

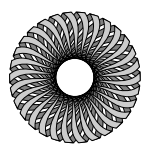
Tom Sefton, Sarah Byford, David McDaid, John Hills ja Martin Knapp

# Taloudellinen arviointi sosiaaalialalla

Englanninkielinen alkuteos

Contemporary research issues

Making the most of it -  
Economic evaluation in the social welfare field



**Hyvät käytännöt**

**Menetelmä-käsikirja**

Julkaisua voi tilata maksutta osoitteesta:  
STAKES  
Sosiaali- ja terveysalan tutkimus- ja kehittämiskeskus  
FinSoc/Hilkka Nyström  
PL 220, 00531 Helsinki

Puhelin: (09) 39671 ja (09) 3967 2313  
Faksi: (09) 3967 2278  
Sähköposti: [hilkka.nystrom@stakes.fi](mailto:hilkka.nystrom@stakes.fi)

© Stakes 2004  
Hyvät käytännöt -logo © Rainer Tenhunen 2004

# SISÄLLYS

<b>LUKIJALLE .....</b>	<b>7</b>
<b>OSA I: TAUSTAA.....</b>	<b>9</b>
<b>1 Johdanto .....</b>	<b>9</b>
Kirjan tarkoitus .....	9
Soveltamisala .....	9
Valittu näkökulma.....	10
Lähestymistapa.....	11
Raportin rakenne .....	11
<b>2 Mitä on taloudellinen arviointi?.....</b>	<b>13</b>
Tarkoitus .....	13
Määritelmä .....	13
Mitä taloudellinen arviointi ei ole.....	14
Taloudellisen arvioinnin periaatteet.....	15
Taloudellisen arvioinnin tyypit .....	15
Rakennuspalikat .....	18
<b>3 Taloudellisen arvioinnin nykyinen tila.....</b>	<b>23</b>
Systemaattisen kirjallisuuskatsauksen menetelmä.....	23
Tulokset.....	23
Tutkimusten laatu.....	26
Johtopäätökset.....	26
<b>OSA II: KÄSITTEELLISIÄ KYSYMYKSIÄ.....</b>	<b>27</b>
<b>4 Taloudellisen arvioinnin haasteita.....</b>	<b>27</b>
Miksi taloudellista arviointia on niin vähän?.....	27
Ohjelmien luonne.....	27
Vaikutusten monimutkaisuus .....	30
Interventioiden monimutkaisuus.....	30
Käyttäjien rooli .....	31
Näkökulma ja soveltamisala .....	31
Analyysiyksikkö.....	32
Aikaulottuvuus .....	32
<b>5 Arvioinnin vaihtoehtoiset näkökulmat.....</b>	<b>35</b>
Erilaisia koulukuntia .....	35
Perinteisten lähestymistapojen vahvuuksia ja heikkouksia .....	37
Ajattelumalli monimutkaisille interventioille .....	38
<b>OSA III: SOVELTAMINEN KÄYTÄNTÖÖN.....</b>	<b>42</b>
<b>6 Tarkoituksen ja tavoitteiden selkeyttäminen .....</b>	<b>42</b>
Arvioinnin keskeisiä näkökohtia.....	42

Arvioinnin tarkoituksen selkiyttäminen.....	45
Arvioitavan ohjelman tavoitteiden selkiyttäminen .....	46
<b>7 Vaikutusten ja kustannusten arviointi .....</b>	<b>50</b>
Vaikutusten valitseminen.....	50
Vaikutusten mittaaminen .....	55
Kustannusten arviointi .....	62
Johtopäätökset.....	66
<b>8 Tutkimusasetelma ja menetelmät .....</b>	<b>68</b>
Tutkimusasetelma .....	68
Täydentäviä tutkimusmenetelmiä .....	85
Muita tutkimusasetelmaan liittyviä kysymyksiä.....	93
<b>9 Tulosten analysointi ja tulkinta .....</b>	<b>98</b>
Tulosten analysointi .....	98
Miten ja miksi ohjelmat toimivat .....	101
Tulosten tulkinta ja esittäminen .....	106
<b>10 Arviointiprosessin hallinta.....</b>	<b>110</b>
Keskeisten tahojen osallistuminen.....	110
Yhteistyö toisten tieteenalojen kanssa .....	111
Rajoittavia tekijöitä arvioinnissa.....	113
<b>11 Yhteenveto ja johtopäätökset .....</b>	<b>118</b>
Yhteenveto .....	118
Johtopäätökset.....	120
<b>LÄHTEET .....</b>	<b>121</b>

## KUVIOT

<b>Kuvio 1.</b> Taloudellisen arvioinnin rakennuspalikat .....	22
<b>Kuvio 2.</b> Taloudelliset arviointitutkimukset .....	24
<b>Kuvio 3.</b> Tutkimusten määrä sosiaalisen hyvinvoinnin eri alueilla (1996 - 2000).....	24
<b>Kuvio 4.</b> Prosenttiosuudet tutkimustyypeittäin (1991-96).....	25
<b>Kuvio 5.</b> Tutkimusten maakohtaiset prosenttiosuudet (1991-96).....	25
<b>Kuvio 6.</b> Sosiaalialan interventioiden arvioitavuus .....	29
<b>Kuvio 7.</b> Ajattelumalli monimuotoisten interventioiden arvioinnille.....	40
<b>Kuvio 8.</b> Taloudellisen arvioinnin keskeiset vaiheet .....	44
<b>Kuvio 9.</b> Kustannusvaikuttavuuden arvioiminen.....	99

## TARKENTEET

<b>Tarkenne 1.</b> Taloudellisen arvioinnin menetelmät .....	15
<b>Tarkenne 2.</b> Tapaustutkimusten kuvaukset.....	19
<b>Tarkenne 3.</b> Tavoitteiden selkiyttäminen.....	47
<b>Tarkenne 4.</b> Vaikutusten valinta .....	51
<b>Tarkenne 5.</b> Useiden vaikutusten käyttö tutkimuksissa.....	54
<b>Tarkenne 6.</b> Vaikutusmittareiden käyttö “pehmeiden” vaikutusten mittaamiseen.....	57
<b>Tarkenne 7.</b> Kustannusten määrittely.....	66
<b>Tarkenne 8.</b> Esimerkkejä satunnaisotokseen perustuvista koeasetelmista .	70
<b>Tarkenne 9.</b> Esimerkkejä kvasi-kokeellisista tutkimuksista .....	75
<b>Tarkenne 10.</b> Satunnaisotosta muistuttavat asetelmat .....	77
<b>Tarkenne 11.</b> Ongelmia vertailuryhmän muodostamisessa .....	80
<b>Tarkenne 12.</b> Mahdollisten virhelähteiden huomioon ottaminen .....	81
<b>Tarkenne 13.</b> Esimerkki vertailtavaa ryhmää käyttävästä taloudellisesta arvioinnista .....	84
<b>Tarkenne 14.</b> Esimerkki mallintamisesta.....	86
<b>Tarkenne 15.</b> Laadullisia tutkimusmenetelmiä.....	89
<b>Tarkenne 16.</b> Esimerkkejä laadullisten tutkimusmenetelmien käytöstä.....	90
<b>Tarkenne 17.</b> Esimerkki meta-analyysistä .....	92
<b>Tarkenne 18.</b> Muita arvioinnin lähestymistapoja .....	103
<b>Tarkenne 19.</b> Esimerkki teoriaperusteisesta taloudellisesta arvioinnista .	105

## TAULUKOT

<b>Taulukko 1.</b> Sosiaalialan interventioiden arvioitavuuden kriteerit .....	34
<b>Taulukko 2.</b> Vaiheistettu lähestymistapa arviointiin (Safe In The City - tapaustutkimus) .....	41
<b>Taulukko 3.</b> Taloudellisen arvioinnin kustannuksiin vaikuttavia tekijöitä .....	116



## LUKIJALLE

Taloudellista arviointia on tehty sosiaalialalla suhteellisen vähän. Monet sosiaaliset interventiot ovat monimutkaisia, mikä tekee kustannusten ja vaikutusten systemaattisen arvioinnin vaikeaksi. Sosiaalialalla työskentelevät saattavat myös suhtautua taloustieteilijöihin ja heidän suosimiinsa tutkimusmenetelmiin epäillen. Myöskään päätöksentekijät eivät ole teettäneet perusteellisia taloudellisia arviointeja, vaikka resurssien niukkuus onkin sekä päätöksentekijöiden että sosiaalialalla työskentelevien arkipäivää. Usein tyydytään pelkkien kustannusten tarkasteluun, kun varsinainen taloudellinen arviointi vertailee sekä kustannuksia että vaikutuksia. Pelkkien kustannusten vertailu vain lisää sosiaalialalla työskentelevien kriittisyyttä taloudellista näkökulmaa kohtaan.

Tämä kirja on suunnattu niille, jotka pohtivat taloudellisen arvioinnin sovellettavuutta sosiaalialalle. Tavoitteena on lisätä taloudellisen arvioinnin ymmärrystä ja käyttöä ja nostaa esille taloudellisen arvioinnin keskeisiä kysymyksiä ja ongelmia sosiaalialalle sovellettaessa. Tällaisia ovat muun muassa kustannusten ja vaikutusten tunnistaminen ja mittaaminen sekä prosessin merkitys. Aiheen käsittelytapa soveltuu sosiaalitaloustieteen perusopintoihin, mutta auttaa jäsentämään taloudellisen arvioinnin keskeisiä teemoja myös muissa yhteyksissä. Monet käsiteltävät aiheet soveltuvat arviointiin yleisemminkin.

Englanninkielisen alkuteoksen ”Making the most of it – Economic evaluation in the social welfare field” on julkaissut brittiläinen Joseph Rowntree -säätiö. Säätiö tukee sosiaalipolitiikan, hoivan ja asumisen tutkimusta sekä kehityshankkeita, joiden toivotaan edistävän parempia käytäntöjä ja politiikkaa. Alkuteos on tuotos säätiön rahoittamasta tutkimushankkeesta, jonka toteuttivat London School of Economics -yliopiston yksiköt ESRC Research Centre for Analysis of Social Exclusion ja Personal Social Services Research sekä Psykiatrian tutkimuslaitoksen Centre for the Economics of Mental Health. Projektin tarkoituksena oli kehittää sosiaalialan arviointiin soveltuvaa taloudellista arviointia. Koska projektin tuloksena syntyneeseen kirjaan on koottu sosiaalialan taloudellisen arvioinnin keskeisiä kysymyksiä, kirja palvelee erinomaisesti myös aiheesta kiinnostuneita suomalaisia lukijoita. Se, että kirja on kirjoitettu brittiläisessä kontekstissa, ei muuta perustavaa laatua olevia kysymyksiä kustannusten ja vaikutusten tunnistamisesta ja mittaamisesta. Sama koskee myös kysymystä prosessin merkityksestä niin tavoitteiltaan kuin vaikutusmekanismeiltaan monimutkaisen sosiaalialan arvioinnissa.

Esitämme suuret kiitokset Joseph Rowntree -säätiölle oikeudesta kääntää kirja ja julkaista se Suomessa. Erityisesti kiitämme alkuperäisteoksen kirjoittajia Tom Seftonia, Sarah Byfordia, David McDaidia, John Hillsiä ja Martin Knappia. Sekä käännöstyön kohteena oleva kirja että edellä mainitun tutkimusryhmän muut julkaisut edistävät taloudellisen näkökulman huomiointia ottamista sosiaalialalle soveltuvalla tavalla myös Suomessa.

Käännöstyö on tehty Diakonia-ammattikorkeakoulun Sosiaalitalouden tutkimuskeskuksessa osana Kuopion yliopiston kanssa yhteistyössä toteutettavaa ESR-rahoitteista hanketta. Työn ovat tehneet dosentti Olli Pusa (YTT), dosentti Keijo Piirainen (YTT) ja tutkija Aija Kettunen (VTT). Tekijät kiittävät Diakonia-ammattikorkeakoulun tutkimuspäällikkö Sakari Kainulaista, sillä ilman hänen innostavaa suhtautumistaan käännöstyöhön tuskin olisi ryhdytty. Kiitokset kuuluvat myös tutkimusassistentti Sari Kaasiselle teknisestä avusta ja lehtori Jukka Määttäselle kieliasuun liittyvistä kommentteista.

Ansio siitä, että käännös julkaistaan Suomessa, kuuluu Sosiaali- ja terveysalan tutkimus- ja kehittämiskeskuksen Sosiaalipalvelujen evaluaatioryhmälle (FinSoc). Kiitämme erityisesti tutkimuspäällikkö Ilse Julkusta ja tutkija Tuija Lindqvistiä. Myös julkaisun ulkoasu on viimeistelty Stakesissa, kiitokset Mira Koivusillalle.

### Suomentajat





# OSA I: TAUSTAA

## 1 Johdanto

### Kirjan tarkoitus

Tämä raportti on päätulos kaksivuotisesta projektista, jonka tavoitteena oli lisätä taloudellisen arvioinnin ymmärrystä ja käyttöä sosiaalialalla. Taloudellisen arvioinnin tarvetta ei ole aina tiedostettu. Silloinkin kun se on tiedostettu, eroavat arviointitutkimuksen teettäjien odotukset usein suuresti siitä, mitä todellisuudessa taloudellisella arvioinnilla voidaan saada aikaan. Tämä raportti ei pyri olemaan ohje käytetyistä menetelmistä, vaan se pyrkii nostamaan esille joitakin keskeisiä kysymyksiä niille, jotka arvioivat sosiaalista hyvinvointiin liittyviä ohjelmia. Raportissa kuvataan myös sitä, miten esille nostettuihin kysymyksiin voi tarttua.

Raportti on suunnattu sosiaalialan ohjelmien evaluointiin osallistuville talous- ja ei-taloustieteilijöille, jotka ovat turhautuneita hyvää taloudellista arviointikäytäntöä haittaaviin ongelmiin. Oletuksena on, että raportin lukijoilla on perustiedot taloudellisesta arvioinnista, ja tämän vuoksi avainkäsitteet määritellään vain lyhyesti. Olemme pitäneet taloudellisen ammattislangin mahdollisimman vähäisenä. Raportti keskittyy lähestymistapojen soveltamiseen sosiaalialan kontekstissa. Täydentävä raportti (Osa II) tarjoaa yksityiskohtaisempia teknisiä neuvoja siitä, kuinka taloudellisen arvioinnin periaatteita voidaan soveltaa erilaisissa tilanteissa (Byford ym., 2003). Monet raportissa käsiteltävät aiheet eivät liity vain taloudelliseen arviointiin, vaan niillä on merkitystä myös arvioinnista yleisemminkin kiinnostuneille.

### Soveltamisala

Taloudellinen arviointi määritellään systemaattiseksi yritykseksi tunnistaa, mitata ja vertailla vaihtoehtoisten interventioiden kustannuksia ja vaikutuksia. Kirja keskittyy käytännön toiminnan taloudelliseen arviointiin eikä laajempiin tai strategisempiin toimintapoliittisiin kysymyksiin.

Määritelmämme sosiaalialasta käsittää hoivan, varhaisen puuttumisen ohjelmat, asumiseen liittyvät kysymykset, kuntoutuksen, yhteisöjen kehittämisen, lasten, nuorten ja perheiden kanssa tehtävän työn ja työssä selviytymisen tukemiseen liittyvän toiminnan sekä vastaavat toimialueet. Tämä laaja-alainen määritelmä sopii yhteen Joseph Rowntree -säätiön tutkimusprioriteettien kanssa (Joseph Rowntree Foundation 2001 a). Koska terveydenhuoltoa koskeva taloudellinen arviointi on jo kehittynyt hyvin, olemme jättäneet sen tarkastelun ulkopuolelle. Emme myöskään käsittele tavanomaista koulutusta, ympäristökysymyksiä tai kuljetustoimintaa, vaikka monet taloudellista arviointia koskevat kysymykset ovatkin yhteisiä. Tarkastelemissamme toiminnoissa keskitymme monimutkaisempiin interventioihin,

joissa taloudellinen arviointi on kehittymättömintä ja jossa panoksemme voisi olla arvokkaampaa.

## **Valittu näkökulma**

Aloitamme pohtimalla, miksi sosiaalialalla on suhteellisesti ottaen vähän taloudellista arviointia. Tarkastelemme syitä siihen, miksi sosiaalialan ohjelmien taloudellinen arviointi on niin haastavaa. Koska taloudellinen arviointi on terveydenhuollossa yleisempää, tutkimme myös terveydenhuollossa käytettyjä menetelmiä selvittääksemme, voidaanko samanlaisia lähestymistapoja siirtää muun hyvinvoinnin alueelle. Tämä lähtökohta selittää sen, miksi kokeelliset lähestymistavat korostuvat. Näihin lähestymistapoihin luetaan tässä satunnaisotokseen perustuvat kontrolloidut koeasetelmat ja muut pragmaattisemmat koeasetelmat. Uskomme, että nämä ovat yleensä ottaen sopivampia sosiaalialalle kuin ekonometriset mallit tai mallintamistekniikat, jotka ovat hallitsevia muilla arvioinnin kohdealueilla kuten kuljetustoiminnassa. Joka tapauksessa on tärkeää pohtia, kuinka kokeellisia lähestymistapoja olisi muokattava sosiaalialan kontekstissa, koska sosiaalialan interventiot ovat usein hyvin erilaisia kuin terveystaloustieteilijöiden arvioinnin kohteet.

Emme usko, että laajoihin satunnaistettuihin kontrolliasetelmiin perustuvat kvantitatiiviset analyysit ovat ainoita tekemisen arvoisia tutkimuksia, tai että taloustieteillä ei olisi tarjottavaa ilman edellisen kaltaisia analyysejä. Mutta emme myöskään usko, että täsmälliset kvantitatiiviset arvioinnin lähestymistavat olisivat yleisesti ottaen mahdottomia tai sopimattomia sosiaalialalle. Katsomme, että taloustieteilijöiden pitää yrittää rakentaa systemaattisia kustannusten ja vaikutusten arviointeja ja että tutkimusasetelma on tärkeä, vaikkakaan ei ainoa huomioon otettava asia. Muut asiat kuten ohjelman tavoitteiden selkeyttäminen ja oikeiden vaikutusten valinta ovat myös keskeisiä, kriittisiä ja joskus laiminlyötyjä taloudellisen arvioinnin kysymyksiä.

Näkökulmamme taloudelliseen arviointiin on käytännöllinen ja se tunnistaa monimutkaisten sosiaalialan interventioiden arviointiin liittyvät käytännölliset ja metodologiset ongelmat. Tämän vuoksi uskomme, että täydentävien tutkimusmetodien käyttö voisi parantaa tutkimusten validiteettia (pätevyyttä) ja auttaa tulosten tulkinnassa. Raportti on lähestymistavaltaan pikemminkin moninainen kuin tiettyä oppisuuntaa noudattava nostaessaan esiin useita arvioinnin näkökulmia mukaan lukien ne, jotka ovat tuttuja muualta kuin taloustieteellisestä kirjallisuudesta.

Kirjoittajille selvisi jo varhaisessa vaiheessa, että kysymyksessä on kaksisuuntainen oppimisprosessi. Taloustieteilijöillä on tärkeä osuus sosiaalialalla tehdyssä arvioinnissa, mutta heillä on myös paljon opittavaa muista arvioinnin lähestymistavoista. Uskomme vahvasti, että monitieteinen näkökulma on paras tapa edetä, vaikka ideaalin saavuttamiseen onkin pitkä matka.

## **Lähestymistapa**

Kävimme läpi arviointikirjallisuutta tunnistaaksemme tärkeimmät käsitteelliset ja metodologiset kysymykset, jotka koskevat taloudellista arviointia ja arviointia yleisemminkin. Keskityimme erityisesti näiden merkitykseen sosiaalialan interventioiden ja etenkin monimutkaisempien ohjelmien arvioinnissa.

Emme kuitenkaan halunneet tämän raportin olevan pelkästään teoreettista pohdiskelua, ja siksi teimme useita tapausarviointeja. Nämä pakottivat meidät kohtaamaan monia taloudellisen arvioinnin esteitä sosiaalisen hyvinvoinnin alueella. Tapaus tutkimukset, joiden yhteenveto on seuraavassa luvussa, kattavat monia hyvinvoinnin alueita: yhteisöjen kehittämisen, nuorten asunnottomuuden ehkäisyn, sijaishuollon ja varallisuuden huolehtia asunnon lämmityksestä. Tapausten avulla havainnollistetaan monia tärkeänä pidettäviä asioita; ne myös antavat meille käsitystä taloudellisen arvioinnin prosessista, jota emme voi saada aikaisempia tutkimusraportteja lukemalla.

Lisäksi teimme sähköisen tiedonhaun, jonka avulla etsimme tieteellisissä aikakauslehdissä julkaistuja tutkimuksia saadaksemme todisteita taloudellisen arvioinnin nykytilasta sosiaalialalla. Tiedonhaun avulla olemme pyrkineet myös yksilöimään interventioiden laajasta kirjosta esimerkkejä hyvistä käytännöistä.

## **Raportin rakenne**

Luvussa kaksi määritellään taloudellinen arviointi ja hahmotellaan sen päätyypit. Siinä tunnistetaan vaikeus toteuttaa täydellinen taloudellinen arviointi monilla sosiaalialan toiminta-alueita ja yksilöidään rakennuspalikat, joita tarvitaan täsmällisen taloudellisen arvioinnin toteuttamisessa. Tämä luku on suunnattu ensisijaisesti muille kuin taloustieteilijöille, ja ne, jotka tuntevat hyvin taloudellisen arvioinnin periaatteet, voivat hypätä luvun yli. Kolmannessa luvussa esitellään tuloksia taloudellisia arviointitutkimuksia koskevista systemaattisesta kirjallisuuskatsauksestamme.

Osa II (Luvut 4 ja 5) kuvaa tiettyjä käsitteellisiä kysymyksiä ja niiden vaikutuksia taloudelliselle arvioinnille. Luvussa 4 tarkastellaan niitä haasteita, joita kohdataan pyrittäessä arvioimaan monimutkaisia sosiaalialan interventioita. Luvussa 5 esitellään joitain arviointia koskevia vaihtoehtoisia näkökulmia, jotka ovat muovanneet raportissa esittämäämme taloudellisen arvioinnin lähestymistapaa.

Kolmannessa raportin osassa (luvut 6-10) keskustellaan taloudellisen arvioinnin käytännön soveltamisesta sosiaalialalle käyttäen esimerkkeinä tapaus tutkimuksiamme ja muita julkaistuja tutkimuksia. Nämä luvut ovat loogisessa järjestyksessä, mutta myös erityisistä arvioinnin näkökohdista kiinnostuneet voivat löytää niistä suositeltavia lähdeviitteitä.

Sen jälkeen, kun luvussa 6 on lyhyesti esitelty taloudellisen arvioinnin keskeisiä kysymyksiä, tarkastellaan arvioinnin erilaisia päämääriä ja niiden vaikutuksia taloudellisen arvioinnin toteuttamiseen. Luvussa korostetaan myös, kuinka tärkeää on arvioinnin kohteena olevan ohjelman tavoitteiden selkeyttäminen perustavanlaatuisena ja olennaisena osana mitä tahansa arviointia.

Luvussa 7 annetaan ohjeita olennaisten kustannusten ja vaikutusten yksilöimiseksi ja arvioimiseksi. Siinä painotetaan, kuinka tarpeellista on tarkasteltavien vaikutusten huolellinen valitseminen ja käydään läpi keinoja vaikutusten mittaamiseksi kun on otettava huomioon myös sellaiset vaikutukset, joita ei luonnostaan ole helppo mitata. Luvussa myös väitetään, että taloudellisen arvioinnin kustannuspuoleen on tarpeen kiinnittää huomiota enemmän kuin usein tapahtuu. Kustannuksiin on otettava laajempi näkökulma ja kustannusanalyysiä tulisi pitää aiempaa olennaisempana osana arviointia.

Luvussa 8 käsitellään taloudellisen arvioinnin päälähestymistapoja ja keskitytään kokeellisiin asetelmiin, joissa vertailuryhmä valitaan satunnaisotoksella tai ilman satunnaistamista. Mitä monimutkaisempia interventiot ovat, sitä todennäköisempää on, että pelkät kokeelliset lähestymistavat eivät yksinään riitä. Sen vuoksi tarkastelemme myös muiden, esimerkiksi kvalitatiivisten menetelmien käyttöä, ja uskomme, että näillä on tärkeä rooli taloudellisessa arvioinnissa. Myös muita asetelmakysymyksiä, kuten otoskokoa ja seuranta tarkastellaan.

Luvussa 9 keskustellaan taloudellisen arvioinnin viimeisistä vaiheista, joissa kustannuksia ja vaikutuksia koskeva tieto analysoidaan ja, milloin mahdollista, yhdistetään arvioksi intervention kustannusvaikuttavuudesta. Tässä luvussa paneudutaan yksityiskohtaisemmin vahvaan, koko raportin läpäisevään teemaan: taloustieteilijöiden velvollisuuteen panna enemmän painoa sen ymmärtämiseen, kuinka ja miksi arvioitavat hankkeet toimivat tai eivät toimi. Luvun viimeinen osa tarjoaa joitain yleisiä suuntaviivoja taloudellisten arvioinnin tulosten esittämiseen.

Luvussa 10 käsitellään erilaisia arviointiprosessiin liittyviä kysymyksiä, jotka joskus jäävät arvioinnin tekijöiltä huomaamatta heidän syventyessään arviointimenetelmiin. Tapaustutkimuksemme ovat olleet erityisen hyödyllisiä tuotaessa esille sitä, miten tärkeää on ottaa arvioitavan hankkeen muita toimijoita mukaan arviointiin arvioinnin eri vaiheissa. Hyödyllisiä ne ovat olleet myös tarkasteltaessa sitä, miten tärkeää on työskennellä yhdessä muita tieteenaloja edustavien arvioijien kanssa, ja miten merkityksellisiä rahalliset ja muut arviointia rajoittavia tekijät ovat. Nämä arviointiprosessiin liittyvät kysymykset tulevat esille raportin lopussa, koska tekstissä viitataan aiempien lukujen keskusteluihin. Kuitenkin, arviointitutkimusta tehtäessä näihin kysymyksiin on syytä tarttua varhaisessa vaiheessa.

Luvussa 11 tehdään lopuksi johtopäätöksiä tästä raportista, jossa kehitettiin taloudellisen arvioinnin käyttöä sosiaalialalla kentällä.

## 2 Mitä on taloudellinen arviointi?

### Tarkoitus

Taloudellisen arvioinnin tarkoituksena on informoida päätöksentekoa rajallisten resurssien parhaasta käytöstä. Resurssien käyttö tiettyyn interventioon merkitsee että muihin tarkoituksiin jää käytettäväksi vähemmän resursseja. Palvelu voi olla tehokasta palvelun käyttäjien tarpeiden tyydyttämisessä, mutta se ei ole kustannustehokasta, jos samanlaiset vaikutukset voitaisiin saavuttaa vähemmällä kustannuksilla, tai jos paremmat vaikutukset saavutettaisiin samoilla kustannuksilla. Suurin osa arvioinnista keskittyy vaikutuksiin. Taloudellinen arviointi ottaa huomioon sekä intervention kustannukset että vaikutukset ja pyrkimyksenä on tunnistaa tehokkain tapa toiminnalle asetettujen tavoitteiden saavuttamiseksi.

Koska kaikkiin toimintaa koskeviin päätöksiin liittyy poliittisista, eettisistä ja käytännöllistä harkintaa, on taloudellisella arvioinnilla potentiaalisesti arvokas rooli tässä prosessissa. Tällaisia päätöksiä on tehtävä kaikilla tasoilla. Kysymys voi olla menojen jakamisesta ohjelma-alueiden kesken, uuden järjestelmän toimeenpanosta tai sen päättämisestä, minkä tason palvelua yksittäiselle palvelun käyttäjille tarjotaan. Taloudellisen arvioinnin tarkoituksena on mahdollistaa näiden päätösten tekeminen aiempaa rationaalisemmin. Tämä on yhdenmukaista sen kanssa, että sekä hallinnon sisällä että sen ulkopuolella on alettu entistä enemmän tähdentää näyttöön perustuvaa toimintaa (ks. esim. Centre for Management and Policy Studies, 2001).

### Määritelmä

Tavallisimman määritelmän mukaan, jota myös me käytämme, taloudellinen arviointi on systemaattinen yritys tunnistaa, mitata ja verrata vaihtoehtoisten interventioiden kustannuksia ja vaikutuksia (Drummond ym. 1997; HM Treasury 1997). Määritelmästä voidaan johtaa taloudellisen arvioinnin perusvaatimukset:

- keskittyminen tiettyyn interventioon
- kustannusten ja vaikutusten systemaattinen arviointi
- jonkinlainen arviointikriteeri tai vertailuasetelma
- yritys yhdistää kustannukset ja vaikutukset lopullisessa analyysissä.

Taloustieteilijöillä on tärkeää annettavaa kaikkiin arviointiprosessin vaiheisiin. He muun muassa:

- auttavat tavoitteiden selkiyttämisessä ja niiden muuntamisessa mitattaviksi vaikutuksiksi
- tekevät selvän eron prosessin, panosten, tuotosten ja vaikutusten kesken
- rohkaisevat systemaattisempaan ja täsmällisempään kustannusten ja vaikutusten arviointiin, painottaen erityisesti tilastollisesti pätevien tulosten aikaan saamista

- tuovat esille tarvetta tarkastella myös sitä, mitä olisi tapahtunut, jos arvioinnin kohteena olevaa interventiota ei olisi toteutettu
- tukevat yhteiskunnallisen tai moninaisen näkökulman omaksumista, joka takaa ohjelman vaikutusten kokonaisvaltaisen arvioinnin
- auttavat selittämään miten tavoiteohjelmat tai muut ohjelmat toimivat, erityisesti ne, jotka ovat luonteeltaan taloudellisia tai joissa on taloudellinen elementti (mm. verokannustimet, joilla rohkaistaan ihmisiä työllistymään).

Näin ollen vaikutusten arviointi ei mielestämme ole taloustieteen toiminta-alueen ulkopuolella eikä taloustieteilijöiden roolia tule rajoittaa vaikutusten arvottamiseen. Väitämme, että taloustieteilijöiden pitäisi olla mukana auttamassa vaikutusten arvioinnissa ja että hyvä taloudellinen arviointi voidaan toteuttaa ilman, että vaikutukset välttämättä arvotetaan rahamääräisinä.

## **Mitä taloudellinen arviointi ei ole**

Taloudellisessa arvioinnissa ei ole kysymys ohjelman tai siitä vastuussa olevan organisaation taloudellisesta elinkelpoisuudesta. On mahdollista, että organisaatio, joka kamppailee selviytyäkseen taloudellisesti, toteuttaa kustannus-tehokasta ohjelmaa, kun vain käyttäjille ja yhteiskunnalle koituvat hyödyt laajasti otetaan huomioon.

Toinen väärinkäsitys on, että taloudellisessa arvioinnissa on kysymys vain kustannuksista. Tämän väärinkäsityksen johdosta taloustieteilijöiden rooli on usein rajoittunut erillisten kustannustutkimusten tekemiseen. Kustannuksia koskeva informaatio on keskeinen osa taloudellista arviointia. Kustannustutkimukset voivat olla myös hyödyllisiä sinänsä (Holtermann 1998), esimerkiksi osana budjetointia tai informoitaessa laajaa keskustelua toimintaohjelmasta (ks. esim. Holtermann 1995). Kuitenkaan ne eivät kerro, onko jokin tietty ohjelma kannattava; tämän vuoksi on välttämätöntä katsoa samanaikaisesti sekä kustannuksia että vaikutuksia.

Pyrimme myös erottamaan taloudellisen arvioinnin ja muuntyyppiset taloudelliset analyysit, vaikka käytännössä ero voi ollakin melko epäselvä. Mikrotaloustiede on kiinnostunut ihmisten ja organisaatioiden taloudellisesta käyttäytymisestä: miten nämä tekevät tiettyjen hyödykkeiden tai palvelujen kuluttamista tai tuottamista koskevia valintoja, ja millaisia vaikutuksia näillä yksilöllisillä päätöksillä on. Taloustieteilijät ovat esimerkiksi kehittäneet teorioita ja malleja, jotka auttavat ymmärtämään yksilön työnetsintään liittyvää käyttäytymistä. Tällaisella taloudellisella analyysillä on vaikutuksia tietyn tyyppisten interventioiden kustannus-vaikuttavuuteen ja se voi tämän johdosta tukea arviointia (esim. tuen piiristä työhön -suunnitelma). Emme sisällytä sitä kuitenkaan taloudellisen arvioinnin määritelmään, koska se ei keskity tietyn intervention arviointiin.

## Taloudellisen arvioinnin periaatteet

Taloudellinen arviointi tukeutuu eräisiin hyvinvoinnin taloustieteen perusperiaatteisiin:

- *Vertailtavuus*: Koska taloustiede on kiinnostunut niukkojen resurssien parhaasta käytöstä, taloudellisessa analyysissä on vertailtava vaihtoehtoisia toimintatapoja.
- *Yhteiskunnallinen näkökulma*: Resurssien jakoa koskevien päätösten olisi perustuttava ohjelman kaikkien mahdollisten kustannusten ja hyötyjen arviointiin kaikille yhteiskunnan sektoreille (Johannesson 1995) sekä niiden jakovaikutusten arviointiin. (Voi kuitenkin olla olosuhteita, joissa kapeampi näkökulma on tarkoituksenmukainen, esimerkiksi pyrittäessä ymmärtämään, miten kannustimet vaikuttavat tiettyyn toimintaan.)
- *Lopulliset vaikutukset*: Hyvinvoinnin taloustieteen tavoite on yhteiskunnan hyvinvoinnin maksimoiminen. Tämän vuoksi taloustieteilijät pitävät parempana keskittymistä lopullisiin vaikutuksiin (ts. muutokseen elämänlaadussa) kuin välittäviin tuotoksiin (esim. parempiin kotipalveluihin) tai prosessia kuvaaviin mittareihin (esim. useampiin asukaskonsultaatioihin).
- *Mieltymyksiin tukeutuminen*: Kustannus-hyöty -analyysi (ja hyvinvoinnin taloustiede yleisemminkin) määrittää hyvinvoinnin sen perusteella, miten kuluttajan mieltymykset tulevat tyydytetyiksi (Hausman ja McPherson 1994). Näin ollen taloustieteilijät panevat suuren painon yksilöiden preferensseille arvioitaessa erilaisten vaikutusten arvoa.

## Taloudellisen arvioinnin tyypit

On olemassa erilaisia taloudellisen arvioinnin tyyppejä. Vaikka ne kaikki pyrkivät systemaattisesti arvioimaan asiaankuuluvia kustannuksia ja vaikutuksia, ne eroavat toisistaan suhtautumisessaan vaikutusten mittaamiseen. Taloudellisen arvioinnin päätyypit on kuvattu tarkenteessa 1.

### Tarkenne 1. Taloudellisen arvioinnin menetelmät

#### Kustannusten minimointi -analyysi

Tämä on yksikertaisin taloudellisen arvioinnin muoto, mutta sitä pitäisi käyttää vain, kun on vahvaa näyttöä siitä, että tarkastelun kohteena olevat interventiot ovat yhtä vaikuttavia. Tällaisessa arvioinnissa vertaillaan vain kustannuksia, ja pyrkimyksenä on tunnistaa vähiten kustannuksia aiheuttava vaihtoehto.

### **Kustannus-vaikuttavuus -analyysi**

Kustannus-vaikuttavuus -analyysi (cost-effectiveness analysis CEA) on tavallisin taloudellisen arvioinnin lähestymistapa terveydenhuollossa. Siinä hyötyjä mitataan 'luonnollisissa' yksiköissä kuten saavutettuina elinvuosina, toiminnanvajavuudesta vapaina päivinä tai masennuksen tasoina. Sosiaalialan tutkimuksessa luonnollisiin yksiköihin voi sisältyä esimerkiksi perheen toiminnan paranemisen ja nuorten itsearvostuksen lisääntymisen kaltaisia vaikutuksia. Kahden tai useamman interventiön hyödyt yhdistetään kustakin aiheutuneisiin kustannuksiin, jolloin saadaan kustannukset tuotettua hyöty-yksikköä kohden. Ensisijalle pannaan sellaiset interventiot, joitten kustannukset ovat saavutettua vaikutusyksikköä kohden alhaisimmat.

### **Kustannus-utiliteetti -analyysi**

Kustannus-utiliteetti -analyysi on kustannus-vaikuttavuus -analyysin erityistapaus. Kuten kustannus-vaikuttavuus -analyysissä, siinä annetaan mittari vaikutukselle, mutta vaikutukset mitataan utiliteettina (tai hyvinvointina). Yksi esimerkki utiliteettimittarista on laatu painotettujen elinvuosien mittaaminen (Quality Adjusted Life Year, QALY). Siinä käytetään painottamista, jonka avulla otetaan huomioon eroja koetussa elämänlaadussa, kun terveydentia on erilainen. Painoarvot perustuvat yleensä tavallisen väestön mieltymyksiä kartoittaviin kyselyihin.

