

Tulevaisuuteen kohdistuvan Delfi-tutkimuksen reliabiliteetti

*Sama artikkelikäsikirjoitus löytyy myös englanninkielisenä versiona **Reliability of Delphi Technique***

Jari Metsämuuronen, RN, EdD
tutkija, SoTeKeKo-projekti

Tässä artikkelissa tarkastellaan *Sosiaali- ja terveysalan työn tulevaisuuden osaamistarpeet* – ESR-ennakointihankkeen selvityksiin ja empiiriseen aineistoon perustuen Delfi-tekniikalla tehtyjen tutkimusten reliabiliteettia. Delfi-tutkimusten toistettavuus on yleisesti ottaen epämääräinen: emme tiedä antaisiko jokin toinen raati oleellisesti toisenlaisen tuloksen tai antaisiko sama raati samanlaisen tuloksen pienen ajan kuluttua. Artikkelissa esitellään kolme erilaista tapaa arvioida Delfi-tutkimuksen reliabiliteettia. Empiirisenä aineistona esitellään *Sosiaali- ja terveysalan työn tulevaisuuden* –ennakointihankkeen kvalifikaatiovaatimusten luotettavuutta.

Asiasanat: Kvalifikaatiot, Reliabiliteetti, Validiteetti, Delfi-tekniikka, Sosiaali- ja terveysala, Luotettavuus, Kvalifikaatiovaatimukset

1. Johdanto

Tulevaisuutta ei varsinaisesti ole olemassa, eikä näin ollen mitään todellista tietoa-kaan tulevaisuudesta voi olla. Koska kuitenkin voimme kuvitella tulevaisuuden ja tiedämme, että elämässä ja yhteiskunnassa on olemassa säännönmukaisuutta, voimme helposti kuvitella, millaista tulevaisuus voi olla. Toisaalta voimme hyvin tarkasti laskea todennäköisyyksiä erilaisille tulevaisuuden tiloille, jos oletetaan, että muutos ei ole täysin kaoottista. Esimerkiksi almanakka on tulevaisuutta koskeva ennuste, sillä kahta samanlaista planeettavuotta ei löydy, koska painovoima aiheuttaa kaoottisuutta planeettojen liikkeisiin. Toisaalta almanakan ennuste on sään ennustamiseen verrattuna äärimmäisen tarkka, sillä planeettojen dynamiikka sallii hyvän ennustettavuuden toisin kuin sää, jolla on vahvasti sekoittavia ominaisuuksia.

Tulevaisuustietoa voidaan hankkia monella tavalla. Yksinkertaisten voidaan sanoa, että tulevaisuutta koskevaa tietoa voidaan hankkia joko laskennallisesti (kuten trendianalyysit, tulevaisuuden mallitus, tilastotiedot) tai asiantuntija-arvioiden perusteella (kuten delfi-tekniikka, haastattelut, barometrit, strategiatyöskentely, skenaariotyöskentely) (vrt. Millett & Honton 1991, joilla jako kolmeen: trendianalyysit, asiantuntija-analyysit ja monivaihtoehtoiset analyysit).

Asiantuntijamenetelmissä kysytään asiantuntijoilta heidän subjektiivista kannanottoaan siitä, millaisena he näkevät tulevaisuudesta. Oletetaan, että asiantuntija, joka on oman tiedon alansa kärjessä, pystyy ennakoimaan asioita paremmin kuin henkilö, jolla ei ole yhtä paljon tietoa kuin asiantuntijalla. Delfi-tekniikka on eräs asiantuntijamenetelmistä. Delfi-tekniikalle tyypillistä on se, että tietoa kerätään kahden tai useamman kierroksen aikana, jolloin asiantuntijat antavat mielipiteensä tulevaisuutta koskevista asioista toisten mielipiteistä tietämättä, mutta voivat muuttaa mielipidettään kun tutkija tuo koko asiantuntijaraadin mielipiteet kyseisille asiantuntijoille uudelleen arvioitavaksi. Kun Delfi-kierrokset on käyty läpi ja on saatu asiantuntijaraadin yhteinen mielipide, emme voi olla varmoja olisiko toinen asiantuntijaraati antanut aivan toisenlaisen mielipiteen. Toisaalta emme voi tietää, antaisiko sama raatikaan samanlaisen mielipiteen pienen ajan kuluttua.

Tässä artikkelissa pohditaan keinoja arvioida Delfi-tutkimuksen luotettavuutta, sen reliabiliteettia ja validiteettia. Lopuksi esitellään empiirisenä aineistona sosiaali- ja terveysalan työssä tulevaisuudessa tarvittavaa osaamista koskevan tutkimuksen reliabiliteetti ja validiteettitarkasteluja.

2. Delfitekniikka ja sen kritiikkiä

Delfi-tekniikka

Perinteisen Delfi-tekniikan idea on siinä, että kerätään tietoa asiantuntijoiden mielipiteistä. Kun asiantuntijoiden mielipiteet on koottu, ne lähetetään uudelleen samoille asiantuntijoille arvioitaviksi. Delfi -kierroksia käydään läpi niin monta kertaa, että asiantuntijoiden voidaan sanoa antaneen yhden yhteisen mielipiteen.

Terveystieteissä Delfi-menetelmää on käytetty vuodesta 1971 lähtien selvitetäessä mm. hoitotyön tutkimuksen prioriteetteja (Lindeman 1975, Ventura & Waligora-Serafin 1981, Bond & Bond 1982, Daly ym. 1996), koulutuksen kehittämistä (Spivey 1971, Tanner & Lindeman 1987, French ym. 1996), pätevyys- ja kvalifikaatiovaatimuksia (Duffield 1993) sekä hoitohenkilökunnan työn kehittämistä (Butterworth & Bishop 1995, Carney ym. 1996, Procter & Hunt 1994).

Vaikka tekniikka olikin tunnettu jo 1920-luvulla (Whitehead 1925) ja 1940-luvulla (Churchman 1948), sen varsinaisena lähtökohtana pidetään vuonna 1953 RAND-yhtiöiden toimesta tehtyä Yhdysvaltain puolustus-strategiaan liittyvää tutkimusta. Tutkimus tietenkin oli salainen, mistä syystä tekniikkaa koskeva kritiikki julkaistiin vasta 1963 tutkimuksen suorittajan Olaf Helmerin toimesta (Helmer 1963). Delfi-tekniikka kehitettiin siis alunperin strategisiin tarkoituksiin, mutta jo varhain sitä hyödynnettiin tulevaisuuden taloustutkimukseen (Helmer & Quade 1963).

Delfi -tekniikan iskulauseita ovat anonymiteetti ja konsensus. **Anonymiteetillä** tarkoitetaan sitä, että panelistit eivät tiedä, kuka osallistujista on sanonut minkäkin mielipiteen. Näin ollen menettely mahdollistaa panelistien tasavertaisuuden; vahvojen persoonien liiallinen vaikutus on mahdollista eliminoida (Linstone & Turoff 1975, Williams & Webb 1994, 181). **Konsensus** on mahdollista saavuttaa, mikäli Delfi-kierroksia kasvatetaan. Alkuaikojen tutkimuksille oli tyypillistä vaatimus yksimielisyydestä. Tästä ollaan luopumassa tai jo luovuttu (Turoff & Hiltz 1996, 56).

