59,4 prosenttia ja maaseudun asukkaaksi 4,3 prosenttia. Vastaavasti Nauvossa yksikään vastaaja ei kokenut itseään kaupunkilaiseksi ja maaseu-

dun asukkaaksi itsensä luokitteli 58,3 prosenttia. Dikotomian käyttö näiden tapausten kohdalla vaikeuttaa perustellulta.

Taulukon 2 totuustaulut kuvaavat eri tavoin jäsennettyjä luokituksia selitettävästä muuttujasta. Toimenpide voitaisiin tehdä myös selittäville muuttujille, joista erityisesti työttömyysaste ja korkea keski-ikä ovat tulkinnallisia. Kuitenkin jo selitettävän muuttujan säätely osoittaa menetelmän antamien tulosten tulkinnallisen ongelman. Tarkastelussa siis katson, kuinka QCA:n tuottaman

yksinkertaistetun totuustaulun tulkinta muuttuu, kun sumeiden luokkien määrää kasvatetaan.

Alueellisten erityispiirteiden erottaminen yksittäiseksi selittäviksi tekijöiksi on vaikeaa. Taulukon 2 totuustauluissa käytetyt selittävät muuttujat kuvaavat erilaisia alueita luonnehtivia tekijöitä. Luokitukset 0 ja 1 kuvaavat määräkäsitteitä *suuri* tai *pieni* sekä *paljon* tai *vähän*. Luokkien rajanvedon tulisikin perustua selvästi poikkeavien määrien erottamiseen. Tapauksissa, joissa suhteelliset osuudet ovat lähellä toisiaan, luokittelu on tehty mahdollisimman tasapainoista jakaumaa silmälläpitäen. Esimerkkinä on työttömyysasteen arvottaminen joko ”suureksi” tai ”pieneksi”. Aineistossa työttömyysasteen keskiarvo on noin 13 prosenttia. Tähän perustuen vedin rajan arvojen 12 (pieni) ja 13 (suuri) väliin. Rajanveto kuvaa selkeästi luokitteluongelmaa, joka usein kohdataan muuttaessa kvantitatiivinen muuttuja kvalitatiiviseksi ryhmittelyksi.

Tulokset ja vertailu

QCA-työkaluna käytin QCA 3.0 -ohjelmaa, joka on perustyökalu Boolean yhtälöiden käsittelyssä. Taulukon 2 neljä totuustaulua jakaantuvat kolmeen pääluokkaan: ehtoihin, tuloksiin ja tapauskuntiin. Selittäviksi muuttujiksi valitsin viisi tekijää, jotka takaavat riittävän monimuotoisuuden muuttujien välillä. Totuustaulun käsittelyssä ohjelmistossa sovelletaan yleensä Quine–McCluskey-algoritmiin perustuvaa minimointimenetelmää. Menetelmä varmistaa pienimmän mahdollisen lopputuloksen loogisten yhdistelmien etsinnässä, mutta sen haittapuolena on laskennallinen hitaus suurien aineistotaulujen tapauksissa (Drass 1992: 30).

Taulukkojen minimointi lopputilan 1 suhteen antaa tulokseksi seuraavat Boolean yhtälöt (kirjainsymbolien selitykset taulukossa 2):

- (1) $A = KEMT + KEMi$
- (2) $A = KEMT + KEMi - \text{don't care} = -$
 $A = KEMIT + KEMit - \text{don't care} = 1, 0$
- (3) $A = KEMT - \text{dонт't care} = -$
 $A = KEMIT - \text{don't care} = 1, 0$
- (4) $A = KEMT - \text{don't care} = -$
 $A = KEMIT - \text{don't care} = 1, 0$

