

Mielipidetiedustelun pätevyden arviointi kvantitatiivisin ja kvalitatiivisin menetelmin

ERKKI PAHKINEN

ABSTRACT

Qualitative and Quantitative Evaluation of a Validity Problem in Research on Public Opinion Polls

Data from political opinion polls are often presented as cross-classified frequency tables on which various statistical tests of homogeneity or independence are performed. To ensure then scientific validity of inference, these frequency tables should fulfill certain internal and external conditions. In this paper, two such external conditions are considered by qualitative and quantitative methods for cases where the data has been collected by quota (nonprobability) sampling and the measurement instrument has been a structured questionnaire. A case study of voting behaviour of Finnish youth is presented.

1. Tutkimusongelma

Mielipiteitä mittaavat tutkimukset ovat yleisiä käyttäytymis- ja yhteiskuntatieteissä. Niissä tutkimusasetelmana on usein otantatutkimus, jossa mielipiteitä mitataan asenneväittämällä. Yleensä vastaukset kerätään pitkälle jäsennehtynä lomakehaastatteluna tai joissain tapauksissa vapaamuotoisempana teemahaastatteluna. Näin kerätyn empiirisen aineiston analyysin tulisi täyttää tietyt tieteelliset pätevyysvaatimukset, jotta saadut tulokset olisivat luotettavia. Erityisesti kvalitatiivisen analyysin, kuten teemahaastattelun, tuloksia on kyseenalaistettu sekä itse metodin että johtopäätösten osalta. Sama on voimassa kvantitatiivisessä analyysissä, joskin sen sisäinen pätevyys, kuten valmisohjelmiin perustuva laskentarutiini on yleensä kunnossa. Sitä vastoin ulkoiset pätevyysvaatimukset, etenkin lähtöoletusten paikkansapitevyys, on syytä usein kyseenalaistaa.

Tässä tutkimuksessa sovelletaan Steinar Kvalen (1989) perustelemaa kyseenalaistamisen tekniikkaa (engl. *questioning*) arvioitaessa kvalitatiivisen analyysin tieteellistä pätevyyttä (engl. *validity*). Pätevyys tarkoittaa tässä lähinnä sitä, että läh-

töoletusten paikkansapitevyys on varmistettu ja niistä johdetut päätelmät tehty oikealla tavalla. Kvalen mielestä tutkimusprosessin kaikkia etene-misvaiheita on kyseenalaistettava tutkimustulosten tieteellisen pätevyuden varmistamiseksi.

Valittuna esimerkkinä tarkastellaan kahden muuttujan ristiintaulukon analysointia, kun aineiston keruussa on todennäköisyysotannan sijasta käytetty harkintaotantaa (kiintiöpoiminta) ja mittarina on ollut asenneväittämään annettu monivaltavastaus lomakehaastattelussa. Tässä kyseenalaistetaan ristiintaulukon kvantitatiivista analyysiä ulkoisten sovellusehtojen osalta. Niitä tarkastellaan poimintamenetelmän osalta teoreettisesti ja mittausongelman osalta empiirisen yksittäistapa-uksen pohjalta. Jälkimmäisessä kyseenalaistetaan lomakehaastattelussa toteutunutta haastattelijan ja haastatellun välistä pitkälle jäsennehtyä vuoropuhelua, jonka tulkintaa syvennetään teemahaastattelulla (vrt. Jääsaari ja Martikainen 1991).

Kvalitatiivisessä ja kvantitatiivisessä mielipidetiedustelujen analyysissä usein käytetty tekniikka on rakentaa mielipidemuuttujan kanssa mielekkäästi koostettuja ristiintaulukoita, joissa toisena muuttujana on mielipidehajontaa selittävä tausta-

muuttuja. Tekniseltä rakenteeltaan ristiintaulukko näyttäisi taulukon 1 mukaiselta. Siinä mielipidemuuttuja MIEPIDE on ajateltu jakautuvan neljään luokkaan ja sen kanssa ristiintaulukoitu taustamuuttuja sukupuoli SUKUP kahteen luokkaan. Ruuduissa olevat luvut ovat esiintymislukuja n_{ij} .

Taulukko 1. Muuttujien MIEPIDE ja SUKUP ristiintaulukko.

MIEPIDE	SUKUP		M+N
	M	N	
A-vaihtoehto vastaus	n_{11}	n_{12}	$n_{1.}$
B-vaihtoehto vastaus	n_{21}	n_{22}	$n_{2.}$
C-vaihtoehto vastaus	n_{31}	n_{32}	$n_{3.}$
D-vaihtoehto vastaus	n_{41}	n_{42}	$n_{4.}$
A+B+C+D	$n_{.1}$	$n_{.2}$	n

Tällaisen ristiintaulukon tavallisimpia analyysejä ovat erilaiset homogeenisuushypoteesit, kuten ovatko naisten ja miesten mielipidejakaumat keskenään samanlaiset tai, sama riippuvuushypoteesina, riippuko mielipideilmaisuus vastausvaihtoehto vastaajan sukupuolesta. Yleensä päädytään Pearsonin X^2 -testiin. Testin käyttö edellyttää ainakin kahden ehdon voimassaoloa:

- Ehto 1: Poimintamenettelyä on yksinkertainen satunnaisotanta, eli poiminnassa ei ole henkilöiden välistä riippuvuutta.
- Ehto 2: Molempien muuttujien MIEPIDE ja SUKUP luokat ovat muuttujien sisällä toisensa poissulkevat.

Ehdon 1 ollessa voimassa testisuure noudattaa χ^2 -jakaumaa vapausastein $(k-1)(l-1)$, missä k on muuttujan MIEPIDE vastausvaihtoehtojen lukumäärä ja l on taustamuuttujan SUKUP luokkien lukumäärä. Ellei ehto 1 ole voimassa, tuloksena on tilastollisia testejä, joiden hylkäysalueiden koot eli havaitut tilastolliset merkitsevyystasot (ns. p -arvot) poikkeavat olennaisesti nimellisistä ja voivat siten aiheuttaa virheellisiä johtopäätöksiä. Esimerkiksi homogeenisuushypoteesi saatetaan hylätä, vaikka se todellisuudessa pitäisikin jättää voimaan. Jos ehto 2 ei ole voimassa, on seurauksena vakavia tulkinnallisia pulmia.