Delfi-tekniikan edut

Delfi-tekniikalla on neljä etua aikasarja-analyysiin ja komiteatyöskentelyyn verrattuna. Ensinnäkin Delfi-tekniikka soveltuu hyvin sellaisten taite- ja käänneasteiden löytämiseen ja ajoittamiseen, jotka on vaikea päätellä analyttisesti. Tällä tarkoitetaan sitä, että on olemassa analyttisiä tutkimusmenetelmiä (kuten aikasarja-analyysi ja siihen liittyen regressioanalyysi), joilla pyritään ennustamaan tulevaisuutta sen perusteella, mitä viestejä menneisyydestä ja nykyisyydestä tulee. Tällaiset numerotietoon perustuvat analyysimenetelmät ovat kuitenkin kyvyttömiä havaitsemaan ihmisen luovia ja tavoitteellisia ratkaisuja. Delfi-tekniikalla voidaan parhaassa tapauksessa tavoittaa juuri tällaista tietoa. Toiseksi, hyvässä Delfi-tutkimuksessa on mahdollista tuoda yksittäisen raadin jäsenen havaitsemat ”heikot signaalit” ja alan kehittämistavoitteet suuremman joukon arvioitaviksi. Heikkojen signaalien ajoissa kuuleminen, ymmärtäminen ja hyödyntäminen on ennakoivia. Kolmanneksi, hyvä Delfi-tutkimus ehkäisee arvovalta- ja intressiristiriitoja vaikuttamasta tutkimuksen tulokseen. Koska kukin vastaaja vastaa itsenäisesti, toimii tekniikka ”mies-ja-ääni” -periaatteella. Kaikkien mielipiteet ovat yhtä arvokkaita riippumatta siitä, miltä taholta mielipide tulee. Koska konsensusvaatimuksestakin on äärimuodoissaan luovuttu, tämä antaa panelistille mahdollisuuden olla eri mieltä asiasta. Tämä taas johtaa siihen, että panelistien ei tarvitse myöskään sitoutua tulokseen yhtä voimakkaasti kuin esimerkiksi komiteatyöskentelyssä. Neljänneksi Delfi-tekniikassa on erittäin helppo yhdistää kvantitatiivinen ja kvalitatiivinen tutkimusote (Brannen 1994, Hammersley 1994). On selvää, että mitä useampaa menetelmää käyttää, sitä varmempaa on saatu

tieto. Tätä kutsutaan triangulaatioksi (Denzin 1988); samaa ilmiötä tarkastellaan useasta eri suunnasta. Samaa tutkimusta voi siis tehdä sekä positivistisesti että fenomenologisesti, tiedon laatu vain on erilaista. Delfi-tekniikkaan, joka menetelmänä on selkeästi kvalitatiivista tietoa tuottava, on usein liitetty mukaan kvantitatiivisia elementtejä. Niinpä esimerkiksi konsensusta voidaan etsiä kvantitatiivisilla menetelmillä (Whitman 1990, 378; Duffield 1993, 228; Williams & Webb 1994, 183-184; Daly ym. 1996, 148). Tällöin panelistit arvottavat ensimmäisellä Delfi-kierroksella saadut mielipiteet esimerkiksi Likert-asteikolla.

Delfi-tekniikan kritiikkiä ja luotettavuus

Etujensa lisäksi Delfi-tekniikalla on eräitä puutteita, joita erinäinen määrä tutkijoita on ottanut esille. Ankarinta kritiikkiä lienevät esittäneet RAND-yhtiössä palvelleet Harold Sackman (1974) ja Johan Asplund (1979), mutta seuraava kritiikki pohjautuu pääosin terveystieteiden kirjallisuudessa esitettyihin kannanottoihin. Oleellista on se, että seuraavista kriittisistä seikoista huolimatta ovat kritiikkiäkin esittäneet tutkijat käyttäneet Delfi-tekniikkaa. Varovaisesti sanoen mainitut seikat ovat sellaisia, jotka tutkijan tulee tiedostaa ja joihin hänen tulee varautua tehdesään Delfi-tutkimusta. Tutkimuksen luotettavuuden ja erityisesti validiteetin takaamiseksi on tutkijan pohdittava systemaattisesti niitä tekijöitä, jotka saattavat alentaa tutkimuksen luotettavuutta (Cook & Campbell 1979).

Ensimmäinen kriittinen kohta Delfi-prosessissa on asiantuntijoiden valinta (Duffield 1993, 228). Kukaan ei pysty sanomaan kuinka monta asiantuntijointa tulisi olla (Reid 1988; Williams & Webb 1994, 182). On ollut paneeleita, joissa panelisteja on ollut yli tuhat (ks. Reid 1988). Toisaalta Ziglio (1996, 14) pitää 10-15 hengen panelia riittävänä. Yleensä panelistien määrä on vaihdellut 30:n ja 100:n välillä (Turoff & Hiltz 1996, 57). Kukaan ei myöskään pysty sanomaan sitä, kuka olisi ”riittävän asiantuntija” paneeliin (Goodman 1987, 731-732, ks myös Ziglio 1996, 14). Kysymyshän lienee siitä, onko kukaan lokaali asiantuntija todella oman tiedon-alansa kärjessä. Tässä mielessä mielenkiintoinen on professori Olavi Borgin ajatus asiantuntijuudesta, joka koskee liikenneongelmia: liikenneinsinöörin sijaan (tai rinnalla) todellinen asiantuntija saattaakin olla kokenut taksikuski, joka päivittäin joutuu liikenteen kanssa tekemisiin (professori Borg 1997, luento Tulevaisuuden tutkimuksen seminaarissa 19.3.1997 Tampereella). Samoin kiinnostava on Suomen tulevaisuustutkimuksen seuran puheenjohtajan Jarl-Thure Erikssonin viittaus siihen, että tavallisetkin ihmiset voivat olla heikkojen signaalien suhteen herkkiä, mikäli tarkkailevat ympäristöään ja oppivat aistimaan vivahteita ja muutostrendejä (Eriksson 1996, 4). Kysymys kuuluukin: ketä pitäisi käyttää asiantuntijana tai kuka lopultakin täyttää asiantuntijuuden kriteerit.

Toinen kriittinen kohta on ensimmäinen Delfi-kierros. Oleellista on tehdä oikeat kysymykset; väärät kysymykset saattavat johtaa asiantuntijat väärää suuntaan. Eräs Suomen johtavista Delfi-asiantuntijoista Osmo Kuusi on esittänyt, ettei ensimmäisissä tutkimuksissa pidetty tärkeänä oikeiden kysymysten laadintaa. Tähän vaiheeseen hän suosittaa asiantuntijan apua. Kuusi pitää vaihetta krusiaalina tutkimuksen onnistumisen kannalta. (Kuusi 1993, 136.) Ensimmäiseen Delfi-kierrokseen liittyy muitakin ongelmia. Jos kysely suoritetaan postikyselynä, tästä yleensä seuraa katoa. Sosiologisissa tutkimuksissa pidetään kohtuullisena, mikäli 70%:kin

postituskierröksellä mukana olleista vastaisi lomakkeeseen. Se, miten kato vaikuttaa tulokseen, jää jokaisen tutkijan harkittavaksi. Haastattelumenetelmällä kato saattaa jäädä oleellisesti pienemmäksi. Edelleen on huomautettu, että ensimmäisen kierroksen tulosten analysointia ei välttämättä pystytä saamaan samanlaisiksi erilaisten tutkimusten välillä, sillä yhtenäistä metodologiaa ei analyysistä ole (Whitman 1990, 378, Procter & Hunt 1994, 1004).