Taulukon 2 taulun A minimoinnin tuloksena saatu lauseke merkitsee sitä, että kaupunkilaisuuden kokeminen on riippuvaista kunnan kaupunkiasemasta, lyhyistä etäisyyksistä, positiivisesta muuttovoitosta ja matalasta työttömyysasteesta. Vaihtoehtoinen (TAI) tila sisältää kolme ensimmäistä ehtoa ja matalan keski-ään. Saadut tulokset ovat yhteneväiset kvantitatiivisen analy-

sin antamien tulosten kanssa, ja analysoidut kvalitatiiviset tekijät ovat yhteydessä tiheätä asutusta kuvaaviin tilastomuuttujiin (taulukko 3, taulu A). Tilastoanalyysiin verrattuna QCA kuitenkin hävittää tietoa. Laajempi kuva ilmiöstä saadaan tarkasteltaessa käänteistulosta, jolloin totuustaulu minimoidaan arvon 0 suhteen ja tulokseksi saadaan $a = keml + keIt + kEmiT$. Tulos osoittaa käänteisen tuloksen, joka tarkoittaa maaseutuidentiteetin olevan yhteydessä korkeaan keski-ikään, alueen työttömyyteen ja pitkiin etäisyyksiin kun ne esiintyvät samanaikaisesti muuttotappion kanssa.

Taulukon 2 taulusta B eteenpäin suoritin minimoinnin suhteessa sumeiden luokkien vaihtoehtoihin. Tulkinnallisesti ero syntyy vain mikäli *don't care* -luokiksi valitaan ?-luokat. Luokat 1 ja 0 tuottavat samat minimointitulokset. *Don't care* -luokkien lisääminen yksinkertaistaa tulosyhtälöitä. Ristiriitatapauksien määrä vähenee ja malleista tulee väljempää. Minimoidaessa taulu 2B vaihtoehtoisesti arvon 0 suhteen saatu tulos on tismalleen sama kuin taulun 2A tapauksessa. Tämä pätee myös käytettäessä kaikkia vaihtoehtoarvoja eli arvoja -, 1 ja 0.

Taulukon 2 taulujen C ja D kohdalla havaitaan, että sumeiden luokkien lisääminen ei enää muuta minimoitua loppuyhtälöä. Tulosten mukaan olisi perusteltua väittää, että sumeiden luokkien lisääminen vähentää tulkinnallisia vaihtoehtoja. Taulujen 2C ja 2D minimoinnit käänteisarvon 0 suhteen tuottavat triviaalisti de Morganin lain mukaan samat tulokset. Kaikille *don't care* -luokille saadaan Boolean yhtälö $a = keml + keIt$. Tulos tukee olotilan 1 minimoinnista saatavia päätelmiä. Etäisyydet, muuttoliike ja alueellinen työttömyys ovat yhteydessä selitettävään ilmiöön. Pitkien etäisyyksien esiintyminen samanaikaisesti korkean keski-ään kanssa tukee myös kvantitatiivisen analyysin tuottamia tuloksia.

Tilastollisten menetelmien avulla saadut tärkeimmät selittävät alueelliset muuttujat esitettyyn ongelmaan ovat kunnan elinkeinorakenteeseen, ikärakenteeseen ja koulutukseen liittyvät aggregaattimuuttujat. Kyselylomakkeesta saatujen selittävien henkilökohtaisten muuttujien (ikä, koulutus ja tulotaso) ohella myös etäisyyttä kuvaavat muuttujat ovat merkitseviä. Tulokset ovat yhdensuuntaiset QCA:n tuottamien tulosten kanssa. Taulukossa 3 esitetään aluekohtaisten muuttujien avulla tehdyn tilastotestin tulokset. Menetelmänä käytin erotteluanalyysiä (esim. Mäkinen 1978: 191–195; Alkula ym. 1994), joka on soveltuva menetelmä ryhmittäisen ennustemallin laatimiseen havaittujen tapausarvojen pohjalta. Menetelmä soveltuu kahden tai useamman ryhmän (tässä kolmiluokkainen kokemusmuuttuja) erotteluun

selittävien muuttujien suhteen. Selitysmalli perustuu lineaarisiin yhdistelmiin ja sen yleinen muoto muistuttaa lineaarista monimuuttujaregressioyhtälöä. Taulukkoon 3 lisätty graafinen tarkastelu (3C) havainnollistaa alueellisuutta luonnehtivien ryhmäkeskusten sijoittumisen hajanaisuutta.