### **Kustannus-seuraus -analyysi**

Kustannus-seuraus -analyysi poikkeaa kustannus-vaikuttavuus -analyysistä siltä osin, että vaikuttavuusmittareita on enemmän kuin yksi. Joissakin interventioissa olisi epätarkoituksenmukaista keskittyä yhteen vaikutukseen, koska tällöin muut tärkeät vaikutukset jäisivät analyysistä pois. Kustannus-seuraus -tutkimukset esittävät intervention vaikutukset ilman pyrkimystä yhdistää niitä yhteen vaikuttavuuden mittariin. Kustannus-seuraus -analyysi ei mahdollista erilaisten interventioiden asettamista tehokkuuden mukaiseen järjestykseen. Tämä on mahdollista ainoastaan, jos joku interventio ei tule muita kalliimmaksi mutta on kaikkien mitattujen vaikutusten osalta muita parempi. Sen sijaan kustannus-seuraus -analyysi antaa päätöksentekijöille mahdollisuuden tehdä tietoisempia ratkaisuja, joissa he tukeutuvat omiin julki lausumattomiin tai julki lausuttuihin, eri vaikutuksia koskeviin painotuksiinsa.

### **Kustannus-hyöty -analyysi**



Kustannus-hyöty -analyysissä vaaditaan, että sekä kustannukset että hyödyt mitataan rahamääräisesti, mikä mahdollistaa suorien vertailujen tekemisen intervention kustannusten ja hyötyjen välillä. Projekti tai ohjelma on kannattava, jos hyödyt nousevat suuremmiksi kuin kustannukset, ja jos se sopii budjetin raameihin. Tämä lähestymistapa mahdollistaa eri sektoreita kuten terveydenhuoltoa, koulutusta tai puolustusta edustavien interventioiden vertaamisen keskenään. Ongelmana kuitenkin on, kuinka arvottaa hyödyt rahallisesti. Menetelmiä on olemassa, mutta ne ovat kiistanalaisia eikä niitä käytetä kovin laajasti, lukuun ottamatta joitakin erityisiä alueita kuten kuljetustoimintaa.

Joissakin tutkimuksissa arvioidaan ohjelman vaikuttavuutta sen perusteella, miten ohjelma onnistuu vähentämään sosiaalista 'paha' (esimerkiksi rikollisuutta). Niissä arvioidaan myös kyseiseen vaikutukseen liittyviä mahdollisia säästöjä (esimerkiksi rikkosoikeusjärjestelmän kustannusten vähenemistä). Vaikka nämä esitellään usein kustannus-hyöty -analyysinä, ne ovat tarkasti ottaen kustannus-säästöt (cost-savings) tai kustannus-kompensaatio (cost-offset) -analyysinä, sillä tällainen hyötyjen arvottamistapa ei ole yhdenmukainen hyvinvoinnin taloustieteen periaatteiden kanssa. Esimerkiksi rikollisen syyttämistä ja vangitsemisesta aiheutuneet kulut ovat eri asia kuin se, mitä ihmiset olisivat valmiita maksamaan rikollisuuden vähentämisestä (Tämän taloustieteilijät pyrkivät saamaan esille määrittäessään saavutettavan parannuksen rahallista arvoa kustannus-hyöty -analyysissä). Lisäksi nämä tutkimukset ovat harvoin kattavia, koska niillä on taipumus jättää huomiotta hyödyt, jotka eivät tuota konkreettisia säästöjä (esimerkiksi rikosten uhrien psykologiset menetykset). Silti kustannusten säästö -analyysit voivat olla hyödyllisiä joissakin tietyissä ohjelmissa, jos oletettujen säästöjen voidaan osoittaa olevan suurempia kuin oletettujen kustannusten.

Kunkin arviointitavan käyttökelpoisuus riippuu vaikutusten luonteesta ja mahdollisuudesta yhdistää hyötyjä tai muuntaa niitä rahamääräisiksi. Kustannus-seuraus -analyysi on tällä hetkellä vallitseva arviointityyppi sosiaalialalla (ks. luku 3). Kustannus-vaikuttavuus -analyysi ei ole niin tavallinen, eikä sen avulla yleensä ole mahdollista yhdistää sosiaalialan ohjelman vaikutuksia yksittäiseksi vaikutusmittariksi. Kustannus-hyöty -analyysi on vaikea, koska monet vaikutukset, erityisesti sosiaalialalla, on hankala arvottaa rahamääräisesti (ks. alaluku 'Tulosten mittaaminen' luvussa 7).

Kun käyttökelpoisuus voi määrätä lähestymistavan valinnan, valittavan lähestymistavan tulisi riippua myös arviointitutkimuksen tavoitteesta. Esimerkiksi jos arvioinnin tavoitteena on verrata kahta ohjelmaa, joilla on samanlaiset päämäärät, hyötyjen muuntamisesta rahamääräisiksi ei ole paljoa etua; tällaisessa tapauksessa kustannus-vaikuttavuus tai kustannus-seuraus -analyysi on todennäköisesti tarkoituksenmukaisempi. Hyvä esimerkki voisi olla varhaisen lukemian oppimisen ohjelmaa arvioiva tutkimus, jossa ryhmäperusteisen ohjelman vaikutusta testiryhmän lukutaitoon verrataan kalliimman, yksilökohtaisen opettamisen vaikutuksiin (Nicolson ym. 1999). Tulokset osoittivat, että ryhmäperusteisella ohjelmalla saavutettu lukutaidon

paraneminen oli pistemäärillä mitattuna verrattavissa olemassa olevan lukutaidon parantamisohjelman (Reading Recovery Programme) pistemääriin, mutta että kustannukset olivat vain kymmenesosa tämän kustannuksista. Näin kustannus-vaikuttavuus -tutkimus oli riittävä kuvaamaan ohjelman (suhteellista) tehokkuutta verrattuna toiseen vakiintuneeseen ohjelmaan.

Jos tavoitteena olisi arvioida, onko uusi ohjelma hyvää julkisten varojen käyttöä verrattuna jonkin toisen sektorin ohjelmiin, niin silloin kustannus-hyöty -analyysi voisi olla suositeltavampi; edellyttäen, että kaikki tärkeät hyödyt voitaisiin arvottaa rahassa. Hyvä esimerkki on Irlannissa toteutetun, energiatehokkuutta lisänneen kymmenvuotisen toimenpideohjelman arviointi (Clinch ja Healy, 2001). Tämän ohjelman arviointitutkimus sisältää yleisluontoisen, mutta kattavan kustannus-hyöty -analyysin. Se sisältää mahdolliset energiasäästöt, ympäristölliset hyödyt sekä asumismukavuutta ja terveyttä koskevat parannukset, ja kaikki nämä on arvoitettu rahassa. Tulokset osoittavat, että tällaisella ohjelmalla on selviä nettohyötyjä. Kustannus-hyöty -analyysistä on etua arvioitaessa tämänkaltaisia ohjelmia, kun ei ole selvää vertailukohdetta.

Toinen tärkeä erottelu voidaan tehdä sen perusteella, toteutetaanko arviointi ennen ohjelman toimeenpanoa (ex ante), samanaikaisesti ohjelman toimeenpanon kanssa (prospektiivisesti) vai jälkikäteen (retrospektiivisesti). Ex ante -arviointi, jota usein kutsutaan asiantuntija-arvioksi (HM Treasury 1997), on hyödyllinen päätettäessä siitä, aloitetaanko uusi ohjelma ja jos aloitetaan, kuinka ohjelmaa suunnataan siten, että siitä saadaan kustannus-vaikuttavampi. Asiantuntija-arvio ei kuitenkaan korvaa toimeenpannun ohjelman arviointia.

Ihanteellista olisi, että kaikki arviointi toteutettaisiin prospektiivisesti, jolloin ohjelmaa voidaan seurata alusta lähtien ja muokata mikäli tarpeellista. Tämä auttaa myös varmistamaan, että tarpeellinen kustannuksia ja vaikutuksia koskeva tieto tulee kerättyä. Retrospektiivinen arviointi on selvästi parempi kuin se, että arviointia ei toteutettaisi ollenkaan, mutta usein jälkikäteen tehtävissä arviointitutkimuksissa joudutaan tyytymään heikkotasoiseen tietoon. Tiedot on usein kerätty muista syistä (esimerkiksi hallinnollisista) ja ne voivat olla epätarkoituksenmukaisia tai epätäydellisiä arvioinnissa käytettäviksi. Tietyt tutkimusasetelmat, kuten satunnaistetut kontrolliasetelmat, voidaan toteuttaa vain prospektiivisesti.

## **Rakennuspalikat**

On tärkeää tunnistaa vaikeudet, jotka liittyvät täydellisen ja tarkan taloudellisen arvioinnin toteuttamiseen monilla sosiaalisen hyvinvoinnin alueilla. Taloudellisen arvioinnin kehittyminen tällä alueella on vähittäinen prosessi, joka saa vauhtia uusista, aiempien pohjalle rakentuvista tutkimuksista. On hyödyllistä ajatella rakennuspalikoita, joita tarvitaan, jotta taloudellinen arviointi kokonaisuudessaan saadaan lopulta toteutettua (ks. kuvio 1). Näiden arvioinnin vaiheiden tunteminen on hyödyllistä myös päätöksentekijöille.

Teimme osana tätä tutkimusprojektia neljä tapaustutkimusta, jotka kohdistavat huomiota taloudellisen arvioinnin eri näkökulmiin. Vaikka mikään näistä tutkimuksista ei edusta täydellistä taloudellista arviointia, ne yrittävät asettaa paikoilleen joitakin taloudellisen arvioinnin rakennuspalikoita. Tarkenteessa 2 on jokainen tapaustutkimus kuvattu lyhyesti. Yksityiskohtaisempia tietoja tapaustutkimuksista saa projektin nettisivuilta ([www.econeval-social-welfare.org](http://www.econeval-social-welfare.org)). Monia arviointiin liittyviä kysymyksiä, joita tapaustutkimukset nostavat esille, seurataan läpi raportin varsinaista tekstiä täydentävissä tarkenteissa.

## **Tarkenne 2. Tapaustutkimusten kuvaukset**

### **Turvapaikka kaupungissa -tapaustutkimus (Safe In The City, SITC)**

SITC-ohjelma rahoittaa ja seuraa joukkoa projekteja kohteenaan asunnottomuusuhan alaiset nuoret. Hanke on ollut käynnissä kolmisen vuotta. Projektin toinen, vuonna 2001 alkanut vaihe on organisoitu kahdeksaksi maantieteelliseksi projektiryhmäksi. Jokaisessa projektiryhmässä on yksi persoonalliseen kehittymiseen liittyvä projekti, yksi työllistymiseen ja osaamiseen liittyvä ja yksi projekti, jossa työskennellään perheiden kanssa. Kun asiakas on tullut ohjelman piiriin, hän voi hyödyntää kaikkia kyseisen projektiryhmän projekteja.

Tekemämme arviointitutkimuksen tarkoituksena oli auttaa SITC:tä kehittämään arviointistrategia nuorten asunnottomuuden ehkäisyohjelmalle. Tutkimus koostui :

- kirjallisuuskatsauksesta, jossa haettiin asunnottomuutta, nuorisotyötä ja perheille suunnattuja palveluja käsittelevien projektien arviointeja
- yhdeksän SITC:n rahoittaman projektin projektipäällikön ja työntekijöiden strukturoidusta haastattelusta
- keskusteluista SITC:n johtohenkilöiden ja alan asiantuntijoiden kanssa.

Tapaustutkimuksessa selvitetään ohjelman toisen toimintavaiheen kokeellisen arvioinnin toteutettavuutta sekä ohjataan arviointiin parhaiten soveltuvien menetelmien käytössä. Myös tutkimusasetelman laadintaa, vaikutusten mittaamista ja kustannusten arviointia tarkastellaan. SITC-ohjelman luonteesta aiheutuu erityisiä haasteita tämältyypiselle arvioinnille. Sen vuoksi SITC tarjosi oivan tilaisuuden tarkastella edellä mainittuja kysymyksiä, joilla on laajaa merkitystä sosiaalialan ohjelmien arvioinneille.

### **Lämmityspolttoaineen puute –tapaustutkimus (Fuel poverty)**

The Home Energy Efficiency Scheme (HEES; kotien energiatehokkuus -suunnitelma; käytetty myös nimeä ”Warm Front” )

käynnistettiin kesäkuussa 2000, ja se suunniteltiin helpottamaan lämmityspolttoaineen puutetta niiden keskuudessa, jotka kärsivät herkimmin kylmän aiheuttamista terveyshaitoista. Pienituloisilla perheillä ja vammaisilla on oikeus saada korkeintaan 1000 punnan avustus yksilöityyn asunnon lämmöneristystoimenpiteeseen. Pienituloiset vanhukset ovat oikeutettuja korkeintaan 2000 punnan avustukseen lämmöneritykseen lisäämiseen tai lämmitysjärjestelmän parantamiseen. Hanke on suunnattu yksityiselle asuntosektorille, josta löytyy lähes 70 % niistä kotitalouksista, jotka kärsivät riittämättömästä lämmityksestä. Vuoteen 2004 mennessä hallitus on suunnannut ohjelmaan 600 miljoonaa puntaa, ja tähän mennessä ohjelman odotetaan hyödyntävän noin 800 000 avun tarpeessa olevaa kotitaloutta.

Tapaustutkimuksen keskeinen tarkoitus oli mallittaa taloustieteellisellä lähestymistavalla HEES:n vaikutusta käyttämällä aineistona laajaa kotitalouskyselyä, joka sisältää tietoja yli 7000 kotitaloudesta. Mallia käytettiin hankkeen kustannusvaikutavuuden tutkimiseen ja erityisesti sen selvittämiseen, miten hanketta pitäisi suunnata, jotta lämmityspolttoaineen puutteesta kärsivien määrä vähenisi mahdollisimman paljon. Tutkimuksen kolme keskeistä osatekijää ovat:

- ekonometrinen analyysi, jolla selvitetään kotien lämpimyyden ja energiatehokkuuden välistä yhteyttä (lämmön tarve -yhtälö)
- syvähaastattelu kahdeksalletoista HEES-avustusta hakenneelle kotitaloudelle eräiden lämmön tarve -yhtälön taustalla olevien oletusten testaamiseksi
- simulaatiomallin rakentaminen ekonometrisen analyysin tulosten perusteella HEES -ohjelman vaikutusten arvioimiseksi

### **Paikallisyhteisöjen kehittäminen –tapaustutkimus (Gatsby)**

Kysymyksessä on viisivuotinen Gatsby -hyväntekeväisyysjärjestön rahoittama hanke, jonka tarkoituksena on innostaa oma-apu projektien syntymistä pienyhteisöissä ympäri Brittein saaria, etupäässä taantuvilla alueilla. Gatsby -ohjelma tarjoaa lyhyitä kurseja pienyhteisöjen asukkaille. Osallistujat työstävät kurseilla omia yhteisön kehittämisprojekti-ideoitaan, joiden toteuttamiseen he voivat hakea korkeintaan 5000 punnan avustusta. Vuodesta 1996 lähtien tällaiselle kurseille on osallistunut noin 700 ryhmää, ja yli 70 avustusta on myönnetty laajalle valikoimalle erilaisia projekteja. Näihin sisältyy muun muassa puistoja, ilta-päiväkerhoja, yhteisökahviloita ja neuvontapisteitä. Ohjelman tunnusomaisia piirteitä ovat koulutuksen voimakas korostaminen, koulutuksen ja avustusohjelman liittyminen toisiinsa ja tuen tarjoaminen avustuksen saaneille.

London School of Economics (LSE) toteutti tähän tapaustutkimukseen liittyvän toimintatutkimuksen, jonka tavoitteena oli kehittää systemaattisempaa vaikutusten arviointikehikkoa. Tapaustutkimus koostui:

- kirjallisuuskatsauksesta, jossa haettiin vallalla olevia yhteisöjen kehittämisen lähestymistapoja
- teoreettisesta tarkastelusta, jossa käsiteltiin yhteisöryhmien toimintakyvyn taustalla olevia tekijöitä kirjallisuuden sekä niiden kokemusten perusteella, joita LSE sai ryhmien kanssa työskennellessään
- kvalitatiivisten (laadullisia) ja kvantitatiivisten (määrällisiä) menetelmien kehittämisestä yhteisöryhmien arvioimiseksi laadittujen kriteereiden perusteella
- pieneen yhteisöryhmäjoukkoon kohdistuvan lähestymistavan testaamisesta.

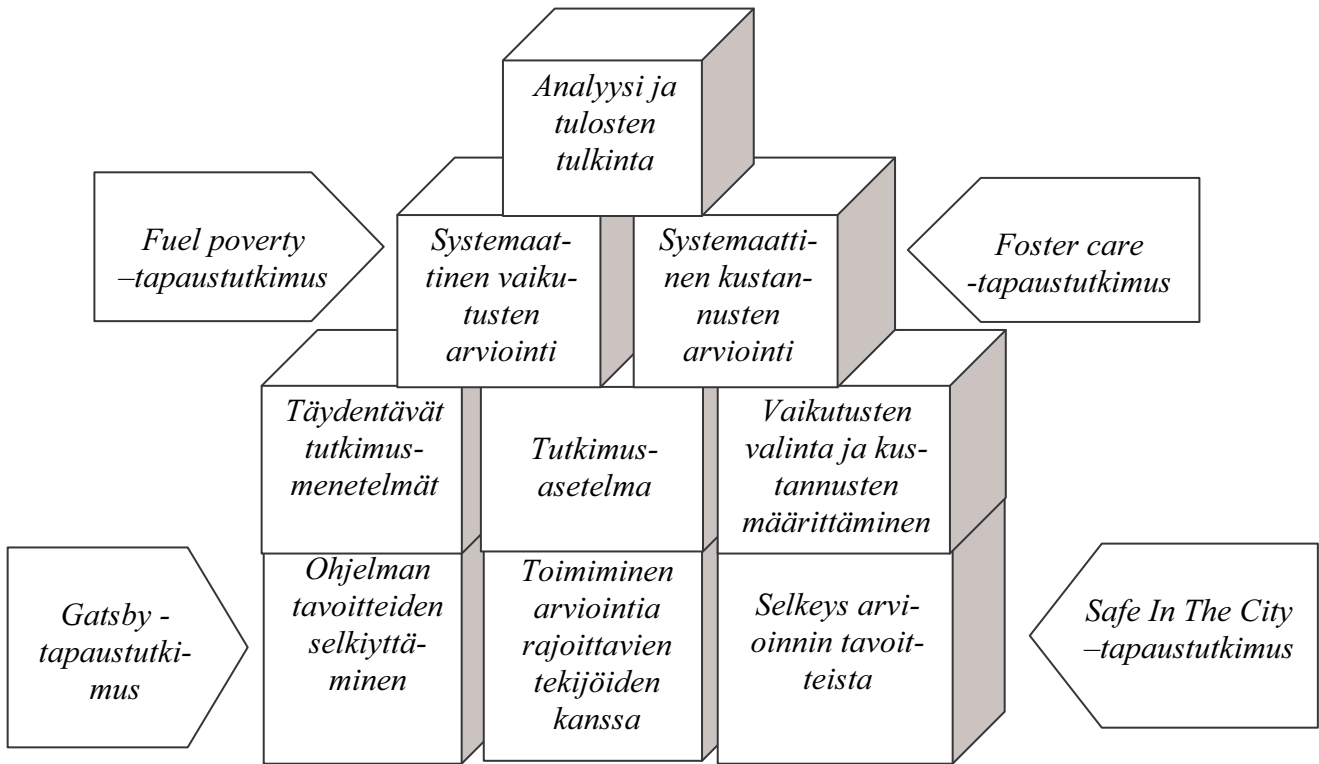
Tutkimuksemme pyrkii yksilöimään merkityksellisiä vaikutuksia ja analysoimaan niitä johdonmukaisemmin, ja tuomaan näin esille yhden keskeisistä ohjelmien arvioinnin osatekijöistä.

### **Sijaishuolto –tapaustutkimus (Foster care)**

Kysymyksessä on kaksivuotinen, sijaishuollossa olleitten lasten seurantatutkimus. Tavoitteena on tutkia ja selittää sijaishuoltoa seuraavaa tuen käyttöä sekä vaikutuksia ja kustannuksia, erityisesti suhteessa sijoitustapaan, jossa nuori oli seurantatutkimuksen alkaessa. Tulokset eivät pysty tarjoamaan selvää vastausta kysymykseen erilaisten sijoitusvaihtoehtojen suhteellisesta kustannusvaikuttavuudesta. Ne tarjoavat kuitenkin näyttöä, jonka avulla valintoja voidaan tulevaisuudessa ohjata. Näytön avulla opitaan myös ymmärtämään paremmin erilaisille nuorille hyviä tuloksia tuottavan ja kustannusvaikuttavuudeltaan paremman sijaishuollon ominaisuuksia. Sen avulla osataan tukea tällaisia palveluja. Tutkimuksen tavoitteina oli:

- kuvata sijoitusten vaikutuksia (missä lapset ovat, miten he voivat sekä miten sijoitus vastasi osallistuvien tahojen toiveita)
- selittää lasten tukiasiakkuusprofiileja ja heidän yhteydenpitoaan tai sen puuttumista sijaisvanhempiin kun sijoitus on päättynyt
- arvioida sijoitusten vaikutusta lasten hyvinvointiin suhteessa sijoitustapaan, sosiaalipalvelujärjestelmältä saatua tukea ja huoltajien asennoitumiseen
- arvioida sosiaalipalvelukokonaisuuksien kustannuksia ja arvioida laajemmin muille yhteiskunnan sektoreille koituvia kustannuksia vaihtoehtoisten sijoitustapojen kustannus-vaikuttavuuden selvittämiseksi.

**Kuvio 1.** Taloudellisen arvioinnin rakennuspalikat



### **3 Taloudellisen arvioinnin nykyinen tila**

Yksi tätä projektia motivoivista tekijöistä oli vaikutelma siitä, että taloudellista arviointia on tehty suhteellisen vähän sosiaalialalla, ainakin Isossa-Britanniassa (UK). Saadaksemme tästä konkreettista näyttöä teimme systemaattisen kirjallisuuskatsauksen sosiaalialalla tehtyihin taloudellisiin arviointitutkimuksiin. Olimme myös kiinnostuneita taloudellisen arvioinnin yleisyydestä hyvinvoinnin eri alueilla. Tiesimme ennestään, että terveydenhuolto on poikkeustapaus. Kirjallisuuskatsaus on raportoitu yksityiskohtaisemmin erillisessä julkaisussa (McDaid ym. 2002).

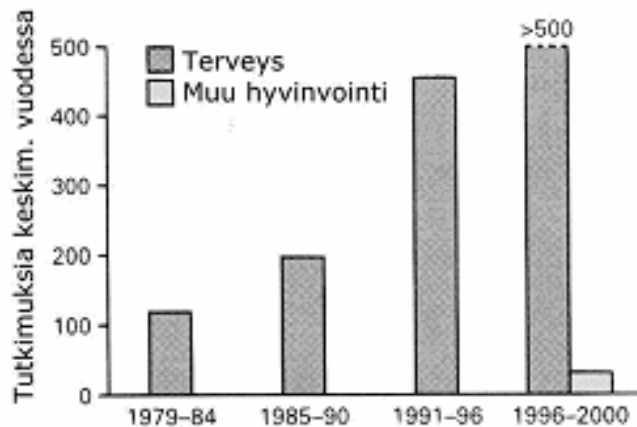
#### **Systemaattisen kirjallisuuskatsauksen menetelmä**

Systemaattisen kirjallisuuskatsauksemme perustui hakuihin, jotka tehtiin merkittävimmät kansainväliset tieteelliset aikakauslehdet vuosilta 1996-2000 sisältäviin elektronisiin tietokantoihin. Haun ulkopuolelle jäivät julkaisemattomat tutkimukset ja sellaiset tutkimukset, joita ei ollut julkaistu tieteellisissä aikakauslehdissä, kuten esimerkiksi hallinnon raportit. Hakustrategiassa yhdistettiin joukko taloudellisen arvioinnin käsitteitä (esim. cost-effectiveness, cost-benefit) joukkoon sosiaalialan termejä (esim. social care, families, housing). Tulokseksi saatu lista tutkimusabstrakteja 'suodatettiin' ja arvioitiin tutkijatiimissä. Tämä tapahtui käyttämällä sovittuja taloudellisen arvioinnin tunniste- ja luokittelu kriteereitä. Informaatio talletettiin taloudellisen arvioinnin tyyppien mukaisiin ryhmiin sekä sosiaalialan toimialueen ja alkuperämaan mukaan.

#### **Tulokset**

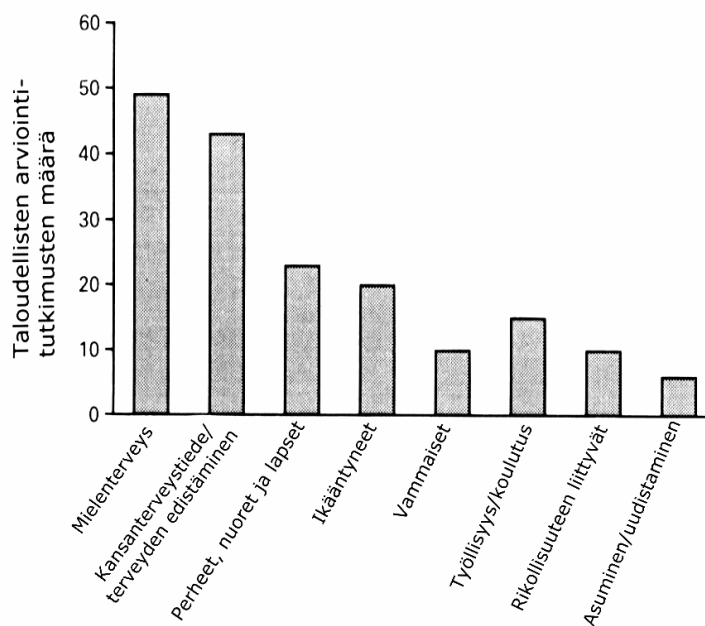
Reilun viiden vuoden tarkastelujakson ajalta löytyi 131 taloudellista arviointitutkimusta. Kun otetaan huomioon, että elektroniset rekisterit eivät olleet viimeisten vuosien osalta täydellisiä, tämä merkitsee keskimäärin kolmeakymmentä tutkimusta vuodessa. Samanlainen haku tuotti terveydenhuollon osalta 450 tutkimusta vuodessa ajanjaksolla 1991-1996 (Elixhauser ym. 1998), ja määrä on todennäköisesti noussut merkittävästi ajanjakson jälkeen. Tähän ja aikaisempaan vastaavalla lähestymistavalla tehtyyn terveydenhuoltoa koskeneeseen kirjallisuuskatsaukseen (Elixhauser ym. 1993) perustuva kuvio 2 osoittaa, että terveydenhuoltoa koskevien taloudellisten arviointitutkimusten määrä kasvaa trendinomaisesti ajan mittaan. Ei ole näyttöä siitä, että hyvinvoinnin muita alueita koskevan taloudellisen arviointitutkimuksen määrä olisi kasvanut, ei ainakaan tarkastelemamme reilun viiden vuoden aikana.

**Kuvio 2.** Taloudelliset arviointitutkimukset



Kuviosta 3 käy ilmi hyvinvoinnin eri alueilla tehtyjen tutkimusten määrä. Ne näyttävät keskittyvän alueille, jotka ovat lähellä terveydenhuollon valtavirtaa, etenkin mielenterveyteen ja kansanterveyteen (joista jälkimmäiseen lasketaan mukaan huumeiden ja alkoholin väärinkäyttöä koskevat tutkimukset). Yhdessä ne kattavat kaksi kolmannesta kaikista tutkimuksista. Huomasimme myös, että monet tutkimukset (noin kolmannes) voitiin sijoittaa useampaan kuin yhteen kategorioistamme. Esimerkkinä voidaan mainita tutkimukset, joiden kohteena ovat mielenterveysongelmallisille suunnatut tuetun työllistymisen ohjelmat ja huumehoito-ohjelmat, jotka pyrkivät myös vähentämään rikollisuutta tai kodittomuutta. Tämä heijastaa monien sosiaalialan toimialueiden keskinäisiä suhteita ja tukee käsitystä tämän alueen ohjelmien monimutkaisuudesta.

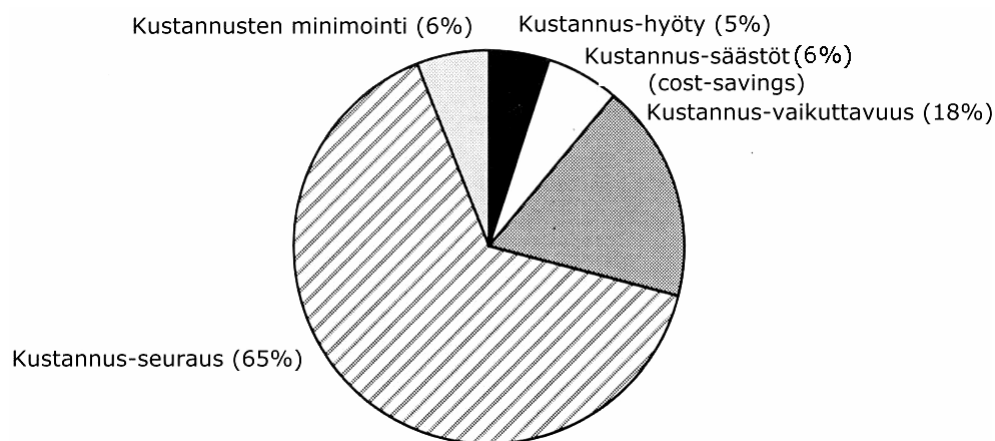
**Kuvio 3.** Taloudellisten arviointitutkimusten määrä sosiaalisen hyvinvoinnin eri alueilla (1996 - 2000)



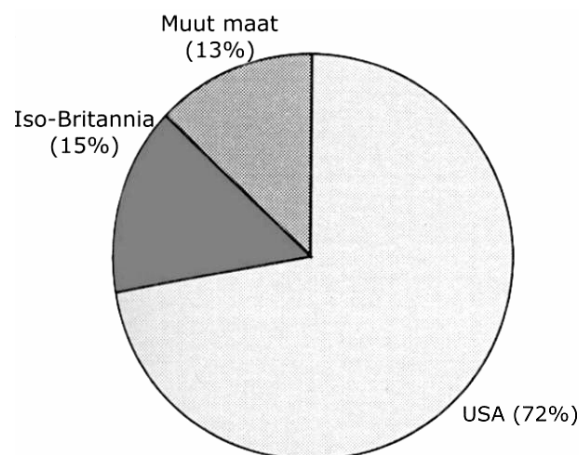


Suurin osa tutkimuksista on odotusten mukaisesti kustannus-seuraus – analyysijä, joissa tarkasteltavia vaikutuksia on useita (ks. Kuvio 4). Kustannus-hyöty -analyysijä on vähemmän kuin yksi kahdestakymmenestä tutkimuksesta (mukaan ei ole laskettu kustannusten säästö - analyysijä; ks. tarkenne 1). On kiinnostavaa, että kustannus-hyöty -analyysien osuus on alhainen myös terveydenhuoltosektorilla ja osuus on edelleen laskenut (vuosien 1979-1984 yli 40 prosentista vuosien 1991-1996 alle 20 prosenttiin). Näyttää todennäköiseltä, että kustannus-vaikuttavuus –analyysien osuus on terveydenhuollossa suurempi kuin sosiaalihuollon kentällä.

**Kuvio 4.** Prosenttiosuudet tutkimustyypeittäin (1991-96)



**Kuvio 5.** Tutkimusten maakohtaiset prosenttiosuudet (1991-96)



Lähes kolme neljännessä kaikista tutkimuksista tehtiin USA:ssa (ks. Kuvio 5). Loput jakaantuvat melko tasaisesti Englannin ja muun maailman kesken. Tässä voi kuitenkin olla hieman vääristymää, sillä pystyimme tarkastelemaan vain englannin-, ranskan- ja espanjankielisiä abstrakteja. Yhdysvalloissa tehtyjen tutkimusten hallitsevuus on sosiaalialalla suurempi kuin terveydenhuollon sektorilla, jossa niukasti yli puolet tutkimuksista oli tehty USA:ssa (ajanjaksolla 1991-1996).

## Tutkimusten laatu

Taloudellisten arviointitutkimusten laatua on hyvin vaikea arvioida pelkästään abstraktien perusteella. Useimmissa tapauksissa meillä oli kuitenkin tietoa otoskoosta ja verrokista, ja tämä antoi jotain mielikuvaa tutkimuksen laajuudesta ja tieteellisestä kurinalaisuudesta. Noin kolmanneksessa tutkimuksista ei ollut systemaattista vertailuryhmää ja yhdessä kuudesta tutkimuksesta otoskoko oli pienempi kuin 50 (vaikka tämä ei ole välttämättä ongelma). Kun abstrakteja oli kaikkiaan 131, näistä neljästäkymmenestä hankimme myös koko artikkelin. Näiden tutkimusten vahvuuksia ja heikkouksia meillä oli mahdollisuus arvioida tarkemmin. Esimerkkejä hyvistä käytännöistä löytyi, ja näihin viittaamme myöhemmissä luvuissa. Kuitenkin tutkimusten laatu oli hyvin vaihteleva. Seuraavassa suurimmat havaitsemamme heikkoudet:

- Monipuolinen kustannusarvio sisältyi harvoin tutkimuksiin. Jopa silloin, kun vaikutuksia oli arvioitu täsmällisesti, kustannusanalyysi oli usein karkea.
- Monissa tutkimuksissa esitettiin vahvoja väitteitä interventioiden kustannus-vaikuttavuudesta. Väitteet menivät usein paljon pidemmälle kuin esitetyn todistusaineiston perusteella olisi voinut päätellä.
- Kun tutkimusasetelma ei ollut ideaalinen, usein perustellusta syystä, tästä aiheutuneista vääristymistä ei monissakaan tutkimuksissa keskusteltu, saati että niitä olisi yritetty kontrolloida.

## Johtopäätökset

Systemaattinen kirjallisuuskatsaus tukee väitettämme, että toisin kuin terveyttä, sosiaalialaa koskevia taloudellista arviointitutkimusta on hyvin vähän (ainakaan tieteellisissä aikakauslehdissä julkaistuna). Suurin osa löytämistämme tutkimuksista on terveydenhuollon ja sosiaalihuollon raja-alueelta ja suurin osa tutkimuksista on tehty USA:ssa. Tutkimusten laatu on myös hyvin vaihteleva.

Jos sosiaalialan taloudellisen arviointitutkimuksen kehitys seuraa terveydenhuollon kaavaa, odotettavissa on nopea tutkimusten määrän kasvu ja asteittainen USA:n osuuden pieneneminen. Mutta tähän mennessä on vähän näyttöä siitä, että sosiaalialaan liittyvien taloudellisten arviointitutkimusten määrä olisi lisääntymässä.

## OSA II: KÄSITTEELLISIÄ KYSYMYKSIÄ

### 4 Taloudellisen arvioinnin haasteita

#### Miksi taloudellista arviointia on niin vähän?

On tärkeätä ymmärtää, miksi taloudellista arviointia on monilla sosiaalialan toimialueilla niin vähän. Ymmärtäminen on edellytys sille, että tämänkaltaisiin esteisiin voitaisiin puuttua. Mahdollisia selityksiä on useita (ks. Kuvio 6):

- Monet sosiaaliset interventiot ovat hyvin monimutkaisia, mikä tekee kustannusten ja vaikutusten systemaattisen arvioinnin vaikeaksi.
- Muut kuin taloustieteilijät tai käytännön työntekijät saattavat vastustaa taloustieteilijöiden suosimia tutkimusmenetelmiä.
- Perusteellinen taloudellinen arviointi voi vaatia paljon resursseja, jos uutta tietoa on kerättävä ja analysoitava pitkältä ajalta.
- Poliitiikan tekijät ja muut päätöksentekijät eivät vaadi taloudellista arviointia tarpeeksi. Terveystieteiden tutkimuksessa tietyn määrän taloudellista arviointia panevat liikkeelle säätelevät ja institutionaaliset rakenteet. Näihin kuuluu viimeaikainen Kliinisen laadun kansallisen instituutin (National Institute for Clinical Excellence) rooli uusien farmaseuttisten ja kliinisten interventioiden arvioinnissa Englannissa ja Walesissa. Samanlaista painetta ei tällä hetkellä ole muilla hyvinvoinnin alueilla. Joissakin tapauksissa taloudellisen arvioinnin tarve on tunnustettu, mutta päätöksentekijät eivät aina ole kovin arvostelukykyisiä (Knapp 1999). Heillä voi olla liian suuria odotuksia taloudellista arviointia kohtaan, mutta heiltä puuttuu asiantuntemusta toimeksiannon laadun arvioimiseen.