Kolmas vaihe, jossa pitää tehdä kriittisiä valintoja, on iteraatiokertojen määrän valinta. On sanottu, että jos kierroksia on enemmän kuin kolme, tulee tutkimuksesta aikaavievä ja kallis (Williams & Webb 1994, 182; Kuusi 1993, 136). On myös esitetty, että kolmannen kierroksen jälkeen ei juurikaan muutoksia tulisi mielipiteisiin (Duffield 1988).

Neljäs keskeinen ongelma on konsensus. Ongelmahan on se, mitä pidetään konsensusena. Terveystieteiden kirjallisuudessa ratkaisuksi konsensusongelmaan on esitetty ”riittävän konsensuksen” löytämistä (Williams & Webb 1994, 184; Duffield 1993, 228, 230; Daly ym. 1996, 148), mikä yleensä perustuu kvantitatiiviseen mielipiteiden arvottamiseen. Täydellistä konsensusta ei pidetä enää toivottavana, sillä ajatellaan, että pyrkimys täydelliseen mielipiteeseen voi johtaa tärkeän informaation katoamiseen (Turoff & Hiltz 1996, 56, Kuusi 1993, 136).

Viides kritiikin lähde on tutkimuksen toistettavuus, joka on erityisesti tämän artikkelin aihe. Delfi-tutkimuksen reliabiliteetti eli toistettavuus on epämääräinen (Reid 1988; Williams & Webb 1994, 182, ks. myös Ziglio 1996, 15-17). Olisivatko toiset asiantuntijat antaneet toisenlaisen lopputuloksen? Tässä mielessä mielenkiintoinen on Duffieldin hoitoalan kompetensseihin liittyvä tutkimus, jossa otettiin kaksi eri paneelia ja verrattiin heidän vastauksiaan. 92,86% tutkituista hoitotyön kompetensseista oli samoja molemmilla ryhmillä. (Duffield 1993, 236.) Toisin sanoen asiantuntija paneelit olivat varsin yhtä mieltä keskeisistä kompetensseista. Koska yhtenäistä metodologiaa materiaalin analyysistä ei ole, on teoriassa mahdollista, että eri analysoijat voisivat saada samasta aineistosta erilaisia tuloksia riippuen taustasitoumuksistaan, lahjakkuudestaan ja vaivannäöstään.

Toisenlainen metodologinen ongelma on kysymys asiantuntijan mielipiteen stabiiludesta eli pysyvyydestä (Metsämuuronen 1997): mielipiteillä on taipumus ajan myötä muuttua. Saattaa olla hankala saada selville sitä, milloin muutos asiantuntijan mielipiteessä johtuu asiantuntijasta itsestään (asiantuntija ei ole osannut päättää, tai mielipide stabiilista ilmiöstä muuttuu). Tällöin voidaan puhua asiantuntijan stabilidesta. Toisaalta pitäisi tietää, milloin muutos mielipiteessä johtuu itse ilmiössä tapahtuvasta muutoksesta (ilmiön stabilius). Delfi-tekniikkaan on sisäänrakennettuna mahdollisuus muuttaa mielipidettään. Ongelma syntyykin siitä, että kun riittävät Delfi-kierrokset on käyty läpi, ja on löytynyt yhteinen mielipide, emme tiedä tulisiko kuukauden, puolen vuoden tai vuoden päästä osittain hyvinkin erilaisia tuloksia, mikäli tehtäisiin uusi Delfi-tutkimus.

3. Tulevaisuustutkimuksen luotettavuuden arviointi

Tulevaisuutta ennakoivan tutkimuksen luotettavuuden kriteeriä ei ole vielä olemassa. Nimittäin tulevaisuustutkimus on tiukasti ajatellen luotettava silloin, jos se tuottaa paikkaansa pitäviä tuloksia. Toisaalta tulevaisuustutkimuksen arvo ei välttämättä olekaan siinä, onko se totta vai ei, vaan siinä miten kiinnostava tai vaikuttava itse tutkimus on. Tulevaisuus tutkimus on myös tulevaisuuteen vaikuttamista.

Tutkimuksen luotettavuutta on perinteisesti arvioitu kahdelta kannalta. Luotettavuus jaetaan reliabiliteettiin ja validiteettiin, joista reliabiliteetti kuvaa toistettavuutta ja validiteetti yhtäältä sitä, kuinka yleistettäviä tulokset ovat (ulkoinen validiteetti) ja toisaalta tutkitaanko sitä, mitä on tarkoituskin tutkia (sisäinen validiteetti). Tässä artikkelissa paneudutaan hieman reliabiliteetin arvioimiseen tulevaisuustutkimuksessa, joka on tehty Delfi-menetelmällä.

Reliabiliteetti

Reliabiliteetti kuvaa sitä, kuinka yhteneviä ja toistettavia tulokset ovat. Eräs keskeinen tapa arvioida mittauksen reliabiliteettia on tarkastella sitä mittarin sisäisen konsistenssin eli yhtenevyyden kautta. Tämän alueen tunnetuin mitta on Cronbachin alfa (Cronbach 1951), vaikka alfa onkin saanut osakseen terävää kritiikkiä (esimerkiksi Tarkkonen 1987). Alfaan idea on, että mikäli osiot todella mittaavat keskenään samantyyppistä asiaa (kuten sosiaalisia taitoja), osiot korreloivat keskenään. Jos alfa on korkea, muuttujat mittaavat samantyyppistä asiaa, sillä kukin vastaaja on vastannut osioihin konsistentisti, yhtenevästi. Cronbach itse kirjoitti (Cronbach 1951, 331), että alfaa voidaan pitää reliabiliteetin alarajana. Tämä johtuu siitä, että alfan laskukaava tekee teoreettisesti mittarista erilaisia puolituksia, *split-half*jeja, joiden välille lasketaan korrelaatio. Mukaan tulee erinomaisia osituksia, mutta myös heikkojakin osituksia. Alfa on näiden erilaisten ositusten korrelaatioiden keskiarvo. Perinteisesti on esitetty, että pienemmät alfan arvot kuin 0.60 pitäisi jättää hyväksymättä (Nunnally & Bernstein 1994, vrt Knapp & Brown 1995, 465-469). Tällöin mittaria pitäisi kehittää luotettavuuden lisäämiseksi esimerkiksi poistamalla heikkoja osioita. Alfaan laskemisen tekniikasta johtuen pitkä mittari on usein luotettavampi kuin lyhyt mittari, sillä muuttujien yhteisvaihtelun kasvaessa myös alfa saa suurempia arvoja.