Taulusta 3A saadaan käsitys niistä muuttujista, jotka ovat tärkeimmässä tilastollisessa yhteydessä selitettävään kokemusmuuttujaan. Näistä ensisijainen on alueen asukastiheys, jolla on suurin yhteisvaihtelu selitettävän muuttujan kanssa. Askeltava menetelmä sisältää neljä vaihetta. Niistä viimeinen sisältää tilastomenetelmän antamat tärkeimmät tekijät, joiden kautta kaupunkilaisuuden kokeminen on ymmärrettävissä. Muuttujat ovat kaupunkiasutukselle tyypillisiä. Pareja vertaamalla havaitaan, että tilastollisesti ryhmien erot ovat merkitseviä. Vastaavasti rakennematriisi kertoo korrelaatiojärjestyksen selittävien muuttujien joukossa suhteutettuna kahteen saatuun ulottuvuuteen eli diskriminanttifunktion.

Verrattaessa tilastollisen testauksen muuttujaryhmiä funktioiden 1 ja 2 sisällä QCA-yhtälöihin havaitaan niiden samansuuntainen käyttäytyminen. QCA ei pysty kuitenkaan kertomaan mitään yhteisvaihtelun voimakkuudesta, joten ensisijaista on yhteyden suunnan määrittäminen: Ilmiö vain joko ”on” tai ”ei ole”. Tulkinta pysyy samana nostettaessa QCA:ssa käytettävien selittävien muuttujien määrää tilastotestausta vastaavaksi. Käyttämässäni esimerkeissä selittäviä muuttujia on QCA:ssa viisi ja tilastomallissa neljätoista kappaletta. QCA:n käyttäminen vaatiikin tutkijalta enemmän taustatyötä käytettävien tekijöiden valinnassa.

Näkökulmia aiemmista soveltavista QCA-tutkimuksista

Pertti Alasutari (1994: 170–171) on arvioinut QCA:n soveltuvuutta kulttuurintutkimukseen todeten menetelmän hyödyn olevan ennen kaikkea totuustaulujen laadinnan kurinalaisuudessa. Yksinkertaistaminen kuitenkin johtaa ongelmiin, jolloin menetelmää tulisi käyttää enemmänkin aineistoa jäsentävänä luokitteluvälineenä. Kausaalinen kysymyksenasettelu on vertailevassa tutkimuksessa usein hankalaa, koska tapauskohteita on yleensä vain muutamia. Alle kymmenen tapauskohteen vertailut aiheuttavat ongelmia kvantitatiivisten menetelmien käytössä (esim. Toivonen 1999).

Olli Kangas (1994) on käsitellyt eri menetelmiä sosiaaliturvan järjestelmien tutkimuksessa. Lähtökohtana on valtiotas, jossa QCA:ta on yleensä käytetty. Tilastollisina menetelminä Kangas käyttää lineaarista pienimmän neliösumman estimoin-

tia ja klusterianalyysia. Lähtökohta on mielenkiintoinen, koska tässä kirjoituksessa valitsin tilastomenetelmäksi erotteluanalyysin, joka on klusterianalyysin tavoin aineistoa luokitteleva menetelmä (ks. Alkula ym. 1994).

Kangas (1994: 360) toteaa QCA:n ja regressioanalyysin tuottaneen kutakuinkin vastaavanlaiset löydökset. Verrattuna tässä kirjoituksessa käsittelemääni aineistoon tulos on siis vastaava. Kvalitatiivisen vertailun ja regressioanalyysin eroksi Kangas esittää QCA:n tehokkaamman kyvyn paljastaa kausaalisia yhteyksiä pienissä aineistoissa. Valtiotasolla näin ehkä onkin. Pienemmille aluetasolle ja kansallisen kulttuurin sisään siirryttäessä kausaalisten syy–seuraus-suhteiden etsiminen kuitenkin vaikeutuu ja ihmisen toiminnan mallipohjainen tai lineaarinen ymmärtäminen ei vaikuta nykypäivän käsityksiä vastaavalta. Tilastomenetelmien ja QCA:n tuottamien muuttujaryhmien merkitys on lähinnä aineiston ryhmittelyssä. Niiden avulla saadaan tietoa mahdollisesti ilmiöön vaikuttavista tekijöistä mutta eksplisiittisen syyn ja seurauksen etsintä paikallistasolla on hyvin hankalaa ellei mahdotonta.