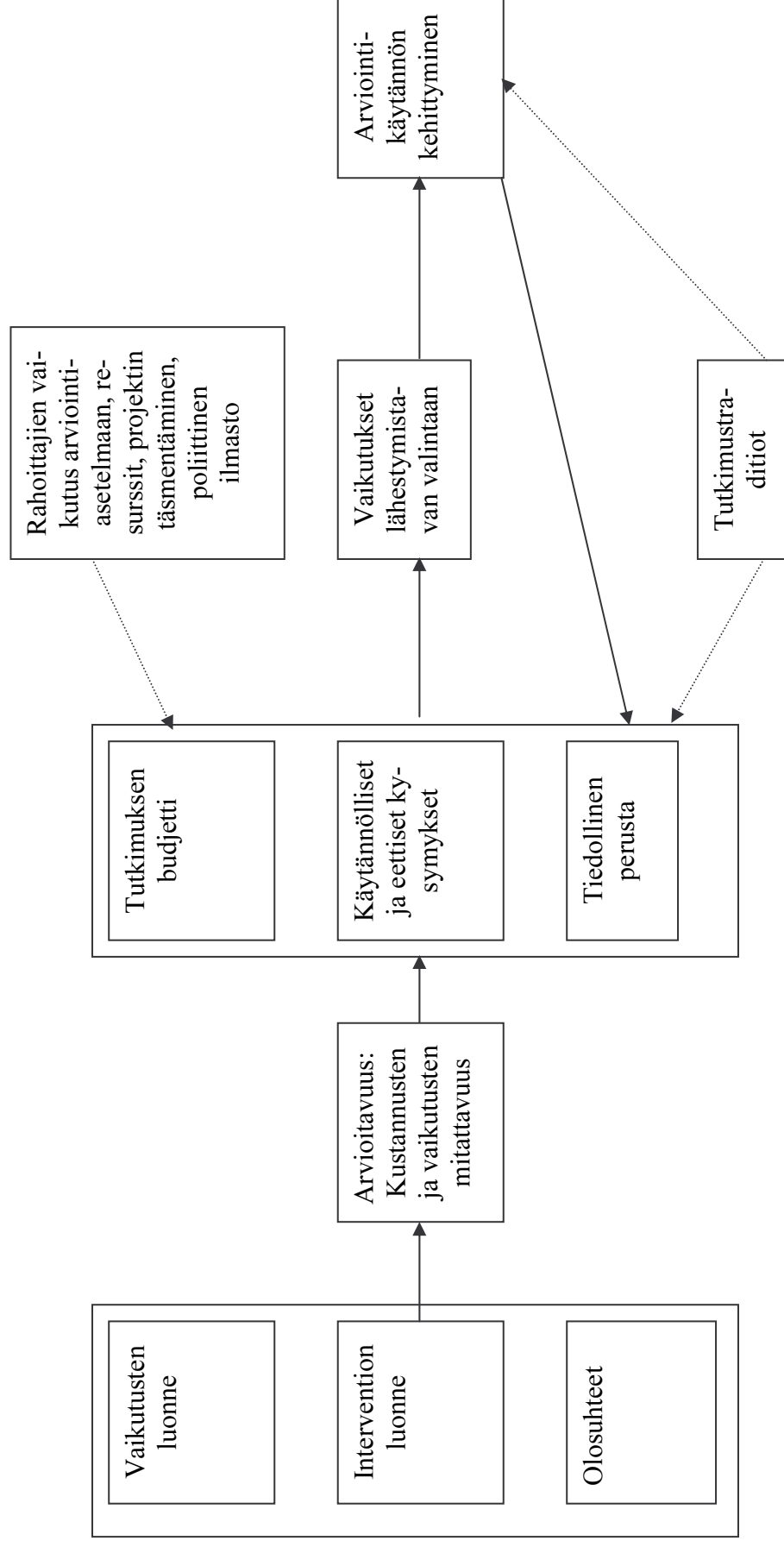
Tämä luku keskittyy edellisen listan ensimmäiseen kohtaan, ja huomiota kiinnitetäänkin monien sosiaalisten interventioiden monimutkaisuuteen ja tästä arvioinnille aiheutuviin vaikutuksiin. Seuraava luku paneutuu listan toiseen kohtaan: siinä tarkastellaan vaihtoehtoisia arvioinnin lähestymistapoja ja tunnistetaan piirteitä, joissa perinteinen taloustieteilijöiden (ja muiden kvantitatiivista arviointia tekevien) näkökulma eroaa muunlaisista arvioinnin lähestymistavoista. Listan kolmannen ja neljännen kohdan aiheista keskustellaan luvussa 10 yhdessä muiden arviointia rajoittavien tekijöiden kanssa.

#### Ohjelmien luonne

Monilla sosiaalialan interventioilla on ominaisuuksia, joiden vuoksi niitä on vaikea arvioida käyttäen perinteisiä taloudellisen arvioinnin lähestymistapoja. Toisaalta vain harvat ovat eri mieltä siitä, että paras tapa arvioida uuden lääkkeen vaikuttavuutta on satunnaistettu kontrolloitu koe. Niinpä on syytä kysyä, mitä ovat ne lääkeintervention ominaisuudet, joiden vuoksi lääkein-

terventiot soveltuvat taloustieteilijöiden suosimaan systemaattiseen ja kvantitatiiviseen arviointitapaan. Kysyä pitää myös kuinka nämä ominaisuudet eroavat muiden terveydenhuollon ja sosiaalialan interventioiden ominaisuuksista. Me käytämme yhtä tapaustutkimuksistamme, Turva kaupungissa –ohjelmaa (The Safe in the City, SITC) havainnollistamaan joitakin näistä seikoista. Vaikka terveydenhuollon ja sosiaalialan interventiot ovat erilaisia, arvioinnin kannalta merkitykselliset kysymykset ovat samoja, mukaan lukien myös monimutkaisempia terveydenhuollon interventioita koskevat kysymykset. Tarkastelemme lisäksi joitakin sosiaalialaan liittyvän taloudellisen arvioinnin käytännöllisiä kysymyksiä.

**Kuvio 6.** Sosiaali-alan interventioiden arvioitavuus



## **Vaikutusten monimutkaisuus**

Lääkeinterventioissa vaikutukset ovat usein suhteellisen helposti mitattavissa, erityisesti jos ne ovat fyysisiä (vastakohtana psykologisille) ja paikallisia. Esimerkiksi tiettyjen lääkeryhmien vaikuttavuutta voidaan arvioida sen perusteella, miten ne vaikuttavat kuolleisuuteen tai yksilöityyn sairastavuuteen (esimerkiksi kolesterolitasoihin, verenpaineeseen), ja näitä on suhteellisen helppo mitata. Monet muut terveydenhuollon interventiot pyrkivät vaikuttamaan olosuhteisiin kokonaisvaltaisemmin, ja olosuhteet vaikuttavat puolestaan moniin yksilön elämän ulottuvuuksiin, esimerkiksi mielenterveysongelmiin. Tällaisia interventioita on paljon vaikeampi mitata ja verrata. Ne soveltuvat huonommin objektiivisilla mittareilla mitattaviksi ja ne ovat myös moniulotteisempia, sillä ne sisältävät psykologisen, sosiaalisen ja fyysisen toimintakyvyn (Nocon ja Qureshi 1996). Laajempi sosiaalihuollon kenttä on tässä suhteessa samanlainen. Esimerkiksi SITC-ohjelma pyrkii vähentämään nuorten asunnottomuutta tukemalla asiakkaiden itsearvostusta, rohkaisemalla heitä ottamaan vastuuta omasta elämästään ja auttamalla heitä selviytymään ristiriidoista kotona ja koulussa. Tällaisia vaikutuksia on paljon vaikeampi arvioida määrällisillä mittareilla.

## **Interventioiden monimutkaisuus**

Lääkeinterventioissa käytetään vakioituja lääkeaineita, joita annetaan tarkka määrä kullekin potilaalle. Joidenkin terveydenhuollon toimenpiteiden, kuten leikkausten, pitäisi myös olla melko yhdenmukaisia, mutta vaihtelua voi todellisuudessa esiintyä, sillä eri sairaaloiden ja asiantuntijoiden kliiniset käytännöt eroavat toisistaan tai palvelua antavien taidoissa tai pätevytyksessä on eroja. Monien muiden terveydenhuollon interventioiden on oltava joustavampia, jotta ne voidaan sovittaa yksittäisten potilaiden tai asiakkaiden vaihteleviin tarpeisiin. Esimerkkinä voidaan mainita terveysneuvonta. Suurin osa sosiaalialan interventioista on tässä suhteessa samanlaisia. Esimerkiksi SITC-intervention osaprojektit räätälöidään yksilöiden tarpeisiin. SITC on myös epäyhtenäinen, sillä se koostuu 24 osaprojektista. Projektit on koordinoitu kahdeksaan klusteriohjelmaan, joissa kaikissa on hieman erilaiset tavoitteet, vaikka yleistavoitteena kaikilla onkin nuorten asunnottomuuden ehkäisy. Tieto osaprojektien vaihtelun laajuudesta ja luonteesta on olennaista pyrittäessä ymmärtämään, miten erilaisuus vaikuttaa intervention vaikutuksiin. Esimerkiksi, jos interventio näyttää olevan vaikuttavampi joillekin asiakkaille kuin toisille, on hyödyllistä tietää, johtuuko tämä siitä, että asiakkaat olivat vastaanottavaisempia vai siitä, että interventiot olivat erilaisia. Samalla tavoin, jos jokin interventio koostuu joukosta toimenpiteitä, on hyödyllistä tietää, onko joillakin toimenpiteillä enemmän vaikutusta kuin toisilla.

Edellisen lisäksi SITC on kehittämishanke, jota on muutettu huomattavasti sen jälkeen kun se käynnistettiin neljä vuotta sitten. Myös osaprojektit ovat muokanneet toimintaansa projektin edetessä. Kuten monet muutkin sosiaali-

seen hyvinvointiin liittyvät interventiot, SITC-ohjelma on arvioijille ”liikkuva kohde”. Vaikka arviointitutkimus kykenisi osoittamaan, että ohjelma kokonaisuus on vaikuttava, voi olla vaikea määritellä, mikä tarkasti ottaen toimi ja miten ohjelma toistettaisiin muualla.

## **Käyttäjien rooli**

Potilaat ovat lääkeinterventiossa melko passiivisia. Lukuun ottamatta hoitoon mukautumista (esim. sitä, että muistaa ottaa lääkettä oikean annoksen), potilaiden merkitys lääkkeen vaikuttavuuteen tai lääkeinterventioon on muilta osin rajallinen. Samoin potilaat ovat melko passiivisia leikkaustilanteissa; vaikka he voivat vaikuttaa paranemisaikaan, heidän vaikutuksensa kirurgisen toimenpiteen onnistumiseen tai leikkaustapahtumaan on vähäinen. Vaikka asiakkaat ovat kaikissa interventioissa mukana jonkin verran, voidaan sanoa, että he eivät ole mukana niin paljon kuin sosiaalialan ja terveydenhuollon moninaisemmissa, yhteisöllisissä interventioissa. SITC:n tapauksessa ohjelman onnistuminen riippuu ratkaisevasti siitä, miten nuoret interventioon reagoivat. Tähän vaikuttavaa moni tekijä, esimerkiksi näkevätkö nuoret intervention hyödyllisenä vai asioihinsa sekaantumisena, miten hyvin he tulevat toimeen projektityöntekijöiden ja muiden ohjelmassa mukana olevien nuorten kanssa sekä tukeeko koulu tai perhe interventiota.

Mitä aktiivisemmin potilaat tai asiakkaat ovat mukana interventiossa, sitä todennäköisemmin heidän asenteensa, käyttäytymisensä ja olosuhteensa vaikuttavat tuloksiin. Näistä muodostuu joukko sekoittavia muuttujia, jotka on otettava huomioon arvioinnissa ja joita on vaikea luokitella tai esittää määrinä samalla tavoin kuin iän ja sukupuolen kaltaisia vakiintuneita taustamuuttujia. Tämä vuorostaan vaikeuttaa arviointitulosten yleistämistä. Jos interventiosta halutaan saada sama vaikutus eri olosuhteissa, on intervention toiminta ymmärrettävä hyvin ja erityisesti se, miten ihmiset siihen reagoivat.

## **Näkökulma ja soveltamisala**

Lääkeinterventioiden vaikutukset rajoittuvat yleensä potilaisiin ja terveyspalveluihin, joten intervention vaikutuksista on suhteellisen helppoa saada kokonaisvaltainen kuva - mitä taloustieteilijät kutsuvat yhteiskunnalliseksi näkökulmaksi. Kuitenkin, esimerkiksi mielenterveysongelmaisten vanhus-ten avo- ja sairaalahoidon vertailun on oltava paljon laajempaa sisältäen terveyspalvelut mutta myös sosiaalipalvelut, vapaaehtoispalvelut, epävirallisen hoivan ja sekä potilaille että heidän perheilleen aiheutuvat kustannukset. Sosiaalialan interventioiden arviointi kattaa monasti useita sektoreita, sillä monet toimijat osallistuvat palvelujen tarjoamiseen ohjelmien kohteena oleville haavoittuville ryhmille. Esimerkiksi SITC-ohjelman arvioinnissa olisi tarkasteltava myös hankkeen mahdollisia kerrannaisvaikutuksia sosiaalipalveluille, kouluille, työvoimapalveluille ja vapaaehtoispalveluille sekä nuorille ja heidän perheilleen.

## **Analyysiyksikkö**

Lääkeinterventiot, kuten suurin osa terveydenhuollon palveluista, on suunnattu yksittäisille potilaille. Mutta jotkut terveydenhuollon interventiot pyrkivät lisäämään terveyttä (ja mahdollisesti myös muuta kuin terveyttä) vaikuttamalla sairauksien rakenteellisiin tai sosioekonomisiin syihin. Esimerkiksi terveyttä edistävä interventio voi pyrkiä vähentämään tupakointia luomalla ympäristöjä, jotka suosivat tupakoimattomuutta (esim. savuttomat alueet, mainonnan kieltäminen). Tällaisten yhteisöllisten interventioiden suosio lisääntyy myös sosiaalialalla. Eräässä toisessa tapaustutkimuksis-tamme, Gatsby -mallissa, pyrkimyksenä on auttaa elvyttämään paikallisyhteisöjä tukemalla yhteisön pienryhmiä. Vaikka lopullisena tavoitteena on parantaa alueella elävien ihmisten elämänlaatua, interventio kohdistaa huomionsa pikemminkin yhteisöryhmiin kuin yksittäisiin asukkaisiin. Tämän-tyyppisiä interventioita on vaikeampi arvioida, sillä vaikutuksia on arvioitava eri tasoilla, myös koko organisaation tai yhteisön tasolla.

## **Aikaulottuvuus**

Monet lääkeinterventiot vaikuttavat melko nopeasti. Niinpä vaikutuksia voidaan arvioida suhteellisen pian, eikä tärkeitä vaikutuksia jää huomiotta, vaikka mahdolliset sivuvaikutukset voivatkin vaatia pidemmän seuranta-ajan. Vastaavasti jotkut sosiaalialan interventiot liittyvät akuuttiin kysymyk-siin, esimerkiksi aterioiden tarjoamiseen vanhukselle, jonka omaishoitaja (informal carer) tarvitsee levähdystauon. Toiset terveydenhuollon ja sosiaa-lihuollon ongelmat ovat kuitenkin luonteeltaan pysyvämpiä, ja siksi kustan-nuksia ja vaikutuksia on seurattava paljon kauemmin. Lisäksi, minkä tahansa vaikutuksen kehittymisaikaan voi liittyä epävarmuutta, minkä vuoksi arviointiaikaa saatetaan joutua pidentämään, ja tästä aiheutuu lisäkustan-nuksia. Tämä on erityinen ongelma ehkäisevissä projekteissa kuten SITC:ssä, jossa vaikutusten syntyvät pitkän ajan kuluttua tai kumuloituvat asteittain ajan kuluessa. SITC:n projektit, jotka on suunnattu 13-vuotiaisiin lapsiin, pyrkivät vähentämään kyseisten lasten riskiä jäädä asunnottomiksi 16-18-vuotiaina.

Taulukossa 1 esitetään hyvin karkea menetelmä sen hahmottamiseksi, miten helppo erilaisia ohjelmia on arvioida (ohjelmien arvioitavuus). Aiemmin esillä olleiden kysymysten lisäksi tauluko sisältää myös muita arviointiin yleisesti liittyviä ongelmia. Sosiaalialan interventiot, kuten myös terveydenhuollon interventiot kattavat koko kirjon yksinkertaisista monimutkaisiin. Monilla sosiaalialan interventioilla ja monimutkaiste sekä pitkäkestoisten terveysongelmien hoidolla on kuitenkin keskenään enemmän yhteistä kuin niillä kummallakaan on lääkeinterventioiden kanssa. Monia sosiaalialan ja pitkäkestoisiin terveysongelmiin liittyviä interventioita luonnehtii käyttäjien voimakas osallistuminen, toiminnan kohteiden suuri vaihtelu, monimutkaiset ja pitkäaikaiset vaikutukset ja monien toimijoiden mukana olo. Mikä tahansa edellä luetelluista ongelmista saatetaan kohdata suurimmassa osassa minkä tahansa toimialueen (policy area) arviointeja. Mutta todennäköistä on, että sosiaalialan ohjelmia ja monimutkaisempia terveydenhuollon inter-



ventioita arvioitaessa joudutaan samassa tutkimuksessa käsittelemään useita arvioinnin ongelmia, mikä vaikeuttaa tilannetta. Esimerkiksi, luonteeltaan laadulliset vaikutukset voidaan usein arvioida käyttämällä määrällisiä mitta-asteikkoja. Mutta jos tämän lisäksi vaikutusten odotetaan olevan pieniä ja niihin vaikuttavat monet ulkopuoliset muuttujat, ohjelman vaikutusta voi olla hyvin vaikea mitata tarkasti.

Vaikka edellä kuvatut monimutkaisuudet eivät estäisikään hyvälaatuisen taloudellisen arvioinnin tekemistä, ne vaikeuttavat sitä merkittävästi. Seurauksena on, että arviointitutkimukset jäävät kauas ideaalisista. Tämän raportin III osassa tarjotaan käytännön ohjeita, joiden avulla ongelmiin voidaan tarttua.

**Taulukko 1.** Sosiaalialan interventioiden arvioitavuuden kriteerit

<b>Vaikutusten luonne</b>	Onko tavoitteet määritetty hyvin?	Ovatko ohjelman tavoitteet epäselvät?
	Lyhyt vai pitkä ilmenemisaika	Onko ohjelman lopullisia vaikutuksia vaikea arvioida käytettävissä olevan ajan puitteissa?
	Yksi vai monta	Onko ohjelman vaikutukset epätarkoituksenmukaista ilmaista yhtenä tai tärkeimpänä vaikutuksena?
	Määrällinen vai laadullinen	Ovatko vaikutukset luonteeltaan laadullisia tai jostain muusta syystä vaikeasti mitattavia?
<b>Ohjelman luonne</b>	Moniaineisuus	Onko ohjelma erilaisten interventioiden tai projektien sekoitus?
	Toimeenpano	Onko todennäköistä, että toimeenpano vaihtelee merkittävästi ohjelman sisällä?
	Analyysiyksikkö	Kohdistuuko ohjelma alueisiin tai yhteisöihin?
<b>Konteksti</b>	Mittakaava	Onko odotettavissa oleva vaikutus todennäköisesti pieni suhteessa ongelman suuruuteen?
	Ulkoiset vaikutukset	Onko ohjelman odotettavissa oleva vaikutus todennäköisesti pieni suhteessa siihen, miten paljon muut ulkoiset tekijät voivat vaikuttaa samaan lopputulokseen?
	Aktiiviset vai passiiviset asiakkaat	Onko ohjelman vaikuttavuus hyvin herkkä sille, miten asiakkaat päättävät reagoida?

Lähde: Perustuu Sefton (2000)

## **5 Arvioinnin vaihtoehtoiset näkökulmat**

### **Erilaisia koulukuntia**

Toinen syy siihen, miksi taloudellinen arviointi on vakiinnuttanut hitaasti asemaansa eräillä sosiaalisen hyvinvoinnin alueilla on se, että monet taloustieteilijät eivät ole kovin hyviä työskentelemään yhdessä muita tieteenaloja edustavien arviointitutkijoiden kanssa. Tässä luvussa tarkastellaan erilaisten arvioinnin koulukuntien välisiä eroja. Vaikka tarkastelu on hyvin yksinkertaistettua ja lähestymistavoista esitetään vain ääripäät, se auttaa havaitsemaan, kuinka perinteisen taloudellisen arvioinnin lähestymistapoja pitäisi mukauttaa sosiaalialalla käytettäessä.

### **Summatiivinen vai formatiivinen**

Useimpien taloudellisten arviointitutkimusten tavoitteena on testata hyvin määriteltäviä interventioita, joiden odotetaan pysyvän melko vakaina koko intervention toteuttamisen ajan. Jos interventio muuttuu merkittävästi arvioinnin aikana, on vaikeaa sanoa tarkkaan, mikä on ollut vaikuttavaa tai kustannusvaikuttavaa, vaikka arvioinnin tulokset olisivat lupaavia. Arvioitavan intervention määrittely tarkoittaa tarjottavien palvelujen yksilöintiä: kenen toimesta, kenelle ja miten paljon palvelua tarjotaan. Ihanteellista olisi, että arvioitava interventio olisi standardisoitu selvästi kirjoitettujen ohjeiden avulla ja tarkistettu siten, että voitaisiin olla varmoja käytännön toteutuksen ja standardoinnin vastaavuudesta. Tämä ei tarkoita sitä, että asiakkaiden olisi saatava samaa palvelua riippumatta heidän olosuhteistaan, vaan sitä että jokaisen variaation on oltava järjestelmällinen. Kun taloudellinen arviointi pyrkii siis olemaan summatiivinen, muut arvioinnin lähestymistavat ovat yleensä formatiivisia. Oletuksena formatiivisissa lähestymistavoissa on, että suurin osa interventioista kehittyy ajan kuluessa reaktiona ohjelman toiminnasta saatuun kokemukseen, muuttuneisiin olosuhteisiin tai yksinkertaisesti ohjelman toimeenpanoprosessiin, erityisesti silloin kun mukana on eri toimijoita. Formatiiiviset arvioijat ovat kiinnostuneita siitä, kuinka ohjelmat muuttuvat ajan kuluessa ja, toimintatutkimuksen tapauksessa, kuinka he auttavat aktiivisesti muovaamaan tätä prosessia (Greenwood 1998).

### **Määrällinen vai laadullinen**

Taloudellinen arviointi on lähes aina luonteeltaan määrällistä osittain sen vuoksi, että se vastaa taloustieteilijöiden ajattelutapaa. Se on määrällistä myös siksi, että on paljon helpompi verrata määrällisiä kustannuksia ja vaikutuksia. Silloin, kun vaikutukset eivät ole luonteeltaan määrällistä mittaamista suosivia, voidaan käyttää vaikutuksiin liittyviä mitta-asteikkoja tai välillisiä mittareita. Kaiken kaikkiaan tutkimusasetelma suunnitellaan palvelemaan ohjelman kustannusten ja yksilöityjen vaikutusten tarkkaa mittaamista. Toisaalta laadulliset tutkijat suhtautuvat paljon epäilevämmiin numeroihin, jotka heidän mukaansa voivat antaa karkean ja joskus harhaanjohta-

van kuvan ohjelman tuloksista. Numerot eivät kykene kovinkaan hyvin vangitsemaan ihmisten kokemusten rikkautta ja erilaisuutta. Lisäksi on olemassa vaara, että keskitytään vain niihin muuttujiin, jotka voidaan helposti mitata. Määrällisessä analyysissä on myös helppo olla välittämättä olosuhteista, joissa ohjelmaa toteutetaan. Tällä seikalla voi kuitenkin olla merkittävä vaikutus tuloksiin. Laadullisessa lähestymistavassa on mahdollista kertoa tarinoita, päästä pintaa syvemmälle ja ymmärtää paremmin, kuinka ohjelma todella vaikuttaa yksilöön tai instituutioon.

### **Mitä vai kuinka/miksi**

Satunnaisotoksella valittuun kontrolliryhmään perustuvat koeasetelmat ja muut kokeelliset arvioinnin lähestymistavat muodostavat perustan monille taloudellisille arviointitutkimuksille, erityisesti terveydenhoidossa ja hoidossa. Nämä lähestymistavat keskittyvät panosten ja vaikutusten väliseen suhteeseen ja kiinnittävät usein vähän huomiota siihen, kuinka mitattavat vaikutukset ovat syntyneet. Laadulliset tutkijat ovat kiinnostuneempia ymmärtämään, kuinka ohjelmat toimivat, ja heidän metodinsa ovat sopivampia selvittämään tämän tyyppisiä kysymyksiä. Ohjelman toimintaa ei katsota erilliseksi asiaksi, vaan osaksi ohjelman vaikutuksen ymmärtämistä. Kiinnostus teoriaperusteisen arvioinnin niitä lähestymistapoja kohtaan, joissa pannaan enemmän painoa sekä vaikutusten että ohjelmien toimintaa kuvaavien prosessien tai mekanismien ymmärtämiselle, on kasvamassa. (esim. Pawson ja Tilley 1997.) Prosessien ja mekanismien ymmärtäminen ei kuitenkaan kuulu kiinteänä osana useimpiin taloudellisiin arviointeihin.

### **Positivismi vai konstruktivismi**

Taloudellinen arviointi liittyy positivistiseen maailmankatsomukseen. Siinä oletetaan, että on olemassa 'oikeita vastauksia' ja niiden esiin saamiseksi tarvitaan vain oikeiden metodien käyttöä. Tämä näkyy positivismiin nojautuvien tutkijoiden arviointinäkemyksessä. Sama asennoituminen näkyy myös siinä, että nämä ovat mieltyneitä arviointitutkimuksiin, jotka toteutetaan riittävän irrallaan ihmisistä, jotka ovat vastuussa ohjelman suunnittelusta tai toimeenpanosta. Tavoitteena on, että arvioija voi olla niin objektiivinen kuin mahdollista. Osallisten mukaan ottamiseen suhtaudutaan välttelevästi, koska heidän pelätään voivan vaikuttaa tuloksiin. Positivismi selittää myös huolta mahdollisesta tulosten vinoutumisesta (eli mistä tahansa sellaisesta, mikä voi antaa vääristyneen kuvan 'todellisesta' vaikutuksesta). Toinen äärimmäinen lähestymistapa on konstruktivismi, jossa katsotaan, että ei ole olemassa objektiivista totuutta (esim. Guba ja Lincoln 1989). Ohjelman arviointi vastaa tämän näkemuksen mukaan enemmän maalauksen tai kirjan arvostelua kuin esineen korkeuden tai pituuden mittaamista: Toisin sanoen totuus tai todellisuus on katsojan silmässä. Tällaisessa tilanteessa arvioijan tehtävänä on etsiä eri näkökulmien väliltä yksimielisyyttä tai neuvottelun tuloksena saatavaa yhteisymmärrystä tai - mikä realistisempaa - kuunnella erilaisia näkökohtia ja yrittää yhdistää niitä (Kushner 1996). Tällainen arviointi on paljon osallistuvampaa toimintaa kuin positivistinen arviointi.

## **Individualismi vai institutionalismi**

Yksi perinteisen taloudellisen lähestymistavan kritiikki, jota esittävät myös jotkut taloustieteilijät, moittii perinteistä taloustiedettä liian kapeasta keskittymisestä yksilöllisiin vaikutuksiin ja laajemman sosiaalis-taloudellisen ympäristön muutosten unohtamisesta. Ympäristöllä tarkoitetaan tällöin yhteisön arvoa, infrastruktuurin tukea ja sosiaalista pääomaa. Kritiikin esittäjät väittävät, että on siirryttävä pois puhtaasti yksilöllisestä (tai utilitaristisesta) tarkastelutavasta sellaiseen tarkasteluun, joka tunnistaa yksilöiden ja heidän ympäristönsä välisten monimutkaisten vuorovaikutussuhteiden olemassaolon. Tämä helpottaa yksilöiden käyttäytymisen ymmärtämistä. Tämän seurauksena perinteiset taloudellisen arvioinnin lähestymistavat sopivat huonosti sellaisten ohjelmien arviointiin, joissa kiinnostuksen kohteena ovat pikemminkin yhteisöt kuin yksilöt - esimerkiksi kouluperusteiset interventiot, jotka pyrkivät vähentämään koulukiusaamista parantamalla kouluympäristöä (esim. vahvistamalla positiivisia normeja). Tällaisten ohjelmien vaikutuksilla kestää kauemmin tulla esiin, koska aikaa tarvitaan muutokseen vaadittavien olosuhteiden luomiseen (mutta vaikutusten pitäisi toisaalta vaikuttaa pitempään): Siksi arvioinnilla tulee olla pidempi aikajänne, jos täydet hyödyt halutaan sisällyttää tuloksiin. Jos tällaiset ohjelmat onnistuvat, ne jättävät yhteisölle paremmat edellytykset tulevaisuudessa hoitaa omia ongelmiaan. Ne myös lisäävät yhteisön kykyä reagoida ulkopuolelta tuleviin aloitteisiin - molemmat ovat tärkeitä vaikutuksia, joita ei ehkä saisi esiin yksilöihin keskittyvässä arvioinnissa. Institutionaalisen taloustieteen edustajat (ja muut) uskovat, että yhteisöperusteisen intervention arvioinnissa olisi pyrittävä tarkastelemaan sekä yhteisön muutoksia että vaikutuksia yksilöille ja että olisi tutkittava myös näiden välisiä (kaksisuuntaisia) suhteita. Muussa tapauksessa tällaisten interventioiden arviointitutkimukset ovat järjestelmällisesti vääristyneitä (Shiell ja Hawe 1996).

## **Perinteisten lähestymistapojen vahvuuksia ja heikkouksia**

Erot edellä esiteltyjen erilaisten lähestymistapojen kesken ovat liioiteltuja, mutta todellisia. Uskomme kuitenkin, että molemmilla näkökulmilla on paikkansa sosiaalialan interventioiden arvioinnissa. Määrällisellä arvioinnin lähestymistavalla on seuraavia etuja:

- Kokeellisista arvioinneista saatu todistusaineisto näyttää poliitikkojen käytössä olevan vallitsevaa.
- Ne ovat systemaattisempia, jolloin ne takaavat tutkimusten suuremman yhtenäisyyden ja mahdollistavat ohjelmien välisen vertailun.
- Ne vaativat ohjelmien tavoitteilta suurempaa selkeyttä ja täsmällisyyttä, koska tavoitteet on kyettävä kuvaamaan määrällisesti mitattavina vaikutuksina.
- Sekoittavien tekijöiden vaikutusta on helpompaa kontrolloida.
- Määrälliset tulokset on helpompaa esittää tiiviisti.

Mutta määrällisen arvioinnin lähestymistavoilla on myös rajoituksensa, jotka ovat luultavasti sitä suurempia, mitä monimutkaisempia ovat interventiot:

- Tärkeät vaikutukset voivat olla vaikeasti mitattavia, minkä vuoksi ne saatetaan jättää huomiotta tai mitataan sopimattomalla tavalla.
- Kokeellisten tutkimusten toteuttamisessa voi olla käytännöllisiä ja eettisiä rajoituksia.
- Näyttö saatetaan saada vasta ohjelman päätyttyä (tai ennalta määriteltynä välianalyysivaiheissa), mikä voi olla liian myöhäistä, jos halutaan saada aikaan muutoksia vireillä olevassa ohjelmassa. Tämän vuoksi määrällisen lähestymistavan anti hyvien käytäntöjen kehittämiseksi on todennäköisesti lyhyellä tähtäimellä heikko.
- Määrällinen lähestymistapa pakottaa käyttämään melko joustamattomaa arviointikehystä, joten lähestymistapa voi sopia huonosti, kun arvioidaan nopeasti kehittyviä ohjelmia.
- Tällaiset arvioinnit eivät kerro paljoa siitä, kuinka ohjelma toimii ja miksi se on vaikuttava (tai ei ole).
- Ne voivat aliarvioida yhteisöllisten vaikutusten merkityksen.

Joissakin tapauksissa nämä potentiaaliset puutteet voidaan ratkaista huolellisella suunnittelulla tai niihin voidaan tarttua laajemman arviointistrategian muilla osatekijöitä. Mutta taloustieteilijät ovat myös saaneet haasteen ajatella lähestymistapaansa arviointiin kokonaisvaltaisemmin. Erityisesti on esittelemiemme menetelmien valinnassa kiinnitettävä huomiota suurempaan joustavuuteen ja moninaisuuteen, vaikka nämä eivät aina helposti sopisikaan positivistiseen tai yksilölliseen maailmankuvaan (Mannion ja Small 1999). Oma vastauksemme tähän haasteeseen heijastuu siinä taloudellisen arvioinnin lähestymistavassa, minkä kuvaamme tämän kirjan loppuosassa. Pyrkimyksenä on ottaa mukaan muita lähestymistapoja kuitenkin uhraamatta perinteisen taloudellisen arvioinnin vahvuuksia.

## **Ajattelumalli monimutkaisille interventioille**

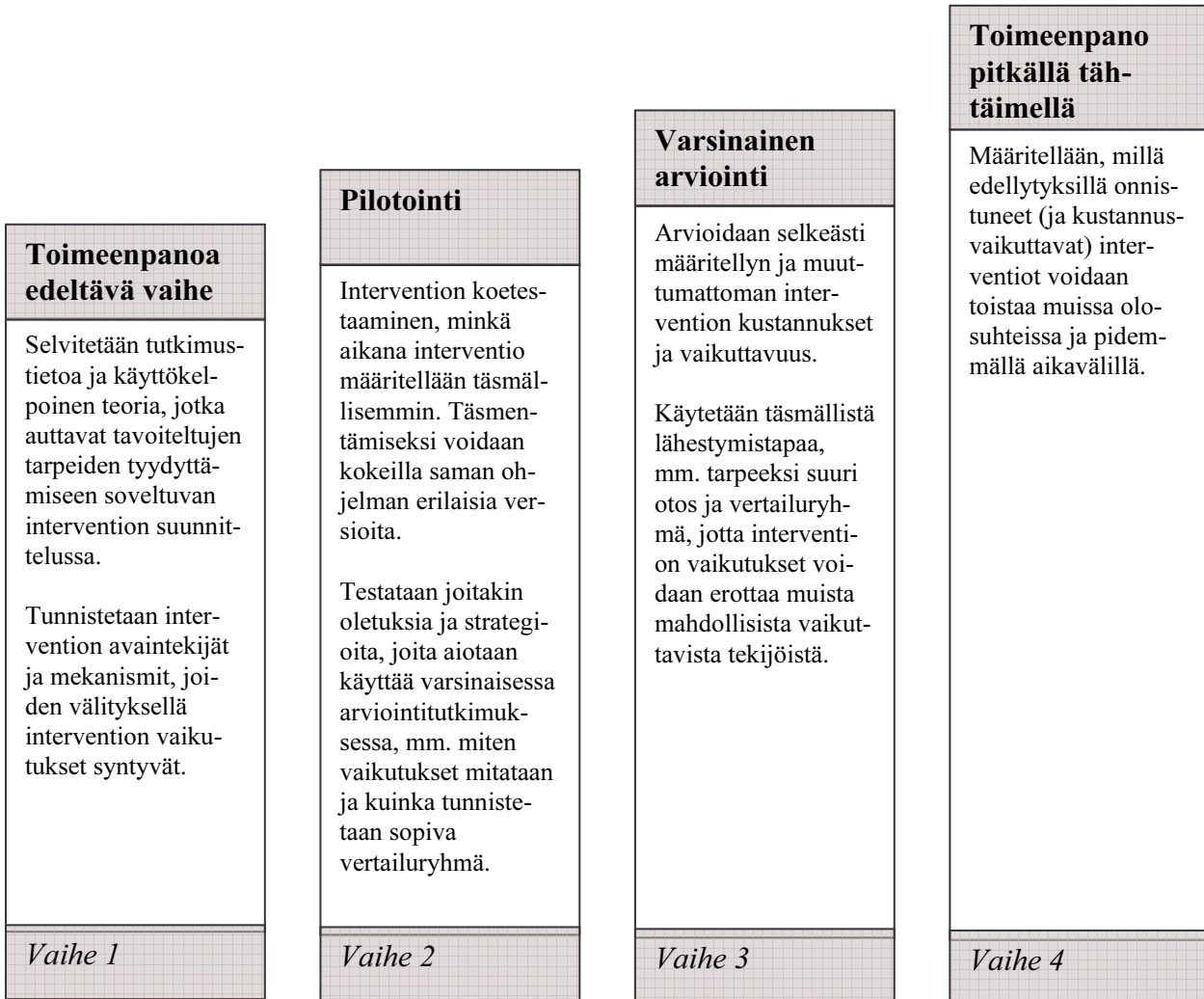
Tässä kappaleessa esitellään monimutkaisten interventioiden arvioinnissa hyödyllinen ajattelumalli, joka on sovellutus uusien lääkkeiden arvioinnin vaiheista (Medical Research Council 2000). Emme ehdota, että sosiaalialan interventiot olisi testattava samalla tavalla kuin lääkkeet tai että kaikki arviointi olisi vietävä kohti laajoja kontrolloituja kokeita. Mutta vaiheistettu näkökulma arviointiin on hyödyllinen yhdistettäessä erilaiset arviointitoiminnot yleiseksi strategiaksi, jonka tavoitteena on tuottaa todistusaineistoa intervention vaikuttavuudesta ja kustannus-vaikuttavuudesta. Lähtökohdaksi olisi yhtä hyvin voitu ottaa jollakin muulla hyvinvoinnin alueella kehitetty samantapainen ajattelumalli. (Ks. esimerkiksi 'viisitasoinen näkökulma arviointiin', jonka on kuvannut Pecora ym. 1995).