Tulevaisuuskuvan toistettavuus Delfi-tekniikalla arvioituna on siis epävarma. Jos asiantuntijoita on tutkimuksessa kohtuullisen paljon, se antaa teknisen mahdollisuuden jakaa asiantuntijajoukko kahteen ryhmään, joille on mahdollista tehdä rinnakkaismittaus. Tämä käytännössä tarkoittaa sitä, että muodostetaan samalla tekniikalla kaksi riippumatonta tulevaisuuden kuvaa. Mikäli osaamistarpeet ovat kahdella ryhmällä samanlaiset (tai ainakin hyvin samanlaiset) on Delfi -tekniikka tuottanut luotettavaa, reliabelia, tietoa. Tässä yhteydessä luotettavuus ei tarkoita sitä, että tulevaisuus tulisi olemaan juuri kuvatun kaltainen, vaan sitä, että asiantuntijat ovat olleet yksimielisiä tulevaisuuden osaamistarpeista. Mikäli taas kahden asiantuntijaryhmän esittämät osaamistarpeet eroavat toisistaan huomattavasti, siitä seuraa, että päätelmien suhteen tulee olla varovaisempi. Viimeksi mainittu tilanne johtaa tietenkin myös maltillisempiin toimenpiteisiin, mikä lieneekin järkevää oletetussa tilanteessa. Todennäköistä kuitenkin on, että pieniä eroja lukuunot-

tamatta kahden ryhmän konsensus on hyvin pitkälle samanlainen (näin esimerkiksi Duffield 1993).

Mikäli käytetään kahden raadin tekniikkaa ja kvantitatiivista arvotusmenetelmää, on mahdollista antaa toiselle tai molemmille raadeista vastattavaksi oman lomakkeen myös toisen raadin lomake. Tällöin kyetään vertailemaan kohtuullisen tarkasti molempien raatien eroja kolmella tavalla. Ensimmäiseksi asiantuntijaraatien yhtäpitävyyttä voidaan arvioida sillä perusteella, kuinka paljon erilaisia, uniikkeja, osioita ilmeni toisella asiantuntijaraadilla. Tämä tapahtuu vertaamalla kummankin raadin omia lomakkeita toisiinsa. Toiseksi raatien yhtäpitävyyttä arvioidaan sitä, oliko raatien välillä aidosti eroja kun vertailtavana on sama lomake molemmissa ryhmissä. Kolmanneksi arvioidaan toisen asiantuntijaraadin omaa konsistenssia. Tällöin voidaan verrata toisen raadin tai molempien raatien vastaamia kahta lomaketta toisiinsa.

4. Empiirinen esimerkki sosiaali- ja terveysalan kvalifikaatiotutkimuksesta

Tausta

31.7.1998 päättyi sosiaali- ja terveysministeriön, opetushallituksen, sosiaali- ja terveysalan tutkimus- ja kehittämiskeskuksen (Stakes) ja Suomen Kuntaliiton vuonna 1995 käynnistämä SoTeKeKo –yhteistyöprojekti, jonka tavoitteena oli sosiaali- ja terveysalan työelämän ja koulutuksen yhteistyön kehittäminen. Tähän projektiin liittyi myös osittain Euroopan Sosiaalirahaston (ESR) rahoittamana *Sosiaali- ja terveysalan työn tulevaisuuden osaamistarpeet* –ennakointihanke, jonka tavoitteena oli selvittää, millaista osaamista tarvitaan sosiaali- ja terveysalalla tulevaisuudessa.

Tutkimustekniikan kuvaus

Tutkimus suoritettiin kahden kierroksen Delfi-tekniikalla. Asiantuntijoina käytettiin sosiaali- ja terveysalan työelämän ja ammatillisen koulutuksen edustajia. Otanta oli stratifioitu: mukaan haluttiin sosiaalialan, terveysalan ja sosiaali- ja terveysalan edustajia. Lisäksi asiantuntijoiden tuli edustaa työelämää, koulutusta ja opiskelijoita. Tutkimus aloitettiin normaaliin tapaan postikyselynä, jolloin mukaan saatiin noin 187 sosiaalialan ja terveysalan työelämän ja koulutuksen asiantuntijaa sekä 18 järjestöjen ja tutkimusalan edustajaa. Asiantuntijoista muodostettiin kaksi raatia, joista varsinainen asiantuntijaraati koostui työelämän ja koulutuksen edustajista ja kontrolliraati järjestö- ja tutkimusalan ihmisistä. Ensimmäisen kierroksen avoimet vastaukset analysoitiin sisällön analyysi- ja käsitekarttatekniikalla. Toiselle kierrokselle valittiin varsinaisesta asiantuntijaraadista argumentaatioanalyysin perusteella parhaat osaamistarpeiden argumentoijat. Näin toiselle kierrokselle tuli mukaan oleellisesti pienempi joukko asiantuntijoita (yhteensä 43 varsinaiseen raatiin ja 18 kontrolliraatiin) niistä vastaajista, jotka jo olivat mukana ensimmäisellä kierroksella.

Toisella kierroksella asiantuntijat saivat vastattavakseen lomakkeen, johon heidän piti arvottaa 6-portaisella Likert-asteikolla kukin osaamistarve sillä perusteella, tarvitaanko osaamista tänään ja tarvitaanko osaamista tulevaisuudessa. Näiden kahden numeron perusteella arvioitiin aluksi, oliko kyseinen osio ns. ”turha” osaamistarve, eli jotain sellaista, mitä asiantuntijat valtaosin eivät pitäneet tärkeänä osaamisena nyt eikä tulevaisuudessa. Toiseksi arvioitiin ns. muutosintensiteettiä, ts. sitä, kuinka paljon enemmän tai vähemmän kyseistä osaamista tarvitaan tulevaisuudessa kuin tänään. Kvalifikaatioiden hyväksymisessä mittariin käytettiin kolmea kriteeriä. Kvalifikaatio pidettiin mittarissa, jos joku seuraavista kriteereistä täyttyy: 1) Kvalifikaation painoarvo (keskiarvo) oli suurempi kuin 5 (eli asiantuntijat olivat täysin yksimielisiä tai melko yksimielisiä kvalifikaation tärkeydestä), 2) muutosintensiteetti oli suurempi kuin 0.5 joko sosiaalialan tai terveystalouden edustajilla (eli kyseessä on todelliset tulevaisuuden kvalifikaatiot) ja 3) tilastollisesti merkitseviä eroja ei esiintynyt sosiaalialan ja terveystalouden eikä koulutuksen ja työelämän edustajien välillä (eli että mielipiteen olivat yhteneviä eri ryhmissä). Ne osiot siis poistettiin, joissa kaikki kolme kriteeriä jäivät toteutumatta.

Toisella Delfi-kierroksella kato jäi varsinaisessa asiantuntijaraadissa pieneksi (5 %) ja toisessakin raadissa kohtuulliseksi (28 %).