Ragin (1994b) on itse käyttänyt menetelmäänsä eläkejärjestelmien tutkimuksessa. Lähtökohtana on myös ollut kvantitatiivisen analyysin yhdistäminen QCA:han. Menetelmäksi on valittu jälleen klusterianalyysi. Se toimii aineiston jäsentäjänä, jonka avulla intervallisteikolliset muuttujat ovat muutettavissa Boolean algebraan soveltuviksi dikotomioiksi. Kankaan analyysistä poiketen Ragin (1994b: 337) käyttää myös *don't care* -luokkia totuustaulussa, jota käytetään lopulliseen analyysiin. Luokat viittaavat niihin tapauksiin, jotka eivät sovi Esping-Andersenilta (1990) lainattuun hyvinvointivaltioiden ”regiimiluokituksen”. Esping-Andersenin mukaan hyvinvointivaltiot voidaan luokitella institutionaalisten hyvinvointijärjestelmien perusteella kolmeen pääluokkaan, jotka ovat universaali (pohjoismaainen tai sosiaalidemokraattinen), korporatiivinen (keskieurooppalainen) ja perusturvalähtöinen (angloamerikkalainen tai markkinavetoinen) regiimi. Näiden lisäksi voidaan vielä erottaa eteläamerikkalaiseen ja aasialaiseen perinteeseen nojaavia järjestelmiä. Maantieteilijän näkökulmasta esitetty jako muistuttaa hieman kulttuuripiirierottelua.

Boolean algebran luokitusongelma paljastuu Raginin omasta QCA-sovelluksesta. Esping-Andersenin ”korporatiivisen regiimin” tyypillinen esimerkkimaa on Saksa. Se kuitenkin jäi klusterianalyysin tuottaman totuustaulun ulkopuolelle. Lisäksi perusturvalähtöistä sosiaalipolitiikkaa toteuttanut Britannia ei mahtunut luotuihin klustereihin. Mielenkiintoinen nyanssi on Suomen sijoittaminen

Taulukko 3. Erotteluanalyysin keskeiset tulokset. (A) Rakennematriisi. (B) Pareittainen vertailu.
Table 3. Key results of the discriminant analysis. (A) Structure matrix. (B) Comparison of pairs.

A) Rakennematriisi A) Structure matrix	Funktio 1 Function 1	Funktio 2 Function 2
Alueen asukastiheys Population density	-0,938*	0,199
Alueella koulutettuja (%) Regional education level (%)	-0,871*	-0,009
Työttömyysaste (a) Unemployment rate (a)	-0,827*	0,110
Akateemisesti koulutettuja (%) Academic education (%)	-0,815*	0,229
Alueen väkiluku (N) Population (N)	-0,764	0,378
Maatalousyrittäjiä (%) Workforce in primary industry (%)	0,762*	0,019
Palkansaajia (%) Employed workforce (%)	-0,733*	0,014
Keskitulot Average income	-0,654*	-0,199
Yrityksen työntekijöiden työmatka (keskimäärin) (a) Average distance to work for an employer (a)	0,577*	-0,020
Työntekijän työmatka (keskimäärin) Average distance to work for an employee	-0,539*	0,018
Hyvätuloisia (%) (a) People with high income (%) (a)	-0,307*	-0,281
Kotitalouden tulot Household income	-0,098*	-0,065
Asukkaiden keski-ikä Average age	0,186	0,415*
Etäisyydet kotikunnassa Experienced distances in home community	-0,377	0,394*
Koulutus Personal education level	0,140	0,145*

Muuttujat on järjestetty korrelaation mukaiseen järjestykseen funktion sisällä.
Variables ordered by correlation within function.