Kuviossa 7 kuvataan vaiheet, jotka tarvitaan hankittaessa todistusaineistoa uuden intervention tueksi. Arvioinnin tarkoitus ja menetelmät ovat joka portaalla erilaisia, mutta jokainen askelma informoi seuraavaa. Ainakin jotkut

arvioinnin erilaiset näkökulmat voidaan tämän kehyksen puitteissa sovittaa yhteen. Arvioinnin varhaisissa vaiheissa formatiivisella lähestymistavalla on selkeä rooli määriteltäessä interventiota, tutkittaessa toimiiko se tarkoitetulla tavalla ja autettaessa hiomaan ohjelmaa sellaiseksi, mitä voidaan arvioida täsmällisemmin. Kvantitatiivisten menetelmien rooli muuttuu tärkeämmäksi siirryttäessä myöhäisempiin arviointiprosessin vaiheisiin, joissa hankitaan 'tieteellistä' näyttöä ohjelman vaikuttavuudesta ja kustannusvaikuttavuudesta. Kvantitatiivisia menetelmiä voidaan kuitenkin joutua täydentämään laadullisilla, jotta saataisiin esiin vaikeasti mitattavia vaikutuksia (tai vaikutusten piirteitä). Lopuksi, laadullisten menetelmien avulla saatava parempi ymmärrys siitä, kuinka ja miksi interventio toimii, voi helpottaa sen arvioimista, kuinka hyvin keskeisimpiä arviointitutkimuksen tuloksia voidaan yleistää muihin tilanteisiin.

Nämä vaiheet on hahmoteltu lääkekokeilujen yhteydessä. Sosiaalialalla prosessi on todennäköisesti paljon "sumeampi", sillä vastaavat institutionaaliset rakenteet interventioiden testaamiselle puuttuvat. Mutta politiikan tekijöille on silti hyödyllistä pohtia, missä he tällä jatkumolla ajattelevat olevansa. Arvioijien on puolestaan hyödyllistä pohtia, mitkä tekniikat ovat tarkoituksenmukaisia milläkin askelmalla ja kuinka niitä voidaan hyödyntää myöhemmissä arviointivaiheissa. Käyttämällä esimerkkinä yhtä tapaustutkimustemme taulukossa 2 listataan arviointikysymyksiä, jotka voivat olla mielekkäitä arviointistrategian eri vaiheissa. Taloudellinen arviointi näyttää tulevan mukaan vasta myöhemmissä arviointistrategian vaiheissa, mutta todellisuudessa monet taloudellisen arvioinnin rakennuspalasista olisi asetettava paikoilleen jo varhaisemmissa vaiheissa.

**Kuvio 7.** Ajattelumalli monimuotoisten interventioiden arvioinnille



Lisääntyvän näytön jatkumo →

(Lähde: Sovellettuna Medical Research Council 2000.)



**Taulukko 2.** Vaiheistettu lähestymistapa arviointiin (Safe In The City - tapaustudkimus)

<i>Tarkoitus</i>	<i>Erityiset arviointikysymykset</i>
<b>Vaihe 1: Toimeenpanoa edeltävä vaihe</b>	
Dokumentoidaan ohjelman tarpeellisuus	Mitkä ovat nuorten asunnottomuuden riskitekijät? Kuinka yleisiä ovat riskitekijät alueilla, joilla Safe In The City –ohjelmaa toteutetaan? Mitä palveluja näillä alueilla on jo käytettävissä ja mitkä ovat järjestelmän puutteet?
Määritetään ohjelman peruspiirteet	Mitkä palvelut tarjotaan ohjelman toimesta ja kenelle ne suunnataan?
Osoitetaan havaittujen tarpeiden ja ohjelman yhteensopivuus	Mikä teoria on palvelujen takana, eli millä perusteella palvelut on valittu ja kuinka niiden oletetaan tuottavan haluttua vaikutusta?
<b>Vaihe 2. Pilotointi</b>	
Selkeytetään projektin pyrkimykset ja tavoitteet	Kuinka projektissa mukana olevat eri toimijatahot ymmärtävät asunnottomuuden ehkäisyn? Onko toimijatahoilla selvästi ilmaistut tavoitteet ja ovatko ne yhdenmukaisia ohjelman yleisten tavoitteiden kanssa?
Testataan kuinka ohjelma toimii käytännössä ja pyritään tekemään parannuksia?	Kuinka ohjelman eri osat toimeenpannaan? Mikä muiden projektien perusteella näyttää olevan ‘parasta käytäntöä’?
Kehitetään arviointistrategia seuraavaa vaihetta varten	Mitkä kustannukset ja vaikutukset pitäisi arvioida ja miten? Mikä voisi olla sopiva vertailuryhmä ja kuinka se voitaisiin hankkia?
<b>Vaihe 3. Varsinainen arviointi</b>	
Arvioidaan saavuttavatko yksittäiset projektit tavoitteensa	Saavuttavatko yksittäiset projektit tavoitteensa? Mitä mieltä projektien asiakkaat ovat saamastaan palvelusta?
Arvioidaan ohjelman vaikuttavuutta ja kustannusvaikuttavuutta verrattuna muihin ohjelmiin, joilla on samankaltaiset tavoitteet	Mikä on ohjelman nettovaikutus ottaen huomioon muut tuloksiin vaikuttavat tekijät? Onko ehkäisy kustannus-vaikuttavaa verrattuna vaihtoehtoisiin toimintatapoihin (esimerkiksi ‘kriisi-interventiot’)?
<b>Vaihe 4. Toimeenpano pitkällä tähtäimellä</b>	
Ehdotetaan ohjelma- tai projektimalleja, joita kannattaa toistaa	Minkä tyyppiset interventiot näyttävät toimivan parhaiten, mille kohderyhmälle ja missä olosuhteissa?

## OSA III: SOVELTAMINEN KÄYTÄNTÖÖN

### 6 Tarkoituksen ja tavoitteiden selkeyttäminen

#### Arvioinnin keskeisiä näkökohtia

Taloudellinen arviointi käsittää määritelmämme mukaan intervention kustannusten ja vaikutusten systemaattisen arvioinnin. Taloudellisen arvioinnin keskeiset osatekijät on esitetty kuviossa 8.

Tässä luvussa keskustellaan arvioinnin alkuvaiheista tai valmistelevista vaiheista - kysymys on kahdesta ylimmästä laatikoista kuviossa 8. Luvuissa 7-10 tarkastellaan muita taloudellisen arvioinnin tekijöitä kuten kustannusten ja vaikutusten arviointia, tutkimusasetelmaa, analyysiä ja tulosten tulkintaa sekä arviointiprosessin läpi vientiä.

#### Arviointitutkimuksen kriteerit

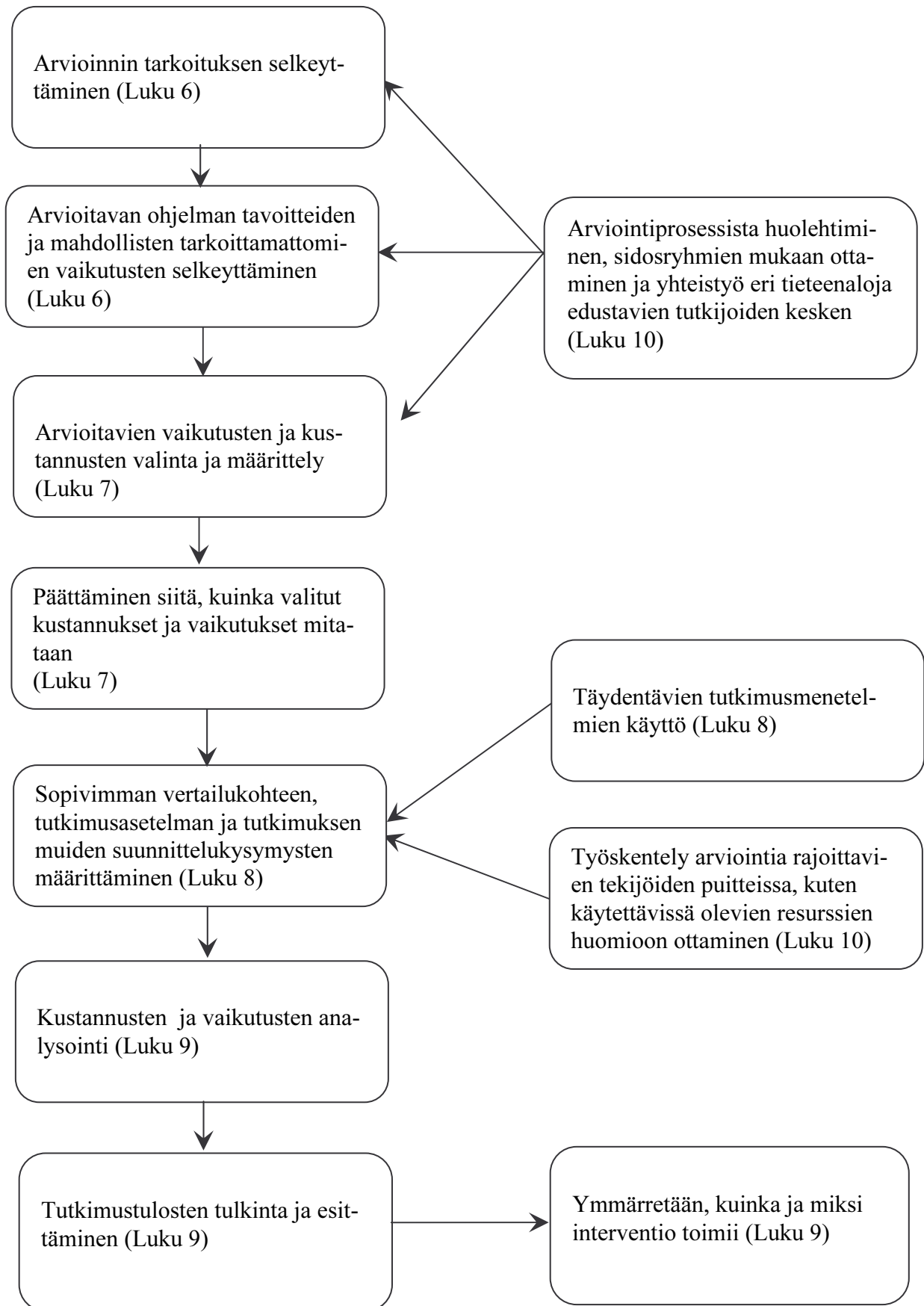
Ennen edellä mainittujen kysymysten käsittelyä on syytä tarkastella niitä kriteerejä, joiden perusteella arviointitutkimusta ja -asetelmaa arvioidaan:

- *Validiteetti* (pätevyys): miten virheettömän tiedon kustannuksista ja vaikutuksista arviointi antaa. Taloustieteilijät (ja muut kvantitatiiviseen arvioinnin tekijät) keskittyvät tavallisesti validiteetin tilastolliseen määrittelyyn eli siihen missä määrin käytetty menetelmä tarjoaa vinoutumattoman (unbiased) arvion kustannuksista ja vaikutuksista (esim. Moser ja Kalton 1993). Kvalitatiiviseen tutkimukseen keskittyneille tutkijoille validiteetti riippuu siitä, miten luotettavasti tutkimus välittää ihmisten kokemuksia interventioista (esim. National Centre for Social Research 2000).
- *Yleistettävyyden*: tässä on kysymys tutkimustulosten laajemmasta sovellettavuudesta muissa olosuhteissa tai muille asiakasryhmille. Kvantitatiiviset arviointitutkijat keskittyvät siihen, kuinka edustava tutkimuksen otos kohderyhmästä on tilastollisesti (esim. Moser ja Kalton 1993). Laajempi määrittely voi sisältää muunlaista tietoa, kuten tietoa siitä miksi eräät interventiot toimivat tietyissä olosuhteissa ja tietyillä ihmisryhmillä.
- *Relevanssi*: tässä on kysymys siitä, miten hyödyllisiä tulokset ovat päätöksentekijöille. Suurin osa taloudellisista arviointitutkimuksista on suunniteltu vastaamaan yksilöityyn kysymykseen: käyttääkö arvioitava interventio resursseja tehokkaasti verrattuna vaihtoehtoisin resurssien käyttötapoihin. Mutta päätöksentekijät voivat olla kiinnostuneita paljon laajemmista arviointikysymyksistä, kuten siitä, miten arvioitavaa ohjelmaa voitaisiin parantaa.

- *Toteutettavuus*: missä määrin ehdotettu arviointistrategia voidaan käytännössä panna toimeen, kun otetaan huomioon arviointia rajoittavat tekijät.

Kaikki nämä kriteerit ovat tärkeitä. Esimerkiksi ei ole paljon mieltä arvioida tarkasti ohjelman vaikutusta tietyissä olosuhteissa, jos olosuhteet poikkeavat siitä, missä ohjelma normaalisti toimeenpannaan. Samoin, ei myöskään ole kovin järkevää suunnitella hienostunutta arviointiasetelmaa, jos se tulisi liian kalliiksi tai ei vastaisi päätöksentekijöitä eniten askarruttaviin kysymyksiin.

**Kuvio 8.** Taloudellisen arvioinnin keskeiset vaiheet



## Arvioinnin tarkoituksen selkiyttäminen

Ennen arviointistrategian suunnittelua on tärkeää olla selvillä arvioinnin tarkoituksesta. Lopullisena tavoitteena on löytää interventio, joka tuottaa halutut vaikutukset ja tarjoaa vastinetta rahalle paremmin kuin toiset samoihin tavoitteisiin pyrkivät interventiot ja joka voidaan toistaa muualla. Kuitenkin arvioinnilla voi olla myös muita tärkeitä tavoitteita, etenkin ohjelman kehityksen alkuvaiheissa. Tällaisia ovat:

- dokumentointi, mitä ohjelmassa tehdään ja mihin kustannukset käytetään, jotta voidaan huolehtia tilivelvollisuudesta ohjelman rahoittajille
- tarkkailu, toimiiko ohjelma niin kuin ennalta ajateltiin, erityisesti kun mukana on useita toimijoita
- keinojen etsiminen ohjelman kehittämiseksi, tavoitteena karsia ‘las-tentauteja’ ja parantaa ohjelman vaikuttavuutta
- havainnoiminen, miten ohjelma edistyy tavoitteiden saavuttamisessa
- välineiden tarjoaminen pr-toiminnan käyttöön, jotta voidaan helpottaa ohjelman ‘myyntiä’ rahoittajille ja poliitikoille
- projektityöntekijöiden ja asiakkaiden motivointi auttamalla heitä näkemään aikaansaamansa kehitys.

Moni sosiaalialalla tehtävistä arvioinneista keskittyy edellä mainittuihin kysymyksiin (Newburn 2001). Vaikka taloustieteilijöillä voi olla vähemmän annettavaa näille formatiivisille arviointityypeille, tällaisella arvioinnilla on kuitenkin useita vaikutuksia taloudelliseen arviointiin. Ensinnäkin, jotkut näistä voivat tuottaa aineistoa taloudelliselle arvioinnille, minkä vuoksi taloustieteilijöiden olisi oltava mukana jo varhaisessa vaiheessa. Esimerkiksi informaation keräyssysteemit, joita pystytetään seuranta- tai kirjanpitar-koituksiin, voivat myös tarjota taloudellisen arvioinnin tarvitsemää tietoa. Samoin, ohjelman dokumentointiprosessi ja sen keskeisten osatekijöiden tunnistaminen voi auttaa seurattavien vaikutusten valinnassa.

Toiseksi, on tärkeää tunnistaa, että arvioijan eri roolit voivat olla keskenään ristiriitaisia. ”Safe in The City” -tutkimuksessa (Turvapaikka kaupungissa; nuorten kodittomuuden ehkäisy -hanke), yksittäiset projektit katsoivat arvioinnin hyödylliseksi keinoksi esitellä asiakkaille aikaan saamaansa kehitystä. Mutta jos arviointia käytetään tällä tavoin, se todennäköisesti suosii myönteisten vaikutusten rekisteröintiä, mikä vääristää tuloksia. Sen vuoksi olisi tarpeen pitää tämä prosessi erillään taloudelliseen arviointiin tarvittavan aineiston keräämisestä. Safe in The City -projektin Lontoon alueen koordinointiorganisaatiolla oli erilainen näkökulma. Vaikka se myöhemmin on muuttanut näkemystään, se katsoi aluksi, että oli keskityttävä mittaamaan vaikutusta nuorten asunnottomuuteen, koska asunnottomuus oli ohjelman ‘ylätason’ tavoite, vaikka mittaamiseen liittyikin ongelmia (katso Taulukko 4 Luvussa 7).

Kolmanneksi, on tärkeää, että taloustieteilijät tunnustavat edellä mainittujen muiden tavoitteiden olemassaolon ja hyväksyvät sen, että niihin on kiinnitettävä huomiota taloudellisen arvioinnin rinnalla. Siten taloudellinen arviointi olisi nähtävä vain yhdeksi arviointistrategian muodostaman kokonaisuuden osaksi. Yhtenä ongelmana on tietää, milloin taloudellista arviointia kannattaa toteuttaa. Ei ole todennäköisesti tarkoituksenmukaista toteuttaa kontrolloitua koetta ohjelmassa, joka on kehittelyn alkuvaiheessa ja jota muutetaan jatkuvasti. Samalla on pidettävä mielessä, että uusien ohjelmien kustannus-vaikuttavuutta on tärkeää arvioida niin aikaisessa vaiheessa kuin mahdollista, jotta voidaan havaita ovatko ohjelmat jatkamisen arvoisia tai hylättäviä.

## **Arvioitavan ohjelman tavoitteiden selkiyttäminen**

Suurin osa arvioinnista, kuten myös taloudellinen arviointi, on tavoiteperusteista. Oletetaan, että kaikilla ohjelmilla pitäisi olla selvästi määritellyt tavoitteet, jotka asetetaan etukäteen ja joita vasten niiden toteutumista voidaan arvioida (esim. HM Treasury 1997). Mutta näin ei useinkaan ole laita sosiaalisen hyvinvoinnin alueella:

- Tavoitteet eivät välttämättä ole kovin selkeitä tai ne voidaan ilmaista hyvin yleisesti.
- Tavoitteet voivat muuttua ajan kuluessa, usein pikkuhiljaa ohjelman kehittyessä.
- ‘Operatiiviset’ tavoitteet, jotka heijastavat tapaa, jolla ohjelma toimii käytännössä, voivat erota ennalta ilmaistuista tavoitteista, erityisesti jos eri toimijat osallistuvat ohjelman toteuttamiseen.
- On mahdollista, että tavoitteita on useita ja ne ovat keskenään ristiriitaisia, esimerkiksi siksi, että mukana on erilaisia toimijoita. Aina ei myöskään ole selvä, mikä on eri tavoitteiden tärkeysjärjestys.
- Jotkut ohjelmat ovat tarkoituksellisesti joustavia ja kehittyviä. Usein perusteluna on se, että ohjelmien on kyettävä mukautumaan paikallisiin tarpeisiin, ja paikalliset toimijat on saatettu kutsua esittämään paikallisia ratkaisuja paikallisiin tarpeisiin. Rahoittaja ei ohjaile yksittäisten projektien suunnittelua, itse asiassa ne voivat pyrkiä rohkaisemaan monimuotoisuuteen edistääkseen innovatiivisia toimintatapoja (Martin ja Sanderson 2000).

Jotkut taloustieteilijät voivat pitää huonosti määriteltyjä tavoitteita heijastumana heikosta toiminnan suunnittelusta eikä niinkään luonnollisena seurauksena monimuotoisesta päätöksentekoprosessista. Toisaalta tavoitteisiin keskittynyt arviointi myöntää, että päämäärät tai tavoitteet ovat harvoin niin selviä, kuin monet arvioijat toivoisivat. Tavoitteiden selkiyttämisestä on tehty keskeinen osa arviointia (siis päämäärä sinänsä) eikä vain lähtökohta (Peled ja Spiro 1998). Tavoitteiden selkiyttäminen on toistuva (iteratiivinen) ja vuorovaikutteinen prosessi, ja sen avulla tunnistetaan ennalta päätetyt tai ilmaistut tavoitteet. Niitä verrataan toimintaan osallistuvien kanssa keskusteltaessa esille tuleviin operatiivisiin tavoitteisiin, ja tavoitteita tarkistetaan,

jos se on tarpeen. Vasta tämän jälkeen arvioinnissa voidaan edetä tarkastelemaan tavoitteiden saavuttamista.

Edellä esitetty ei kiistä selkeästi määriteltyjen tavoitteiden tärkeyttä taloudellisen arvioinnin ‘kulmakivinä’, mutta sillä on vaikutusta tapaan, jolla taloustieteilijät ajattelevat tavoitteista. Monissa sosiaalialan interventioissa tavoitteet eivät ole ennalta ‘annettuja’; itse asiassa, tavoitteiden selkiyttämisprosessi on perustavanlaatuinen ja olennainen osa arviointitutkimusta. Vaikka vastuu ohjelman tavoitteiden selkeästä ilmaisemisesta on loppujen lopuksi niillä, jotka vastaavat ohjelman toimeenpanosta, arvioijat voivat, ja heidän pitääkin, avustaa tässä prosessissa (Shiell ja Hawe 1996). Tarkenteessa 3 tarkastellaan näitä kysymyksiä kahden tapaustutkimuksen osalta.

### **Tarkenne 3. Tavoitteiden selkiyttäminen**

#### **Turvapaikka kaupungissa –tapaustutkimus (Safe In The City)**

Yksi syy siihen, miksi tavoitteiden selkiyttäminen oli vaikeaa hankkeen uutta projektiryhmiin perustuvaa vaihetta suunniteltaessa, oli hankkeen koostuminen 24:stä eri projektista omine tavoitteineen. Tämä ei ole yllättävää, kun otetaan huomioon, että projektien toteutuksesta vastasivat riippumattomat toimijat, eikä yhtenäisyyteen pyrkivää koordinaatiojärjestelmää ollut vielä käytössä. Tutustuimme yhdeksään päätyypit kattavaan projektiin (työllisyys ja osaaminen, persoonallinen kehittyminen ja perheen tukeminen) saadaksemme selville, mitä niissä haluttiin saada aikaan ja millaisia vaikutuksia meidän sen vuoksi pitäisi pyrkiä arvioimaan.

Oli mielenkiintoista, että projektityöntekijät eivät puhuneet kovinkaan paljoa nuorten asunnottomuuden ehkäisemisestä. Asunnottomuuden ehkäisy ei sisällynyt myöskään heidän omiin arvioihinsa projektien vaikuttavuudesta. Vaikutukset keskittyivät lyhyen tähtäimen tavoitteisiin. Näitä olivat nuorten auttaminen ottamaan enemmän vastuuta omasta elämästään ja perheiden rohkaiseminen käsittelemään ongelmiaan.

Ensisilmäyksellä eri projektien tavoitteet näyttivät eroavan huomattavasti. Keskustelut projektityöskentelijöiden kanssa auttoivat havaitsemaan merkittävää yhdenmukaisuutta siinä, mihin projektit pyrkivät, vaikka ne toimivat erilaisin keinoin. Esimerkiksi työllisyyteen ja osaamiseen kohdistuvat projektit tunnistiivat persoonallisen kehityksen (esim. itsearvostus) tärkeäksi projektin tulokseksi ja myös välineeksi muiden vaikutusten (esim. parempien työllistymismahdollisuuksien) aikaan saamisessa. Erot projektien välillä liittyivät siis enemmän painotuksiin kuin sisältöihin. Monitoimijainen rakenne ja viimeaikaiset sovitut toimintatavat lisäävät yhdenmukaisuutta ohjelman tavoitteissa.

## **Lämmityspolttoaineen puute –tapaustutkimus (Fuel poverty)**

HEES:n (The Home Energy Efficiency Scheme; kotien energiatehokkuussuunnitelma) HEES:n tavoitteena on lämmityspolttoaineen puutteessa olevien kotitalouksien lukumäärän vähentäminen. Mutta lämmityspolttoaineen puutteen määrittely on josain määrin epäselvää, ja sen vuoksi epäselvä on myös HEES:n tavoite. Yleisin määritelmä lämmityspolttoaineen puutteelle ilmaistaan sen avulla, onko lämmitykseen varaa. Toisin sanoen: lämmityspolttoaineen puutteessa olevalla ei ole mahdollisuutta lämmittää kotiaan riittävästi, ilman että lämmitykseen joutuu käyttämään enemmän kuin 10% kotitalouden tuloista. Merkittävä ero on kuitenkin siinä, miten paljon kotitalouden pitäisi käyttää rahaa asunnon riittävään lämmitykseen ja miten paljon se tähän tarkoitukseen todella kuluttaa. Eräät kotitaloudet ovat määritelmän mukaan lämmityspolttoaineen puutteessa, mutta pitävät kotinsa lämpimänä ja käyttävät lämmittämiseen suuren osan nettotuloistaan. Toisilla taas olisi vara lämmittää kotejaan (käyttämättä enempää kuin 10% tuloistaan), mutta ne eivät silti lämmitä riittävästi. Siis lämmityspolttoaineen puute ei (määritelmän perusteella) ole synonyymi kylmälle asunnolle, vaikka näiden välillä onkin merkittävää päällekkäisyyttä. Hallitus keskittyy vähentämään polttoaineen puutteessa olevien haavoittuvimpien kotitalouksien lukumäärää. Mutta hallitus ja muut toimijat ovat terveyshaittojen ja, äärimmäisissä tapauksissa, kylmiin talviin liittyvien kuolemantapausten vuoksi selkeästi huolestuneita myös kylmissä asunnoissa asuvien määrästä (joista joillakin olisi varaa lämmittää).

Lisäksi HEES:llä on mahdollisia, erityisesti hiilidioksidipäästöjen vähentämiseen liittyviä ympäristöhyötyjä. Vaikka näitä ei oltu mainittu hallituksen HEES –ohjelmaa käsittelevässä mietinnössä (Department of the Environment, Transport and the Regions, 2000), energiatehokkuus on tärkeä osa hallituksen ilmastomuutosstrategiaa. Energiatehokkuuden parantuminen palvelee sekä sosiaalisten että ympäristöllisten tavoitteiden saavuttamista: kotitaloudet pystyvät lämmittämään asuntojaan paremmin ja samanaikaisesti vähentävät lämmitysenergian kuluusta tai siirtyvät ympäristöystävälliseen energianlähteeseen. Kuitenkin, näiden kahden tavoitteen välillä on mahdollinen ristiriita, koska suurin energian säästö saataisiin hyvätuloisilta kotitalouksilta, joiden energiankulutus on suurinta. Siksi ohjelman kohderyhmä riippuu erilaisten poliittisten tavoitteiden painotumisesta.

Taloudellisen arvioinnin pitäisi tarkastella ohjelman kaikkia tavoitteita sekä näiden välisiä mahdollisia ristiriitoja. Vaikka ympäristöhyödyt eivät kuuluneet HEES:n julki lausuttuihin tavoitteisiin, ne olisi silti otettava huomioon ohjelman täydellisessä taloudellisessa arvioinnissa.





## **7 Vaikutusten ja kustannusten arviointi**

### **Vaikutusten valitseminen**

Arvioitavan ohjelman tavoitteiden pitäisi olla tarkkailtavien vaikutusten pääasiallinen määrittäjä (katso Luku 6). Mutta on olemassa kaksi tärkeää seikkaa, joissa taloudellinen arviointi eroaa puhtaasti tavoiteperusteisesta arvioinnista. Ensinnäkin, kuten jo aiemmin on esitetty, taloustieteilijät suosivat, mikäli mahdollista, lopullisten vaikutusten tarkastelua (esimerkiksi kun käytetään elämänlaatumittareita), kun taas ohjelman tavoitteet on voitu ilmaista muulla tavoin. Esimerkiksi lähiöhoitoprojektin tavoitteet voivat kohdistua välituloksiin (esim. nopeampiin korjauksiin) tai jopa prosessia kuvaaviin toimiin (esim. asukkaiden kuulemistapoihin), eikä niinkään lopulliseen vaikutukseen, joka olisi asukkaiden elämänlaadun paraneminen.

Toiseksi, taloudellinen arviointi pyrkii olemaan kattava, minkä vuoksi on tärkeää tarkastella sellaisiakin vaikutuksia, ehkä odottamattomia, joita ei ole sisällytetty tavoitteisiin. Tämä pitää sisällään vaikuttamisen muihin, kuin pääasialliseen kohderyhmään. Tämä on tunnistettu hyvin joillakin sosiaalisen hyvinvoinnin alueilla. Esimerkiksi rikollisuuden torjunnassa yleinen käytäntö on tarkastella torjuntatoimenpiteiden seurannaisvaikutuksia lähi-alueille (esim. Welsh ja Farrington 1999). Samoin tuen-piiristä-työhön -ohjelmien arvioinnissa on vakiintunut käytäntö tarkastella syrjäyttämisen- ja sijaisvaikutuksia, eli sitä, että jotkut ohjelmaan osallistujat työllistyvät muiden työhakijoiden kustannuksella (Gardiner 1997). Hoivasektorilla on jo kauan tarkasteltu niin hoivaajiin kuin asiakkaisiinkin kohdistuvia vaikutuksia (esim. Davies ja Challis 1986; Knapp ym. 1992).

### **Vaikutusten arvioitavuus**

Mitattavia vaikutuksia valittaessa on kuitenkin myös syytä ottaa huomioon käytännöllisiä näkökohtia. Aina ei ole tarkoituksenmukaista valita tarkasteltavaksi 'ilmeisintä' vaikutusta: tämän mittaaminen tarkasti voi olla vaikeaa tai muutosten odotetaan ilmenevän vasta pitkän ajan kuluttua - myöhemmin, kuin mikä on realistista arviointiaikataulun mukaan. Lisäksi intervention vaikutuksia voi olla vaikea erottaa muista ulkoisista vaikutuksista, erityisesti jos odotettu vaikutus on suhteellisen pieni. Näistä syistä voi olla tarkoituksenmukaista valita muita vaikutuksia, esimerkiksi väli- tai lyhyen tähtäimen vaikutuksia, joihin interventiolla on suurempia yhteyksiä. Joitakin tapaustutkimuksissamme esiintyneitä esimerkkejä tarkastellaan tarkenteessa 4.

### **Paikallisyhteisöjen kehittäminen –tapaustutkimus (Gatsby)**

Tämä tapaustutkimus havainnollistaa monimutkaisen ohjelman arviointiin soveltuvien vaikutusten valintaprosessia. Hankkeen budjetista käytetään puolet yhteisöryhmien asukaskoulutukseen, jota pidetään hyödyllisenä ja joka on edellytys pääsille mukaan hankkeeseen. Rahoituksen avulla käynnistetyt projektit ovat osa ryhmien oppimisprosessia, minkä vuoksi tuki kohdistetaan melko uusille ryhmille ja tuki on pitkäkestoista. Ohjelmaan liittyy siis vahva ryhmien suorituskyvyn parantamiselementti, mikä on otettava huomioon arvioitavien vaikutusten valinnassa.

Tarkastelimme kolmea tasoa, joilla ohjelmaa voitaisiin arvioida:

- sosiaaliset ja taloudelliset vaikutukset paikallisyhteisöön
- yksittäisten, ohjelman kautta rahoitettujen projektien onnistuminen
- yhteisöryhmien vaikuttavuus

Lupaavimmaksi näistä arvioitiin kolmas vaihtoehto. Yhteisöryhmiin kohdistuva vaikutus on välitön, ja siksi se pitäisi olla helpompi (vaikkei silti helppo) tunnistaa kuin laajempaan yhteisöön kohdistuvat vaikutukset. Nämä vaikutukset voivat olla liian pieniä, jotta niitä voitaisiin mitata tai erottaa muista tekijöistä. Ryhmien vaikuttavuus kattaa myös kaksi ohjelman keskeistä elementtiä; koulutus vaikuttaa ryhmän taitoihin, sitoutumiseen sekä tulevaisuudenkuviin ja avustukset auttavat ryhmiä toteuttamaan ideoitaan käytännössä. Avustukset lisäävät ryhmän tietoja ja taitoja myös välillisesti projektien toteuttamisen avulla (tekemällä oppimista). Lisäksi, koska prosessin merkityksellisyys tunnustetaan, lähestymistapa on yhdenmukainen yhteisön kehittämisen periaatteiden kanssa (autetaan paikallisyhteisöjä auttamaan itseään parantamalla yhteisöryhmien suorituskykyä).

### **Turvapaikka kaupungissa –tapaustutkimus (Safe In The City)**

Tämä tapaustutkimus on hyvä esimerkki siitä, miksi vaikutusten valinta ei ole niin yksinkertaista, kuin aluksi saattaa näyttää. Ilmeinen arvioitava vaikutus tässä ohjelmassa olisi asiakkaiden asunnottomuuden väheneminen, joka on ohjelman ”ylätason” tavoite. On kuitenkin monia syitä, miksi asunnottomuuden väheneminen ei sovellu arvioitavaksi vaikutukseksi:

- Asunnottomuuteen kohdistuvan vaikutukset sekä toteutuvat pitkällä aikavälillä että ovat epävarmoja. Tästä syystä on vaarana, että tähän vaikutukseen keskittyvä tutkimus ei pysty havaitsemaan yhtään merkittävää muutosta todennäköisen arviointiajan (2-3 vuotta) puitteissa.
- Vaikka arviointiaikaa pystyttäisiin pidentämään riittävästi, voi olla erittäin vaikea selvittää aikaisempien asi-

akkaiden tilannetta. Tapausten vähenemisen johdosta tutkimuksen validiteetista (pätevyydestä) saatetaan joutua tinkimään melkoisesti. Näin käy erityisesti, jos ne joilla on eniten vaikeuksia asunnon säilyttämisessä, ovat vaikeimmin jäljitettävissä.

- Vaikka nämä ongelmat voitaisiin ratkaista, ohjelman vaikutus asunnottomien nuorten määrään voi olla liian pieni havaittavaksi. Pieni vaikutus ei tarkoita sitä, että ohjelma ei olisi toteuttamisen arvoinen. Pienikin vaikutus voi olla riittävä perustelemaan ohjelman toteuttamisen, jos nuorten asunnottomuudesta aiheutuvat kustannukset ovat suuria ja pitkäkestoisia. Mutta melko laaja-alaan aineistoon (esim. 300-500 vastaajaa) perustuva arviointitutkimus ei kykenisi havaitsemaan vaikutuksia.
- Viimeisimpänä, mutta ei vähäisimpänä, mielestämme olisi viisasta, että kaikkea huomiota ei keskittäisi vain yhden vaikutuksen mittaamisen, erityisesti kun otetaan huomioon tavoitteiden moninaisuus projekteissa, joissa vierailimme.

Edellisistä syistä olisi mielestämme liian uskaliaasta keskittyä pelkästään asunnottomuuteen tai pitää sitä edes keskeisimpänä ohjelman vaikutuksia mitattaessa. Sen sijaan suosittelimme muutamien lyhyemmällä aikavälillä toteutuvien vaikutusten mittaamista. Tällaisia ovat esimerkiksi nuorten emotionaalisten ja käytösongelmien sekä perheen sisäisten ristiriitojen väheneminen ja nuoren elämänlaadun paraneminen, vaikkakin on järkevää tarkastella myös asumiseen liittyviä vaikutuksia.

### **Matalan kynnyksen toimipiste (Revolving Door Agency)**

Hieman erilainen on tilanne silloin, kun arvioitava interventio on yksi monista samaan kohderyhmään vaikuttavista interventiosta. Esimerkiksi Revolving Door Agency (RDA) pyrkii parantamaan poliisin kanssa tekemisissä olleiden ja kaoottista elämää elävien ihmisten pääsyä mielenterveyspalvelujen piiriin. (Revolving Door Agency, 2000). Lopullinen tavoiteltu vaikutus on asiakkaiden mielenterveyden parantuminen, mutta siihen vaikuttaa myös paikallisten sosiaali-, terveys- ja asumispalvelujen laatu, ja nämä ovat pääosin RDA:n vaikutusmahdollisuuksien ulkopuolella. Ohjelman vaikutus asiakkaiden mielenterveyteen onkin välillinen. Niinpä RDA:n onnistumista pitäisikin ilmeisesti arvioida mieluummin sen perusteella, miten voimakkaasti se rohkaisee suurempaan tai paremmin soveltuvien palvelujen käyttöön kohteena olevassa asiakasjoukossa. Näihin asioihin sen vaikutus on välitön.