Mittareiden reliabiliteetti

Mittareiden reliabiliteettia mitattiin neljällä tavalla. Ensiksi reliabiliteettia tutkittiin varsinaisen raadin mittarien sisäisen yhtenevyyden eli konsistenssin avulla. Osaamistarvemittareiden reliabiliteetti on kohtuullinen ja vaihtelee välillä 0.65 - 0.93 (taulukko 1). Uniikkiluokille ei ole mahdollista laskea alfan arvoa.

Faktorianalyysin avulla olisi mahdollisesti löydetty parempia mittareita, mutta aineiston pienuus (n=41) ei mahdollista kvalifikaatioiden eksploratiivista faktorointia. Tällaisinaankin kvalifikaatioluokat ovat riittävän konsistenteja jatkoanalyysien kannalta; luokkien nimet kuvaavat riittävällä tarkkuudella luokan muodostavia kvalifikaatioita.

Taulukko 1. Osaamistarveluokkien sisäinen konsistenssi Cronbachin alfalla mitattuna

Kvalifikaatioluokka	Kvalifikaatioita luokassa (lkm)	Indikaattorilauseita 1. Delfi-kierroksella	α
Sosiaalinen osaaminen	10	149	0.93
Ihmisen kohtaamisosaaminen	15	149	0.90
Kansainvälisyysosaaminen	5	110	0.66
Muuttumisosaaminen	15	103	0.86
Teknologinen osaaminen	3	93	0.75
Tiedonhankintaosaaminen	6	85	0.75
Asiakaspalveluosaaminen	10	60	0.87
Työhön ja työssä vaikuttamisosaaminen	15	59	0.80
Oman persoonan käytön osaaminen	12	58	0.92
Yhteistyöosaaminen	7	55	0.85
Innovatiivisen työnotteen osaaminen	6	55	0.76
Yhteiskuntaosaaminen	5	51	0.72
Päätöksenteon osaaminen	9	45	0.81
Yrittäjyysosaaminen	3	43	0.73
Itsenäisyysosaaminen	2	34	0.76
Aktiivisen muuttamisen osaaminen	6	28	0.72
Arvo-osaaminen ja eettinen osaaminen	6	26	0.86
Äänetön ja pehmeä osaaminen	10	18	0.86
Projektityöosaaminen	3	17	0.65
Voimakkaampi vastuunotto	1	16	*
Johtamisosaaminen	3	16	0.81
Itsepuolustusosaaminen	1	3	*

* reliabiliteettia ei laskettu, sillä luokassa oli vain yksi osaamistarve

Toiseksi verrattiin ensimmäisen raadin lomaketta toisen raadin omaan lomakkeeseen. Ensimmäisen asiantuntijaraadin ilmaisemia kvalifikaatioita oli lukumäärällisesti 181 ja toisella raadilla 91. Toisen raadin osaamistarpeista 17 oli uniikkia kvalifikaatiota, joiden indikaattoreita ei kyetty luokittelemaan ensimmäisen raadin luokitusten mukaisesti. Voidaan siis laskea, että toisella asiantuntijaraadilla oli 9.4 % sellaisia kvalifikaatioita, joita ensimmäisellä raadilla ei ollut. Kaikki toisen raadin uniikit kvalifikaatiot olisi voitu luokitella ensimmäisen asiantuntijaraadin olemassa oleviin 22 kvalifikaatioluokkaan.

Kolmanneksi raatien yhtäpitävyyttä arvioitiin sitä, oliko raatien välillä aidosti eroja, kun vertailtavana on ensimmäisen asiantuntijaraadin laaja lomake; verrataan siis ensimmäisen raadin lomaketta molemmissa ryhmissä. Ryhmien keskiarvojen välistä eroa tutkittiin t-testillä alkuperäisten kvalifikaatioiden suhteen (217 osaamistarvetta). Kvalifikaatioiden painoarvojen suhteen eroja ryhmien välillä on niin niukasti, että voidaan sanoa molempien raatien asiantuntijoiden olevan yksimielisiä osaamistarpeista. 217 osaamistarpeesta vain 8:ssa ryhmien välillä oli tilastollisesti merkitsevä ero. Näin ollen 217:sta kvalifikaatiosta samoja oli 209 eli 96.3 %. Luku vastaa Duffieldin saamaa arvoa (Duffield 1993, 236). Kiintoisia vertailuja oli 8. Bonferroni-korjauksen jälkeen ei yksikään t-arvo ollut tilastollisesti merkitsevä.

Taulukko 2.Toisen asiantuntijaraadin sisäinen konsistenssi

korre- laatio	stabi- teetti	osaamistarve ”Tulevaisuudessa tarvitaan...”
0,81	-	Korkeaa ammattitaitoa
0,41	0.38 ¹	Taitoja (todellista osaamista)
0,44	0.77	Laaja-alaisuutta
0,97	-	Erytisoaamista
0,97	-	Erikoistumista
0,87	-	Usean ammatin hallintaa
0,76	-	Taito toimia moniammatillisesti
0,02	0.61	Taito toimia hoitoketjuissa
0,27	0.77	Kommunikaatiotaitoja
0,27	0.77	Vuorovaikutustaitoja
0,63	-	Yhteistyökykyä
-	0.69	Yhteistyöosaamista
0,76	-	Verkostoitumistaitoja
-0,08	0.85	Taitoa toimia moniammatillisessa
1,00	-	Yhdessä tekemisen taitoa
1,00	-	Yhteisöllistä toimintakykyä
0,64	-	Sosiaalisia taitoja
0,42	0.85	Ihmisen kohtaamisen taitoa
0,68	-	Taitoa huomioida asiakkaan
0,64	-	Taitoa vastata asiakkaiden
-	0.85	Havainnointitaitoa
0,82	-	Taito jäsentää monipuolisesti
0,76	-	Taito kuunnella
0,46	0.69	Taito myötäelää
0,55	-	Huolenpidon taidot
0,37	0.54	Taitoa olla ihminen ihmiselle
0,63	-	Halu ja kyky muuttua
0,39	0.80	Kyky siirtyä joustavasti työyhteisöstä
0,90	-	Taito vaikuttaa yhteiskunnallisiin
0,81	-	Rohkeutta puuttua epäkohtiin
-0,23	0.61	Työn kehittämishalua
0,13	0.69	Työn kehittämistaitoja
0,77	-	Reflektointikykyä
0,81	-	Voimakkaampaa sitoutumista
0,73	-	Taito vaikuttaa aktiivisesti
0,72	-	Monikulttuurisuustaidot
0,91	-	Tuntemusta muista kulttuureista
0,94	-	Tuntemusta muista uskonnoista
0,78	-	Tietoa muista lainsäädännöistä
0,64	-	ATK:n käyttötaito
0,69	-	Teknisten laitteiden hallintaa
1,00	-	Tiedon hankintataito
0,40	0.85	Kyky soveltaa uutta tietoa
0,70	-	Elinikäisen oppimisen taidot
0,76	-	Halu itsensä jatkuvaan kehitt äminen
0,79	-	Taitoa itsensä jatkuvaan kehitt .
0,87	-	Ymmärrystä yhteiskunnallisista
0,89	-	Yhteiskunnallisten syy-seuraussuhteet
0,52	-	Kokonaisuuksien ymmärtäminen
0,19	0.69	Kyky itsenäiseen työsk entelyyn
0,67	-	Voimakkaampaa vastuunotto
0,88	-	Taloudellinen ajattelu
0,48	0.61	Yrittäjyystaidot
0,47	0.53	Markkinointitaidot
0,43	0.77	Palveluhenkisyys
0,55	-	Ohjaustaidot