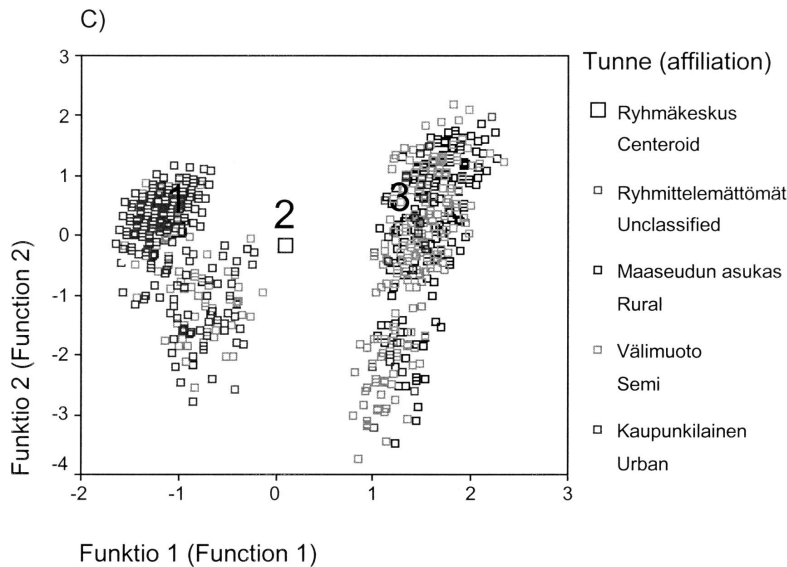
* Suurin korrelaatio muuttujan ja diskriminanttifunktioiden välillä
* Largest absolute correlation between each variable and any discriminant function

(a) Muuttujaa ei ole käytetty analysissä
(a) This variable not used in the analysis.

B) Parittainen vertailu B) Comparison of pairs				
Askel Step	Tuntee itsensä Affiliation	Kaupunki Urban	Välimuoto Semi	Maaseutu Rural
1	Kaupunkil. Urban		250,434	732,741
	Merkitsevyys (significance)		.000	.000
	Välimuoto Semi	250,434	.000	187,235
2	Maaseutu Rural	732,741	187,235	.000
	Merkitsevyys (significance)	.000	.000	.000
	Kaupunkil. Urban		127,420	385,451
3	Merkitsevyys (significance)		.000	.000
	Välimuoto Semi	127,420	.000	103,640
	Maaseutu Rural	385,451	103,640	.000
4	Merkitsevyys (significance)	.000	.000	.000
	Kaupunkil. Urban		87,839	266,338
	Merkitsevyys (significance)		.000	.000
5	Välimuoto Semi	87,839	.000	71,742
	Maaseutu Rural	266,338	71,742	.000
	Merkitsevyys (significance)	.000	.000	.000
6	Kaupunkil. Urban		66,076	203,390
	Merkitsevyys (significance)		.000	0,000
	Välimuoto Semi	66,076	.000	56,195
7	Maaseutu Rural	203,390	56,195	.000
	Merkitsevyys (significance)	.000	.000	.000

Taulukko 3. Erotteluanalyysin keskeiset tulokset. (C) Kanoniset diskriminanttifunktiot. (D) Analysoitavat muuttujat (askeltava menetelmä).

Table 3. Key results of the discriminant analysis. (C) Canonical discriminant functions. (D) Variables in the analysis (stepwise method).