Tästä valinnasta seuraa kuitenkin uusi ongelma. Samalla kun projektin yhteyshenkilöt ovat rohkaisseet asiakkaitaan käyttämään itselleen sopivia julkisia palveluja, ovat projektin yhteyshenkilöt ottaneet myös ohjausroolin, ja suhde joihinkin asiakkaisiin on muodostunut läheiseksi. Tämän seurauksena ohjelmalla on myös välittömiä vaikutuksia asiakkaiden mielenter-

veyteen, ja nämä jäävät huomiotta, jos arvioinnissa selvitetään vain muutosta palvelujen käytössä.

### **Useita vaikutuksia**

Sosiaaalialalla on yleensä tarkoituksenmukaista analysoida useita vaikutuksia seuraavista syistä:

- Joillakin ohjelmilla on monia tavoitteita.
- Tavoite voi olla moniulotteinen, kuten ikääntyneiden hyvinvoinnin parantaminen. Tämän vuoksi voi olla helpompaa jakaa tavoite osatekijöihin (esim. liikkuvuuteen, itsenäisyyteen, ulkonäköön, sosiaalisuuteen).
- Interventiot voivat vaikuttaa eri tasoilla. Voi olla tärkeää arvioida vaikutusta sekä yksilöihin että yhteisöihin tai organisaatioihin, jossa interventio toteutetaan. Näin saadaan selville ohjelman kokonaisvaikutus.
- Monien vaikutusten tarkkaileminen on hyödyllistä tutkimuksen pätevyuden (validisuuden) osoittamiseksi erityisesti, jos on epäilyksiä vaikutusmittareiden luotettavuudesta (reliabiliteetista). Vastaavista syistä on joskus hyödyllistä tarkastella samaa vaikutusta eri näkökulmista. Esimerkiksi, jos seurattava vaikutus ilmenee nuoren henkilön käyttäytymisessä, kannattaa ehkä kysyä nuorten oman käsityksen lisäksi vanhempien, opettajien tai sosiaalityöntekijöiden mielipidettä.
- Kun arvioidaan mahdollisia lumipallo- tai leviämisaikutuksia, voidaan käyttää monia vaikutuksia. Tämä lisää arvioinnin kokonaisvaltaisuutta.
- Välittäviä tai prosessiin liittyviä vaikutuksia voidaan seurata siksi, että ne auttavat tulkitsemaan varsinaisissa vaikutuksissa tapahtuvia muutoksia, erityisesti jos niiden väliset yhteydet tunnetaan (tai yhteyksille on teoreettisia perusteita).

Tarkenne 5 osoittaa, kuinka useita vaikutuksia on käytetty joukossa julkais-  
tuja tutkimuksia.

## **Tarkenne 5.** Useiden vaikutusten käyttö tutkimuksissa

### **Esimerkki 1**

Kysymyksessä on ala- ja yläasteella tapahtuvaa koulusosiaali-  
työtä koskeva kokeellinen tutkimus. Tietoja kerättiin monista  
muuttujista, kuten koulupinnauksesta, itse raportoiduista varka-  
uksista, kiusaamisesta, huumeiden käytöstä, ilkeistä sekä  
käytöshäiriöstä johtuvista koulusta erottamisista. Monien vaiku-  
tusmittareiden käyttäminen on tarkoituksenmukaista, koska tut-  
kittavan projektin tavoitteena oli vaikuttaa kaikkiin edellä mai-  
nittuihin asioihin. Myös se, että tutkijat pystyivät osoittamaan  
positiivisia muutosta kuudessa muuttujassa seitsemästä, auttoi  
lisäämään tulosten pätevyyttä (validisuutta). Tulos oli paljon  
vakuuttavampi kuin jos tutkimuksessa olisi valittu vain yksi  
edellä mainituista muuttujista, jossa olisi osoitettu myönteistä  
kehitystä.

(Lähde: Bagley ja Pritchard, 1998)

### **Esimerkki 2**

Tämä taloudellinen arviointi vertailee ns. Darlingtonin-mallia,  
jossa tarjottiin intensiivistä ja koordinoitua tukea huonokuntoi-  
sille vanhuksille avohuoltona, ja pitkäaikaista laitoshoidon (sai-  
raalahoitoa). Tällaisella interventiolla on todennäköisesti vaiku-  
tuksia moniin vanhuksen elämänlaadun ulottuvuuksiin. Siksi oli  
tärkeää tarkastella monia vaikutuksia, ja käyttää sekä psykolo-  
gisia (esim. yleinen elämäntyytyväisyys, masennus, yksinäi-  
syys) että käyttäytymistä kuvaavia muuttujia (esim. fyysinen  
vajaakuntoisuus, apatia, kommunikaatiovaikeudet). Myös kuol-  
leisuusastetta seurattiin molemmissa potilasryhmissä. (Lähde:  
Challis ym., 1995)

### **Esimerkki 3**

Arvioinnissa, jossa tarkasteltiin aboriginaalien vastasyntyneiden  
seulontahanketta, katsottiin suorien terveysvaikutusten lisäksi  
myös muita tapoja, joilla interventio saattoi vaikuttaa aborigi-  
naaliyhteisöön. Näitä ovat mm. positiiviset muutokset käsityk-  
sissä, joita yhteisöllä on terveydenhoitopalveluja kohtaan, esim.  
halukkuus lääkärikäynteihin. Näitä epäsuoria vaikutuksia pidet-  
tiin sinänsä arvokkaina, koska niiden oletettiin parantavan yh-  
teisön edellytyksiä hyötyä terveyspalveluihin investoiduista re-  
sursseista. (Lähde: Jan, 1998)

### **Esimerkki 4**

Arvioinnin kohteena on Brittiläinen tuen-piiristä-työhön –hanke  
(RESTART), joka tarjoaa neuvontaa ja jatkuvaa tukea työttö-  
mille. Vaikka pääasiallisena vaikutusmittarina oli korvausha-  
kemusten väheneminen, informaatiota kerättiin myös useista  
muista muuttujista. Näihin kuuluivat työn laatu, jota kuvattiin  
palkkojen avulla, ja prosessia kuvaavat mittarit, esim. työhake-

musten määrä työnhakuaktiivisuuden kuvaajina. Edellä mainitut auttoivat tulkitsemaan tutkimuksen tuloksia ja tekemään erilaisia oletuksia siitä, miten hanke vaikutti, esim. työnsikö se ihmisiä alempiarvoisiin töihin. (Lähde: White ja Lakey, 1992)

Taloustieteellisestä näkökulmasta, monien vaikutusten ottaminen tarkasteluun ei yleensä mahdollista interventioiden järjestämistä siististi kustannusvaikuttavuuden mukaan. Tämä on kuitenkin mahdollista, jos yksi interventio ei ole muita kalliimpi ja jos se suoriutuu paremmin kaikilla tulostuloksilla mitattuna (tai on halvempi ja eikä toimi huonommin minkään vaikuttavuusmittarin mukaan). Terveystieteellisessä on tavanomaista nimetä ennalta ensisijainen vaikutus, mikä mahdollistaa kustannusvaikuttavuuden analyysin tekemisen. Mutta tämä on soveliaista vain, jos muita suuremman painoarvon antamiselle tietyille vaikutukselle on hyvät perustelut.

### **Vaikutusten täsmentäminen**

Ennen kuin siirrytään vaikutusten mittaamiseen, voi olla tarpeellista rajata vaikutukset tarkemmin. Esimerkiksi, mitä tarkasti ottaen tarkoitetaan 'paremmalla' tai 'varmemmalla' työpaikoilla tuen-piiristä-työhön ohjelmissa. Mitä tarkasti ottaen tarkoitetaan paremmalla kouluympäristöllä kouluperusteinen sosiaalityö -intervention tapauksessa?

Joissakin tapauksissa täsmälliset vaikutukset saadaan suoraan ohjelman tavoitteista, jos tavoitteet ovat selkeät ja riittävän yksityiskohtaiset. Esimerkiksi kouluperusteisen sosiaalityön ohjelma tarkensi tavoitteitaan siten, että tavoitteena on vähentää kiusaamista ja syrjäytyneiden määrää. Joissakin muissa tapauksissa vaikutukset on määriteltävä alusta alkaen. Gatsby-tapaustutkimuksessa meidän piti selvittää, mikä on vaikuttava yhteisöryhmä, ennen kuin pystyimme arvioimaan ryhmiä näiden kriteereiden avulla. Vaikutusten täsmentämisprosessiin liittyy keskustelu toimintaan osallistuvien kanssa, erityisesti niiden kanssa, jotka ovat vastuussa ohjelman tavoitteiden asettamisesta. Gatsby-tutkimuksessa kävimme läpi myös kirjallisuutta löytääksemme näyttöä ja teoriaa vaikuttavista yhteisötason organisaatioista.

### **Vaikutusten mittaaminen**

Jotkut vaikutukset sopivat paremmin mitattavaksi kuin toiset. Esimerkiksi arvioitaessa tuen-piiristä-työhön -ohjelmia Iso-Britanniassa mittana on nettovaikutus työllisyyteen, kun taas rikollisuuden torjuntaan liittyvissä ohjelmissa mitataan eri tyyppisen rikollisuuden vähenemistä. Mutta edes näissä tapauksissa vaikutusmittarit eivät ole itsestään selviä. Esimerkiksi tuen-piiristä-työhön -ohjelmien arvioinnissa pitäisi ihanteellisessa tapauksessa ottaa huomioon työpaikkojen laatu (eli palkka, työsuhteen varmuus) yhtä hyvin kuin luotujen työpaikkojen lukumäärä. Myös mittaamisongelmia voi olla. Tämä huoli liittyy erityisesti rikostilastointiin, jossa virallisten tilastojen tiedetään aliarvioivan rikollisuutta. Jos asiaa kysytään suoraan rikosten tekijöiltä, heillä on selkeät kannustimet kertoa todellisuutta pienempi rikosten määrä. Lisäksi tuloksia vääristää, jos kerättäviä vaikutusmittareita käytetään

tään myös organisaation suoriutumisen arviointiin. Vääristymiä voi syntyä erityisesti sellaisissa tapauksissa, kun asioiden kirjaamisessa tilastoihin on vapausasteita. (Esimerkiksi, mikäli poliisilla on paineita vähentää rotuenakkoluuloihin perustuvaa rikollisuutta, tällaisia rikoksia kirjataan poliisin rekisteriin todennäköisesti vähemmän.) Jopa silloin, kun 'kovaa' tietoa on saatavilla, nämä voivat olla paremminkin indikaattoreita kuin vaikuttavuuden mittareita. On torjuttava kiusaus käyttää mitä tahansa saatavilla olevaa tietoa, jos tieto ei kuvaa hyvin arvioitavaa vaikutusta.

Monia vaikutuksia ei ole luonnostaan helppo mitata. Näitä kuvataan toisiinsa 'pehmeiksi' vaikutuksiksi (esim. vanhempien vaikutus lapseen, ikään-tyneen hyvinvointi, nuoren käyttäytyminen). Monilla sosiaalialalla työskentelevillä, mukaan lukien muut kuin määrällisen lähestymistavan arviointitutkijat, näyttää olevan yleinen käsitys, jonka mukaan tällaisia vaikutuksia ei voi mitata. Määrälliseen lähestymistapaan perustuva arviointi onkin usein hylätty tällä perusteella. Jopa termi 'pehmeä' viittaa siihen, että tällaiset vaikutukset ovat jotenkin vähemmän todellisia kuin 'kovat' vaikutukset. Esimerkiksi, SITC -ohjelmassa todettiin, että 'pehmeät' vaikutukset, kuten asiakkaiden itsearvostuksen lisääntyminen, olivat tärkeitä ja näitä haluttiin seurata; tuntui kuitenkin siltä, että poliitikot halusivat tietää pääasiassa kovista vaikutuksista, kuten nuorten asunnottomuuden alenemisluvuista, jota alun perin oli painotettu.

Näkemyksemme on, että monia näistä vaikutuksista voidaan mitata jollain tavalla (katso seuraava tarkenne), mikä mahdollistaa kaikenlaisten interventioiden arvioinnin määrällisesti. Mitattaessa hyvin monimutkaisia vaikutuksia on kuitenkin tarpeen olla käytännöllinen. Nämä vaikutukset voivat vaatia laadullista arviointia, mutta niitä ei saa jättää huomiotta.

### **Vakiintuneet vaikutusmittarit**

Joillakin tieteenaloilla, kuten psykologiassa, on vahvat perinteet määrällisten mittareiden käytössä arvioitaessa vaikutuksia, jotka eivät luonnostaan sovellu mitattaviksi. Tällaisia vaikutuksia ovat esimerkiksi perheen toiminta, masentuneisuus tai itsearvostus (esim. Department of Health 2000). Mittarit perustuvat tavallisesti kyselyihin, jotka voidaan pisteyttää numeerisesti. Tarkenteessa 6 on esimerkki tällaisesta mittarista (Strengths and Difficulties Questionnaire, SDQ), jota on käytetty sijaishuoltoa koskevassa tapaustutkimuksessa. Jos vakiintunutta vaikuttavuusmittaria ei ole käytössä, voidaan kehittää uusi mittari. Tarkenteessa 6 on esimerkki arviointia varten kehitettyä työkalusta, jonka kokosimme Gatsby -tapaustutkimuksen yhteydessä helpottamaan yhteisöryhmien vaikuttavuuden arviointia (katso Esimerkki 2). Tämän kirjan II osassa tarjotaan enemmän tietoa vaikuttavuusmittareiden kehittämisestä ja esitetään lisäesimerkkejä olemassa olevista mittareista, jotka voivat olla hyödyllisiä sosiaalialalla (Byford ym., 2003).

Tehtäessä täsmällisiä kokeellisia arviointitutkimuksia tulisi lähtökohtana olla se, että vakiintuneita vaikuttavuusmittareita, kuten SDQ-mittaria, käytetään aina, kun se on mahdollista. Tällaiset mittarit on testattu monilla tavoil-



la, mikä auttaa parantamaan tulosten pätevyyttä (validisuuden). Vakiintuneiden mittareiden käyttö lisää myös tutkimusten vertailtavuutta.

Vakiintuneita mittareita on kuitenkin käytettävä varovaisuutta noudattaen. Arvioijan on tärkeää olla varma, että mittarit mittaavat oikeaa asiaa. Esimerkiksi jotkut mittarit, kuten EuroQol, on kehitetty mittaamaan terveyteen liittyvää elämänlaatua ja ne keskittyvät enemmän fyysiseen kuin psyykkiseen tai sosiaaliseen toimintaan (EuroQol Group 1990). Vaikka pätevyys mitata terveyteen liittyviä muutoksia on testattu, mittarit eivät välttämättä sovellu mittaamaan muutosta laajemman sosiaalisen hyvinvoinnin näkökulmasta. Sopivien mittareiden puute voi vaikeuttaa monimutkaisten mielenterveys- ja sosiaalialan interventioiden vaikutusten mittaamista. Suurin rajoite näyttää kuitenkin olevan kelvollisia mittausvälineitä koskevan tietoisuuden ja ymmärryksen puute. Safe in The City -tapaustutkimuksessa oltiin huolissaan siitä, että vakiintuneet mittarit kadottaisivat tärkeitä vaikuttavuuden piirteitä, joten projektityöntekijöiden kanssa neuvottelemalla rakennettiin arviointia varten oma pehmeä indikaattoripaketti. Käytännössä paras valinta olisi luultavasti se, että käytettäisiin molempien mittarityyppien yhdistelmää, jonka ominaisuudet määräytyvät sen mukaan, mitkä ovat resursit ja minkä verran tietojen keruuseen voidaan käyttää aikaa. SITC:n tapaustutkimuksessa oltiin erityisen huolissaan siitä, ettei asiakkaita rasitettaisi liian monilla kysymyksillä, erityisesti intervention alkuvaiheessa.

Toinen ongelma on, että monet vaikuttavuusmittarit on kehitetty enemmän diagnosointitarkoituksiin (esim. tunnistamaan lasten ongelmakäyttäytymistä) kuin arviointitarkoituksiin. Sen vuoksi ne eivät ole kovin herkkiä muutokselle (Pecora ym. 1995). Lisäksi, mittari voi olla luotu poimimaan ääripäiksi, jolloin se ei ole hyvä saamaan esille muutoksia asteikon alapäässä.

**Tarkenne 6.** Vaikutusmittareiden käyttö “pehmeiden” vaikutusten mittaamiseen

#### **Esimerkki 1**

Vahvuuksien ja vaikeuksien kysely (The Strengths and Difficulties Questionnaire, SDQ) on laajasti käytetty ja validi väline lasten ja aikuisten emotionaalisten ja käyttäytymiseen liittyvien ongelmien seulonnassa. Kyselystä on erilaisia versioita lasten vanhemmille, opettajille ja 11 – 16-vuotiaille nuorille. Mittari sisältää 25 väittämää tai kysymystä, jotka pisteytetään skaalalla 0 - 2. Vastaajia pyydetään valitsemaan vaihtoehto sen perusteella, miltä heistä on tuntunut viimeisen kuuden kuukauden aikana. Kyselylomakkeeseen vastaamiseen kuluu aikaa noin kymmenen minuuttia. Ensimmäiset kymmenen kysymystä ovat:

	Ei ole totta	Jokseenkin totta	Varmasti totta
--	--------------	------------------	----------------

1 Yritän olla ystävällinen ihmisille, välitän heidän tunteistaan	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2 Olen rauhaton, en voi istua pitkään paikallani	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3 Minulla on usein päänsärkyä, vatsakipuja tai pahoinvointia	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4 Lainaen yleensä tavaroitani (ruokaa, pelejä, kyniä, jne.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 Vihastun ja menetän usein malttini	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6 Teen tavallisesti kuten käsketään	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7 Olen usein huolissani	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8 Olen auttavainen, jos joku on loukkaantunut, poissa tolaltaan tai tuntee itsensä sairaaksi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9 Liikehdin levottomasti tai kiemurtelen jatkuvasti	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10 Minulla on yksi tai useampi hyvä ystävä	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### Esimerkki 2

Kehitimme oman kyselyn yhteisöryhmätyöskentelyn vaikuttavuuden arviointiin, koska vakiintuneita vaikuttavuusmittareita ei ollut olemassa. Kysely perustuu väittämäsarjoihin, ja ryhmien jäseniltä kysyttiin, olivatko he samaa vai eri mieltä. Väittämät ovat suurimmaksi osaksi lainauksia aikaisemmista keskusteluista muiden vastaavien ryhmien kanssa. Kysely koostuu neljästä osiosta, joissa kussakin on kymmenen kysymystä. Kukaan osioon vastaaminen vie enintään viisi minuuttia. Seuraava osio käsittelee tiimin yhteistyökykyä:

	Täysin samaa mieltä	Samaa meiltä	Ei samaa eikä eri mieltä	Eri mieltä	Täysin eri mieltä
1 On paljon valtataistelua – joskus ilmapiiri on niin kireä, että sitä voisi leikata veitsellä	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2 Onneksi meillä on vahva johtaja	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3 Jotkut yrittävät vetää samoja rooleja ja astuvat toistensa varpaille	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4 Joskus on vaikea nähdä kuka kantaa vastuun	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 Ihmiset ovat halukkaita näkemään vaivaa ryhmän puolesta	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6 Erimielisyyksiä syntyy pikkuasioista	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7 Mieliala on melko matalalla – ihmiset antavat helposti periksi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8 Täydennämme toisiamme ryhmänä - se toimii	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9 Kun asiat käyvät hankaliksi, ihmiset eivät halua tietää	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10 Ryhmässä työskentely on todella antoisaa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### **Kysymysten vaikutus vastaajiin**

On myös tärkeä pohtia kysymysten vaikutuksia vastaajiin, erityisesti haavoittuvampiin asiakkaisiin. Vaikka tätä huolta usein liioitellaan, tutkijoiden olisi varmistettava, että kysymykset ovat kulttuurisesti sopivia ja vahvuuksiin suuntautuneita (eli että ne eivät keskity vain negatiivisiin ulottuvuuksiin). Esimerkiksi SDQ-mittari sisältää sekä sosiaalisuutta että epäsosiaalisuutta esiin tuovat osiot.

### **Omaan raportointiin perustuvat mittarit**

Vaikutusten omatoiminen raportointi on tavallista sosiaalialalla. Se soveltuu, jos vaikutukset ovat luonteeltaan subjektiivisia (esim. yksinäisyyden tai masennuksen tuntemukset) tai jos tutkijan on vaikea seurata niitä suoraan (esim. jonkun juominen tai perheväkivalta). Itseraportointi ottaa huomioon myös käyttäjän näkökulman, mitä rahoittajat kasvavassa määrin arvostavat (esim. Joseph Rowntree Foundation 2001b). Vaikutusten omatoiminen raportointi ei sovi kovin hyvin, jos vastaajat eivät ole sellaisessa tilanteessa, että pystyisivät tekemään arvioin. Heidän tietonsa voivat olla puutteellisia tai heidän on vaikea suhtautua objektiivisesti omaan tilanteeseensa (esim. omiin työllistymismahdollisuuksiinsa) tai heille voi olla edullista antaa epätarkkaa tietoa (esim. rikollinen käyttäytyminen). Näitä näkökohtia on syytä pohtia. Yleisesti ottaen ihmisten on vaikeampi arvioida tapahtuneita muutoksia kuin absoluuttisia tasoja. Jos mahdollista, olisi parempi mitata seurattava asiaa kahtena tai useampana ajankohtana ja verrata tuloksia kuin pyytää yksilöitä arvioimaan ajan kuluessa tapahtuneita muutoksia.

### **Validiteetti ja reliabiliteetti**

Vaikuttavuusmittareiden validiteettia (pätevyyttä) voidaan parantaa useilla eri tavoilla:

- suositetaan vakiintuneita (validoituja) mittareita itse tehtyjen vaikutusmittareiden sijaan tai niiden rinnalla
- seurataan useita erilaisia - joskin toisiinsa yhteydessä olevia - vaikutuksia, jotta voidaan nähdä, ”kertovatko ne yhdenmukaisen tarinan”

- kerätään aineistoa erilaisista lähteistä (esimerkiksi asiakkaiden kumppaneilta tai hoitajilta)
- käytetään subjektiivisten (esim. juomisen itsearviointi) ja objektiivisten mittareiden (esim. verikokeet veren alkoholimäärän mittaamiseksi) yhdistelmiä
- arvioidaan vaikutuksia laadullisesti, ehkä osalle asiakaista, rinnan määrällisten vaikutusmittareiden kanssa.

## Hyötyjen arvottaminen

Kustannus-hyöty -analyysissä sekä vaikutukset että kustannukset on arvioitava rahamääräisinä. Etuna on, että intervention hyötyjä voidaan tällöin verrata suoraan kustannuksiin.

Jotkut vaikutukset on suhteellisen helppo arvottaa. Esimerkiksi tuenpiiristä-työhön -ohjelmien hyödyt voidaan arvioida rahassa seuraamalla osallistujien ansiotuloja (jotka kilpailluilla työmarkkinoilla rinnastetaan heidän työnsä taloudelliseen arvoon). Samoin energiatehokkuuteen liittyvän toimintamallin rahallista hyötyä voidaan osittain arvioida arvioimalla potentiaaliset polttoainesäästöt markkinahintaan. Kuitenkaan monien sosiaalialan interventioiden vaikutuksille ei voida määrittellä hintaa, koska ne ovat hyödykkeitä tai palveluja, joita ei tarjota markkinoilla. Kuinka esimerkiksi arvottaisit rahaksi vaikkapa sen, että uhka tulla ryöstetyksi vähenee tai turvallisuuden tunne kotona lisääntyy?

Taloustieteilijät ovat kehittäneet hienostuneita tekniikoita, joiden avulla joitakin tällaisia vaikutuksia voidaan arvioida rahassa. "Revealed preference" -tekniikoissa (RP) tutkitaan, miten ihmiset käyttävät rahaa arvotettavan hyödykkeen kaltaisiin hyödykkeisiin, ja tämän tulkitaan kertovan halukkuudesta maksaa vastaavista hyödykkeistä tai palveluksista, joita ei ole tarjolla markkinoilla. Esimerkiksi, kun arvioidaan turvallisuuden rahallista hyötyä, voidaan seurata, kuinka paljon ihmiset käyttävät rahaa murtohälyttimiin tai kuinka paljon enemmän he ovat valmiit maksamaan asunnoista sellaisilla alueilla, joilla on vähän rikollisuutta. Vaihtoehtoinen tapa on kysyä ihmisiltä suoraan, mitä he olisivat valmiita maksamaan myönteisestä muutoksesta (tai millaista kompensatiota he vaatisivat hyväksyä negatiivisen muutoksen). Näissä "stated preference" -tekniikoissa (SP) on tavallista, että edustavaa joukkoa ihmisiä pyydetään tekemään hypoteettisia valintoja kahden tai useamman selvästi määritellyn vaihtoehdon välillä. Esimerkiksi ihmisille, jotka asuvat korkean rikollisuuden alueella, voidaan antaa kuvaus siitä, millainen alue olisi, jos rikollisuutta olisi vähemmän. Seuraavaksi kysytään, olivatko he valmiita maksamaan tietyn summan enemmän kunnallisveroa, jotta muutos saavutettaisiin. SP-tekniikoiden etuna on, että ne ovat hyvin joustavia ja niitä voidaan käyttää melkein minkä tahansa asian rahalliseen arvottamiseen.

Arvottamisen metodit ovat kuitenkin kiistanalaisia ja arvottamista käytetään harvoin osana taloudellista arviointia. Mittareiden pätevyyttä (validiteettia) pidetään kyseenalaisena, ja tutkimusten loppukäyttäjät voivat olla halutto-

mia hyväksymään rahamääräisiä vaikutusmittareita (Hutton 1992; Johannesson ym. 1996). Lisäksi arvottamisen menetelmiin liittyy usein suurelle yleiselle suunnattuja kyselyitä, jotka ovat asianmukaisesti toteutettuina aikaan vieviä ja kalliita.

Tekemässämme kirjallisuuskatsauksessa havaitsimme, että sosiaalialaa koskevissa taloudellisissa arviointitutkimuksissa on käytännössä hyvin vähän sellaisia, joissa vaikutusten rahalliseen arvottamiseen olisi käytetty RP tai SP –tekniikoita. Yhdessä tutkimuksessa käytettiin karkeita arvioita elämän arvosta sen selvittämiseksi, ovatko vanhempainloman hyödyt (pienenevänä lapsikuolleisuutena) suurempia kuin vanhempainloman kustannukset (Ruhm 2000). Toisessa tutkimuksessa arvoitettiin lisääntyneen energiatehokkuuden aikaansaaman asumismukavuuden parantumista arvioimalla, kuinka paljon mahdollisista säästöistä kotitaloudet aiemmin kuluttivat lisälämmitykseen (Clinich ja Healy 2001). Tiedämme myös, että Home Office on tutkimassa edellä kuvattujen tekniikoiden käyttöä rikollisuuden alenemisesta saatavien hyötyjen arvottamiseen (Healey 2001). ("The Home Office is the Government department responsible for internal affairs in England and Wales. We work to build a safe, just and tolerant society, to enhance opportunities for all, and to ensure that the protection and security of the public are maintained and enhanced.")

Mutta useimmissa kustannus-hyöty –analyysiksi luokittelemissamme tutkimuksissa käytettiin paljon karkeampia menetelmiä ohjelmien hyötyjen rahalliseen arvottamiseen. Esimerkiksi alentuneen rikollisuuden hyötyjä arvoitettiin tyypillisesti arvioimalla rikosoikeusjärjestelmän mahdollisia säästöjä ja joissakin tapauksissa menetetyt omaisuuden arvoa (esim. Ekblom ym. 1996; Forrester ym. 1990; Robertson ym. 2001). Näillä tutkimuksilla, jotka ovat oikeastaan kustannus-säästöt –analyysijä (cost-savings), on taipumus aliarvioida rikosten ehkäisyyn ”todellisia” hyötyjä. Ne jättävät huomiotta psyykkiset kustannukset, joita koituu sekä uhreille että niille, jotka pelkäävät rikollisuuden vuoksi. Kuitenkin ne voivat olla varsin hyödyllisiä joillekin tietyille ohjelmille, jos ne osoittavat potentiaalisten säästöjen ylittävän aiheutuneet kustannukset.

### **Pisteyttäminen ja painottaminen**

Pisteyttäminen tai painottaminen on toinen menetelmä yhdistettäessä erilaisia vaikutuksia yhdeksi indeksiksi. Vaihtoehtojen asiantuntija-arviointi ("option appraisal") on yksi tällainen tekniikka, jota on käytetty esimerkiksi terveyspalveluissa arvioitaessa erilaisia pääoman sijoittamishdotuksia (NHS Executive 1994) ja menetelmää voidaan käyttää myös sosiaalialaa koskevassa arvioinnissa (Holtermann 1998). Tässä lähestymistavassa listataan kaikki sellaiset tekijät (kriteerit), jotka ovat tärkeitä arvioitaessa tietyn tyyppistä ehdotusta. Seuraavaksi jokaiselle arvioitavalle vaihtoehdolle annetaan pistemäärä jokaisella arviointikriteerillä, ja jokaiselle arviointikriteerille annetaan painoarvo, joka kuvaa sen suhteellista tärkeyttä. Lopuksi pistemäärät kerrotaan kriteereiden painoarvoilla ja nämä pisteet yhteen laskemalla saadaan jokaiselle vaihtoehdolle kokonaispistemäärä.

Tämän lähestymistavan etuna on, että sen avulla vältytään vaikutusten rahamääräiseltä arvottamiselta, mihin monet muut kuin taloustieteilijät ja myös jotkut taloustieteilijät suhtautuvat epäluuloisesti. Samaan tapaan kuin rahamääräinen arvottaminen, lähestymistapa mahdollistaa kilpailevien ohjelmien tai projektien suoran vertaamisen keskenään (mutta vain jos ne voidaan arvioida samoilla kriteereillä). Vaikka todelliset pistemäärät tai painoarvot voivat olla jossain määrin mielivaltaisia, lähestymistavan kannattajien mukaan se auttaa vähintäänkin tekemään rahan käyttöä koskevat päätökset aiempaa avoimemmin ja rationaalisemmin. Tekniikoita voidaan parantaa ottamalla mukaan asiantuntijoita, myös palveluiden käyttäjiä, ja sopimalla käytettävistä painoarvoista tai pistemääristä. Tekniikoita voidaan parantaa myös tutkimalla tulosten eli ehdotetun paremmuusjärjestyksen herkkyyttä erilaisille valinnoille.

## **Kustannusten arviointi**

Tässä luvussa tarkastellaan kustannusten määrittelyä ja eritä kustannusten mittaamiseen liittyviä ongelmia. Lisäksi käsitellään joitakin erityisesti sosiaalisten interventoiden kustannuksiin liittyviä haasteita. Yksityiskohtaisempia kustannuksiin liittyviä ohjeita on tämän raportin toisessa osassa (Byford ym., tulossa). Tavallisesti kustannuksiksi ymmärretään se rahasumma, joka käytetään tuotteen tai palvelun tuottamiseen. Poiketen yleisestä ajattelusta taloustieteilijät ovat kiinnostuneita vaihtoehtokustannuksista. Niillä tarkoitetaan hyötyjä, jotka menetetään kiinnittämällä voimavarat tiettyyn tarkoitukseen, sillä tällöin voimavarat eivät ole käytettävissä vaihtoehtoiseen tarkoitukseen. Tällaisia taloustieteilijöiden ymmärtämiä kustannuksia voi aiheutua, vaikka suoranaisia rahallisia kuluja ei aiheutuisikaan. Esimerkiksi omaishoitajien tai vapaaehtoisten työstä ei välttämättä makseta mitään, mutta heidän työllään on vaihtoehtokustannus, koska he olisivat voineet käyttää aikansa huvituksiin tai muuhun työntekoon. On asioita, jotka voivat olla ilmaisia palvelun tuottajalle, mutta silti niillä on vaihtoehtokustannus (esim. käytössä voi olla ilman vuokraa huone tai rakennus, joka muutoin olisi jonkun muun käytössä).

Kustannukset, jotka taloudellisessa arvioinnissa pitäisi käsitellä, voidaan jakaa neljään laajaan luokkaan

- *Ohjelmakustannukset*: Palvelun tuottamisen välittömät kustannukset, joka tarkoittaa palvelun kaikkien osatekijöiden kustannuksia, kuten henkilöstökustannuksia vapaaehtoistyöntekijöiden aikaa, rakennuksia, laitteita, kuljetuksia, tukipalveluja jne.
- *Ei-ohjelmakustannukset* (tai säästöt): Kaikki palvelun seurauksena syntyvät vaikutukset kustannuksiin, kuten säästöt jotka voivat syntyä jonkin vaihtoehtoisen palvelun tarpeen vähenemisestä.
- *Potilaan ja tämän perheen kustannukset*: Kaikki palvelun käyttäjän tai omaisen (tai muun huoltajan) kustannukset, jotka aiheutuvat palvelun käyttämisestä, kuten matkat palvelupaikalle ja takaisin tai kustannukset lastenhoidon järjestämisestä.
- *Tuottavuuskustannukset*: Kustannukset, jotka aiheutuvat sairauden tai vammaisuuden vuoksi heikentyneestä työkyvystä, menetykset ta-

loudellisessa tuottavuudessa ennenaikaisen kuoleman seurauksena tai omaishoitajien vähentäessä työssäoloaikaansa tai kokonaan luopuessa työnteosta hoitaakseen omaisiaan.

Useimmissa arviointitutkimuksissa ei tarkastella lainkaan kustannuksia, ja niissäkin joissa kustannuksia tarkastellaan on taipumuksena keskittyä lähinnä ”välittömiin” ohjelmakustannuksiin. Osittain tämä johtuu siitä, että arvioijat eivät ole täysin tietoisia sisällöstä, jonka taloustieteilijät antavat kustannusten käsitteelle. Sen vuoksi arvioijat laiminlyövät sellaiset kustannukset, joihin eli liity välitöntä varojen käyttöä. Mutta on myös useita muita syitä, joihin palaamme myöhemmin.

### **Suppea näkökulma**

Sosiaalialan taloudelliseen kirjallisuuteen kohdistuva analyysimme perusteella näyttää siltä, että monet arviointitutkimukset omaksuvat ohjelmaa rahoittavan tai hallinnoivan organisaation näkökulman. Ne lienevät kiinnostuneita vain vaikutuksista omiin talousarvioihinsa. Mutta kuten aikaisemmin totesimme, taloudellisen arvioinnin tulisi pyrkiä omaksumaan yhteiskunnallinen näkökulma. Tuolloin huomioon otetaan aiheutuneet kustannukset myös kaikille niille, joihin ohjelma epäsuorasti vaikuttaa. Osittaiset analyysit voivat johtaa harhaanjohtaviin johtopäätöksiin tietynlaisten ohjelmien kustannus-vaikuttavuudesta. Esimerkiksi henkilökohtaisen avun vähentäminen vanhusten avopalveluissa voi vaikuttaa houkuttelevammalta kuin se todellisuudessa on, jos arviointi ei ota huomioon mahdollista muiden julkisten palvelujen käytön kasvua tai omaishoitajille aiheutuvaa lisätaakkaa.

### **Kustannusten määrittämisen menetelmät**

Eräät tutkimukset keskittyvät mittaamaan vaikutuksia ja tarkastelevat kustannuksia vain lyhyesti, usein liitettynä arviointiraportin loppuun. Silti nämä tutkimukset saattavat esittää vahvoja väitteitä arvioimiensa interventioiden kustannus-vaikuttavuudesta. Kustannusten arviointia pidetään usein virheellisesti yksinkertaisena tehtävänä, kun se voi olla hyvin monimutkainen ja se tehdään harvoissa tutkimuksissa kunnolla. Tarkkojen kustannustietojen hankkiminen edellyttää erillistä resurssien ja yksikkökustannusten tarkastelua. Riippumatta arvioitavasta interventiosta kustannusarvioinnin avainkohdat ovat seuraavat (Beecham 2000):

- kuvataan ohjelman osatekijät mukaan lukien tilat, henkilöstö ja muut resurssit, joita käytetään ohjelman osina
- tunnistetaan ohjelmaan liittyvät toiminnot: tavoitteena on listata kaikki, mitä tehdään ja päättää sopivasta mittayksiköstä (esim. hoitotuntien määrä tai potilaan hoitokertojen määrä). Moniin sosiaalialan interventioihin liittyy useita toimintoja, joiden kustannukset saataan joutua arvioimaan erikseen.
- arvioidaan kuhunkin toimintaan käytettyjen resurssien määrä
- lasketaan yksikkökustannukset (esimerkiksi annetun hoitotunnin kustannus).

## **Tunnistamisongelmia**

Palveluntuottajilla pitäisi olla kirjanpito siitä, miten paljon asiakkaat käyttävät heidän palvelujaan, mutta tämäkin perustieto puuttuu usein. On löydettävä tasapaino sen suhteen, mikä on arvioinnin näkökulmasta toivottavaa ja mikä toimii käytännössä. Käytännön toimijoilla on usein kiire eivätkä he välttämättä pidä tunteesta, että heitä tarkkaillaan.