Seuraavissa aliluvuissa tarkastellaan, mitä mahdollisuuksia on arvioida ehtojen 1 ja 2 voimassaoloa, kun mielipidetiedustelun havaintoaineisto on

kerätty kiintiöpoiminnalla ja mittarina on lomakehaastattelun monivalintakysymys.

2. Kiintiöpoiminta

Kiintiöpoiminnalla tarkoitetaan sellaista otantamenetelmää, jossa ennalta kiinnitetään kutakin perusjoukon osajoukkoa edustavien otantayksikköiden määrä eli kiintiö. Tämä lukumäärä saadaan lopulliseen otokseen esimerkiksi siten, että kustakin osajoukosta arvotaan yksinkertaisella satunnaisotannalla lähtöosoitteiksi tarvittava henkilömäärä. Jos haastattelija ei tapaa lähtöosoitteeksi valittua henkilöä, hän etenee samassa perusjoukon osajoukossa harkinnanvaraisesti, kunnes haluttu lopullinen otosyksikköjen määrä on saavutettu.

Mielipidetiedustelussa voidaan lähteä väestön demografisesta luokittelusta esimerkiksi ikäluokan, asuinpaikan, sukupuolen ja siviilisäädyn suhteen. Luokittelujen kautta perusjoukko tulee jaettava K:hon osajoukkoon, joista jokaiseen kiintiöidään tietty otoksen osajoukko eli kiintiö. Ongelmat kiintiöpoiminnassa tulevat jo poimintavaiheessa. Esimerkiksi poliittisia mielipiteitä tiedusteltaessa haastattelija lähtee liikkeelle lähtöosoitteesta. Jos hän tapaa siinä otokseen arvotun haastateltavan, hän suorittaa haastattelun. Jos hän ei tapaa, hän etenee samassa osajoukossa seuraavan tavattavan henkilön luokse, josta suorittaa lähtöosoitehenkilön tilalle sijaishaastattelun. Otos kuitenkin vääristyy koskemaan henkilöitä, jotka ovat tiettyyn kellonaikaan eli tavanomaiseen haastatellaan kotona. Näin ollen jos kotona tavoitettavien poliittiset mielipiteet ovat erilaiset kuin liikkuvien ihmisten (ei-tavoitettavat), kyseenalaistuvat kiintiömenetelmällä kerättyjen aineistojen tilastolliset analyysit.

Kiintiöpoiminnan tapauksissa analysointihankaluus tulee siitä, että toteutunut otos ei ole todennäköisyysotos, jossa kaikilla perusjoukon henkilöillä olisi ollut etukäteen tunnettu positiivinen sisältymistodennäköisyys otokseen. Esimerkiksi mielipidetiedustelun tapauksessa positiivinen sisältymistodennäköisyys on vain kotona tavattavilla henkilöillä. Jos heillä on samat mielipiteet kuin muilla henkilöillä, ei tietenkään synny yleistämishetettä analysoitavaksi saadun otoksen perusteella. Tätä virhettä on vaikea arvioida tavanomaisella edustavuustarkastelulla koko perusjoukon ja sen osajoukkojen suhteen, koska kiintiöpoiminta perusajatukseltaan takaa, että otos on kiintiöinnissä käytetyn taustainformaation osalta perusjoukkonsa tarkka pienoiskuva.

Kiintiöpoiminnan tapauksessa otanta-asetelmaan perustuvaa keskivirhettä ei voida määritellä. Keskivirhe on tarpeellinen rakennettaessa tilastollisia testejä ja perusjoukon parametrien luottamusvälejä (95 %, 99 %, 99.9 %), kun halutaan tehdä yleistyksiä taustalla olevaan perusjoukkoon eli tässä yleistää otoksessa havaittu jakauma perusjoukon jakaumaksi.

Kiintiöpoiminnan ongelmat heijastuvat tilastolliseen analyysiin siinä, että kukin haastattelija pyrkii joko tietoisesti tai tiedostamatta homogenisoimaan omaa vastaajajoukkoaan. Tästä pulmasta ovat hyvin tietoisia haastattelututkimuksia toimeksiantoina tekevät laitokset. Kuten Hirvonen (Suomen Gallup Oy) asian ilmaisee: »Käytännön ongelmia syntyy myös tutkimusten toteutusvaiheessa. Erityisesti haastattelijoiden vaikutus otoksen rakenteeseen on merkittävä» (vrt. Hirvonen 1989).

Esimerkiksi yhden haastattelijan ristiintaulukko saattaisi näyttää taulukko 2:lta.

Taulukko 2. Haastattelijan (c) keräämä tutkimusaineisto ristiintaulukkona.

MIEPIDE	SUKUP		M+N
	M	N	
A-vaihtoehto vastaus	$n_{11(c)}$	—	$n_{1(c)}$
B-vaihtoehto vastaus	—	—	—
C-vaihtoehto vastaus	—	—	—
D-vaihtoehto vastaus	—	$n_{42(c)}$	$n_{4(c)}$
A+B+C+D	$n_{1(c)}$	$n_{2(c)}$	$n_{(c)}$

Tässä tapauksessa haastattelija (c) on saanut kerättyä havaintoaineiston, joka taustamuuttujan SUKUP osalta edustaa annettua kiintiöintitavoitetta ja siten myös perusjoukkoa oikealla tavalla. Mitatun mielipidemuuttujan MIEPIDE osalta hän on tavoittanut vain miehiä, jotka ovat valinneet A-vaihtoehdon, ja naisia, jotka ovat valinneet vain D-vaihtoehdon. Jos oletetaan, että todellisuudessa mielipiteiden pitäisi jakautua tasan eri vaihtoehtojen kesken, niin haastattelija valinnoillaan vähentää havaintojen luonnollista vaihtelua. Voidaan tulkita, että haastattelija muodostaa haastateltavistaan rypään, joka mielipidekysymysten osalta on homogeeninen. Tämä ryvä voi kuitenkin poiketa toisen haastattelijan valitsemasta haastateltujen rypästä, joka kuitenkin on jälleen sisäisesti homogeeninen.