korre- laatio	stabili- teetti	osaamistarve
0,28	0.54	Opetustaidot
0,39	0.61	Taito antaa tukea
0,50	-	Ongelmanratkaisukyky
0,32	0.69	Päätöksentekotaito
0,70	-	Johtamistaito
0,69	-	Suunnittelutaito
0,82	-	Taito tuntee tunteita
0,92	-	Taito tunnistaa omia tunteita
0,92	-	Vanha itsetunto
0,83	-	Vahva itsetuntemus
0,36	0.61	Usko omiin kykyihin
0,40	0.54	Rohkeus elää aidosti
0,52	-	Oman elämän hallinnan taito
0,15	0.54	Oman persoonallisuuden luova käyttö
0,00	0.69	Erilaisuuden hyväksymisen taito
0,52	-	Arvo-osaaminen
0,43	0.77	Eettinen osaaminen
0,77	-	Enemmän sosiaalitieteellistä tietoa
0,77	-	Enemmän käyttäytymistieteellistä t..
0,65	-	Enemmän lääketieteellistä tietoa
0,65	-	Enemmän luonnontieteellistä tietoa
0,51	-	Lisää mielenterveytyön osaamista
0,71	-	Lisää tietoa yrittäjyydestä
0,85	-	Vankka kielitaito
0,91	-	Kansainvälisyystaidot

1) stabiliteettikerroin on laskettu vain silloin, kun korrelaatiokerroin on ollut pienempi kuin 0.50.

Stabiliteettikertoimen laskussa ei ole käytetty korjauskaavaa, vaan laskettiin suoraan samamielisenä pysyneiden asiantuntijoiden osuus kaikista asiantuntijoista

Neljänneksi arvioitiin toisen asiantuntijaraadin omaa konsistenssia. Tällöin verrattiin toisen raadin vastaamia kahta lomaketta toisiinsa. Osa kvalifikaatioista oli raadin vastaamissa lomakkeissa samoja (81 kvalifikaatiota). Näitä osioita vertaamalla voidaan päätellä, kuinka johdonmukaisesti toisen asiantuntijaryhmän jäsenet ovat olleet arvotuksissaan. Toinen lomake lähetettiin raadille viikon kuluttua ensimmäisen lomakkeen lähettämisestä. Lomakkeiden vastaamisen välinen aikaväli oli siis viikko, mitä on perinteisesti pidetty riittävän lyhyenä testi-uusintatesti – reliabiliteetin mittaamiselle (Nunnally 1978, 234-236, Tarkkonen 1987, 25). Mittauksen ja uusintamittauksen välinen korrelaatio on joiltain osin suurempi kuin 0.9, mutta ilmeni, että osaamistarpeiden joukossa on joitain osioita, joiden reliabiliteetti korrelaatiolla laskettuna on erittäin matala – jopa negatiivinen (taulukko 2). Joitain korrelaatioita ei laskuproseduuri kyennyt laskemaan lainkaan, sillä asiantuntijat olivat olleet täydellisen yksimielisiä osaamistarpeen tärkeydestä ja näin ollen keskipoikkeama tuli nol-laksi, mikä estää korrelaation laskun. Matalat korrelaatiot (pienemmän kuin 0.50) tarkistettiin stabiliteettikertoimella, jonka on todettu antavan tarkempia tuloksia erityisesti tilanteessa, jossa korrelaatiokerroin redusoituu lähelle nol-laa teknisistä syistä (Metsämuuronen 1997). Yhtä lukuun ottamatta kaikissa tapauksissa stabiliteetti osoitti mielipiteiden säilyneen mittauksetojen välillä samoina, vaikka

korrelaatio ilmoittikin toista.

Kahden mittauksen välisen korrelaation ja stabiliteetin perusteella vain yhden osaamistarpeen suhteen asiantuntijat eivät olleet konsistentteja: osaamistarve *tarvi-taan taitoja (todellista osaamista)* oli sellainen, jonka suhteen saattoi olla suuriakin poikkeamia asiantuntijan mielipiteessä mittauskertojen välillä. Itse osaamistarve onkin diffuusi ja kohdentumaton, mikä ilmeisesti on syynä epämääräisyyteen myös arvotuksissa. Muiden osaamistarpeiden suhteen asiantuntijat olivat erittäin konsistentteja (korrelaatio tai stabiliteetti 0.90 – 1.00, yhteensä 11 osaamistarvetta), melko konsistentteja (korrelaatio tai stabiliteetti 0.70 – 0.89, yhteensä 35 osaamistarvetta) tai kohtuullisen konsistentteja (korrelaatio tai stabiliteetti 0.50 – 0.69, yhteensä 34 osaamistarvetta).

Mittareiden luotettavuuden yhteenveto

Kaiken kaikkiaan erot kahden eri asiantuntijaraadin välillä olivat pienet. Toisella asiantuntijaraadilla oli 17 uniikkia osaamistarvetta, jotka kuitenkin olisi helposti voitu sisällyttää ensimmäisen raadin pohjalta luotuihin 22:een osaamistarveluokkaan. Yhteisiä osaamistarpeita raatien välillä oli 96.3 %. Vain kahdeksan osaamistarpeen suhteen erot raatien välillä olivat tilastollisesti merkitseviä. Bonferroni-korjauksen jälkeen yksikään ero ei ollut tilastollisesti merkitsevä. Asiantuntijat olivat myös valtaosin vastanneet kyselyyn johdonmukaisesti. Näin voidaan päätellä siitä, että mittauksen ja uusintamittauksen väliset korrelaatiot ja stabiliteetit olivat kaikki suurempia kuin 0.50 ja valtaosin suurempia kuin 0.70.

On ilmeistä, että nimenomaan toisen asiantuntijaraadin suhteen Delfi-kierrokset jäivät keskeneräisiksi. Kvalifikaatioita tuli ensimmäisessä raadissa huomattavan paljon enemmän kuin jälkimmäisessä raadissa. Näyttää siltä, että asiantuntijoiden lukumäärä ei ollut riittävä pienemmässä raadissa. Delfi-kierroksia olisi pitänyt lisätä ja saada lisää asiantuntijoita, tai aineisto olisi pitänyt kerätä toisella tavalla, esimerkiksi haastattelemalla.