Interventiosta epäsuorasti aiheutuvien ei-ohjelmakustannusten ja seurausten tunnistaminen on luultavasti vieläkin vaikeampaa. Ensiksikin, voi olla vaikeaa ennustaa näiden vaikutusten luonnetta, erityisesti jos on kyse vaikutuksista, joita ei oltu suunniteltu. Tunnistaakseen mahdolliset kerrannaisvaikutukset (lumipalloefektit) arvioinnin tekijöiden on ehkä tarpeen tehdä erillinen tutkimus arvioitavan toiminnan vaikutusalueesta esimerkiksi keskustelemalla käytännön toimeenpanijoiden ja asiakkaiden kanssa. Toiseksi, jotkin näistä vaikutuksista voivat olla hyvin epäsuoria ja siksi vaikeasti kytkettävissä ohjelmaan. Tästä syystä, jotta vertailu intervention kohderyhmän kanssa olisi mahdollista, on tärkeää kerätä sama tieto myös kontrolli- tai vertailuryhmältä. Kolmanneksi, tieto epäsuorista ei-ohjelmakustannuksista tulee tuskien kerättyä osana ohjelman hallinnointia, ja siksi erillisen selvityksen tekeminen voi olla tarpeen.

## **Tietoihin liittyviä ongelmia**

Jos arviointi suoritetaan jälkikäteen, ei ole mahdollisuutta vaikuttaa siihen, mitä tietoa kerätään, eikä jotakin perustietoa kustannuksista ehkä ole käytettävissä. Tietojenhankinnasta aiheutuvat kustannukset aiheuttavat toisen mahdollisen rajoitteen. Tällöin arvioinnin tekijä saattaa turvautua ”ylhäältä-alas” tehtävään kustannusten määrittämiseen, missä kokonaiskustannukset ositetaan toiminnoille tai palveluille, vaikka tämä lähestymistapa voi aliarvioida kustannuksia (Beck ym., 1999). Ihannetilanteessa tulisi tehdä perusteellisempi ”alhaalta-ylös” -analyysi kaikista käytetyistä resursseista. Monissa sosiaalialaan liittyvissä arvioinneissa tämä vaatisi käyttäjien seuraamista sekä heidän käyttämiensä palvelujen ja heidän kussakin palvelussa käyttämänsä ajan rekisteröimistä. Käytännön esimerkki tästä lähestymistavasta on Client Service Receipt Inventory (CSRI, asiakkaan palveluun liittyvien kuittien läpikäynti), joka alunperin suunniteltiin sosiaali- ja terveydenhuoltoa varten, mutta jota voidaan soveltaa myös muihin tilanteisiin. CSRI-menetelmää voidaan käyttää myös tunnistettaessa ohjelmien eirahamää-räisiä kustannuksia, kuten omaishoidon käyttöä tai vaikutusta ihmisten työkykyyn (Beecham ja Knapp, 1992).

Kun budjetit ovat rajallisia, arvioijat voivat joutua pohtimaan, miten yksityiskohtaista kerättävän tiedon on oltava. Esimerkiksi, jos halutaan tuottaa tietoa laajaa poliittista keskustelua varten, paikallisten toimintamallien yksityiskohtainen kustannusanalyysi voi olla tarpeetonta. Sen sijaan, että kerätään suuria määriä kustannustietoa tietyistä toimintamalleista, jotka eivät kuitenkaan ole edustavia muissa tilanteissa, voi olla järkevää tehdä joitakin yleisiä oletuksia. Joissakin tapauksissa on myös kustannuseriä, jotka voi-



daan turvallisesti jättää huomiotta. Tällaisia ovat kustannukset, jotka ovat suhteellisen pieniä tai joiden osalta on epätodennäköistä, että ne eroavat vertailtavissa ohjelmissa. Lisäksi arvioijat voivat käyttää olemassa olevaa tietoa ja mallintamista arvioidessaan kustannuksia ja tehdessään niistä päätelmiä.

### **Arvottamiseen liittyviä kysymyksiä**

Tärkeimpien toimintojen (niiden jotka kattavat merkittävän osan ohjelman kokonaiskustannuksista tai ovat tärkeitä arvioinnin näkökulmasta) yksikkökustannuksia arvioitaessa suosittelimme ”alhaalta - ylös” lähestymistapaa. Tällöin kootaan tiedot kaikista kustannuksista, jotka aiheutuvat palvelun tuottamisen vaatimista olennaisista panoksista, kuten henkilökunnan ajasta, rakennuksista, välineistä ja hallinnon yleiskustannuksista. Jotkut palveluntuottajat, vaikka eivät suinkaan kaikki, ovat ehkä laskeneet tuottamiensa palvelujen yksikkökustannukset, ja myös valtakunnalliset yksikkökustannukset ovat käytettävissä, mikäli yksityiskohtaisempaa informaatiota ei ole saatavilla. Valtakunnallisia yksikkökustannuksia suositellaan myös arvioitaessa kustannuksia toiminnoille, joiden merkitys palvelun kokonaiskustannuksille on vähäinen (katso esim. Netten ja Curtis, 2000 sosiaali- ja terveydenhuollon yksikkökustannuksista).

Joidenkin kustannusten arvottaminen voi olla erityisen ongelmallista. Tällaisia ovat esimerkiksi omaishoitajien tai vapaaehtoistyöntekijöiden ajankäytön kustannukset. Ongelmallisuus selittää sitä, miksi tällaiset kustannukset jätetään usein pois arvioinneista. Palkattomien omaishoidon ja vapaaehtoistyön arvon määrittelyssä on kuitenkin tapahtunut edistystä, ja tällaiset työpanokset ovat oletettavasti merkittäviä monissa sosiaalialan interventioissa. Tällä hetkellä käytettävissä olevat menetelmät eivät kuitenkaan ole ongelmattomia. Käytännössä toimijoiden voi esimerkiksi olla vaikea ilmoittaa tarkasti sitä aikaa, jonka he käyttävät hoitotoimiin. Omaishoidolle ei ole myöskään helppo määrittellä sopivaa yksikkökustannusta. Omaishoitajien tai vapaaehtoistyöntekijöiden vaihtoehtokustannuksena on usein käytetty tiettyä palkkatasoa, kuten minimipalkka, tai kotiavustajan palkkaamisesta aiheutuvia kustannuksia. Kuitenkin jokaisen yksilön ajankäytöllä on erilainen vaihtoehtokustannus, koska he käyttäisivät ajan erilaisilla tavoilla. Tätä kysymystä on käsitelty vain muutamassa arvioinnissa, jos missään, ja käytetyt menetelmät sekä tutkimusten perusteella tehdyt oletukset vaihtelevat (McDaid 2001).

## Johtopäätökset

Kustannusten tarkka arvioiminen on olennainen osa taloudellista arviointia ja kuitenkin harva tutkimus tarjoaa kattavaa tietoa kustannuksista. Osaksi tämä johtuu siitä, että tutkijoita ei kannusteta tarkastelemaan kustannuksia laajemmin, ja osaksi aikaisemmin käsittelemistämme menetelmällisistä ongelmista. Kustannusten selvittäminen vaatii taloudellisessa arvioinnissa enemmän huomiota kuin se yleensä saa, ja sitä pitäisi käsitellä kiinteämpänä arviointitutkimuksen osana eikä vain ”vaihtoehtoisena” lisukkeena. Tarkenteessa 7 tarkastellaan kustannusten määrittämistä kahdessa arviointitutkimuksessa, joista toinen on sijaishuoltoa koskeva tapaustutkimus.

### **Tarkenne 7. Kustannusten määritteleminen**

#### **Sijaishuolto –tapaustutkimus (Foster care)**

Sijaishuoltoon kohdistuva arviointi on laaja tutkimus, joka käsittelee seitsemän paikallista viranomaista ja useita muita hankkeissa mukana olevia tahoja. Koska toimintaan kuuluu monia palveluja ja koska arviointiin (esim. tietojen keräämiseen) on käytettävissä rajallisesti resursseja, kustannusten arvioiminen on hankala tehtävä. Vaikka rekistereistä olisi saatavilla melko yksityiskohtaista tietoa sosiaalipalvelujen käytöstä, tietoa juuri laitos- ja sijaishuollon käyttöpäivistä voi olla rajoitetusti. Koska tietoa kerätään menneistä tapahtumista, paikallisten viranomaisten sijoituksista ei välttämättä ole saatavana yksityiskohtaisia tietoja, eivätkä arviointiin käytettävät resurssit anna mahdollisuutta kerätä tietoja jokaista sijoitettavaa koskevista hoitomaksuista erikseen. Yksityiskohtainen, jokaisen 500 yksittäistapaoksen (tai niin monen kuin sijoitettuja sattuu olemaan) todellisiin kustannuksiin perustuva kustannusten laskeminen ei siksi ole mahdollista. Sen sijaan voidaan käyttää tietoja keskimääräisistä paikallisille viranomaisille aiheutuvista kustannuksista. Tämä on tärkeä rajoittava tekijä, sillä sijais- ja laitoshuollon kustannuksissa on suurta vaihtelua eri sijoitusvaihtoehtojen ja eri paikallisten viranomaisten kesken.

Muiden palvelujen käyttöä koskevien tietojen tarkkuus on vielä heikompaa, sillä tiedot on kerätty postikyselyllä. Vastaajia oli kyselyssä pyydetty pelkästään kertomaan, kuinka monta kertaa he olivat käyttäneet tiettyjä terveys-, sosiaali- ja koulutuspalveluja sekä olleet tekemisissä rikosoikeusjärjestelmän kanssa. Tämä kattaa esimerkiksi psykiatrisista tai muista syistä johtuvat sairaalassaolopäivät, tai sosiaalityöntekijän tai koulupsykologin tapaamiskerrat. Näiden kustannusten arviointiin voidaan käyttää valtakunnallisia keskimääräisiä yksikkökustannuksia, koska kustannusten tarkempaan määrittämiseen ei ole käytettävissä riittävästi tietoja.

**Päihteitä väärinkäyttäviin lapsiin ja nuoriin kohdistetut avohuollon tukitoimet (Home-based social work interven-**

**tion for children and young people who have deliberately poisoned themselves)**

Tämä sosiaalityöhön liittyvä interventio on esimerkki yksityiskohtaisemmasta kustannusten määrittelemisestä monimutkaisessa toiminnassa, johon osallistuu myös monia toimijoita. Tiedot terveys-, koulu-, sosiaali- ja vapaaehtoispalvelujen käytöstä kerättiin lasten ja nuorten vanhemmilta kuuden kuukauden seuranta-aikana käyttämällä erityisesti tätä tarkoitusta varten suunniteltua kyselykaavaketta. Tämän lisäksi, koska eräiden keskeisten kustannuselementtien, kuten sairaalassa käyntien, virheettömyys haluttiin varmistaa, terveyspalvelujen käyttö varmennettiin vielä terveydenhuollon rekistereistä.

Koska tutkimus keskittyi sosiaalityöhön ja sairaalapalveluiden käyttöön, näiden kustannukset pyrittiin laskemaan tarkasti. Eniten huomiota kiinnitettiin sosiaalityön kustannusten määrittämiseen. Asiakkaan arvioimisistuntojen (assessment sessions) ja interventioistuntojen yksikkökustannukset laskettiin sen perusteella, kuinka paljon terapeutin arvioitiin käyttäneen aikaa sekä asiakas- että muihin kontakteihin, matkoihin ja ohjaukseen jokaisessa yksittäistapauksessa. Tuntikustannus laskettiin asiaan kuuluvan palkkataulukon keskipalkasta, johon lisättiin sivukulut sekä pääoma- ja hallintokulut. Sairaalapalvelujen yksikkökustannukset kerättiin paikallisista sairaaloista, ja se että saadut kustannustiedot olivat niin tarkkoja kuin mahdollista, varmistettiin jokaisen sairaalan rahatoimistosta.

Tutkimuksen näkökulmasta vähemmän keskeiset palvelut ja sellaiset palvelut, jotka lisäsivät kokonaiskustannuksia vain vähän, hinnoiteltiin karkeammilla menetelmillä. Esimerkiksi yleislääketieteen yksikkökustannukset perustuivat vuosittain julkaistaviin valtakunnallisiin arvioihin (julkaisija The Personal Social Service Research Unit, kts. Netten ja Curtis, 2000). (Lähde: Byford ym., 1999)

## 8 Tutkimusasetelma ja menetelmät

### Tutkimusasetelma

Määrällisen arvioinnin kannattajat, kuten taloustieteilijät, suosivat yleensä kokeellisia lähestymistapoja. Tällöin vaikutuksia ja taloudellisen arvioinnin tapauksessa kustannuksia pyritään mittaamaan ryhmälle, joka on ”altistettu” interventiolle ja kontrolli- tai vertailuryhmälle, johon interventio ei ole kohdistunut. Kahden ryhmän välisiä vertailuja käytetään erottamaan interventioon vaikutukset muista tekijöistä, joilla voi olla arviointiajan puitteissa merkitystä vaikutuksiin ja kustannuksiin. ”Klassinen” kokeellinen asetelma on satunnaisotokseen perustuva kontrolloitu koe (Randomised controlled trial, RCT), jossa potilaat on satunnaisesti jaettu niihin, joihin arvioitava interventio (esimerkiksi uusi lääke) kohdistuu ja niihin jotka saavat vaihtoehtoista hoitoa (esimerkiksi lumelääkettä). Terveystieteiden ulkopuolella yleisempiä ovat kuitenkin muunlaiset tutkimusasetelmat, kuten kvasikokeelliset tutkimukset ja vertailuryhmät, joita ei ole valittu satunnaisotoksella. Myös näitä tullaan tarkastelemaan. Jos kokeellinen lähestymistapa ei ole toimiva tai sovelias, ohjelman vaikutuksia on joskus mahdollista mallintaa. Taloustieteilijät käyttävät harvoin laadullisia menetelmiä, mutta tarkastelemme myös niitä. Uskomme, että näillä menetelmillä pitäisi olla selvästi nykyistä merkittävämpi rooli monimutkaisten sosiaalialan interventioiden taloudellisessa arvioinnissa.

Jokaista edellä mainittua lähestymistapaa tullaan tarkastelemaan yksityiskohtaisemmin. Koska erilaisia lähestymistapoja voidaan usein (ja pitäisikin) käyttää yhdessä, haluamme välttää sitä, että esittäisimme lähestymistavat keskenään kilpailevina. Tässä yhteydessä ei ole myöskään hyödyllistä ajatella hierarkiaa, jossa satunnaisotokseen perustuva kontrolloitu koeasetelma (RCT) on huipulla, ja sen jälkeen tulevat vähemmän ”tiukat” lähestymistavat käytettäväksi silloin, kun RCT ei ole mahdollinen. Vaikka terveystaloustieteilijät ovat laajasti RCT:n kannalla (esim. NHS Centre for Reviews of Research and Dissemination, 2001), emme usko, että RCT on siirrettävissä sosiaalialalle. Näkemyksemme perusteiden toivomme tulevan tämän raportin lukijoille selviksi. Lisäksi haluamme välttää keskittymästä liikaa vain yhteen tutkimusasetelmaulottuvuuteen. Usein kiinnitetään liikaa huomiota vertailuryhmän valintaan muiden arvioinnin osatekijöiden kustannuksella. Vaikka arviointitutkimus perustuisi RCT-asetelmaan, se voi olla huono, jos esimerkiksi vaikutusten valintaa tai niiden mittaamista ei pohdita riittävästi.

### Vertailun tarve

Vertailuryhmän olemassaolo on tärkeää, koska yleensä ei voida olettaa, että vaikutukset olisivat samat, jos interventiota ei olisi tehty. Ilman tyydyttävää vertailuryhmää on paljon vaikeampaa väittää, että halutut muutokset johtuvat juuri interventiosta eivätkä muista ”ulkoisista” tekijöistä, mikä selvästi heikentää tutkimuksen tulosten luotettavuutta. Tämä on erityisen tärkeää,

jos on syytä olettaa, että muutokset voisivat tapahtua ”luonnollisesti” ajan kuluessa. Esimerkiksi arvioitaessa tuen-piiristä-työhön -ohjelmaa on luultavaa, että osa mukana olleista olisi löytänyt työpaikan ilman ohjelman apua-kin (Gardiner, 1997). Samoin kriisitilanteessa olevat perheet lähtevät luultavasti mukaan sovitteluun tai muuhun perheille tarkoitettuun tukipalveluun, mutta kriisillä on taipumus rauhoittua myös ilman hoitoa (Pecora ym., 1995). Ilman vertailua olisi mahdotonta erottaa tällaisia luonnollisia muutoksia intervention vaikutuksista.

Ennen kuin päätetään siitä, kuinka vertailuryhmä hankitaan, on tärkeää pohdita, mihin arvioitavaa interventiota pitäisi vertailla. Ihanteellinen vertailukohde on seuraavaksi paras toimintavaihtoehto, joka yleensä on ”ei-tehdä-mitään” tai nykyinen vallitseva käytäntö (toisin sanoen olemassa oleva vaihtoehto), vaikka vertailukohde riippuu kin myös täsmällisestä arviointikysymyksestä. Esimerkiksi alkoholistien kuntoutuskeskuksen arvioinnissa halutaan selvittää, vähentääkö hoito-ohjelman lyhentäminen merkittävästi hoidon vaikuttavuutta. Tällöin on järkevää verrata lyhyempää hoitokäytäntöä nykyiseen käytäntöön eikä ”ei-tehdä-mitään” -vaihtoehtoon (Long ym., 1998).

Joissakin tapauksissa saattaa olla, että ei ole mahdollista määritellä selkeää vertailuasetelmaa. Esimerkiksi ”Safe In The City” tapaustutkimuksessa on luultavaa että useimmat asunnottomuusriskin kohteena olevat nuoret käyttävät jotain palveluja. Näin olisi lähes varmasti, jos vertailuryhmä hankittaisiin sosiaalipalvelujen, koulukuraattorien tai muiden järjestelmien kuten ”Connexions” -ohjelman kautta, koska nämä nuoret järjestelmä on jo poiminut (Connexions on Iso-Britanniassa toteutettava ohjelma, jonka avulla tuetaan nuorten kehitystä). Niitä, jotka ovat pudonneet verkon läpi ja jotka muodostaisivat ”puhtaamman” vertailuryhmän, poimisi tuskin tutkimusryhmäkään. Siispä kaikki toteutettavissa olevat vertailut tehdään SITC-ohjelman ja muiden alueiden ”nykykäytäntöjen” välillä. Nykykäytäntö itsessään vaihtelee valtavasti ja riippuu nuorten olosuhteista ja asuinpaikan paikallisista viranomaisista, joten vertailu on ”sotkuinen”. Kun nykyinen käytäntökin voi joissakin tapauksissa olla melko samanlainen SITC:n kanssa, on entistä vaikeampi tunnistaa ohjelman lisävaikutuksia. Ei siis ole mahdollista arvioida, saako ohjelma aikaan haluttuja vaikutuksia. Voidaan vain arvioida, onko se merkittävästi vaikuttavampi (tai kustannusvaikuttavampi) kuin nykyinen palvelutarjonta muualla Lontoossa.

Edelleen on mahdollista, että vertailuryhmän lisäksi on useampia interventioryhmiä, vaikka tämä edellyttääkin suurempaa otosta. Ryhmien avulla voi testata ohjelman erilaisia muunnelmia (esimerkiksi enemmän tai vähemmän intensiiviset versiot samasta palvelusta tai palvelujen erilaiset yhdistelmät).

### **Satunnaisotokseen perustuva kontrolloitu koeasetelma (Randomised controlled trial, RCT)**

RCT-asetelmaa käytetään yleisesti testattaessa lääkkeiden ja muiden terveydenhoidon interventioiden kliinistä vaikuttavuutta. Terveystaloustieteilijät ovat kasvavassa määrin käyttäneet asetelmaa taloudellisessa arvioinnissa

lisäämällä kustannustiedot uusiin tai aiemmin tehtyihin RCT-tutkimuksiin (Sassi, 2000). Asetelman käyttö on leviämässä moniin muihin terveydenhoidon interventiotyyppeihin ja muille sosiaalisen hyvinvoinnin alueille (katso tarkenne 8 joitakin esimerkkejä).

RCT-asetelmaa pidetään terveydenhuollon arvioinneissa ”kultaisena standardina”, koska hyvin toteutetut RCT-tutkimukset tuottavat tarkkoja (tai vääristymättömiä) arvioita kustannuksista ja vaikutuksista. Interventoryhmän ja vertailuryhmän väliset erot seurattavissa vaikutuksissa voidaan luotettavasti liittää suoritettuun interventioon, koska kaikki muut erot ryhmien välillä ovat satunnaisia ja pieniä, edellyttäen että otos on riittävän suuri. Useiden sosiaalialaa koskevien arviointitutkimusten tarkastelu on osoittanut, että ei-satunnaisotokseen perustuvat vertailuryhmät näyttävät järjestelmällisesti liioittelevan interventioiden vaikuttavuutta (Oakley ja Fullerton, 1996). Tämän vuoksi RCT-tutkimuksilla saatua näyttöä arvostetaan yleensä korkeammalle kuin näyttö muunlaiseen asetelmaan perustuvista arviointitutkimuksista. Terveydenhuollossa tarkastajat jättävät usein kokonaan huomiotta tutkimukset, jotka eivät perustu RCT-asetelmaan (esim. NHS Centre for Reviews of Research and Dissemination, 2001). Esimerkki 1 tarkenteessa 8 tuo hyvin esille laajan kontrolloidun kokeen etuja. Ilman satunnaistamista ei olisi ollut mahdollista tunnistaa pientä, mutta tärkeää ”New Deal for Lone Parents”- ohjelman vaikutusta, mikä riitti oikeuttamaan tämä ohjelma kustannus – hyödyllisyyden perusteella.

#### **Tarkenne 8.** Esimerkkejä satunnaisotokseen perustuvista koeasetelmista

**Esimerkki 1:** Yksinhuoltajille suunnatun New Dealin – hankkeen arviointi (New Deal for Lone Parents, NDLP)

NDLP on osa Britannian hallituksen tuen-piiristä-työhön – ohjelmaa. Sen tavoitteena on rohkaista yksinhuoltajia työelämään, ja osallistuminen on vapaaehtoista. Yksinhuoltajat, joilla on kouluikäisiä lapsia, kutsutaan henkilökohtaiseen haastatteluun. Haastattelussa annetaan tietoa ja neuvoja työllistymismahdollisuuksista. Työssä käyvien äitien on mahdollista saada myös taloudellisesta tukea. Pilottihanke käynnistyi 1997 ja se arvioitiin ennen kuin hanke levitettiin koko maahan. Vaikka kysymyksessä ei tarkasti ottaen ole satunnaistettu koeasetelma, asetelmaa jäljiteltiin niin paljon kuin mahdollista.

Arviointiryhmän avainkysymys oli, kuinka tunnistaa hankkeesta saatu lisähyöty. Toisin sanoen, miten otetaan huomioon se, että monet hankkeen aikana työpaikan löytäneet yksinhuoltajat olisivat työllistyneet hankkeesta huolimattakin. Ihannetapauksessa verrattaisiin arvioitavaan hankkeeseen osallistuvia niihin, jotka eivät hankkeeseen osallistu ja vakioitaisiin kaikki muut olosuhteet. Onneksi hankkeen asetelma antoi mahdollisuuden tämän tyyppiseen järjestelyyn. Kriteerit täyttävien yksinhuoltajien määrä oli niin suuri, että kaikkia ei voitu kutsua osallistumaan hankkeeseen sen alkaessa. Siksi kutsut porrastettiin ja

porrastus perustui sosiaaliturvatunnuksen (National Insurance number) viimeiseen numeroon, joka on satunnainen. Tutkijat pystyivät arvioimaan hankkeen vaikutusta vertaamalla muutoksia työllistymisessä. Niitä, jotka olivat saaneet kutsun hankkeeseen aikaisessa vaiheessa verrattiin vastaavaan ryhmään yksinhuoltajia, joille kutsu lähetettiin keskimäärin kuutta kuukautta myöhemmin. Jälkimmäinen ryhmä toimi kontrolliryhmänä. Tiedot yli 15 000 tapauksesta otettiin hallinnollisista tietokannoista, jotka sisältävät informaatiota kaikkien yksilöiden korvaustilanteesta (työtön tarvitsee toimeentulotukea).

Todennäköisyys sille, että yksinhuoltajat, jotka oli kutsuttu osallistumaan ohjelmaan, tarvitsevat kuuden kuukauden kuluttua toimeentulotukea oli merkittävästi alhaisempi kuin sen ryhmän todennäköisyys toimeentulotuen tarpeeseen, joita eivät vielä oltu kutsuttu. Pidemmälle ajalle arvioituna voidaan sanoa, että niiden yksinhuoltajien määrä, jotka tarvitsevat toimeentulotukea vielä vuoden kuluttua, laskisi NDLP:n ansiosta 85 %:sta 83 %:iin. Tutkimus pystyi siten tunnistamaan suhteellisen pienen, mutta merkittävän vaikutuksen tuen tarvitsijoiden määrään. (Lähde: Hasluck ym., 2000)

**Esimerkki 2:** Synnytyksen jälkeistä tukea tarjoavan tukihenkilötoiminnan kustannukset ja hyödyt

Kysymyksessä on synnytyksen jälkeistä tukea tarjoavan hankkeen arviointi Sheffieldissä, jossa 623 vastasyntyttynyttä äitiä tutkittiin satunnaistetulla kontrolliasetelmalla. Äidit, jotka päättivät osallistua hankkeeseen, jaettiin satunnaisesti kontrolliryhmään ja interventioryhmään. Kontrolliryhmä sai tavanomaista kättilön ja kotisairaanhoidajan palvelua. Interventioryhmälle tarjottiin tämän lisäksi tukityöntekijän palveluja. Tukityöntekijä vieraili enintään kymmenen kertaa äidin luona ensimmäisten 28 päivän kuluessa synnytyksestä, ja työntekijä antoi sekä käytännöllistä että emotionaalista tukea. Tuki räätälöitiin äidin henkilökohtaisiin tarpeisiin, mutta se sisälsi myös kotitaloustöitä, ohjausta rintaruokinnassa ja apua muiden lasten hoitamisessa.

Pääasiallinen vaikutusmittari oli äidin käsitys yleisestä terveydentilastaan kuuden viikon kuluttua synnytyksestä. Tätä mitattiin niin sanotulla SF-36 mittarilla. Lisäksi käytettiin toissijaisia vaikutusmittareita. Näitä olivat mm. ns. synnytyksen jälkeisen depression Edinburgh mittari (Edinburgh Postnatal Depression Scale) ja rintaruokinnan määrä. Terveys- ja sosiaalipalvelujen käyttöä arvioitiin samalta kuuden viikon ajalta asiakaskyselyn avulla. Seurantakysely kustannuksista ja vaikutuksesta tehtiin kuuden kuukauden kuluttua.

Tulokset analysoitiin kustannus-seuraus –menetelmällä. Tutkijat eivät löytäneet merkittäviä eroja äidin kokemassa terveydentilassa kuuden viikon kuluttua synnytyksestä. Jonkin verran

näyttöä löydettiin siitä, että toissijaiset vaikutusmittarit soveltuisivat paremmin kontrolliryhmän arviointiin. Kuitenkin äidit, jotka saivat tukea, suhtautuivat siihen erittäin myönteisesti, myönteisemmin kuin muihin saamiinsa äitiyspalveluihin. Intervention kokonaiskustannukset olivat äitiä kohti noin 180 £ eikä intervention havaittu vaikuttavan muiden terveyspalvelujen käyttöön.

(Lähde: Morrell ym., 2000b)

### *Käytännöllisiä ja eettisiä kysymyksiä*

Monilla sosiaalisen hyvinvoinnin alueilla sekä käytännön toimijat että alueelle erikoistuneet arvioijat vastustavat RCT-asetelman käyttämistä (Oakley ja Fullerton, 1996). Useimmat RCT- menetelmää vastustavat argumentit liittyvät käytännöllisiin tai eettisiin kysymyksiin, vaikka niitä ei aina olekaan selkeästi ilmaistu:

- Eräät kokeet ovat epäeettisiä (esim. pidättäytyminen mahdollisesta ihmisenelämän pelastavasta tai elämäntilannetta muuttavasta palvelusta) tai ne voivat näyttää epäeettisiltä (esim. pääsyn kieltäminen hyvin järjestettyyn palveluun, vaikka sen vaikuttavuudesta ei olekaan todisteita).
- Otos voi olla liian pieni, jotta satunnaisotos olisi toimiva (esim. hyvin pienissä ohjelmissa tai alueisiin perustuvissa hankkeissa, joissa alueet pitäisi valita yksilöiden sijasta satunnaisotoksella).
- Arvioijilla ei ole ehkä mahdollisuutta kontrolloida, ketkä saavat tiettyä palvelua. Toisaalta ne arvioijat, joilla on mahdollisuus vaikuttaa palvelun saajiin, voivat olla haluttomia käyttämään satunnaisotosta, koska heillä on erilaiset prioriteetit (esim. halu kohdistaa palvelu niille, jotka sitä eniten tarvitsevat).
- Satunnaistettu kokeiluasetelma voi ”saastua” (esim. kontrolliryhmään kuuluvat löytävät muita tapoja päästä vastaavaan palveluun) tai tulla toimimattomiksi (esim. suuri osa kieltäytyy osallistumasta tutkimukseen tai keskeyttää kokeen).
- Eräät kokeet voivat olla hyvin kalliita (tai niitä voidaan pitää kallina).
- Politiikka voi tulla mukaan kuvioihin: ne, jotka teettävät tutkimuksen, voivat olla enemmän kiinnostuneita saamaan ”ammuksia” tiedotustoimintaansa varten kuin tiukoista näytöistä, jotka eivät ehkä tue heidän ”asiaansa”.

Joissakin tapauksissa nämä ongelmat ovat todellisia ja sulkevat tehokkaasti pois RCT-lähestymistavan. Mutta jotkut näistä huolista ovat joko virheellisiä tai liioiteltuja, ja niistä voidaan selvittää ja käytännössä on selvittykin, jos tahtoa tähän on ollut. Jotkin yllä olevista argumenteista voidaan myös kääntää ylösalaisin. Esimerkiksi lienee eettisempää tutkia interventiota, jonka vaikutuksia ei ole selvitetty, kuin jatkaa testaamattoman intervention käyttöä laajemmassa väestöjoukossa. Siitä huolimatta, että menettely saattaa tarkoittaa palvelun saannin estämistä joiltakin ihmisiltä jonkin aikaa (Williams, 1992). Muut ongelmat voidaan välttää huolellisella suunnittelulla



(esim. minimoimalla otoksen kato) tai muovaamalla koeasetelmaa. Esimerkiksi satunnaisotoksen hyväksyttävyyttä voidaan lisätä, jos kontrolliryhmään kuuluville henkilöille tarjotaan jokin vaihtoehtoinen interventio ”lumelääkkeen” sijasta, vaikka tämä samalla vaikeuttaa sen osoittamista, että interventio on vaikuttavampi. Vaihtoehtoisesti, voi olla mahdollista jäljitellä RCT-asetelmaa ilman satunnaistamista. Esimerkkejä ”lähes satunnaisotoksesta” tarkastellaan kvasikokeellisia tutkimuksia käsittelevässä luvussa.

### *RCT-asetelman soveltuvuus*

On myös tärkeää pohtia onko RCT aina käyttökelpoinen vaikka käytännölliset ja eettiset ongelmat voitaisiin tyydyttävästi voittaa. Ensiksi, satunnaistaminen voi vaikuttaa tutkimustulosten pätevyyteen (validiteettiin) ja yleistettävyyteen. Satunnaisotos voi vääristää vaikutuksia, jos ihmiset tietävät, mihin ryhmään he kuuluvat ja jos mieltymykset määräävät vaikuttavuutta merkittävästi. Esimerkiksi ne, jotka on sijoitettu kontrolliryhmään, voivat olla masentuneita, mikä itsessään voi heikentää vaikutuksia. Kuuluminen kontrolliryhmään voi myös kannustaa heitä menestymään hyvin vastoin oletusta, missä tapauksessa vaikutukset tämän ryhmän osalta ovat keinotekoisesti liioittelevia. Tämä ongelma on suhteellisen helppoa ratkaista lääkerekokeissa (esim. käyttämällä etiketöimiä lääkepulloja), mutta ongelma paljon vaikeampi palvelujen arvioinnissa.

Satunnaistamisella näyttää myös olevan vaikutusta siihen, kuka on halukas osallistumaan kokeeseen. Esimerkiksi RCT-asetelmalla toteutettaviin terveydenhuollon ennaltaehkäiseviin interventioihin osallistuvat potilaat ovat yleensä paremmin toimeentulevia, paremmin koulutettuja ja todennäköisemmin terveellisiä elämäntapoja noudattavia kuin potilaat, jotka kieltäytyvät osallistumasta (Mc Kee ym., 1999). Mitä vähemmän edustavaksi RCT-otos tämän seurauksena muodostuu, sitä vähemmän tulokset ovat yleistettävissä laajempaa väestöön.

Toiseksi, kokeellinen lähestymistapa aiheuttaa arvioinnille myös eräitä yleisempiä ongelmia. Käytämme synnytyksen jälkeisen tuen interventiota havainnollistamaan joitakin näistä kysymyksistä (katso esim. esimerkki 2 tarkenteessa 8). Tämä oli hyvin toteutettu RCT-tutkimus, joka julkaistiin Health Technology Assessment -raportissa (Morrell ym., 2000a) ja British Medical Journal:ssa (Morrell ym., 2000b), joten kritiikkimme kohdistuu kokeellisen lähestymistavan rajoituksiin eikä niinkään tutkimuksen täsmällisyyteen. Tietenkin on aina mahdollista väittää, että tutkimus olisi voitu suunnitella toisella tavalla, ja että kritiikki ei kohdistu kokeellisen arvioinnin sovellettavuuteen sinänsä. Väitämme kuitenkin, että ongelmat ovat säännöllisesti toistuvia ja että on sinisilmäistä ajatella paremman suunnittelun aina ratkaisevan nämä ongelmat.

Muiden kokeellisten tutkimusten tapaan myös tässä vaikutuksia, kuten äidin yleistä terveydentilaa ja masentuneisuuden astetta, mitataan määrällisillä mitta-asteikoilla. Arvioitaessa tutkimusta onkin syytä suhtautua varoen vaikutusmittareiden soveltavuuteen. Kuuden viikon jälkeen terveydentilassa ei voitu osoittaa positiivisia muutoksia, ja verrattaessa tukipalvelua saaneita

niihin, jotka eivät sitä saaneet, palvelua saaneessa ryhmässä havaittiin joitakin negatiivisia vaikutuksia. Silti osallistujat olivat erittäin tyytyväisiä ohjelmaan. Vaikutusmittarit saattoivat olla tarkkoja, mutta on mahdollista, että mittarit eivät poimineet juuri niitä positiivisia vaikutuksia, joita äidit arvosivat. Vaikka tämä epäselvyys voitaisiin ratkaista jatkotutkimuksissa käyttämällä erilaisia vaikutusmittareita, esimerkki havainnollistaa määrällisten mittareiden rajoituksia, vaikka mittarit olisivatkin standardoituja (validoituja).

Kokeelliset tutkimukset keskittyvät nettovaikutuksiin, mutta olisi ollut hyödyllistä tietää, miksi tämä interventio ei näyttänyt toimivan; se auttaisi arvioimaan, olisiko sama interventio vaikuttava toisenlaisissa olosuhteissa tai tarkemmin määritellylle kohderyhmälle ja kuinka toiminnasta saataisiin vaikuttavampaa. Tutkijat esittävät toimintamallin epäonnistumisen syyksi sitä, että se ei kehittänyt ihmisten omaa selviytymiskykyä, mutta tämä teoria on spekulatiivinen. Kokeelliseen asetelmaan perustuva arviointitutkimus ei tuota näyttöä siitä, miksi toimintamalli ei näytä toimivan. Kuitenkin tämä on erittäin tärkeää, jos halutaan arvioida toimintamallin mahdollisuuksia kattavasti.

Kuten useimmissa kokeellisissa tutkimuksissa, vaikutukset mitattiin tiettyinä ajankohtina, tässä tapauksessa kuuden viikon ja kuuden kuukauden kuluessa. Kyselyjen ajoitus on usein melko mielivaltaista, vaikka sillä voi olla merkittävä vaikutus tuloksiin. Tässä tutkimuksessa ensimmäinen seuranta-kysely tehtiin pian intervention päättymisen jälkeen. Kirjoittajat ehdottavat, että äidit ovat ehkä kokeneet ”vieroitusoireita” interventiosta, mikä voisi selittää havaittuja negatiivisia vaikutuksia. Mutta myös tämä on spekulatiivista; kokeellinen arvioinnin lähestymistapa ei helposti sovellu tämälähtöiseen kysymykseen.