Haastattelijan vaikutus kiintiöpoiminnan tapauksessa kohdistuu siis ensijaisesti haastateltavien valintaan. Haastattelijan vaikutusta suoraan

haastatteluvastauksiin on paljon tutkittu (vrt. Cochran, 1977) ja päädytty homogenisuutta mittaavan sisäkorrelaation käsitteeseen, joka on ryväotannon tärkeä parametri. Koska haastateltavien harkintaperusteinen valinta homogenisoi vastaajajoukkoa, on myös tässä syytä tulkita toteutunut havaintoaineisto ryväotannaksi. Tällöin aiemmin esitetty ristiintaulukon analysoinnin ulkoinen ehto 1 ei ole voimassa. Tämä merkitsee sitä, että tilastolliset tunnusluvut on laskettava aivan kuin kysymyksessä olisi ositettu yksiasteinen ryväotanta. Yksinkertaisessa laskentatilanteessa esimerkiksi keskiarvon keskivirhe on

$$s(\bar{y}_{ryv}) = s(\bar{y}_{yso}) \sqrt{1 + (\bar{m} - 1)\omega},$$

missä $s(\bar{y}_{ryv})$ on ryväotannon keskivirhe, $s(\bar{y}_{yso})$ on yksinkertaisen satunnaisotannon keskivirhe, ω on haastattelijan valinnasta aiheutuvaa homogeenisuutta mittaava sisäkorrelaatio ja \bar{m} keskimääräinen haastateltavien lukumäärä haastattelijaa kohden.

Yleensä sisäkorrelaatio on positiivinen, mikä merkitsee sitä, että yksinkertaisen satunnaisotannon keskivirheen käyttö aliarvioi todellisen hajonnan. Tästä edelleen seuraa, että suoraan aineistosta laskettu Pearsonin X^2 -testisuure antaa väärän tuloksen. Tilanne on kuitenkin korjattavissa huomioimalla ryvästyminen laskelmissa. Toimenpiteeseen käytettäviä tekniikoita on kehitelty otantateorian alueella 80-luvulla (kts. Rao ja Thomas 1989). Ne hyödyntävät pääsääntöisesti yksinkertaisen satunnaisotoksen ja ryväotannon keskivirheiden neliöityä suhdetta eli *deff*-lukua, joka on

$$deff(\bar{y}) = s^2(\bar{y}_{ryv}) / s^2(\bar{y}_{yso}).$$

Yleensä ryväotannassa *deff* > 1, mikä heijastaa positiivista sisäkorrelaatiota. Näiden laskeminen alkuperäisestä havaintoaineistosta on toteutettavissa merkitsemällä haastattelijan järjestysluku eli tunnus (c) uudeksi muuttujaksi kullekin haastatellulle. Tämän tyyppistä asetelmatunnusta eivät suuret valmishjelmat, kuten SPSS-X, BMDP ja SAS, tunne, joten analyysivaiheessa on käytettävä erikoistuneita ohjelmia kuten SUDAAN tai SUPERCARP (ohjelmien käytöstä kts. Skinner ym. 1989). Tuloksena on taulukko, jossa on mukana kaikki haastatellut henkilöt (esiintymisluvut) ja lisäksi haastattelijan valinnasta aiheutuneet *deff*-luvut.

Edellä *deff*-luvuin täydennetyistä ristiintaulukosta on mahdollista laskea korjattu Pearson X^2 -testisuure. Eräs yksinkertaisimmista korjauksista on

Taulukko 3. Muuttujien MIEPIDE ja SUKUP ristiintaulukko. Haastattelijavalinnasta aiheutuneen homogeenisuuden vaikutusta kuvaavat deff-luvut mukana.

MIEPIDE	SUKUP		M+N
	M	N	
A-vaihtoehto vastaus	n_{11}, d_{11}	n_{12}, d_{12}	$n_{1.}, d_{1.}$
B-vaihtoehto vastaus	n_{21}, d_{21}	n_{22}, d_{22}	$n_{2.}, d_{2.}$
C-vaihtoehto vastaus	n_{31}, d_{31}	n_{32}, d_{32}	$n_{3.}, d_{3.}$
D-vaihtoehto vastaus	n_{41}, d_{41}	n_{42}, d_{42}	$n_{4.}, d_{4.}$
A+B+C+D	$n_{.1}, d_{.1}$	$n_{.2}, d_{.2}$	n

ns. Fellegi-korjaus, jossa ristiintaulukon ruuduissa olevista deff-luvuista (d_{ij}) lasketaan keskiarvo, jolla jaetaan tavanomaisesti laskettu X^2 -testisuure. Siten jos haastattelijan valinta homogeenisoi hänen vastaajajoukkoa, se näkyy ykköstä suurempina deff-lukuina, joiden keskiarvokin on siten ykköstä suurempi. Kun tällä jaetaan alkupe-
räinen X^2 -testisuure, tuloksena voi olla päättely-
tulos, jossa hylättäväksi katsottu hypoteesi jääkin voimaan, kun päättely perustetaan korjatulle testisuurelle. Tästä on olemassa lukuisia kansainvälisiä esimerkkejä. Suomen oloissa vastaavia korjaustilanteita on raportoitu kansanterveyden ja koulusaavutustutkimusten alueelta (Pahkinen ja Lehtonen 1989).

Kiintiöpoiminnalla tuotettua havaintoaineistoa ei voida täysin korjata haastattelijan harkinnasta aiheutuneesta virheestä. Esimerkiksi tilanteessa, jossa kaikki haastattelijat valitsevat oman kiintiön-sä puiteissa samalla lailla kotoa helposti tavoitettavia, tuloksiin jää kaksi virhettä; haastattelijain valinnasta aiheutuva homogeenisuus, jonka korjaamista käsiteltiin edellä, ja systemaattinen virhe, koska toteutuneet haastattelut edustavat tietyn-tyyppistä vastaajajoukkoa. Jälkimmäistä virhettä on vaikeampi korjata. Se edustanee eräänlaista yksikkökadon aiheuttamaa vääristymää (kts. katorvirheen korjaamisesta Little ja Rubin 1988).