D) Analysoitavat muuttujat (askeltava menetelmä)				
D) Variables in the analysis (stepwise method)				
Askel Step		Toleranssi Tolerance	F poistoraja F to Remove	Wilksin Lambda Wilks' Lambda
1	Alueen asukastiheys Population density	1.000	372,207	
2	Alueen asukastiheys Population density	0,746	159,353	0,753
	Keskitulot Average income	0,746	12,415	0,598
3	Alueen asukastiheys Population density	0,727	168,115	0,751
	Keskitulot Average income	0,736	9,826	0,586
	Kotitalouden tulot Household income	0,972	8,518	0,585
4	Alueen asukastiheys Population density	0,456	127,232	0,702
	Keskitulot Average income	0,158	11,907	0,588
	Kotitalouden tulot Household income	0,972	8,496	0,579
	Palkansaajia (%) Employed workforce (%)	0,113	5,065	0,576

korporatiivisen järjestelmän piiriin pois sosiaali-demokraattisesta järjestelmästä. Makrotasoisessa (valtiotason) yhteiskuntajärjestelmien tutkimuksessa analyysin karkeus vielä menettelee, mutta nopeasti käy ilmi, että luokitus ja samalla QCA-menetelmän käytön järjestyminen voidaan kyseenalaistaa. Esimerkiksi Kankaan (1994: 358) klusterianalyysi vuoden 1985 OECD-maiden terveysvakuutusjärjestelmästä sijoittaa Suomen samaan

ryhmään muiden Pohjoismaiden kanssa. Ero Reginin ja Kankaan klusterianalyysin tuloksissa selittyy varmasti sillä, että niissä käytettiin eri aineistoja, mutta havainto osoittaa aineistolähtöisyyden niin klusterianalyysillä kuin QCA:llakin tuotetun tiedon tulkinnassa.

QCA:n suurin ongelma liittyy binaarisen logiikan käyttöön yhteiskunnallisten ilmiöiden selittämisessä. Kati Rantala ja Eeva Hellström (2001)

ovat pyrkineet laajentamaan QCA:n mahdollisuuksia esittämällä *hermeneuttisen tiedon käsitteilyn* osana QCA:ta. Heidän mukaansa (2001: 87–90) QCA muuttuu prosessiksi, joka parhaimmillaan muistuttaa hermeneuttista kehää. Kehäajattelun mukaisesti Hellström (2001) on väitöskirjassaan jakanut QCA:n kahteen osaan eli ymmärtävään ja empiiristypologiseen (mekaaniseen). Mekaaninen osa toimii aineiston muovaajana, jonka tulosten avulla luodaan hienosyisempi (hermeneuttinen) analyysitaulu. Aineiston kaksinkertainen käsittely ei kuitenkaan poista ensisijaista ongelmaa, joka binaarilogiikkaan liittyy. Tämä ongelma on yhteiskunnallisten ilmiöiden monimuotoisuus, ei kaksiluokkaisuus. Tapaus tutkimuskohtaisessa analyysissä QCA vaikuttaa toimivan paremmin kuin yleisluonteisissa tutkimuksissa. Tässä on ristiriita, koska QCA:lla pyritään induktioon.

Johtopäätökset

Käsitlemiäni aineistojen antamat tulokset osoittavat QCA-yhtälöiden tulkinnan pysyvän samankaltaisena riippumatta sumeiden luokkien määrästä. Oleellista on saada käsitys niistä tekijöistä, jotka ilmenevät tutkimusongelman kanssa samanaikaisesti. Esitetyt yhtälöt sisältävät tyypillisiä kaupunki-alueeseen liittyviä tekijöitä ja saadut tulokset osoittavat, että QCA toimii parhaiten aineiston jäsentäjänä. Kausaalisia päätelmiä on tietysti mahdollista tehdä, mutta niiden järjestyminen on kyseenalaista. Tätä osoittavat erityisesti *don't care* -luokkien kasvattamisen aiheuttamat muutokset malleissa.

Kausaalisen tutkimuslähtökohdan näkökulmasta QCA toimii teoriallasolla hyvin. Analyttinen datamatriisin redusointi tapahtuu tehokkaasti ja useimmiten tulokset ovat selkeästi tulkittavissa myös intuitiivisesti. Empiirisen aineiston analyysissä ongelmat kuitenkin nousevat esiin. Erilaisin aineistomuotoiluina ja uusien *tukevien muuttujaluokkien* luominen auttaa esimerkiksi riittävän vaihtelun aikaansaamiseksi, mutta menetelmän perustavanlaatuisen lähtökohdan ei muutu. Kysymys ”Onko tekijä A seurauksen B syy vai ei?” on vaikeasti perusteltavissa joko–tai-dikotomian keinoin.