Kokeellinen asetelma käsittelee interventiota yhtenä homogeenisena palveluna, vaikka se koostuisi useista hyvin erilaisista toiminnoista. Tämä lähestymistapa ei siis kykene antamaan informaatiota päätettäessä erilaisten toimintojen sopivasta yhdistämisestä tai toiminnan ajoituksesta.

Näistä syistä uskomme, että RCT-asetelmien ylivoimaisuuteen on syytä suhtautua varauksella tietyn osin. Silloinkin, kun asetelma on tarkoituksenmukainen ja käyttökelpoinen, arviointiin on usein tarvetta yhdistää muita tutkimusmenetelmiä. Esimerkiksi aiemmin esitellyssä tutkimuksessa olisi voitu käyttää laadullisia menetelmiä tunnistettaessa niitä toimintamallin vaikutuksia, jotka olivat äideille tärkeitä. Tätä tietoa olisi voinut käyttää arvioinnissa mitattavien vaikutusten valinnassa. Esimerkkinä käytetty arviointitutkimus päätyi ehdottamaan, että äitejä tukeva toimintamalli ei ollut kustannus-vaikuttava, mutta tarkoituksenmukaisempien vaikutusmittareiden käyttö olisi ehkä tuottanut toisenlaisen johtopäätöksen. Lisäksi laadullista tutkimusta olisi voitu käyttää selittämään RCT-tutkimuksen tuloksia ja testaamaan joitakin kirjoittajien omia tulkintoja, joiden he itsekkin myöntävät olevan spekulatiivisia – esimerkiksi kokivatko äidit vieroitusoireita intervention loppuessa. Lopuksi, mutta ei vähäpätöisimpänä, laadulliset menetelmät olisivat auttaneet tunnistamaan niitä toimintamallin osia, joita eri

henkilöt arvostivat eniten. Tätä informaatiota olisi voitu käyttää toimintamallin kehittämiseen ja tehokkaampaan kohdentamiseen. Laadullisia tutkimusmenetelmiä tarkastellaan myöhemmin tässä luvussa.

### **Kvasikokeelliset asetelmat**

Kvasikokeelliset asetelmat muistuttavat RCT-asetelmia, mutta kontrolli ja verrokkiryhmiä ei valita satunnaisotoksella. Tarkenteessa 9 on kaksi esimerkkiä kvasikokeellisista tutkimuksista ja toinen niistä on sijaishuoltoa koskeva tapauksittomus. Kvasikokeellinen lähestymistapa voi olla RCT-asetelmaa hyväksyttävämpi, koska se välttää joitakin satunnaisotokseen perustuvien asetelmien sekä todellisista että ongelmiksi mielletyistä ongelmista. Joissakin tapauksissa tämä on ainoa mahdollinen lähestymistapa, koska satunnaisjärjestely voi olla epäkäytännöllinen.

Milloin satunnaisjärjestely ei ole mahdollinen, arvioijat voivat käyttää hyväkseen ”luonnollista” vaihtelua ohjelmien toimeenpanossa löytääkseen sopivan vertailuryhmän. Esimerkiksi koulua, joka on ottanut käyttöön uuden toimintatavan, voidaan verrata sellaiseen, joka ei ole vielä niin tehnyt. Jos olosuhteet ovat oikeat, voidaan saavuttaa lähes satunnaisjärjestelyä muistuttava tilanne (katso tarkenne 10, esimerkkejä satunnaistamista muistuttavista järjestelyistä).

### **Tarkenne 9. Esimerkkejä kvasi-kokeellisista tutkimuksista**

#### **Esimerkki 1: Darlingtonin avohuoltoprojektin arviointi (Darlington Community Care Project)**

Darlington –projekti oli yksi monista 1980-luvulla aloitetuista pilottiprojekteista, jotka rahoitettiin hoivaa yhteisöissä –aloitteen (Care in the Community Initiative) puitteissa. Hanke tarjosi avohuoltona intensiivistä ja koordinoitua tukea ikääntyneille, jotka muuten olisivat joutuneet jäämään sairaalan pitkäaikaispotilaiksi.

Tutkijat käyttivät kvasikokeellista tutkimusasetelmaa hallinnollisista syistä (esim. satunnaistettua asetelmaa ei olisi todennäköisesti hyväksytty poliittisesti), mutta myös siksi että he halusivat arvioida sekä projektia kokonaisuutena että sen vaikutuksia yksilöihin. Vertailuryhmä muodostettiin viereisen terveydenhoitoalueen pitkäaikaisessa hoidossa olevista ikääntyneistä. Vertailuryhmä osoittautui hyvin samanlaiseksi projektiryhmän kanssa. Interventoryhmän ja vertailuryhmän välisten erojen kontrolloimiseksi tehtiin myös tilastollisia vakiointeja. Tutkimukseen osallistui hieman yli 200 ikääntynyttä.

Osallistujat haastateltiin lyhyesti heidän kotiututtuaan sairaalasta (tai vertailuryhmän tapauksessa, silloin kun he olisivat voineet kotiutua) ja sen jälkeen kuuden ja 12 kuukauden kuluttua. Vaikutusmittarit sisälsivät mittareita subjektiivisesta hyvinvoin-

nista (esim. yleinen tyytyväisyys, henkinen vireys, masentuneisuus ja yksinäisyys) ja useita käyttäytymisindikaattoreita (esim. fyysinen toimintakyvyn aleneminen, apatia, kommunikointivaikeudet ja sosiaaliset ongelmat) sekä joitakin hoivan laadun indikaattoreita (esim. koettu hoivan lisäämisen tarve). Myös kuolleisuusastetta seurattiin molemmissa ryhmissä. Lisäksi arvioitiin hoitajiin kohdistuvia vaikutuksia (esim. hoivatehtäviä ja rasittumista).

Tulokset osoittivat merkittävää paranemista yhtä lukuun ottamatta kaikissa subjektiivisen hyvinvoinnin indikaattorissa sekä hoidon laadun indikaattoreissa. Suurimmassa osassa käyttäytymisindikaattoreista parannusta ei havaittu, mikä ei ollut yllätys. Nämä myönteiset vaikutukset saavutettiin lisäämättä hoitajien rasittuneisuutta. Vaikka kuuden kuukauden kuluttua interventoryhmässä oli tilastollisesti merkitsevästi korkeampi kuolleisuusaste, eroa ei voitu havaita enää kahdentoista kuukauden kuluttua.

Darlingtonin -malli oli myös halvempi sairaalahoitoon verrattuna (noin £30 halvempi viikossa tavanomaisten kustannusoletusten mukaan). Siten projektin mukainen toimintamalli näyttää olevan kustannus-vaikuttavampi vaihtoehto sairaalahoitoon verrattuna, ja hyödyttää sekä potilaita että hoitajia alhaisemmilla kustannuksilla.

### **Esimerkki 2:** Sijaishuolto -tapaustutkimus (Foster care)

Lasten sijoittaminen kodin ulkopuolelle lastensuojelussa on siinä mielessä epätavallinen sosiaalialan toimenpide, että satunnaistamista sijoituspaikkaa valittaessa selvästikään voidaan käyttää. Vaikkakin käytännöllisestä näkökulmasta kaikki sijoitustyytit olisivat mahdollisia kaikille sijoitusta tarvitseville nuorille, näin ei kuitenkaan ole. Olisi epäeettistä sattumanvaraisesti asettaa nuori elinolosuhteitaan koskevaan "lottoarvontaan".

Olisi esimerkiksi kohtuutonta pyytää nuorta ihmistä palaamaan takaisin biologisten vanhempiansa luokse, jos hän on lähtenyt kotoaan ongelmien vuoksi, joita ei ole pystytty ratkaisemaan. Ei myöskään ole mahdollista taata, että sijoituspaikka ei vaihdu sijoituksen aikana. Esimerkiksi vakaata ja pitkäkestoista sijaishuoltosuhdetta ei synny, jos nuori itse tai sijaisperhe ovat tyytymättömiä.

Keskeisimpänä kiinnostuksen kohteena tässä tutkimuksessa ovat sijaishuollon vaikutukset sijoitettuihin lapsiin. Kun ei voida välttyä siltä, että nuorten sijoitusmuoto vaihtuu tutkimuksen aikana, syntyy mahdollisuus vaihtoehtoisten elinolosuhteiden vertailuun. Erilaisten sijoitustyyppien vaikutusten ja kustannusten vertailu on tässä tilanteessa kuitenkin täynnä virhelähteitä. Eri tilanteissa olevia (biologiseen perheeseen palauttamista tukevassa yksikössä, sijaisperheessä, adoptiossa, laitossijoitukses-

sa ja tiheästi muuttavien) nuorten ominaisuudet vaihtelevat todennäköisesti merkittävästi. Esimerkiksi adoptoidut voivat olla nuorempia ja heillä saattaa olla vähemmän käytösongelmia kuin muilla otokseen kuuluvilla; niillä, jotka palaavat kotiin, on voinut olla vähemmän ongelmia sijoitusvaiheessa; ja toistuvasti sijoituspaikkaa vaihtavat voivat olla häiriintyneimpiä.

Vaikka eri sijoitusmuotojen suora vertailu ei tule kysymykseen, eroja voidaan tutkia tilastollisilla (monimuuttuja) analyyseillä. Nämä mahdollistavat ryhmien välisten erojen kontrolloimisen. Tilastollista vakiointia voidaan käyttää, kun nuorten ominaisuuksista, sijoituksista ja huoltajista on paljon tietoja. Lisäksi kvalitatiivinen aineisto auttaa tulkitsemaan ja selvittää kvantitatiivisten analyysien tuloksia.

## **Tarkenne 10. Satunnaisotosta muistuttavat asetelmat**

### **Kontrolliryhmä odotuslistalta**

Kontrolliryhmä voidaan joskus muodostaa niistä, jotka odottavat hankkeeseen pääsyä. Esimerkiksi vanhempien kasvatustehdävää tukevan kokeilu arvioimiseksi vanhemmat sijoitettiin kotiaan lähinnä sijaitsevaan valmennusryhmään. Kun ryhmä täytyi, vanhemmat jäivät jonotuslistalle, josta samalla muodostui kontrolliryhmän (Thompson ym., 1996). Tämä on hyvä lähestymistapa, jos halukkaita osallistujia on enemmän kuin voidaan ottaa mukaan ja edellyttäen että osallistujia ei valikoida vaan he pääsevät mukaan ilmoittautumisjärjestyksessä. Tällöin ryhmä muistuttaa satunnaisotosta. Yksi ongelma arvioijille on, että odotuslistalla olijat voivat päästä mukaan ohjelmaan ennen arviointijakson päättymistä. Tämän vuoksi edellä kuvattu lähestymistapa soveltuu lyhytvaikutteisten tai sellaisten hankkeiden arviointiin, joiden jonotuslista on pitkä.

### **Porrastettu toimeenpano**

Kun hankkeen aloittaminen on porrastettu, voidaan verrata niitä, jotka saavat palvelua aiemmin niihin, joiden osalta hanke alkaa myöhemmin. Esimerkkinä on yksinhuoltajien työllisyyttä tukevan ohjelman, New Deal for Lone Parents (katso Tarkenne 8) arviointi. Kutsut lähetettiin sosiaaliturvatunnukseen perustavassa satunnaisessa järjestyksessä. Kuitenkin, koska yksinhuoltajat saattoivat hakeutua hankkeeseen myös oma-aloitteisesti ennen kuin olivat saaneet henkilökohtaisen kutsun, vertailtavat ryhmät eivät muodostuneet täysin satunnaista. Jotkut sosiaaliturvatunnuksensa mukaan kontrolliryhmään kuuluvat yksinhuoltajat olivatkin hyötyneet ohjelmasta, ja seurauksena on ohjelman vaikutusten hienoinen aliarviointi analyysituloksissa. (Lähde: Hasluck ym., 2000)

### **Luonnolliset erot**

Jossain tilanteissa arvioijat voivat käyttää hyväksi suunnittelemattomia tai "luonnollisia" eroja hankkeen toimeenpanon. Nämä voivat olla melko satunnaisia. Esimerkkinä tästä on nuorille rikoksentehtäjäille kohdistetun Nokomis-hankkeen arviointi (De-

schenes ja Greenwood, 1998). Vertailuryhmä muodostettiin nuorista rikoksentehtäjäistä, jotka lähetettiin rangaistuslaitokseen, koska eräät tuomarit kieltäytyivät ohjaamasta nuoria rikollisia uuteen ohjelmaan. Toinen esimerkki on Toinen-toiselle lasten lukutaito-ohjelman arvioinnista (Farkas, 1998). Jotkut lapset ottivat muita vähemmän osaa istuntoihin aikatauluongelmien vuoksi, eivätkä aikatauluongelmat ilmeisesti liittyneet lasten ominaisuuksiin, toisin sanoen vähäisempi osallistuminen kohdistui lapsiin melko satunnaisesti). Hankkeen vaikutusta arvioitiin vertaamalla paljon istunnoista poissa olleiden - jotka toimivat vertailuryhmänä - lukutaidon kehitystä niihin, jotka olivat osallistuneet koko kurssiin.

### **Kohorttien vertailut**

Yksi lähestymistapa on verrata palvelujenkäyttäjryhmiä, kohortteja, jotka ovat saaneet palvelua ennen toimintatavassa toteutettua muutosta ja sen jälkeen. Tätä lähestymistapaa käytettiin arvioitaessa uudistetun alkoholistien hoito-ohjelman vaikutuksia. Perinteisen hoito-ohjelman vaikutuksia hoitoyksikön asiakasryhmälle verrattiin asiakasryhmään, joka käytti palveluja sen jälkeen, kun ohjelmaa oli lyhennetty viidestä viikosta kahteen viikkoon (Long ym. 1998). Joissakin kroonisissa ongelmisissa voi olla mahdollista verrata samaa yksilöä koskevia tuloksia hoidossa tapahtuneita muutoksia seuraavina ajankohtina.

### *Pragmaattiset (käytännölliset) lähestymistavat*

Käytännön näkökulmaa korostavissa tutkimusasetelmissa tehdään usein ihanteellisista asetelmaa koskevia kompromisseja. Esimerkiksi ohjelmaa hallinnoivat haluavat ehkä suosia tiettyjä asiakkaita palvelun saajina ja siksi kontrolli- ja vertailuryhmien valinnassa tingitään ideaalista. Esimerkkinä tästä on vanhempana olemista tukevan hankkeen arviointi (Thompson ym., 1996). Vertailuryhmän arvioinnissa muodostivat ne, jotka odottivat pääsyä hankkeen tarjoaman tuen piiriin - olivat odotuslistalla (tarkenne 10). Kuitenkin sellaiset odotuslistalla olevat vanhemmat, jotka tunsivat tarvitsevansa välitöntä apua, pääsivät jonon ohi (ja pois vertailuryhmästä). Joissakin tapauksissa arvioijat saatetaan pakottaa tällaisiin kompromisseihin. Koulujen aamiaisklubitoiminnan kansallinen arviointi perustui siihen, että rahoitus jaettiin ohjelman ensimmäisessä vaiheessa satunnaisotoksella valituille kouluille, ja ne koulut, jotka eivät saaneet rahoitusta, muodostivat vertailuryhmän. Kuitenkin osa vertailuryhmän kouluista sai rahoitusta hankkeen toisessa vaiheessa ja aloitti aamiaisklubitoiminnan ennen arviointijakson päättymistä (Mugford, 2001). Niinpä sen sijaan että arvioinnissa olisi verrattu kouluja, joissa on aamiaisklubitoimintaa niihin joissa ei toimintaa ole, verrataankin kouluja sen perusteella, otettiinko aamiaisklubitoiminta käyttöön aikaisemmin vai myöhemmin.

### *Vertailuryhmän hankinta*

Joissakin tapauksissa on suhteellisen helppoa tunnistaa mahdollinen vertailuryhmä, joka on samanlainen kuin se johon interventio kohdistetaan. Mutta

asia ei aina ole näin yksinkertainen (katso tarkenne 11). Vertailuryhmän hankkiminen on vaikeampaa jos:

- kohderyhmää ei ole määritelty kovin tarkasti (esim. asunnottomuus-riskin kohteena olevat nuoret), erityisesti jos valintakriteerien käytännön soveltamiseen jää harkintavapautta.
- kohderyhmän ominaisuudet eivät ole helposti havaittavissa (esimerkiksi mielenterveysongelmaiset tai lukemisvaikeuksista kärsivät lapset), jolloin jonkinlainen seulontaprosessi voi olla välttämätön
- palvelu on universaali tai hyvin laajasti käytetty, jolloin voi olla vaikeaa löytää ihmisryhmää, joka ei jo käytä palvelua (esim. uusi saattohoitolaki)
- kohderyhmä on hyvin pieni ja/tai hyvin hajanainen, jolloin on vaikeampaa (ja kalliimpaa) löytää riittävästi samanlaisia ihmisiä
- ei ole suoranaisia sopivia alueita tai paikkoja, joista vertailuryhmä voitaisiin rekrytoida, tai alueita on useita, mikä tekee rekrytointiprosessista monimutkaisemman
- ohjelmaan osallistuminen perustuu asiakkaan omaan valintaan (esim. vapaaehtoiset) tai on palvelun tarjoajan tarkasti valvomaa. Tämä tuo mukanaan uusia sekoittavia tekijöitä, kuten asiakkaan motivaation, ja vaikeuttaa vertailuparin löytämistä. Sekoittavilla tekijöillä voi olla huomattava vaikutus lopputulokseen.

## Tarkenne 11. Ongelmia vertailuryhmän muodostamisessa

### **Turvapaikka kaupungissa –tapaustutkimus** (Safe in The City , SITC)

Vertailuryhmän muodostaminen oli SITC -hankkeen arvioinnissa erityisen ongelmallista, ja on todennäköistä, että samankaltaiset kysymykset rajoittavat kokeellisen lähestymistavan soveltamista muissakin sosiaalialan arvioinneissa. Osaprojekteissa vastustettiin voimakkaasti sitä, että jotkut projektiin ohjatut nuoret olisi satunnaisesti tai muuten ohjattu vertailuryhmään, jolloin he eivät olisi saaneet palvelua. Kysymys ei olisi ollut niin merkittävä, jos olisi tiedetty, että projekteihin tulee olemaan pitkät jonotuslistat. Hankkeen alkuvaiheessa tätä oli kuitenkin vaikea ennustaa. Kun vierailimme osaprojekteissa, jotkut näistä olivat suhteellisen uusia tai ne olivat vasta laajentaneet palveluaan. Siksi ne tekivät parhaansa kehittääkseen hyviä suhteita mahdollisiin asiakaita ohjaaviin yksiköihin. Projekteissa oltiin huolissaan siitä, että yksiköt voisivat lopettaa asiakkaiden lähettämisen, jos niissä tiedettäisiin, että osa asiakkaista ohjataan vertailuryhmään. Toinen mahdollisuus oli muodostaa vertailuryhmä niistä nuorista, jotka kieltäytyivät ottamasta osaa SITC-hankkeeseen tai jotka keskeyttivät osallistumisensa varhaisessa vaiheessa. On kuitenkin todennäköistä, että nämä nuoret ovat vähemmän motivoituneita kuin ne, jotka ovat mukana koko hankkeen ajan. Ei siis voida olettaa, että se mitä tapahtuu keskeyttäneille edustaisi sitä, mitä tapahtuisi hankkeessa loppuun asti pysyneille, jos he eivät olisi osallistuneetkaan hankkeeseen.

Ainoa toteuttamiskelpoinen mahdollisuus olikin vertailuryhmän muodostaminen toisesta Lontoon kaupunginosasta, sellaisesta, jossa ei hanketta toteutettu. Vertailuryhmän muodostaminen näiltä alueilla vaatii jonkinlaisen “jäljittelevän” ohjausjärjestelmän. Järjestelmässä paikallisia palvelujen tarjoajia, esimerkiksi koulukuraattoreja tai tuetun opetuksen yksiköitä (Pupil Referral Units) pyydetäisiin tunnistamaan ominaisuuksiltaan SITC-ohjelmaan osallistuvia muistuttavat nuoret, ja välittämään heitä koskevat tiedot tutkimusryhmälle. Järjestelmää, jolla nuoria ohjattiin SITC-ohjelmaan, olisi vaikea jäljitellä, koska potentiaalisten asiakkaiden valikointiin käytettyjä kriteereitä ei oltu määriteltä selvästi ja ne vaihtelevat eri osaprojekteissa. Kuitenkin uusi standardoidumpi arviointiprosessi tekisi asian helpommaksi. Tämä lähestymistapa on kuitenkin hyvin riippuvainen alueella toimivien yksiköiden välisestä yhteistyöstä, ja nämä eivät välttämättä sijoita tutkimusta tärkeimpien tehtäviensä joukkoon.



## *Potentiaalisia virhelähteitä*

Ei-satunnaisotokseen perustuvat ryhmät lisäävät tulosten virhemahdollisuuksia. Jos interventioryhmä ja vertailuryhmä eivät ole identtisiä, ainakin osa eroista oletetuissa vaikutuksissa voi olla seurausta ryhmien ominaisuuksien eroavuuksista eikä interventiosta. Tämä ei ole vain teoreettinen mahdollisuus, sillä kun satunnaisotoksen sijasta osallistuminen on vapaaehtoista tai asiakkaat valitaan ohjelmaan, vertailtavista ryhmistä voi monestakin syystä muodostua erilaisia, vaikka eroja on vaikea havaita. Esimerkiksi palvelun tarjoajat, joita arvioidaan asiakkaiden saavuttamien tulosten perusteella, voivat valikoida asiakkaisiin sellaisia, jotka heidän kokemuksensa perusteella kaikkein todennäköisimmin tulevat menestymään hyvin.

Tämän vuoksi on tärkeää tunnistaa, ja aina kun mahdollista, kontrolloida interventioryhmän ja vertailuryhmän väliset erot. Kehittyneemmät kvasikokeelliset tutkimukset pyrkivät eri tavoin minimoimaan tai kontrolloimaan potentiaalisia vääristymien aiheuttajia (katso tarkenne 12). Näillä tekniikoilla onnistutaan todennäköisemmin, jos informaatiota kerätään niin monista olennaisista muuttujista kuin mahdollista (muuttujista, joilla voi olla merkitystä intervention vaikutuksiin). Otoskoosta riippuu, kuinka monta näistä tekijöistä voidaan käytännössä kontrolloida. Kvasikokeellisia tutkimuksia pidetään yleensä seuraavaksi parhaina RCT-asetelmaan perustuvien tutkimusten jälkeen, mutta usein ei keskustella siitä, onko toiseksi paras riittävän hyvä. Tämä riippuu siitä, kuinka hyvin kvasikokeellisessa tutkimuksessa kyetään kontrolloimaan virhelähteitä. Tutkittaessa tukijärjestelmien piiristä työelämään siirtymistä hylätään kvasikokeelliset asetelmat joskus siksi, että asetelmat eivät mahdollista tärkeän virhelähteen, asiakkaiden motivaation tai asenteen kontrolloimista, koska nämä eivät ole määrällisesti mitattava (Payne, 2000). On vähemmän ongelmallista, jos keskeiset ominaisuudet ovat havaittavia ja mitattavia tai mahdollinen virhe on pieni suhteessa ohjelman odotettuihin vaikutuksiin. Valitettavasti tutkimusten tulokset esitetään toistuvasti ilman mahdollisten virhelähteitten pohdintaa. (katso Tulosten tulkinta ja esittäminen luvussa 9).

### **Tarkenne 12.** Mahdollisten virhelähteiden huomioon ottaminen

#### **Esimerkki 1:** Sovitetut eli kaltaistetut parit

Kun kontrolliryhmää ei ole voitu valita satunnaisotannalla, interventioryhmän ja vertailuryhmän välisiä eroja voidaan minimoida käyttämällä sovitettuja pareja. Hyvä esimerkki tästä on alkoholiongelmaisille suunnatun amerikkalaisen hoito-ohjelman arviointi. Siinä arvioitiin hoidon lyhentämisen merkitystä vertailemalla pitkään hoitoon osallistunutta asiakasryhmää tai kohorttia (vertailuryhmä) asiakasryhmään, joka osallistui hoito-ohjelmaan sen merkittävän lyhentämisen jälkeen (interventioryhmä). Suora ryhmien välinen vertailu olisi voinut olla harhaanjohtava. Esimerkiksi, jos interventioryhmässä olisi enemmän asiakkaita, joiden alkoholiriippuvuus olisi vähäisempää, tämä vinouttaisi tuloksia interventioryhmän eduksi. Sen sijaan tutkijat yrittivät löytää jokaiselle lyhennettyyn ohjelmaan osallistuvalla interventioryhmään kuuluvalla samankaltaisen parin

vertailuryhmään kuuluvista. Perusteena parien muodostamisessa oli sukupuoli, ikä ja alkoholiriippuvuuden vakavuus. Näin saatiin 75 paria eli mukaan tutkimukseen saatiin 150 henkilöä alkuperäisestä 212 henkilön otoksesta. Samankaltaisista pareista muodostuvien ryhmien vertailu osoitti, että vuoden seurantajakson jälkeen hieman yli puolet asiakkaista kummassakin ryhmässä raitistui tai alkoholin käyttö muuttui kohtuulliseksi. Hoito-ohjelman lyhentyminen ei siis näyttänyt vaikuttavan tuloksiin, mutta lyhyemmän hoitojakson kustannukset olivat merkittävästi alhaisemmat. Näin ollen tutkijat totesivat, että uusi hoito-ohjelma oli kustannusvaikuttavampi.

(Lähde: Long *ym.*, 1998)

### **Esimerkki 2:** Tilastolliset menetelmät

Seuraavana esiteltävä kahden nuorille rikoksentekijöille suunnatun intensiivisen yhteisöperusteisen intervention arviointi on esimerkki tilastollisten menetelmien käytöstä mahdollisen virhelähteiden kontrolloimiseksi. Tutkimuksessa tarkasteltiin interventioiden vaikutusta rikosoikeudellisiin kustannuksiin ja tätä verrattiin tavanomaiseen koeaikaan. Toisin sanoen, onnistuivatko intensiiviset interventiot ehkäisemään rikoksia ja niiden aiheuttamia rikosoikeudellisia kustannuksia. Kaikki kolme nuorten rikoksia käsittelevää tuomioistuinta Missisipissä osallistui ohjelmaan, joten tutkittavien valitseminen satunnaisesti ei ollut mahdollista. Lisäksi keskeyttäneiden määrä oli hyvin korkea, hieman alle 50 %, mikä oli toinen potentiaalinen virhelähde.

Kvasikokeellista tutkimusasetelmaa vahvistettiin tilastollisesti. Ensinnäkin, tutkijat estimoivat osallistujien ominaisuuksien avulla regressioyhtälön, jolla voitiin ennustaa keskeyttämistodennäköisyyttä. Muuttujina käytettiin ikää, aiempia kontakteja rikosoikeusjärjestelmään ja erilaisia käyttäytymiseen liittyviä indikaattoreita. Yhtälö auttaa estämään sellaista tulosten vinoutumista, joka johtuu tietyn tyyppisten asiakkaiden keskeyttämisistä ja putoamisista pois arviointitutkimuksesta.

Tämän jälkeen tutkijat estimoivat toisen yhtälön ennustamaan rikosoikeudellisia kustannuksia ohjelmaan osallistuneiden asiakkaiden ominaisuuksien ja saaman ohjauksen funktiona. Jos kustannukset riippuvat vain asiakkaiden ominaisuuksista, ei asiakkaan saama ohjausta - interventiota - kuvaavalla muuttujalla ole merkitystä kustannuksille. Kuitenkin, kun erot nuorten ominaisuuksissa oli otettu tilastollisesti huomioon, tulokset osoittivat toisen kahdesta intensiivisestä interventiosta, kognitiivisen käyttäytymisterapian, vähentävän 18 kuukauden aikana merkittävästi rikosoikeudellisia kustannuksia. Arvioidut säästöt olivat hieman alle \$3,000 osallistujaa kohti, joka on noin kaksi kertaa enemmän kuin arvioidut interventioista aiheutuvat lisäkustannukset. Tällainen analyysi parantaa myös tutkimustulosten uskottavuutta, jos keskeisimpien muuttujien kertoimet ovat yhdenmukaisia tehtyjen ennako-oletusten kanssa, esimerkiksi,

jos nuori on ollut aiemmin tekemisissä rikosoikeusjärjestelmään kanssa, kuluja aiheutuu enemmän.  
(Lähde: Robertson ym., 2001)

### **Vertailtavat ryhmät**

Kvasikokeellisten tutkimusten pitäisi pyrkiä kontrolloimaan interventioryhmän ja vertailuryhmän ominaisuuksien välisiä eroja. Silloin kun tämä ei ole mahdollista, on ehkä tarkoituksenmukaisempaa puhua vertailtavista ryhmistä erotuksena täysin vastaavista (matched) tai tilastollisesti kontrolloiduista vertailuryhmistä. Vertailtavia ryhmiä käytetään usein ohjelmissa, jotka suunnataan kokonaisille alueille tai instituutioille, jolloin otoskoko on liian pieni tilastolliseen analyysiin tai vastaavan vertailuryhmän löytämiseen. Tarkenteessa 13 on esimerkki, jossa koulussa tehtävää sosiaalityöstä tutkittiin käyttäen vertailtavaan ryhmää.

Tämä voi olla myös tarkoituksellinen valinta, jos tutkittavien alueiden tai instituutioiden määrää halutaan rajoittaa niin, että niitä voidaan tutkia perusteellisemmin. Ratkaisua tulee punnita suhteessa suuremman ja siksi edustavamman otoksen etuihin. Erityisesti monimutkaisemmissa interventioissa on tärkeää, että ei kiinnitetä huomiota vain mitattaviin vaikutuksiin, vaan yritetään myös ymmärtää, miten ohjelma toimii ja miksi se on vaikuttava tai ei sitä ole. Tämä muistuttaa tapaustutkimuksen analyysiä, ja laadullista informaatiota käytetään helpottamaan määrällisen tiedon tulkintaa. Hyvä esimerkki tästä on Priority Estate Project (PEP, ”Suositut lähiöt”) arviointi (Glennister ja Turner, 1993). Suora vertailu Hullin kokeilulähiön ja vertailtavan lähiön välillä ei osoittanut uudenlaisen asumisen hallinnon saavan aikaan paljoa muutosta. Mutta koko lähiöitä koskeva vertailu peitti alleen merkittävän vaihtelun Hullin lähiön sisällä. Osa lähiöstä menestyi hyvin ainakin osittain PEP toimintamallin tuomien hyödyllisten muutosten ansiosta. Samanaikaisesti eräs talo oli rappeutunut huomattavasti, koska sinne oli sijoitettu suuri määrä nuoria ja aikaisemmin kodittomia asukkaita. Näin ollen kaikki PEP:n hyödylliset vaikutukset kuittautuivat pois paikallisviranomaisten sijoituspolitiikan muutoksen johdosta. Tämä on ohjelman kannalta ulkoinen vaikutus, joka tuskin olisi noussut esille suuremman lähiöjoukon pinnallisemmassa analyysissä.

Hyvää vertailtavaa ryhmää voi olla vaikeaa löytää, koska ei ole olemassa kahta samanlaista aluetta tai instituutiota – ja on monia erilaisia huomioon otettavia tekijöitä. Vaikka kaksi aluetta tai instituutiota olisivat hyvin samanlaisia arvioinnin alkaessa, olosuhteet voivat muuttua arviointiajan kuluessa. Vaihtoehtoisesti ”kokeilu” voi saastua tutkittavan alueen tai instituution toimintatapojen muuttuessa. Esimerkiksi PEP:n arvioinnissa erästä vertailtavaa lähiötä hallinnoivat päättivät ottaa käyttöön joitakin PEP-toimintamallin osia.

Koska satunnaisotokseen perustumattomiin ja yleensä pienimuotoisiin vertailtavista ryhmiä käyttäviin tutkimuksiin liittyy paljon virhelähteitä, on vieläkin tärkeämpää, että pohditaan tutkimustulosten vaihtoehtoisia selityksiä. Samalla kun tällainen tutkimus on selvästi parempi, kuin jos vertailuryhmää

ei olisi laisinkaan, tällaista tutkimusta on tuettava muulla näytöllä, jotta interventio ja havaitut vaikutukset saadaan liitettyä toisiinsa. Ei voida olettaa, että kaikki erot interventioryhmän ja vertailijaryhmän välillä johtuvat välttämättä interventiosta.

Pienimittakaavaiset tutkimukset täydentävät tietoa tietyistä interventiosta. Aikaa myöten, jos esimerkiksi toisiaan seuraavat tutkimukset tuottavat samanlaisia tai yhdenmukaisia tuloksia, voidaan muodostaa tarkempi ja täydellisempi kokonaiskuva. Jos tutkimuksia on riittävästi, voi jopa olla mahdollista toteuttaa meta-analyysi (katso myöhemmin tässä luvussa).

**Tarkenne 13.** Esimerkki vertailtavaa ryhmää käyttävästä taloudellisesta arvioinnista

**Kustannus-hyöty –analyysiin perustuva koulusosiaalityön kokeellinen tutkimus** (An experimental study of school social work with cost-benefit analyses)

Kysymyksessä oli kolmevuotinen koulusosiaalityön kokeiluhanke. Ala-asteella, yläasteella ja lukiossa työskenteli yksi osaaikainen ja kaksi kokoaikaista sosiaalityöntekijää. Koulut olivat huonomaineisilla vuokratiloilla Englannin Dorsetissa.

Projektin kahta koulua verrattiin muihin saman alueen ala-asteen ja yläasteen kouluihin ja lukioon. Tietoja kerättiin 1300 lapsen ja nuoren koulupinnauksesta, itse raportoiduista varkauksista, koulukiusaamisesta, huumeiden käytöstä ja ilkeistä sekä koulusta erotettujen nettomääristä – koulusta erotettujen määrää vähennettiin koulutyöhön palanneet. Lisäksi kerättiin laadullista aineistoa haastatteleamalla kaikki projektikoulujen opettajat sekä pieni otos koululaisten vanhempia ja koululaisia.

Tulosten mukaan kuusi seitsemästä koulun ongelmia kuvaavasta indikaattorista osoitti ongelmien vähenemistä, myös erotettujen nettomäärässä aleni. Perustuen arvioon koulusta erotettujen vähenemisestä koituvista kustannussäästöistä tutkimuksen tekijät päättelivät, että kokeilu oli ‘hyvin kustannusvaikuttava’. Kvalitatiivinen analyysi näytti tukevan näitä tuloksia ja antoi viitteitä siitä, että koulussa tehtävä sosiaalityö voi parantaa koulun ilmapiiriä ja opettajien henkistä vireyttä.

Vaikka tulokset olivat kiistämättä lupaavia, tutkimuksessa oli muutamia tiedostettuja ongelmia. Arviointiprosessin aikana osoittautui, että yksi vertailussa mukana olleista kouluista ei ilmeisestikään ollut paras mahdollinen ja jälkiviisaasti - sekä uuden demografisen ja sosiaalipalveluja koskevan tiedon avulla – tämän koulun sijasta tutkimukseen olisi pitänyt valita toinen koulu. Siten muilla kuin kokeilusta johtuvilla eroilla koulujen välillä saattoi olla osuutta siihen, että kokeilussa mukana olleet koulut vaikuttavat suoriutuneet paremmin. Tulosten vahvistami-

seksi tarvitaankin lisää arviointitutkimusta siten, että hanke toteutetaan eri kouluissa ja eri sosiaalityöntekijöiden toimesta ja vertailukoulut valitaan paremmin.  
(Lähde: Bagley ja Pritchard, 1998)

### **Ei vertailtavaa eikä vertailuryhmää**

Monissa arviointitutkimuksissa vaikutukset arvioidaan ennen interventiota ja sen jälkeen, mutta ilman vertailuryhmää. Tämän tyyppiset tutkimukset tosiasiallisesti hallitsevat vaikuttavuutta käsittelevää kirjallisuutta sosiaalityön kaltaisilla alueilla (MacDonald, 1996). Tällaiset tutkimukset tuottavat harvoin tuloksia, joita voidaan käyttää taloudellisen arvioinnin osana, jollei ole vahvaa näyttöä siitä, että muutoksia ei olisi tapahtunut ilman interventiota. Tuloksia pitää tulkita hyvin varovaisesti; parhaimmillaan ne ovat suuntaa antavia. Tulosten luotettavuus voi parantua, jos useat, eri käyttäjillä ja erilaisissa olosuhteissa toteutetut tutkimukset tuottavat samanlaisia tuloksia (MacDonald, 1996). Mutta mieluummin kuin lopullisena näyttönä intervention vaikuttavuudesta tai kustannus-vaikuttavuudesta, tällaisia tutkimuksia tulisi käyttää esitutkimuksen luonteisesti, esimerkiksi perustelevaan perusteellisemmän (ja kalliimman) arvioinnin tarvetta.

### **Täydentäviä tutkimusmenetelmiä**

Silloin kun muita tutkimusmenetelmiä käytetään osana taloudellista arviointia, niitä todennäköisesti käytetään täydentämään jotakin edellä tarkasteltua kokeellista menetelmää. Yleensä, mitä monimutkaisempi interventio on, sitä todennäköisempää on, että puhdas kokeellinen arviointi ei sellaisenaan ole riittävä. Näin oli asianlaita tapaustutkimuksissamme, joihin kaikkiin sisältyi muita tutkimusmenetelmiä.