3. Teemahaastattelu

Teemahaastattelu on verraten vapaamuotoinen mielipiteiden mittaustilaisuus, johon osallistuvat haastateltava ja haastattelija. Lähtökohtana on haastattelijan ennalta työstämä väljä tai rakenteellisesti rajattu aihe. Olennainen ero lomakehaastatteluun on vastausten ja mielipiteiden vapaamuotoinen ilmaus, joka ei ole sidottu ennalta kiinnitettyihin vastausvaihtoehtoihin. Teemahaastattelun

on todettu olevan käyttökelpoinen menetelmä tilanteisiin, joissa haastateltavat tiedostavat heikosti tutkittavaa aluetta eikä heillä ole siitä omakohtaista jäsentynyttä tietoa. Teeman jäsentymisaste ratkaisee, miten vapaasta keskustelutilanteesta lähdetään liikkeelle. Kaikkein väljin ratkaisu on käyttää täysin avointa teemahaastattelua, jossa jäsentymisaste on alhainen. Teemahaastattelun tarkoitus on antaa tutkimuksen kohteena olevasta asiasta joko syvällisempää tietoa tai sitten tietoa täysin toisesta näkökulmasta kuin mitä saataisiin lomakehaastattelulla. Näin ollen teemahaastattelua voisi käyttää lomakehaastattelun mittareiden sisällöllisen pätevyyden tarkistamiseen. Samoin se voisi antaa viitteitä, miten lomakehaastattelun vastausvaihtoehtoja tulisi valita.

Teemahaastattelua kuvataan keskustelutilanteeksi, jossa haastateltava ja haastattelija käyvät vuoropuhelua. Vuoropuheluun sisältyy paitsi puhuttu teksti myös se tapa, millä se on esitetty. Sitten keskustelun tulkinnaassa oma paino tulee tavalle, miten äänenpainot ovat jakautuneet, miten haastateltava ja haastattelija vuorottavat puhettaan, pitävät taukoja tai puhuvat päällekkäin. Tästä syystä teemahaastattelu yleensä nauhoitetaan, jotta sen analysointi olisi tehtävissä keskusteluna, litteroimalla aluksi äänite syvemmän käsityksen saamiseksi puhuttuihin teksteihin. Tällaista lähestymistapaa on korostettu monilla tahoilla, kuten Atkinson ja Heritage (1984) painottavat. Toisaalta monet keskustelut ovat jäykistyneet kirjoitetuiksi teksteiksi, jolloin niiden ainoa analysointimahdollisuus on yksistään tekstianalyysi.

Seuraavassa on esimerkki avoimen haastattelun tuloksena kirjoitetusta tekstistä (kts. Jääsaari 1989). Kysymyksessä on pääkaupunkiseudun nuorten aikuisten poliittisia valintoja koskenut tutkimus. Haastateltava on ollut 26-vuotias nainen.

Ote avoimesta haastattelusta.

»Kai se on ihan silleen, ku sä katot jotain tommosii, tiätsä, 18—20, niil on jotenki, ku on jutellu jengin kanssa, ihan siis omien frendien ja tollai kans, tulee vaan semmonen fiilis, et niin kauan ku ei oo mitään, niin ei oikein kiinnostu paljon äänestäkään, ku ei... elikä tällanen fatalistinen, ku ei kuiteskaan täs tapahdu mitään. Eli, et ku ei uskota yksinkertaisesti tohon jengiin, joka tuol on ehdokkaana, niinku pätäkääkään. Ihan monta kertaa helvetin syystä.»

»Et ei niinku jengii varmaan kiinnosta, tai et showvaa, mitä mä siellä teen, joku, joka tietää asioista paremmin, voi äänestää niinku,

hehheh, tavallaan mun puolesta. Mä oon yritänyt kyl sillai pitää, et mä vaikka väkisin meen, ihan ku mä en keksi mitään, et jotain. Periaatteesta.»

Avoimen haastattelun tieteellinen hyödyntäminen edellyttää sen analysointia. Tähän on kehitelty erilaisia tekstianalysejä, joiden kehitystyötä on tehty varsinkin kielentutkimuksen parissa. Kuitenkin laajempi käsitys keskustelun analyysistä vaatii myös käyttäytymistieteiden, kuten psykologian, hyödyntämistä. Näiltä alueilta muita mielenkiintoisempi on van Dijkn (1980) diskurssianalyysi, jossa lähtökohtana on perustekstin jako virkeisiin eli mikropropositioihin. Näistä edetään yleisimpiin tekstin teemoihin käyttäen johtamiseen makrosääntöjä, jotka ovat seuraavat: poisto-, valinta-, yleistys-, konstruktio-, siirto- ja arviointisääntöt. Yleensä riittää kolmen ensimmäisen makrosääntönnön käyttö, jolloin alkuperäisteksti tiivistyy muutamaaan makropropositioon. Lisäksi tekstillä voi olla oma superstruktuurinsa, joka van Dijkn mukaan ei ole yksistään johdettavissa makropropositioista vaan mukaan on otettava aiemmin mainittua tekstin litterointia ja huomattava myös keskustelun kognitiivisten, sosiaalisten ja kulttuuristen ulottuvuuksien käsittely.

Tekstianalyysin käytännön toteutusta varten on kehitelty joukko tietokoneohjelmia, joiden käyttöalueista ja saatavuudesta on yleinen selvitys Fieldingin ja Leen (1991) toimittamassa kirjassa. Nämä ohjelmat helpottavat tutkijan työtä varsinkin suuren haastattelumassan analysoinnissa.

Teemahaastattelu on tuotu tähän lomakehaastattelun täydennykseksi eikä sen kilpailijaksi. Jatkossa on tarkoitus katsoa yhden lomakekysymyksen ja sen teemaan liittyvän yhden avoimen haastattelun keskinäistä suhdetta. Tavoitteena on selvittää, miten hyvin yksinkertaisessa ristiintaulukon tapauksessa ulkoinen analysointiehto 2 on voimassa. Tämä edellyttää lomakekysymyksen vastausvaihtoehtojen sisällöllistä arviointia samaan aiheeseen liittyvän avoimen haastattelun antaman informaation avulla.