5. Pohdintaa

Tässä artikkelissa on pohdittu Delfi-tekniikan mukanaan tuomia luotettavuus-ongelmia, jotka ovat tyypillisiä kvalitatiiviselle aineiston käsittelylle. Eräänä mahdollisuutena ratkaista erityisesti reliabiliteetin laskemisen ongelma esitettiin kahden raadin tekniikkaa johon kytketään kvantitatiivinen mittauskierros. Mikäli jompikumpi tai molemmat raadeista saavat vastattavakseen oman lomakkeensa lisäksi toisen raadin lomakkeen, on näiden lomakkeiden avulla mahdollisuus laskea reliabiliteetti neljällä tavalla: (1) laskemalla mittareiden sisäinen konsistenssi esimerkiksi Cronbachin alfan avulla, (2) laskemalla korrelaatio- tai stabiliteettikerroin raadin kahden eri mittarin yhteisten osioiden välille (testiuusintatesti korrelaatio), (3) laskemalla kuinka paljon aidosti erilaisia osioita raatien

välillä tulee vertaamalla molempien ryhmien omia lomakkeita toisiinsa sekä (4) laskemalla kuinka yksimielisiä raadit ovat vertailemalla eri raatien vastaamia samoja lomakkeita.

Kuten Delfi-tutkimusten reliabiliteetti myös validiteettikin – erityisesti sisäinen validiteetti – voidaan asettaa kyseenalaiseksi. Nimittäin käsitteiden suhteen sisäinen validiteetti ei ole aivan ilmeinen. Kvalitatiiviselle tutkimusotteelle tyypillisesti osa luokitteluista olisi voitaisiin tehdä toisin ja epäilemättä toinen tutkija saattaa yhdistellä indikaattorilauseita toisella tavalla erilaisiksi osaamistarpeiksi ja edelleen erilaisiksi osaamistarveluokiksi. Numeerinen tapa analysoida kvalifikaatioita ei välttämättä anna oikeaa kuvaa asiasta, sillä osa väitelauseista ei välttämättä ole yksikäsitteisiä. Emme siis voi tietää ovatko asiantuntijat tulkinneet sanat ja sanamuodot samalla tavalla. Monitulkintaisuusongelmaa voidaan hieman pienentää tekemällä ennen toista Delfi-kierrosta esimerkiksi fenomenografinen tutkimus käsitteiden erilaisista tulkinnoista (Uljens 1991, 80). Fenomenografinen tutkimuksen tuloksista huolimatta on kuitenkin ainakin kaksi suurta ongelmaa. Toisaalta ongelmallista on, että käsitykset voivat muuttua oppimisen tai lisäinformaation myötä, toisaalta eri ihmisillä on aidosti erilaiset käsitykset asioista. Niinpä jos olisimme saaneet ensimmäisen kierroksen yhteydessä kartoitettua asiantuntijoiden käsitykset vaikeista termeistä, emme yhtäältä tiedä olisivatko käsitykset muuttuneet siihen mennessä, kun toinen kierros alkaa. Toisaalta jos olisimme saaneet selville erilaisia käsityksiä, millä perusteella olisimme valinneet jonkun olemassa olevista tulkinnoista tutkimuksessa käytettäväksi ”lopulliseksi” tulkinnaksi.

Eräs mielenkiintoinen ongelma syntyy siitä, että Delfi-tutkimuksen tuloksia on vaikea saada istumaan valtakunnallisiin skenaarioihin. Ilmeisesti jokaisessa maassa luodaan erilaisia vaihtoehtoisia tulevaisuuden maailmoja, jotka toteutuessaan edellyttävät tiettyjä toimenpiteitä. Erityisesti sosiaali- ja terveysalan tulevaisuuden vertaaminen valtakunnallisiin skenaarioihin on hankalaa. Tämä saattaa johtua siitä, että valtakunnalliset skenaariot perustuvat usein taloudessa tapahtuviin muutoksiin. Sosiaali- ja terveysalan muutokset kyllä perustuvat yhteiskunnalliseen kehitykseen, mutta houkuttelevaa olisi ajatella, että ihmisen muuttuminen olisi se primääri tekijä, joka vaikuttaa yhteiskuntaan, teollisuuteen tai markkinoihin. Ensisijaista ei olisikaan se, kuinka armotonta menoa on tulevaisuus markkinavoimien myllerryksessä, vaan se, antaako ihminen materiaalien arvojen vallita ja hallita, vai kehittykö uudenlainen ”pehmeille” arvoille perustuva yhteiskunta. Sosiaali- ja terveydenhuollon kannalta ei siis välttämättä ensisijaista olekaan se, millainen on tulevaisuuden yhteiskunta julkisen rajoituksen suhteen, vaan se, kuinka ihmiset yksilöinä muuttuvat huomioimaan omaa terveyttään ja omia omiaisiaan.

Vaikka mittaus ja mittarit olisivatkin luotettavia, Delfi-tekniikalla tehtävä tulevaisuustutkimus seisoo ja kaatuu asiantuntijoiden valinnan ja heidän argumenttiansa mukana. Tässä artikkelissa ei pohdittu luotettavan ja monipuolisen argumentoinnin merkitystä luotettavuudessa. On kuitenkin selvää, että mikäli asiantuntijat eivät kykene argumentoimaan olemassa oleviin realiteetteihin perustuen aukottomasti mielipidettään tulevaisuudesta, mielipide ei ole minkään arvoinen luotettavana tulevaisuustietona. Sillä saattaa olla oma paikkansa aavistuksena, intuition tai science fictionina, muttei sen varaan voi rakentaa

mitään.

Yhteenveto

Delfi-tutkimusten luotettavuus ei ole aivan ilmeinen. Toisaalta Delfi-tutkimusta soveltuu nimenomaan tulevaisuustutkimukseen, sillä sen avulla on mahdollista löytää yhtäältä tulevaisuutta koskevia heikkoja signaaleja ja toisaalta se soveltuu erinomaisesti ennustamaan sellaisia tulevaisuuden taitekohtia, joita ei löytäisi pelkän numeerisen aineiston perusteella. Erityisesti sosiaali- ja terveysalan tulee ottaa jo hyvissä ajoin kantaa siihen, kuinka tarjotaan hoivaa ja hoitoa toisen maailman sodan jälkeen syntyneille ns. suurille ikäluokille. Tässä mielessä Delfi-tutkimus puolustaa paikkaansa muiden tulevaisuustutkimuksen menetelmien joukossa.

Mikäli Delfi-tutkimuksen luotettavuutta – ja erityisesti sen reliabiliteettia – halutaan tutkia, se on mahdollista, mutta se vaatii numeerisen tiedonhankintatavan. Likert-asteikollinen arvottaminen soveltuu hyvin reliabiliteetin laskemiseen. Mikäli käytettävissä on suuri joukko asiantuntijoita, on mahdollista käyttää kahden raadin tekniikkaa. Tämä reliabiliteetti on mahdollista laskea usealla erilaisella tavalla, joista tässä artikkelissa esiteltiin 4.