QCA tuottaa pääasiallisesti samankaltaisia tuloksia kuin tilastolliset menetelmät. Aineiston käsittely suurimmalla yksinkertaistuksen asteella ei tuota tulkinnallisesti eroavia päätelmiä verrattuna tilanteeseen, jossa joko–tai-luokkien määrää on nostettu. Saatujen tulosten mukaan on perusteltua arvioida, että vertailevan menetelmän ja tilastollisen analyysin yhteistuloksena saadaan kattavampi kokonaiskuva aineiston käyttäytymisestä kuin käytännöllä ainoastaan yhtä menetelmää. Lisäongelmana

QCA:n käytössä on lopulliseen analyysiin valittavien tekijöiden rajaaminen. Tilastollisen menetelmän ja QCA:n yhteiskäyttö vaikuttaa toimivalta ratkaisulta aineiston oleellisten tekijöiden etsinnässä.

Soveltavassa tutkimuksessa QCA:n ongelma on voimakas aineistoriippuvuus. Aineistoa on tietysti mahdollista muovata erilaisten muunnosten avulla (ks. esim. Ragin 1987; Hellström 2001: 74–75). Vaarana tällöin on kuitenkin luotettavuuden kärsiminen. Raginin QCA-menetelmä tuo mieleen faktorianalyysin. Saatujen faktoreiden tulkintaongelmista johtuen menetelmää voidaan käyttää ensisijaisesti aineiston jäsentäjänä eikä kausaalisen mallinnuksen välineenä. Vaikka QCA:ta voidaan käyttää myös muussa kuin kausaalissa kysymyksenasettelussa, sen kaksiluokkainen logiikka ei voi johtaa täydellisiin päätelmiin yhteiskunnallisista ilmiöistä. Tämä liittyy aineistolähtöisyyteen. Kvalitatiivisten tekijöiden muuttaminen dikotomiseen muotoon johtaa aina tiedon menetykseen. Häviääkö tietoa liikaa johtopäätelmien kannalta, on oma kysymyksensä.

Boolean minimointiin pohjautuvaa vertailevaa tutkimusta arvostellaan useasti sen yksinkertaisuudesta ja joustamattomasta luonteesta. Käytetyt muuttujat ovat irrallisia laajemmista asiayhteyksistään, ja niiden yksittäinen käyttäminen kausaalianalyysiin on arveluttavaa. Kvantitatiivisen analyysin tukena menetelmä toimii kuitenkin hyvin. Saadut tulokset ovat karkeita yksinkertaistuksia, mutta käytännössä kaikki systemaattisia yhteyksiä etsivät menetelmät kohtaavat epistemologiaan liittyvän ongelman yleisen ja yksittäisen suhteesta. Yleistettävä tieto edustaa makronäkökulmaa. Nämä seikat huomioon ottaen kvantitatiivinen vertaileva tutkimus tarjoaa yhden tavan olemassa olevan maailman ymmärtämiseen.

Intuitiivisella tasolla ymmärtävän tulkinnan suorittaa kuitenkin tutkija omien näkökulmiensa kautta. Boolean algebran tapauksessa analytytyyden harha voi usein olla liian voimakas, mikä johtaa aineiston liialliseen muovaamiseen toivottavien tulosten löytämiseksi. Lisäksi analysoitavat aineistot ovat usein sellaisia, että niitä on mahdollista käsitellä joko sisältöanalyysin tai kvantitatiivisten menetelmien avulla. Käyttämäni tilastollinen menetelmä tuotti vähintään yhtä perustellun kuvan ilmiöstä. Kausaalisiin päätelmiin kumpikaan menetelmä ei esimerkitapauksessa anna aihetta. Aluemaantieteellisesti katsoen QCA tarjoaa kuitenkin mielenkiintoisen lähestymistavan tavoiteltaessa ymmärrystä spatiaalisista prosesseista. Esimerkiksi paikkatietojärjestelmällä tuotettavan tiedon epistemologiset lähtökohdat maisemayksikköjen rajaamisessa ovat lähtökohdil-

taan varsin lähellä kategorisoivia menetelmiä, joihin myös QCA kuuluu (esim. Inkinen 1999). Maantieteellisten tietojärjestelmien yhdistäminen QCA:lla saataviin tuloksiin voisikin olla yksi potentiaalinen jatkosovelluskohde monitieteisen tiedon tuottamisessa.