4. Tapausesimerkki

Nuorten poliittista valintaa ovat tutkineet Jääsaari ja Martikainen (1991). Työ liittyy yleisen valtio-opin alueeseen, jossa kohteena on nuorten aikuisten äänestämiskäyttäytyminen pääkaupunkiseudulla. Havaintoaineisto on kerätty kolmesta lähteestä:

- 1) Täydellinen vaalirekisteri 1987 eli kaikki äänioikeutetut (N=395 428)
- 2) Kiintiöpoiminnalla poimittu otos (n=705)
- 3) Teemahaastatteluun valittu osajoukko (n=30)

Tutkitaan päätelmien pätevyteen vaikuttavia seikkoja sekä kvantitatiivisesti että kvalitatiivisesti. Kvantitatiivinen tarkastelu samastuu tilastolliseksi päättelytilanteeksi ja kvalitatiivinen mitta-iden ominaisuuksien tarkasteluun. Yksilöitynä tapauksena ovat ristiintaulukossa homogeenisuus- ja riippuvuushypoteesin analysointiehdot 1 ja 2.

4.1. Täydellinen vaalirekisteri 1987 (N=395 428)

Koko perusjoukon kattava rekisteriaineisto ei sisällä aineiston keruujärjestelmästä, kuten otannasta, johtuvaa satunnaisvirhettä. Tällöin ristiintaulukoitujen muuttujien jakaumavertailut antavat suoraan prosenttilukujen erotuksina mielenkiintoisia yhteyksiä ristiintaulukoitujen muuttujien väliltä. Ei tarvita tilastollisia yhteensopivuus- ja homogeenisuustestejä. Koska virallisessa rekisteriaineistossa henkilö- ja muut taustamuuttujatiedot ovat viranomaisten kirjaamia ja perustuvat luotettaviin lähteisiin, mittareiden validiutta on vaikea kyseenalaistaa. Näin ollen tämän aineiston kanssa voidaan päätelmiä tehdä verraten luotettavasti, jos mahdolliset aineistoon sovitettavat mallit ovat titeenalohtaisesti perusteltavia.

4.2. Kiintiöpoiminnalla kerätty aineisto (n=705)

Osa havaintoaineistosta on kerätty kiintiöpoiminnalla pääkaupunkiseudun nuorten aikuisten joukosta. Kvantitatiivisessa analyysissä, kuten ristiintaulukkojen homogeenisuus- ja riippuvuustesteissä, tulisi huomioida poiminnasta aiheutuva haastattelijattavien homogeenoituminen haastattelijattain. Tätä ongelmaa käsiteltiin jo aliluvussa 2, jossa osoitettiin, että kiintiöpoiminnan tuloksissa saattaa homogenisoituminen pienentää vastausten luonnollista vaihtelua. Samaten voi tuloksiin tulla muuta systemaattista virhettä, joka johtuu haastattelun joukon valikoitumisesta. Kumpaakin virhelähdettä voidaan paikallistaa otantateoriassa olevilla korjaustekniikoilla. Niiden vaikutukset voivat näkyä taulukon 3 mukaisissa ristiintaulukoissa ykkösen ylittävänä *deff*-lukuina (*deff* > 1).

Mittauksen ongelma liittyy siihen, miten hyvin mielipidekysymysten lomakkeella tarjotut vastaus-

vaihtoehdot ovat toisensa poissulkevia. Kyselyn keräystilanteessa tämä toteutuu varmasti, jos haastattelija merkitsee vastaukseksi yhden vaihtoehdon. Tulos ei kuitenkaan kerro täsmälleen, ovatko vastausvaihtoehdot todellisuudessa toisensa poissulkevia, mikä mielipidejakaumasta tehtävien johtopäätösten kannalta olisi tärkeä tieto. Tämä koskee ristiintaulukon analysoinnissa ehtoa 2.

Mielipidemittareiden ominaisuuksia voidaan tutkia kvantitatiivisin menetelmin erilaisina reliabiliteettikertoimina. Syvemmän käsityksen jonkin mielipidekysymyksen vastausvaihtoehdoista voi saada käyttämällä lomakehaastattelun rinnalla kvalitatiivista tutkimusstrategiaa, kuten nyt esillä olevassa tapausesimerkissä on tehty. Vapaa- muotoisessa keskustelussa haastateltava voi perustella erilaisia vastausvaihtoehtoja tai jopa löytää niiden tilalle uusia.

4.3. Teemahaastattelun aineisto (n=30)

Teemahaastattelun aineisto on Jääsaaren tutkimuksesta (1989) ja käsittää 30 nuorta aikuista pääkaupunkiseudulla. Heidät on haastateltu jäsentämättömällä avoimella haastattelulla, joka on toimeenpantu vuosina 1987—1989. Haastateltavat on valittu harkinnanvaraisin kriteerein. Alkuperäisenä tavoitteena on ollut hyödyntää haastattelujen antamaa informaatioa kohdan 4.2 kiintiöpoimintaan perustuvassa poliittista valintaa koskevassa tutkimuksessa. Tästä aineistosta on poimittu yksi haastattelutapa, joka on tässä analysoitu van Dijkn diskurssianalysysillä, ja verrattu sen antamaa tulosta vastaavan aihealueen lomakekysymykseen ja sen tarjoamiin vastausvaihtoehtoihin. Ennakkokäsityksenä pidetään sitä, että mielipidekysymyksen vastausvaihtoehdot ovat toisensa poissulkevia, eli aiemmin esitetty ehto 2 ristiintaulukon analysoinnista olisi voimassa.

Teemahaastattelu tähtää yleensä rajatun ongelman empiiriseen tarkasteluun. Yleistettävyyden sijaan perusjoukkoon, josta haastatellut on poimittu, ei sinällään ole kiinnostavaa eikä edes mahdollistakaan. Teemahaastattelun aineiston analyysille voidaan asettaa tietysti omat tieteelliset pätevyysvaatimukset, joita mm. Mäkelä (1990) on luetellut. Yksi niistä on aineiston riittävyyteen ja toinen toistoon liittyvä. Riittävyys viittaa siihen, että tapausten tulokinnassa alkaa toistua sama säännönmukaisuus eivätkä lisätapaukset tuo uusia näkökulmia vaan vahvistavat jo aiemmin todettuja. Toistettavuus vuorostaan viittaa siihen, että analyysimenetelmä kirjataan niin selvästi, että toinen

tutkija toistaessaan saman analyysin päätyy samoihin tuloksiin.