Lähteet

- Asplund, J. 1979:** *Theorier om framtiden*. Liberförlag, Stockholm
- Bond, S. & Bond, J. 1982:** A Delphi survey of clinical research priorities. *Journal of Advanced Nursing* 7, 563-575
- Brannen, J. 1994:** Combining qualitative and quantitative approaches: an overview. Teoksessa Brannen, J. (toim.) 1994: *Mixing Methods: qualitative and quantitative research*. 3. painos. Cedric Chivers Ltd., Bristol. 3-38.
- Butterworth, T. & Bishop, V. 1995:** Identifying the characteristics of optimum practice: findings from a survey of practice experts in nursing, midwifery and health visiting. *Journal of Advanced Nursing* 22, 24-32.
- Carney, O., McIntosh, J. & Worth, A. 1996:** The use of Nominal Group Technique in research with community nurses. *Journal of Advanced Nursing* 23, 1024-1029.
- Churchman, C.W. 1948:** *Theory of Experimental Inference*. Macmillan Co., New York.
- Cook, T.D. & Campbell, D.T. 1979:** *Quasi-experimentation*. Design & Analysis Issues for Field Settings. Houghton-Mifflin Company, Boston.
- Cronbach, L.J. 1951:** Coefficient Alpha and the Internal Structure of Tests. *Psychometrika* 16(3) Sept. 297-334.
- Daly, J., Chang, E.M.L. & Bell, P.F. 1996:** Clinical nursing research priorities in Australian critical care: a pilot study. *Journal of Advanced Nursing* 23, 145-151.
- Denzin, N.K. 1988:** Triangulation. Teoksessa Keeves J.P. (toim.) 1988: *Educational research, methodology, and measurement*. An International Handbook, 511-513.
- Duffield, C. 1988:** The Delfi technique. *The Australian Journal of Advanced Nursing* 6(2), 41-45.
- Duffield, C. 1993:** The Delphi technique: a comparison of results obtained using two expert panels. *International Journal on Nursing Studies* 30(3), 227-237.
- Eriksson, J.-T. 1996:** Ihminen kompleksisessa tietoyhteiskunnassa. *Futura* 4/1996, 4-7.
- French, P., Anderson, J., Burnard, P., Holmes, C., Mashaba, G., Wrong, T. & Bing-hua, Z. 1996:** International comparison of baccalaureate nursing degrees: collaboration in qualitative analysis. *Journal of Advanced Nursing* 23, 594-602.

- Goodman, C.M. 1987:** The Delphi technique: a critic. *Journal of Advanced Nursing* **12**, 729-734.
- Hammersley, M. 1994:** Deconstructing the qualitative-quantitative divide. Teoksessa Brannen, J. (toim.) 1994: *Mixing Methods: qualitative and quantitative research*. 3. painos. Cedric Chivers Ltd., Bristol. 39-56.
- Helmer, O. 1963:** *The Systematic Use of Expert Judgement in Operations Research*. Santa Monica, California. The Rand Corporation, [P]-2795, Sept 1963
- Helmer, O. & Quade, E.S. 1963:** *An Approach to the Study of a Developing Economy by Operational Gaming*. Santa Monica, California. The Rand Corporation, [P]-2718, Mar. 1963
- Knapp, T. R. & Brown J. K. 1995:** Ten Measurement Commandments That Often Should Be Broken. *Research in Nursing & Health* **18**, 465-469.
- Kuusi, O. 1993:** Delphi-tekniikka tulevaisuuden tekemisen välineenä. Teoksessa Vapaavuori, M. (toim.) 1993: *Miten tutkimme tulevaisuutta?* Tulevaisuuden tutkimuksen seura, Acta Futura Fennica no 5. Painatuskeskus, Helsinki. 132-140.
- Lindeman, C.A. 1975:** Delphi survey of priorities in clinical nursing research. *Nursing Research* **24**, 434-441.
- Linstone, H.A. 1978:** The Delphi Technique. Teoksessa Fowles, J. (toim.): *Handbook of Futures Research*. 273-300. Greenwood press, London.
- Linstone, H.A. & Turoff, M. (Eds.) 1975:** *The Delphi Method: Techniques and Applications*. Addison-Wesley, Massachusetts.
- Metsämuuronen, J. 1997:** Asiantuntijoiden mielipiteiden stabiiliuden mittaus tulevaisuustutkimuksessa. *ESR - Työelämän muutosten ja koulutustarpeiden ennakoimien menetelmät käytäntöineen* (toim. K. Mäkelä & K. Heinonen) Osoitteessa <http://www.mol.fi/ennakointi>. Paperiversio on saatavissa kirjoittajalta.
- Millett, S.M. & Honton, E.J. 1991:** *A Managers Guide to Technology Forecasting and Strategy Analysis Methods*. Battelle Press, Columbus.
- Nunnally, J.C. 1978:** *Psychometric Theory*. 2nd edition. McGraw-Hill, New York.
- Nunnally, J. C. & Bernstein, I. H. 1994:** *Psychometric Theory*. 3rd edition McGraw-Hill, New York.
- Procter, S. & Hunt, M. 1994:** Using the Delphi survey technique to develop a professional definition of nursing for analysing nursing workload. *Journal of Advanced Nursing* **19**, 1003-1014.
- Reid, N. 1988:** The Delphi technique: its contribution to the evaluation of professional practice. Teoksessa Ellis, R. (ed.) 1988: *Professional Competence and Quality Assurance in the Caring Professions*. Lontoo, Chapman and Hall.
- Sackman, H. 1974:** *Delphi Critique: Expert Opinion, Forecasting and Group Process*. Lexington, MA: D.C. Heath.
- Spivey, B.E. 1971:** A technique to determine curriculum content. *Journal of Medical Education* **46**, 269-271.
- Standards for Educational and Psychological Testing 1985:** AERA, APA, NCME.
- Tanner, C.A. & Lindeman C.A. 1987:** Research in Nursing Education: Assumptions and Priorities. *Journal of Nursing Education* **26**(2) Feb., 50-59.
- Tarkkonen, L. 1987:** *On Reliability of Composite Scales*. An essay on the measurement and the properties of the coefficients of reliability – an unified approach. Tilastotieteellisiä tutkimuksia 7. Finnish Statistical Society, Helsinki.
- Turoff, M. & Hiltz, S.R. 1996:** Computer-Based Delphi Processes. Teoksessa Adler, M. & Ziglio, E. 1996 (toim.): *Gazing into the Oracle. The Delphi Method and its Application to Social Policy and Public Health*. 56-85.
- Uljens, M. 1991:** Phenomenography – a qualitative approach in educational research. Teoksessa Merenheimo & Syrjälä, L. (toim.) *Kasvatustutkimuksen laadullisia lähestymistapoja*. Oulun yliopisto, Opetusmonisteita ja selosteita 39/1991, 80-107.
- Ventura, M.R. & Waligora-Serafin, B. 1981:** Setting Priorities for Nursing Research. *The Journal of Nursing Administration* **11**(6) June, 30-34.
- Whitehead, A.N. 1925:** *Science and the Modern World*. Macmillan Co., New York.
- Whitman, N.I. 1990:** The Delphi Technique as an Alternative for Committee Meetings. *Journal of Nursing Education* **29**(8), 377-379.
- Williams, P.L. & Webb, C. 1994:** The Delphi technique: a methodological discussion. *Journal of Advanced Nursing* **19**, 180-186.

Ziglio, E. 1996: The Delphi method and its Contribution to Decision-Making. Teoksessa Adler, M. & Ziglio, E. (toim.): *Gazing into the Oracle. The Delphi method and its Application to Social Policy and Public Health*. Jessica Kingsley Publishers, London. 3-33.