KIRJALLISUUS

- Alasuutari, P. (1994). *Laadullinen tutkimus*. 317 s. Vastapaino, Tampere.
- Alkula, T., S. Pöntinen & P. Ylöstalo (1994). *Sosiaalitutkimuksen kvantitatiiviset menetelmät*. 318 s. WSOY, Helsinki.
- Drass, K. A. (1992). *QCA 3.0 qualitative comparative analysis*. 47 s. Centre for Urban Affairs and Policy Research, Northwestern University, Evanston.
- Esping-Andersen, G. (1990). *The three worlds of welfare capitalism*. 248 s. Princeton University Press, Princeton.
- Hellström, E. (2001). Conflict cultures – Qualitative Comparative Analysis of environmental conflicts in forestry. *Silva Fennica Monographs* 2: 2001. 109 s. Helsinki.
- Inkinen, T. (1999). Näkökulmia GIS:n ja digitaalisen telekommunikaation sovelluksiin ja teoriaan. *Terra* 111: 4, 230–232.
- Inkinen, T. (2001). Tieto, kone, alue ja ihminen. Kulttuurimaantieteellinen näkökulma tietoverkkojen yhteiskuntaan. *Finnish Information Studies* 18. 266 s. Turku.
- Kangas, O. (1991). The politics of social rights. Studies on the dimensions of sickness insurance in OECD countries. *Swedish Institute for Social Research* 19. 182 s. Edsbruck.
- Kangas, O. (1994). The politics of social security: on regressions, qualitative comparisons, and cluster analysis. *Teoksessa* Janoski, T. & A. Hicks (toim.): *The comparative political economy of the welfare state*, 346–364. Cambridge University Press, Cambridge.
- Mäkinen, Y. (1978). *Tilastotiedettä biologeille. Tilastotieteen ja tietojenkäsittelyn alkeet*. 4. p. 306 s. Synapsi, Turku.
- Ragin, C. (1987). *The comparative method. Moving beyond qualitative and quantitative strategies*. 185 s. University of California Press, Berkeley.
- Ragin, C. (1994a). *Constructing social research. The unity and diversity of method*. 194 s. Pine Forge Press, London.
- Ragin, C. (1994b). A qualitative comparative analysis of pension systems. *Teoksessa* Janoski, T. & A. Hicks (toim.): *The comparative political economy of the welfare state*, 320–345. Cambridge University Press, Cambridge.
- Ragin, C. (2000). *Fuzzy-set social science*. 352 s. University of Chicago Press, Chicago.
- Rantala, K. (2001). ”Ite pitää keksii se juttu”: Tutkimus kuvataidekasvatuksen ja kasvatettavan kohtaamisesta. *Helsingin yliopiston sosiologian laitoksen tutkimuksia* 239. 226 s. Helsinki.
- Rantala, K. & E. Hellström (2001). Qualitative Comparative Analysis and a hermeneutic approach to interview data. *International Journal of Social Research Methodology. Theory and Practice* 4: 2, 87–100.
- SVT 1999 = Suomen virallinen tilasto, *Kuntafakta 1998 CD-ROM* (1999). Tilastokeskus, Helsinki.
- Toivonen, T. (1999). *Empiirinen sosiaalitutkimus. Filosofia ja metodologia*. 447 s. WSOY, Porvoo.
- Thompson, J. (1994; toim.). *Matematiikan käsikirja*. 436 s. Tammi, Helsinki.