Koska tätä artikkelia sävyttää metodinen lähestymistapa (Miten hyödyntää samassa tilanteessa sekä kvalitatiivista että kvantitatiivista analyysiä?), ei sisällöllinen yleistettävyyden ole pääasia. Tällöin riittävyys tulkitaan toteutuvaksi yhden tapauksen analyysillä. Toistettavuuden varmistamiseksi on avoimen haastattelun analysointimenetelmäksi valittu aiemmin mainittu van Dijkn diskurssianalyysi, jossa analysointisäännöt ovat verraten selkeät. Tarkoitus on etsiä avoimen haastattelun tekstistä, miten siinä esiintyvät yhden mielipidekysymyksen vastausvaihtoehdot, jotka on tarjottu lomakehaastattelussa.

Seuraava asetelma esittää vastaajien (n=705) äänestämiseksi antaman merkityksen. Taustana on tieto siitä, äänestikö vastaaja vai ei vuoden 1987 eduskuntavaaleissa. Vastaukset on kerätty pitkälle jäsenellyllä lomakehaastattelulla. Kysymyksenä on ollut: »Miten kansalaisten tulisi suhtautua vaaleihin ja äänestämiseen?» Itse asiassa käytetty kysymys on kansainvälistä perua, koska se on käännetty Kielin Christian-Albrechts-yliopiston poliittisen tutkimuslaitoksen kehittämästä mielipidetiedustelulomakkeesta. Samainen lomake on käytössä englantilaisessa BGE-tutkimuksessa (British General Election Study), kuten ilmenee Jääsaaren tutkimuksesta (1989).

Äänestäminen on ...	Äänesti v. 1987	Ei ään. v. 1987	Yht. %
— kansalaisvelvollisuus ja oikeus, mikä ilman muuta on käytettävä	57	17	47
— oikeus, jonka hyvällä omallatunnolla ja syyllä voi jättää myös käyttämättä	29	58	36
— tapa, jolla nykyisin ei ole paljoakaan todellista merkitystä	5	11	6
— äänestämällä tulee antaneeksi hyväksymisensä poliitikkojen toiminnalle	4	12	6
— muu tai ei osaa sanoa	5	2	5
Yhteensä % (n=705)	100	100	100

Vastauksista näkyy, miten äänestäneet painottavat eri tavalla äänestämiseksi annettavaa merkitystä kuin ei-äänestäneet henkilöt. Lisäksi äänestäneistä lähes 60 % pitää äänestämistä ehdottomana kansalaisvelvollisuutena. Osa heistä äänestää (5 % + 4 % = 9 %), vaikka äänestäminen nähdään rituaalina tai peräti poliitikkojen legitimointina.

Edellä tehtyjen päätelmien tieteellinen pätevyys riippuu siitä, miten hyvin esitetty kysymys mitataan tavoitteena ollutta ominaisuutta. Vastaajan on edellytettävä ymmärtävän kysymys oikein ja tuottaneen siihen validin vastauksen. Haastattelussa käytetty standardilomake on kirjoitettua kieltä, siis mahdollisimman lähellä kirjakieltä, ja toisaalta vastausvaihtoehdoissa on pyritty toisensa poissulkeviin luokkiin.

Seuraavaksi tarkastellaan sitä, miten vastaava ongelma-alue kartoittuu, kun havaintona on teemahaastattelu, jossa on tulkittava vastaajan puhekieltä ja annettava sille asiayhteyteen liittyvä merkitys. Näyte on edelleen aiemmin mainitusta lähteestä (Jääsaari ja Martikainen 1991, 137—138). Haastateltavana on ollut 26-vuotias nainen, joka tunnuspiirteitten osalta samastuu tutkittuun nuorten aikuisten otokseen, vaikkei olekaan sen otosyksiköitä. Haastateltavan tässä analysoitava teksti käsittelee kaksi kappaletta, jotka olivat jo esimerkiksi teemahaastattelua koskeneessa aliluvussa 3. Kirjoitettu teksti on seuraavassa analysoitu van Dijkin diskurssianalyysiin sisältyvillä makropropositioiden muodostamissäänöillä.

Ensimmäinen vaihe analyysissä on siten propositionalistia teksti ja johtaa makrosääntöjä noudattaen ensimmäisen tason (1) makropropositiot. Ne on esitetty alla (makrot oikealla):

Edellä olevan puhutun puheen sisältö koostui 19 virkkeestä, jotka makrostruktuurin ensimmäisellä tasolla tiivistyivät kolmeksi makropropositioksi. Teeman käsittelyä voisi jatkaa etsimällä tekstin superstruktuuri, joka perustuu ensimmäisen tason makropropositioille. Keskeistä näille kolmelle on, että niissä käsitellään nuorten aikuisten äänestämiskäyttymistä ja esitetään siihen liittyviä perusteluja. Superstruktuuri voisi näyttää seuraavalta kaksitasoiseksi kuvattuna:

TASO 1	TASO 2
<i>M*1 Nuorten äänestämättömyys on epäluottamus politiikoille</i>	
<i>M*2 Poliitikkaa tunteva tuttavani voisi äänestää puolestani.</i>	<i>Nuorten aikuisten ja oma poliittinen vieraantuneisuus ei välttämättä estä äänestämistä.</i>
<i>M*3 Äänestän vaikka periaatteesta.</i>	

Toinen makropropositio ei näytä sopivan ilman tarkennuksia toisen tason superstruktuuriin. Tekstistä kuitenkin ilmenee, että siihen sisältyvä »uusi äänestämismuutostohto» soveltuu henkilölle, jolla itsellään ei ole riittävää asiantuntemusta politiikasta. Äänioikeus toteutuisi tällöin paremmin äänioikeutetunkin kannalta, jos hän voisi valtuuttaa

Haastattelu (mikropropositiot)	Makrosääntö ja -propositio	
1. <i>Kai se on ihan silleen,</i>	POISTO —	
2. <i>Iku sä katot jotain tommosii, täitsä</i>	POISTO —	
3. <i>118—20, niil on jotenki,</i>	YLEIST. — M*1	<i>Nuorten äänestämättömyys on epäluottamus politiikoille.</i>
4. <i>Iku on jutellu jengin kanssa, ihan</i>	POISTO —	
<i>Isiis omien frendien ja tollai kans,</i>	POISTO —	
5. <i>Itulee vaan semmonen fiilis,</i>	POISTO —	
6. <i>let niin kauan ku ei oo mitään,</i>	YLEIST. — M*1	
7. <i>Iniin ei oikein kiinnosta paljon</i>	YLEIST. — M*1	
<i>läänestäkään ku ei... elikä</i>		
<i>Itällanen fatalistinen,</i>		
8. <i>Iku ei kuiteskaan täs tapahdu mitään.</i>	YLEIST. — M*1	
9. <i>Eli, et ku ei uskota yksinkertaisesti</i>	YLEIST. — M*1	
<i>Itohon jengiin,</i>		
10. <i>Ijoka tuol on ehdokkaana, niinku</i>	YLEIST. — M*1	
<i>lpätkääkään.</i>		
11. <i>Ihan monta kertaa helvetin syystä.</i>	POISTO —	
12. <i>Et ei niinku jengii varmaan kiinnosta,</i>	POISTO —	
13. <i>Imitä mä tai et showvaa, siellä teen</i>	KONSTR. — M*2	<i>Poliitikkaa tunteva tuttavani voisi äänestää puolestani.</i>
14. <i>Ijoku, joka tietää asioista paremmin,</i>	KONSTR. — M*2	
15. <i>Ivoi äänestää niinku, hehheh, tavallaan</i>	KONSTR. — M*2	
<i>Imun puolesta.</i>		
16. <i>Mä oon yrittäny kyl sillai pitää</i>	KONSTR. — M*3	<i>Äänestän vaikka periaatteesta</i>
17. <i>let mä vaikka väkisin meen,</i>	KONSTR. — M*3	
18. <i>Iihan ku mä en keksi mitään, et jotain.</i>	POISTO —	
19. <i>Periaatteesta.</i>	KONSTR. — M*3	

sijaisekseen politiikassa asiantuntemusta omaavan tuttavansa. Taustalla on tässäkin selvä peruste tällaiselle käyttäytymiselle.

Mielenkiintoista on nyt verrata, miten puheen makropropositiot suhtautuvat esimerkkitaulukon vastausvaihtoehtoihin. Taustaksi on muistettava, että teemahaastateltu oli 26-vuotias nainen, josta ei tiedetä, mitä vaihtoehtoa hän olisi tarjonnut vastaukseksi. Tiedossa ei myöskään ole se, äänestikö hän vuonna 1987 vai ei. Makropropositioiden tasolla ilmenee, miten hän pohdiskelee ja perustelee kolmea vastausvaihtoehtoa:

-
- Vaihtoehto 1. Äänestäminen on kansalaisvelvollisuus ja oikeus, mikä ilman muuta on käytettävä. Vastinparina on makropropositio M*3 »Äänestän vaika periaatteesta».
- Vaihtoehto 2. Äänestämällä tulee antaneeksi hyväksymisenä politiikkojen toiminnalle. Vastinparina tai oikeammin sen peilikuvana on makropropositio M*1 »Nuorten äänestämättömyys on epäluottamus poliitikoille».
- Vaihtoehto 3. Standardilomakkeella on viimeisenä vaihtoehtona »muu tai ei osaa sanoa», jonka vastinpariksi voi hyvällä syyllä laittaa makroproposition M*2 »Politiikkaa tunteva tuttavani voisi äänestää puolestani»
-

Standardilomakkeella haastateltaessa kaikki kolme vaihtoehtoa ovat mahdollisia vastausvalintoja. On syytä olettaa, että vastausvaihtoehdot eivät todellisuudessa olekaan toisensa poissulkevia, vaan ne mittaavat a) eri ominaisuuksia tai b) samaa ominaisuutta eri suunnista. Tällöin näillä kerättyjen prosenttijakaumien validius ei ole hyvä, koska yksittäinen vastaaja on pakotettu yhteen vaihtoehtoon, vaikka mielessä olisi ollut muitakin, kuten nyt käsitellyssä yksittäisessä teemahaastattelussa.

Mielipidetiedustelussa käytettyä asennemittarin mittaussominaisuutta koeteltiin edellä yhden teemahaastattelun aineistolla. Näinkin suppeassa tarkastelussa ilmeni, että mittari ei sisällä toisensa poissulkevia vaihtoehtoja, joka olisi ollut tilastollisen analyysin, kuten prosenttijakaumien tuottamisen, ehtona. Lisäksi kiinnitetty vastausvaihtoehdot eivät sisältäneet mahdollisuutta neljän määritellyn vaihtoehdon lisäksi muuten kuin yleisen »muu»-vaihtoehdon alla. Teemahaastattelun diskurssianalyysi kyseenalaisti monesta suunnasta käytetyn asennemittarin ja samalla rikasti sen tulkintaa. Oikeastaan vasta sen makropropositiot avasivat näkymiä nuorten aikuisten poliittiseen vie-raantumiseen. (Yleinen piirre, jota on länsimais-

sa voimaperäisesti tutkittu [Miller 1991, Jääsaari ja Martikainen 1991]).

5. Johtopäätöksiä

Mielipidetiedustelun analyysituloksia kyseenalaistettiin ristiintaulukon homogeenisuus- ja riippuvuustestauksen tilanteessa. Monesti käytetyn Pearson X^2 -testin tulosten luotettavuudelle asetettiin kaksi ehtoa, jotka koskivat a) haastateltavien keskinäistä riippuvuutta ja b) mielipidekysymysten vastausvaihtoehtojen toisensa poissulkevuutta. Kummankin ehdon paikkansapitävyyttä voidaan koetella tieteellisin keinoin.

Harkintaotoksena, kuten kiintiöpoimintana, kerätty havaintoaineisto ei sellaisenaan sovi kvantitatiiviseen analyysiin, varsinkaan yleisillä valmisohjelmilla tehtyyn analyysiin. Valmisohjelmat edellyttävät yleensä yksinkertaisena satunnaisotoksena kerättyä havaintoaineistoa, koska sellaiseen tilanteeseen niiden laskentamenetelmät on sovitettu. Kiintiöpoiminnassa harkinta aiheutuu haastattelijain suorittamasta harkitusta haastateltavien valinnasta. Tästä aiheutunutta vastaajien homogenisoituneisuutta käsiteltiin edellä samaan tapaan kuin haastattelijan henkilökohtaisen panoksen homogenisoivaa vaikutusta itse haastattelu-tilanteessa.

Kvantitatiivisen analyysin korjaamiseksi otanta-asetelma on tulkittava haastattelijakohtaiseksi ryväsotannaksi, jonka mukaisia laskentamenetelmiä on sovellettava yksinkertaisen satunnaisotoksen laskentamenetelmien sijasta, jos tuloksia halutaan yleistää lähtöperusjoukkoa koskeviksi.

Kvalitatiivista lähestymistapaa käytettiin tarkistettaessa mielipidemittarin vastausvaihtoehtojen pätevyyttä mittaustuloksina. Yksinkertaisessakin tapauksessa teemahaastattelu voi paljastaa, että standardi lomakehaastattelu ei ole riittävän yksityiskohtainen väline haastattelijan ja haastateltavan välisessä vuoropuhelussa. Esimerkkitapauksen tarkastelu antoi tukea tälle käsitykselle.

Edellä tehty inventointi antaa tukea tutkimusasetelmalle, jossa mittaustilanteessa yhdistetään sekä standardi- että teemahaastattelu. Standardilomakkeella kerätään suurehkosta otoksesta (esim. tässä $n=705$) vastaajien taustatiedot ja vastaukset mielipidekysymyksiin. Samanaikaisesti suunnitellaan suppeahko teemahaastattelu mielipidemittareiden kattamista kysymysalueista. Koska kysymyksessä on mittarin mittaussominaisuuksien tarkastaminen, koehenkilöiden lukumäärän ei tarvitse olla suuri. Esimerkiksi suurten otosten tapauksis-

sa 1—2 % alkuperäisen otoksen otoskoosta saattaa riittää. Teemahaastatteluun valittavat voidaan valita kahdella tavalla, joista yhtenä vaihtoehtona on suorittaa toisen vaiheen otanta alkuperäisen otoksen sisältä. Teknillisesti tällainen kaksivaiheinen otanta on helppo tehdä, ja sen tulosten analyysissä sekä alkuperäisen että sen sisältä poimitun otoksen tietoja voidaan yhdistellä. Menettely parantaa laskentatulosten tilastollista validiutta. Toinen tapa on poimia alkuperäisestä henkilöperusjoukosta pienempi teemahaastatteluun tuleva otos. Otoksen kokoonpanoa harkittaessa olisi syytä käyttää alkuperäisen suurotoksen analyysituloksia, jotta tutkimustulokset saataisiin mahdollisimman tehokkaasti laskettua.

Kiitokset: Tämän artikkelin työstöä sen alkuvaiheissa on auttanut keskustelu professori Steinar Kvalen (Aarhus) kanssa. Otanta-alueen sisällön on tarkistanut FT Risto Lehtonen ja käsikirjoituksen tekstinkäsittelyn viimeistellyt Kauko Saarinen.

LÄHTEET

- Atkinson, J.M. ja J. Heritage (toim. 1984): *Structures of social action. Studies in conversation analysis*. Cambridge University Press, Cambridge.
- Cochran, W.G. (1977): *Sampling Techniques*. Wiley, New York.
- van Dijk, T.A. (1980): *Macrostructures. An Interdisciplinary Study of Global Structures in Discourse, Interaction and Cognition*. Lawrence Erlbaum, Hillsdale.
- Fielding, N.G. ja R.M. Lee (toim. 1991): *Using Computers in Qualitative Research*. Sage, London.
- Hirvonen, M. (1989): Käytännön kokemuksia otantatutkimusten teosta. Teoksessa *Otantateorian nykylinjoja. Toinen kansallinen otantasymposiumi Jyväskylässä 16.—18. 5. 1989*, toim. E. Pahkinen ja R. Lehtonen. Jyväskylän yliopiston tilastotieteen laitos, Jyväskylä.
- Jääsaari, J. (1989): *Individualismi ja poliittisen kulttuurin muutos. Pääkaupunkiseudun nuorten aikuisten poliittinen suuntautuminen*. Lisensiaatintutkimus. Helsingin yliopisto.
- Jääsaari, J. ja T. Martikainen (1991): *Nuorten poliittiset vallinnat*. Gaudeamus, Helsinki.
- Kvale, S. (1989): To Validate Is to Question. Teoksessa *Issues of Validity in Qualitative Research*, toim. S. Kvale. Studentlitteratur, Lund.
- Little, R.J.A. ja D.B. Rubin (1987): *Statistical Analysis with Missing Data*. John Wiley, New York.
- Martikainen, T. ja R. Yrjönen (1991): *Vaalit, puolueet ja yhteiskunnan muutos*. Tilastokeskus, Helsinki.
- Miller A.H. ja O. Listhaug (1991): Ideology and Political Alienation. Paper to be read at Annual Meeting of the Midwest Political Science Association, April 18—20, Chicago.
- Mäkelä K. (toim. 1990): *Kvalitatiivisen aineiston analyysi ja tulkinta*. Gaudeamus Helsinki.
- Pahkinen, E. ja R. Lehtonen (1989): *Otanta-asetelmat ja tilastollinen analyysi*. Gaudeamus, Helsinki.
- Rao, J.N.K. ja D.R. Thomas (1988): The analysis of cross-classified categorical data from complex sample surveys. *Sociological Methodology* 18, s. 213—269.
- Skinner, C.J., D. Holt ja T.M.F. Smith (toim. 1989): *Analysis of Complex Surveys*. Wiley, New York.