### **Mallintaminen**

Malleissa käytetään sekundääristä aineistoa (secondary data) ja muita tietolähteitä, kuten asiantuntijoiden mielipiteitä ennustettaessa tietyn intervention vaikutuksia ja kustannuksia hyvin jäsenetyissä kvantitatiivisissa puitteissa. Kaikki mallit sisältävät panoksia ja oletuksia sekä joukon enemmän tai vähemmän monimutkaisia laskelmia vaikutusten tai kustannusten ennustamiseen. Esimerkiksi mallinnettaessa potilaiden kolesterolitason alentamiseen tähtäävien interventioiden pitkäaikaisia hyötyjä on käytetty laajaa tietokantaa ihmisten terveystiedoista ja kolesterolitasoista (Stinnett ym., 1996). Sosiaalialalla mallintamista on käytetty laajasti arvioitaessa veroetujen muutosten vaikutuksia (Mitton ym., 2000). Yksi tapaustutkimuksistamme käyttää mallia ennustamaan, kuinka erilaiset kotitaloudet voivat reagoida asuntonsa energiatehokkuuden muutokseen (katso tarkenne 14).

Mallintamista voidaan käyttää monissa tilanteissa toisen lähestymistavan taustalla tai sijasta (Buxton ym., 1997):

- osana kartoittavaa tutkimusvaihetta helpottamaan niiden muuttujien määrittelyä, jotka ovat todennäköisesti tärkeimpiä ja joihin pitäisi kohdistaa eniten huomiota tutkimuksen suunnittelussa
- tehtäessä välillisiä vaikutuksia koskevan tiedon avulla päätelmiä lopullisista vaikutuksista (esim. elinajan odotusarvosta, joka perustuu lyhyen ajan sairastavuuteen), tai tehtäessä vaikutuksista päätelmiä arviointijaksoa pitemmälle ajalle (esimerkiksi elinikäinen rikollisuus nuoruusajan rikosrekisterin perusteella)
- tutkittaessa kustannusten ja vaikutusten vaihtelua eri väestöryhmissä (katso kohta ”Tulosten analysointi” luvussa 9)
- kokeellisen arviointiasetelman sijasta silloin kun se on mahdotonta; mallintaminen on erityisen käyttökelpoinen arvioitaessa ohjelmia jälkikäteen tai simuloitaessa uuden ohjelman vaikutusta (tai muutosta olemassa olevaan ohjelmaan) etukäteen.

Mallien etuja ovat joustavuus ja alhaiset kustannukset. Ne kykenevät yhdistämään tietoja monenlaisista lähteistä ja malleja voidaan käyttää testattaessa tulosten herkkyyttä taustaoletusten muutoksille. Mallit eivät ole riippuvaisia suoraan käyttäjiltä kootuista tiedoista, vaan hyödyntävät olemassa olevia tietolähteitä kuten hallinnollisia tietokantoja, laajoja kotitalouskyselyjä ja aikaisemmin julkaistujen arviointien tuloksia. Mikäli tiedoissa on puutteita, ne voidaan korvata oletuksilla siihen asti, kunnes luotettavaa tietoa saadaan.

Joka tapauksessa, mallintamisella saatujen tulosten hyvyys on riippuvainen käytettyjen tietojen laadusta ja taustaoletusten pätevyydestä. Mallit eivät korvaa tietojen puutetta vaikka ne auttavatkin käyttämään rajallisia tietoja mahdollisimman hyvin. Tietojen puute rajoittaa mallintamisen käytettävyyttä monilla sosiaalisen hyvinvoinnin alueilla. Toinen haittapuoli malleista on se, että ne voivat olla vaikeasti ymmärrettäviä niille, jotka eivät ole olleet mukana malleja suunniteltaessa. Tällaisia ovat muun muassa päätöksentekijät, joille mallien taustaoletukset eivät välttämättä tule selviksi (Sassi, 2000).

#### **Tarkenne 14. Esimerkki mallintamisesta**

##### **Lämmityspolttoaineen puute –tapaustutkimus (Fuel-poverty)**

Yhtenä kotien energiatehokkuussuunnitelman (The Home Energy Efficiency Scheme, HEES) tavoitteena on vähentää kylmissä kodeissa asuvien ihmisten määrää. Kuitenkin tiedetään suhteellisen vähän siitä, miten energiatehokkuuden parantaminen vaikuttaa ihmisten tapaan säätää kotinsa lämpötilaa. Ihanteellisessa tapauksessa tietoa kerättäisiin kotien lämpötiloista ennen energiatehokkuudessa tehtyjä parannuksia ja niiden jälkeen. Koska tällaista tietoa ei ollut saatavana, arvioimme todennäköistä vastetta käyttämällä kyselyaineistoa. Kyselyyn vastasi edustava otos Ison-Britannian kotitalouksia, ja kysely sisälsi informaatiota 7000 kotitaloudesta, muun muassa kodin lämpötilan ja energiatehokkuusluvun.

Tuotettu mallimme perustui yksinkertaiselle taloudelliselle teorialle. Kuluttajan päätöksenteon taloudellinen analyysi osoittaa, että tuotteen kysyntä riippuu sen hinnasta ja kuluttajan tuloista, henkilökohtaisista mieltymyksistä ja muista ominaisuuksista. Muiden olosuhteiden pysyessä muuttumattomina hyödykkeen tai palvelun kysyntä lisääntyy, jos sen hinta laskee tai kuluttajan tulot nousevat. Tätä perusrakennetta voidaan soveltaa kodin lämpötilaa koskeviin päätöksiin, siis 'lämmön kysyntään'. Ihmiset päättävät, miten lämpöisestä he kotinsa pitävät ja ratkaisuun vaikuttaa se, miten paljon lämmittäminen maksaa, millaiset heidän tulonsa ovat ja myös muut seikat. Esimerkiksi, jos heillä on pieniä lapsia, he voivat suosia korkeampaa lämpötilaa. Me olimme erityisen kiinnostuneita siitä, miten herkästi ihmisten "lämmön kysyntä" muuttuu heidän kotiensa energiatehokkuuden muuttuessa. Energiatehokkuuden muutos tarkoittaa muutosta 'lämmön hinnassa'. Pienentämällä lämmityskustannuksia energiatehokkuuden paranemisen pitäisi nostaa kotien lämpötilaa, mutta kuinka paljon?

Estimoimme lämmön kysyntäyhtälön käyttämällä tilastollista (regressio-)analyysiä ja aineistona käytettiin Englannin kotitalouksien olosuhdekyselyä (English House Condition Survey) vuosilta 1991 ja 1996. Riippuvana muuttujana eli asiana, jota halusimme selittää, oli olohuoneen lämpötila kyselyajankohtana. Avainselittäjänä oli 'lämmön hinta', arvioitu asunnon lämmittämisen päivittäinen kustannus standardoidulle tasolle 18°C eteisessä ja 21 °C olohuoneessa. Tämä on suorassa suhteessa kiinteistön energiatehokkuuteen. Muut muuttujat, jotka saattavat vaikuttaa ihmisten lämmön kysyntään lisättiin regressioyhtälöön. Näitä olivat kotitalouden tulot, hallintamuoto sekä vanhuksen tai lapsen läsnäolon. Teimme useita regressioajoja, mutta parhaiten toimi yksikertainen lineaarinen regressio olohuoneen lämpötilan ja standardoitujen lämmityskustannusten välillä, jossa dummy – muuttujina olivat tuloluokat ja muut tilastollisesti merkitsevät muuttujat. Sitä, miten luotettavia tulokset olivat, testattiin kattavasti. Kokeiltiin muun muassa epälineaarisia malleja ja analyysistä poistettiin tiettyjä havaintoja, esimerkiksi, jos lomakekyselyyn vastaaja oli ilmoittanut, että lämpötila poikkesi normaalista.

Keskeinen tulos oli odotusten mukaisesti tilastollisesti merkittävä negatiivinen 'hintavaikutus'. Kertoimen suuruusluokka merkitsee sitä, että £1 päivittäinen lämmityskustannuksen pieneneminen liittyy olohuoneen lämpötilan kohoamiseen 0.35°C:lla. Vaikutus on suhteellisen pieni, mutta vastaa lämmön kaltaista välttämättömyshyödykettä koskevia odotuksia. Todisteita ei löytynyt myöskään siitä, että lämpötilaa olisi nostettu enemmän (tai vähemmän) köyhemmissä tai ikääntyneissä kotitalouksissa, hankkeen pääasiallisessa kohderyhmässä. Mallin avulla voidaan arvioida myös energiatehokkuussuunnitelman (HEES:n) vaikutusta kylmien kotien lukumäärään Englannissa.

## Laadullisia tutkimusmenetelmiä

Laadullinen tutkimus kattaa suuren joukon menetelmiä (katso tarkenne 15). Toisin kuin edellä käsitellyissä menetelmissä, laadullisessa tutkimuksessa tarkoituksena ei ole tuottaa numeerisia vastauksia vaan keskittyä yksilöiden ja organisaatioiden kokemuksiin, näkemyksiin, merkityksiin ja vaikuttimiin. Laadullinen tutkimus on hyödyllistä tarkasteltaessa eräitä arvioinnin kysymyksiä: ei ”kuinka monta x:ää” vaan ”mikä x on, kuinka x vaihtelee eri tilanteissa ja miksi?” (Pope ja Mays, 1995). On useita tapoja, joilla laadullisia menetelmiä voidaan käyttää osana taloudellista arviointia (esim. Coast, 1999):

- alustavana vaiheena auttamassa määrällisen tutkimuksen kattavuuden ja asetelman määrittelemistä; esimerkiksi tunnistettaessa vaikutuksia, joita käytännön toimijat tai käyttäjät pitävät tärkeinä (ja jotka siksi pitäisi sisällyttää arviointiin) ja tapoja näiden vaikutusten mitaamiseksi (esim. auttamalla kyselyjen kehittämisessä)
- tarjoamaan laadullisia arvioita monimutkaisista vaikutuksista, joita on vaikea mitata määrällisesti (esim. muutokset organisaatiokulttuurissa tai yhteisöllisyydessä) ja jotka voidaan esittää määrällisten vaikutusten rinnalla
- auttamaan määrällisen tutkimuksen tulosten tulkinnessa (esim. miksi tiettyjä vaikutuksia on havaittu, erityisesti jos tulokset ovat odottamattomia tai ilmeisen yhteen sopimattomia)
- arvioimaan, vaikuttaako interventio toimivan tarkoitetulla tavalla, ja tästä askel eteenpäin, tarjoamaan laadullista näyttöä intervention yhteydestä havaittuihin vaikutuksiin
- ymmärtämään paremmin, miksi interventio toimii (tai ei toimi) tiettyissä olosuhteissa ja näin ollen antamaan käsitystä siitä, pitävätkö määrällisen tutkimuksen tulokset paikkansa laajemmin ja missä olosuhteissa.

Taloustieteilijät käyttävät nykyisin harvoin laadullisia tutkimusmenetelmiä. Vaikka menetelmät tunnettaisiin, kannetaan huolta niiden tieteellisestä täsmällisyydestä. Mutta tämä johtuu pääosin siitä, että laadullisen tutkimuksen tarkoitus on ymmärretty väärin. (Mays ja Pope, 1995a). Tavoitteena ei ole tuottaa tilastollisesti merkitseviä tuloksia vaan tutkia teemoja, rakenteita ja yhteyksiä käyttämällä rikasta ja monimuotoista tietoa-aineistoa. Siinä missä kvantitatiivinen tutkija pyrkii osoittamaan syy-seuraussuhteita etsimällä muuttujien välistä tilastollista korrelaatiota, kvalitatiivinen tutkija etsii selityksiä. Sellaisia selityksiä, joita haastateltavat antavat itse, mutta myös tutkijoiden tekemiä päätelmiä sillä perusteella, mitä ihmiset arvioivat tapahtuneen ja kuinka he ovat reagoineet.

Vaarana on, että taloustieteilijät käyttävät näitä lähestymistapoja liian yksinkertaisesti (Coast, 1999). Mutta asianmukaisesti käytettynä laadullinen tutkimus voi olla avuksi tartuttaessa moniin kokeellisen lähestymistavan puutteisiin. Olemme aiemmin esittäneet joitakin esimerkkejä siitä, kuinka laadullinen tutkimus olisi voinut auttaa äiteihin kohdistetun synnytyksenjäl-



keisen tukitoimintamallin arviointia. Tarkenteessa 16 on lisää tapaustutkimustemme perusteella laadittuja esimerkkejä laadullisten menetelmien käyttömahdollisuuksista. Yksi haaste tällä alueella työskenteleville ekonomisteille on löytää parempia tapoja yhdistää määrällisen ja laadullisen analyysin tuloksia yhteiseen tarkastelukehikkoon.

## **Tarkenne 15. Laadullisia tutkimusmenetelmiä**

### **Syvähaastattelut**

Nämä ovat henkilökohtaisia haastatteluja, joiden avulla kiinnostuksen kohteena olevia asioita voidaan tutkia syvällisesti. Ne perustuvat väljästi ennalta suunniteltuihin avoimiin kysymyksiin, jotka ainakin alustavasti määrittävät tutkittavan alueen ja joista haastattelija tai haastateltava voi tarvittaessa poiketa. Laadulliset haastattelut ovat vuorovaikutteisia ja joustavia. Tavoitteena on päästä pintaa syvemmälle, saada yksityiskohtaisemmin - haastateltavan näkökulmasta - selville, mitä mieltä ihmiset ovat, ja paljastaa uusia, ennalta suunnittelemattomia asioita ja ajatuksia. Laadulliset haastattelut ovat erityisen käyttökelpoisia, kun tutkitaan ihmisten kokemuksia tai käyttäytymistä, mielipiteitä tai uskomuksia, tunteita tai tiettyä asiaa koskevaa tietämystä.

(Lähde: Britten, 1995)

### **Fokusryhmät**

Nämä ovat ryhmähaastattelumenetelmiä, joille on ominaista ryhmän sisäinen vuorovaikutus. Vuorovaikutusta käytetään myös hyväksi haastatteluissa. Sen sijaan, että ryhmän jäseniä pyydetäisiin vuoron perään vastaamaan kysymyksiin, ryhmää rohkaistaan keskustelemaan asiasta, kyselemään, kertomaan tarinoita ja kommentoimaan toisten kokemuksia ja näkemyksiä. Erityisen käyttökelpoinen menetelmä on, kun tutkitaan kokemuksia, esimerkiksi asunnottomuudesta tai tietämystä, esimerkiksi asunnottomien saatavilla olevista palveluista. Sen lisäksi mitä ihmiset ajattelevat, menetelmällä voidaan tutkia myös, miksi he ajattelevat tietyllä tavalla. Ryhmähaastattelu soveltuu hyvin tilanteisiin, joissa haastattelija haluaa rohkaista osallistujia tuomaan esille heille tärkeitä asioita heidän omalla kielellään.

(Lähde: Kitzinger, 1995)

### **Havaintomenetelmää käyttävät tutkimukset**

Näissä tutkija järjestelmällisesti menee tarkkailemaan ihmisiä ja tilanteita tehdäkseen havaintoja käyttäytymisestä ja vuorovaikutuksesta luonnollisissa olosuhteissa, esimerkiksi palvelupisteissä. Joskus tutkijat osallistuvat toimintaan, jota he havainnoivat, mutta voivat myös tehdä tutkimusta huomaamattomina. Merkittävänä etuna havainnoivissa tutkimuksissa on se, että voidaan selvittää eroja ihmisten kertomusten ja todellisen toiminnan välillä. Menetelmän avulla voidaan jopa paljastaa käyttäytymistä ja rutiineja, joista toimijat itsekään eivät ole tietoisia. Erityisen hyvin havaintomenetelmät soveltuvat organisaatioiden toiminnan tutkimiseen.

(Lähde: Mays ja Pope, 1995b)

### **Tapaustutkimukset**

Tapaustutkimukset ovat perusteellisia ja kohdistuvat yhteen tai muutamaankin tilanteeseen, jossa laajaa, monitahoista kysymystä on selvitettävä monimutkaisessa ympäristössä. Tavallista on, että tapaustutkimuksissa käytetään monenlaisia tutkimusmenetelmiä, sekä kvantitatiivisia että kvalitatiivisia. Kvalitatiiviset, laadulliset menetelmät ovat parhaita, kun halutaan tutkia, miksi ja miten interventio onnistui tai ei onnistunut, kun yleisten olosuhteiden oletetaan vaikuttavan lopputulokseen ja kun tutkija ei voi vaikuttaa tapahtumien kulkuun. Tapaukset voidaan valita siten, että ne edustavat erilaisia tilanteita tai testaavat (Lähde: Kriin ja Packwood, 1995)

## **Tarkenne 16. Esimerkkejä laadullisten tutkimusmenetelmien käytöstä**

### **Lämmityspolttoaineen puute –tapaustutkimus (Fuel poverty)**

Vaikka keskeisen osan tutkimuksesta muodosti energiatehokkuuden parantamisen vaikutusta arvioivan mallin kehittäminen (kts. Tarkenne 14), käytimme myös laadullisia tutkimusmenetelmiä tukemaan kvantitatiivisten tulosten tulkintaa ja antamaan niille vahvistusta. Teimme syvähaastattelun 18 kotitaloudelle, jotka olivat osallistuneet energiatehokkuusohjelmaan ja joko odottivat energiatehokkuuden parantamistoimenpiteitä tai ne oli äskettäin tehty.

Haastattelut antoivat tukea kvantitatiiviselle löydökselle, jonka mukaan päätös kodin lämpimyydestä ei ole kovin herkkä kodin energiatehokkuudessa tapahtuville muutoksille. Erittäin harvat olivat halukkaita säästämään rahaa asumisviihtyvyyden kustannuksella, ja tämä näkyi heidän käyttäytymisessään:” Jos on kylmä, kytken lämmityksen päälle... En aio istua tällä palelemassa.” Lisäksi haastattelut auttoivat ymmärtämään paremmin sitä, miten ihmiset tekevät lämmittämistä koskevia päätöksiä. Erityisesti ymmärrys lisääntyi siitä, miten taloudelliset ja muut näkökohdat, joita kvantitatiivinen analyysi ei paljastanut, vaikuttavat lämmittämiseen.

Laadullista aineistoa käytettiin kotitalouksien tyypittelyyn niiden käyttäytymisen mukaan. Esimerkiksi yksi kylmissä rakennuksissa asuva kotitalouksien ryhmä pitää lämmitystä ehdottoman tärkeänä ja tekee kaiken mahdollisen pitääkseen kotinsa lämpimänä. He käyttävät esimerkiksi lämpöpuhaltimia ja lämmittävät pitämällä uunin päällä ja luukkua auki. Tämä ryhmä käyttää keskimääräistä enemmän rahaa lämmitykseen, mutta koska heidän lämmitysjärjestelmänsä ei toimi kunnolla, he joutuvat koko ajan taistelemaan kylmää vastaan. Heillä on mahdollisuus hyötyä eniten energiatehokkuuden parantamisesta sekä huomattavina säästöinä että lisääntyneenä viihtyvyytenä, jota he arvostavat muita ryhmiä enemmän. On myös toinen ryhmä koti-

talouksia, joka asuu energiatehokkuudeltaan huonoissa rakennuksissa, mutta joka käyttää lämmitystä paljon edellistä ryhmää niukemmin lämmittäen ehkä vain yhtä huonetta ja vain silloin kun on ”tarve”. He ovat valmiita tinkimään viihtyvyydestä, mutta taloudelliset näkökohdat eivät kuitenkaan näytä olevan ensisijaisia. Jotkut tällaisista kotitalouksista ovat saattaneet tottua asumaan kylmissä asunnoissa eivätkä siksi tarvitse niin paljon lämmintä. Vaikka tällaiset kotitaloudet hyötyisivät muita enemmän asuntojen energiatehokkuuden parantamistoimista, ne eivät kuitenkaan käytä tätä mahdollisuutta. Ne tyytyvät mieluummin pieneen mukavuuden lisääntymiseen, josta aiheutuu vain vähän tai ei ollenkaan lisäkustannuksia. Tässä esimerkissä laadullinen analyysi auttaa siis yksilöimään sellaisia muuttujia, joita ei voi mitata määrällisesti ja jotka jäävät pois kvantitatiivisista analyyseistä, mutta joilla voi olla merkittävä vaikutus sille, miten ihmiset reagoivat HEES -ohjelmaan.

Laadullinen analyysi viittasi myös, että pääasiallisena vaikuttavuusmittarina käyttämämme olohuoneen mittaushetken lämpötila ei ehkä anna kovin edustavaa kuvaa asunnon lämpötilasta. Ihmisten huoli asunnon lämmityksestä oli yksityiskohtaisempaa. Se saattoi koskea tiettyjä liian kylmiä huoneita tai asunnon osia (tuskin koskaan olohuone), huoneet saattoivat olla kylmiä joinakin aikoina päivästä tai lämpötilan oli epätasainen (esim. eräs nainen kertoi ”palelevansa reunoilla”). Siis hienostuneempia lämpimyyden (tai paremminkin viihtyisyyden) mittareita tarvitaan. Koska käyttämämme mittari kuvasi huonosti asunnon lämpimyyttä, kehittämämme malli todennäköisesti aliarvioi kylmien ja lämpimien asuntojen välistä viihtyvyyseroa ja sen vuoksi voi aliarvioida myös hyötyä, joka saadaan parantamalla huonoimpia asumisolosuhteita.

### **Paikallisyhteisöjen kehittäminen –tapaustudkimus (Gatsby)**

Toteutettuun pilottitutkimukseen kuului jokaisen arvioitavan ryhmän haastattelu. Haastattelimme ryhmiä noin puoli päivää niistä asioista, jotka olimme määritelleet yhteisöryhmien tehokkaan toiminnan kannalta tärkeiksi. Tällaisia olivat esimerkiksi tavoitteen selkeys, kuinka hyvin ryhmä työskenteli tiiminä, miten paikallisyhteisö suhtautui ryhmään ja saiko ryhmä riittävästi tukea muilta toimijoilta. Arvioimme ryhmiä haastattelujen aikana muodostamillamme kriteereillä ja käytimme haastatteluja kyselyyn (kts, Tesktikehys 6) saamiemme vastausten tulkinnassa. Näin ollen laadullisia tutkimusmenetelmiä käytettiin objektiivisten mittareiden puuttuessa arvioitaessa ryhmien tehokkuutta. Haastatteluissa saatujen tietojen avulla parannettiin myös ryhmien arvioinnissa käytettävien kriteerien validisuutta muun muassa kysymällä ryhmiltä, mikä heidän mielestään teki yhteisöryhmästä hyvin toimivan. Kriteerimme pärjäsivät hyvin tässä testissä, ja näytti siltä, että kriteerit kattoivat kaikki ryhmien toiminnan keskeiset osa-alueet.

## **Koulusosiaalityön kokeellinen tutkimus** (An experimental study of school social work)

Koulusosiaalityöhankkeen arviointi (kts. Tesktikehys 13) oli keskeisiltä osiltaan laadullinen. Muun muassa projektiin osallistuvien koulujen opettajat ja otos lapsista ja lasten vanhemmista haastateltiin. Haastattelut antoivat tukea kvantitatiivisten analyysien tuloksille ja tarjosivat todistusaineistoa siitä, että kvantitatiivisesti mitatut positiiviset vaikutukset liittyivät interventioon. Haastatteluissa saatiin myös paljon positiivista palautetta nuorilta, jotka olivat saaneet apua projektin sosiaalityöntekijöiltä. Lisäksi haastattelut toivat esille joukon teemoja, jotka liittyivät käsitykseen onnistuneesta projektista. Näitä olivat muun muassa nuoren yksilöllisyyden tunne, se, miten tärkeää on tietää, että ohjausta on saatavilla, luottamuksellisuus, luottamus ja rauhallisuus. Yksi tekijä, jonka tutkijat tunnistivat, oli sosiaalityöntekijän mahdollisuus puuttua ongelmaan heti sen ilmaantumista. Kaikki nämä esille tulleet asiat auttoivat ymmärtämään, miksi interventio näytti onnistuneen, ja edelleen, mitä asioita myöhemmin toteutettavissa vastaavissa kouluprojekteissa kannattaa tehdä samoin.

### *Meta-analyysi*

Meta-analyyseissä yhdistetään tietoa saman intervention aikaisemmista tutkimuksista. Ne eroavat järjestelmällisistä kirjallisuuskatsauksista siinä, että meta-analyysissä tehdään tilastollisia analyysejä aikaisemmista tutkimuksista saatavilla tiedoilla vastakohtana laadullisemmalle näytön arvioinnille. Tämän lähestymistavan etu on, että se lisää kustannus- ja vaikutusarvioitten pätevyyttä kasvattamalla otoskokoja ja sitä voi käyttää tutkittaessa ohjelman vaikuttavuuden (ja kustannus-vaikuttavuuden) vaihtelua useammanlaisissa olosuhteissa. Meta-analyysi on myös halvempi ja nopeampi kuin täydentävien arviointitutkimusten tekeminen. Meta-analyysi on kuitenkin mahdollinen vain, jos interventioista on jo tehty riittävän hyviä arviointitutkimuksia. Lisäksi meta-analyysillä on taipumus keskittyä julkaistuihin tutkimuksiin, mikä voi vinouttaa tuloksia suosissaan tilastollisesti merkitseviä tuloksia. On myös vaarana sekoittaa ”hyvät ja ”huonot” tutkimukset, ellei heikompia tutkimuksia karsita jollakin tavalla.

Tarkenteessa 17 on esimerkki sosiaalialaa koskevasta meta-analyysistä. Kuitenkin meta-analyysin käyttö on todennäköisesti melko rajallista useimmilla sosiaalisen hyvinvoinnin alueilla, ainakin lähitulevaisuudessa, koska hyvin tehtyjä arviointitutkimuksia, joihin meta-analyysit voisivat perustua, on tehty suhteellisen vähän. Harvat ohjelmat toimivat muuttumattomasti riittävän pitkän aikaa, jotta niistä voitaisiin koota vahvaa näyttöä. Useimmat meta-analyysit tukeutuisivat USA:ssa tehtyihin tutkimuksiin (katso luku 3) ja nämä soveltuvat heikosti Britannian tilanteeseen, koska maiden sosiaalipalvelut on järjestetty eri tavoin.

**Tarkenne 17.** Esimerkki meta-analyysistä

## **Ovatko yksityiset vankilat kustannus-vaikuttavampia kuin julkiset vankilat?**

Kysymyksessä on 33 yksittäisen tutkimuksen meta-analyysi, ja kiinnostuksen kohteena oli USA:ssa sijaitsevien julkisten ja yksityisten miesvankiloiden kustannus-vaikuttavuus. Meta-analyysissä testattiin teoriaa, jonka mukaan yksityiset vankilat ovat julkisia tehokkaampia, mikä johtuu kilpailusta ja vähäisemmästä byrokratiasta. Tutkimus kuitenkin osoitti, että kun muut laitoksen ominaisuudet vakioitiin, yksityiset vankilat eivät olleet sen kustannus-vaikuttavampia kuin julkisetkaan. Tämä esimerkki osoittaa meta-analyysin hyödyllisyyden erityisesti, kun aikaisemmat tutkimukset ovat antaneet ristiriitaisia tuloksia. Aiemmissä tutkimuksissa oli yleensä tarkasteltu muutamaa vankilaa – tavallisesti yhtä yksityistä ja yhtä julkista – ja tämän vuoksi niiden avulla oli vaikea tehdä yleisempiä päätelmiä. Joissakin tutkimuksissa oli myös menetelmällisiä puutteita; niissä ei esimerkiksi otettu huomioon muita tekijöitä, jotka saattoivat vaikuttaa vankia kohti laskettuihin kustannuksiin.

Meta-analyysi onnistui yhdistämään yksittäisten tutkimusten tulokset ja vastaamaan niiden avulla yhteiseen tutkimuskysymykseen. Tarkastelu tehtiin regressioanalyysillä, jossa selitettävänä muuttujan oli vankia kohti laskettu kustannus ja selittäjinä omistusta kuvaava muuttuja (yksityinen tai julkinen) sekä pieni joukko muita muuttujia, kuten rakennuksen ikä, vankien määrä ja turvallisuustaso. Syntyneessä yhtälössä omistuspohjan vaikutusta kuvaava kerroin ei ollut tilastollisesti merkitsevä, mikä tarkoittaa, että julkisten ja yksityisten vankiloiden tehokkuuden (eli kustannukset vankia kohti) välillä ei ole systemaattista eroa. Laitoksen ominaisuudet, kuten vankilan koko, ikä ja turvallisuustaso olivat kaikki tilastollisesti merkitseviä selittäjiä vankikohtaisille kustannuksille.

(Lähde: Pratt ja Maahs, 1999)

## **Muita tutkimusasetelmaan liittyviä kysymyksiä**

Vertailuryhmään kiinnitetään usein niin paljon huomiota, että muut tärkeät tutkimusasetelmaan liittyvät kysymykset tahtovat jäädä vaille riittävää huomiota. Tässä luvussa tarkastellaan joitakin niistä.

### **Otoksen koko**

Otoksen koko on keskeinen kysymys kaikissa kvantitatiivisissa tutkimuksissa. Monet tutkimukset ovat tilastolliselta voimaltaan heikkoja, mikä merkitsee, että otos ei ole kyllin suuri, jotta merkittäviä muutoksia vaikutuksissa ja kustannuksissa pystyttäisiin tunnistamaan. Otoksen koko on pienempi ongelma, jos ohjelman vaikutuksen odotetaan olevan hyvin suuri. Jos taas arvioinnissa yritetään tunnistaa pieni mutta tärkeä vaikutus, tarvitaan suuri otos. Esimerkiksi, jos tutkimus selvittää intervention vaikutusta harvinaisissa, mutta

vakavissa tilanteissa (kuten asunnottomuudessa tai väkivaltarikollisuudessa) ei olisi mielekäästä toteuttaa pienimuotoista kvantitatiivista tutkimusta. Otoksen pitää olla suuri myös silloin, jos yksittäisten käyttäjien kustannukset ja vaikutukset vaihtelevat keskenään paljon. Jos muutama yksilö vastaa suuresta osasta palvelun käyttöä, kustannusjakautumat ovat tyypillisesti vinoja, joten taloudellisiin arviointeihin vaaditaan paljon suurempi otos kustannuserojen käsittelemiseksi kuin pelkästään vaikutuksia selvittäviin tutkimuksiin (Briggs, 2000).

Riittämättömän otoskoon seuraukset riippuvat vertailtavien interventioiden luonteesta. Jos arvioitavan ohjelman tarkoitus on tuottaa parempia tuloksia kuin olemassa olevat toimintatavat, silloin tilastolliselta voimaltaan heikot tutkimukset voivat myönteisten muutoksien tunnistamisessa ja antaa heikosti tukea innovatiivisille tai vaikuttavimmille interventioille. Toisaalta, jos olemassa olevaa toimintaa verrataan halvempaan, vähemmän intensiiviseen ohjelmaan, silloin eivät tilastolliselta voimaltaan heikot tutkimukset helposti tunnista eroa, vaikka vaikutukset olisivat merkittävästi vähäisempiä. Tällöin vaarana on tutkimuksen perusteella päätellä, että vähemmän intensiiviset, kustannuksia säästävät interventiot ovat kustannus-vaikuttavampia. Tilastolliselta voimaltaan heikossa tutkimuksessa päätelmä johtuu kuitenkin vain siitä, että otoskoko on liian pieni tunnistamaan vaikuttavuuden vähenemisen.

Yksi ongelma arviointia suunniteltaessa on se, että otoksen riittävää kokoa on usein vaikea tietää ennalta, ellei samanlaisista interventioista ole tehty aikaisempia tutkimuksia. Kuten todettiin, riittävän otoskoon määrittämiseen tarvitaan tietoa kustannusten ja vaikutusten vaihtelusta, mutta tätä informaatiota on käytettävissä vasta kun ohjelman arviointi on toteutettu. Myös tietojen keruun kustannukset tai ohjelman koko voivat rajoittaa otoksen kokoa.

### **Kadon suuruus**

Vaikka alkuperäinen otos olisi riittävän suuri, tutkimus voi heiketa ihmisten keskeyttäessä arvioinnin kuluessa osallistumisensa. Jotkut voivat pudota pois arvioitavasta ohjelmasta, mutta jatkaa osallistumista arviointitutkimukseen. Arvioinnin kannalta tämä on pienehkö ongelma. Jos ei ole erityistä syytä toimia toisin, nämä ihmiset tulisi pitää mukana interventoryhmässä aineistoa analysoitaessa. Keskeyttäneiden sulkeminen pois vääristäisi tuloksia, koska ne, jotka ovat kyllin kykeneviä ja motivoituneita viemään ohjelman loppuun, menestyvät todennäköisesti paremmin myös tavoitellun vaikutuksen suhteen. Tällainen on asianlaita esimerkiksi tuen-piiristä-työhön - ohjelmissa ja muissa sellaisissa interventioissa, joissa menestys riippuu voimakkaasti osallistujien ponnisteluista.

Haittana siitä, että keskeyttäneet pidetään mukana analyysissä, on ohjelman vaikutuksen ”laimeneminen”, koska tällöin osa interventoryhmästä ei ole ollut täydellisen intervention kohteena. Keskeyttäneiden pitäminen mukana vaikuttaa perustellulta, jos asiakkaiden auttaminen ja rohkaiseminen pysymään mukana ohjelmassa loppuun saakka on osa onnistunutta interventiota. Kuitenkin ihmiset voivat lopettaa sellaisistakin syistä, jotka eivät liity oh-

ojelman vaikuttavuuteen. Esimerkiksi voidaan ottaa matalan kynnyksen toimipiste -malliksi (Revolving Door Agency, 2000) kutsutun mielenterveysintervention arviointi. Eräät poliisin huostasta vapautuneet mahdolliset asiakkaat eivät päätyneet projektin piiriin ainoastaan siitä syystä, että he olivat vapautuneet normaalin työajan ulkopuolella, jolloin projektin toimipiste ei ollut auki. Tällöin ohjelmasta pois jäämisen syy ei ollut se, että projektin työntekijät olisivat epäonnistuneet yhteyden saamisessa asiakkaisiin. Tässä tapauksessa voi olla järkevää jättää kyseiset tapaukset pois analyysistä.

Ongelma on erilainen, kun ihmiset jäävät pois arviointitutkimuksesta. Nämä ovat usein samoja ihmisiä, jotka ovat keskeyttäneet myös intervention, vaikka eivät välttämättä. Usein arvioinnin tekijät haluavat kerätä kustannuksia ja vaikutuksia koskevat tiedot ohjelman päättymisen jälkeen, monasti pitkän ajan kuluttua, jolloin osallistujia on yleensä vaikeaa tavoittaa. Ilmeisin kohtaamispaikka (eli palvelupiste) ei enää ole käytettävissä, ja jotkut ihmiset ovat muuttaneet toiselle paikkakunnalla. Tämä on erityinen ongelma sosiaalialan ohjelmissa, jotka on tarkoitettu tavanomaista kaoottisempaa elämää eläville (esim. mielenterveysongelmallisille tai asunnottomuusuhan alaisille). Vaikutusten seuraaminen voi olla mahdollista pitkiäkin aikoja hallinnollisten tietokantojen avulla, ilman suoraa kontaktia osallistujiin, mutta tämä riippuu kerättävän tiedon tyypistä ja laadusta.

Kato on ongelma osittain siksi, että se pienentää otoskokoa (katso edellä), mutta vieläkin tärkeämpää on, että se saattaa vääristää tuloksia. Ne jotka keskeyttävät, ovat tuskin satunnaisotos kohderyhmästä. Mitä pitempi on tutkimusaika, sitä suurempi tämä ongelma todennäköisesti on. Esimerkiksi sijaishuollon arviointitutkimuksessa on vaikeampi jäljittää nuoria, jotka muuttavat usein. Näissä on suhteellisen paljon vanhempia ja ongelmallisempia nuoria, ja taustalla on keskeytyneitä sijoituksia ja huonot suhteet sekä biologisiin että kasvatusvanhempiin. Jos tämä ryhmä jää pois analysoitavasta otoksesta, sijaishuollon vaikutukset näyttävät todennäköisesti paremmilta kuin jos kaikki nuoret olisivat mukana analyysissä.

Kadosta aiheutuva ongelma ei ole helppo korjata. Tutkimuksen tuloksia esitettäessä olisi vähintäänkin raportoitava keskeyttämisaste, ja keskeyttäneiden tärkeimpiä ominaisuuksia pitäisi verrata niihin, jotka eivät keskeyttäneet (kadon analyysi). Jos ominaisuuksissa ei ole merkittäviä eroja, tämä riittänee. Jos eroja on, mahdollisia vaikutuksia tutkimuksen tuloksiin pitäisi pyrkiä selvittämään; ainakin pitäisi tunnistaa vinoutumien todennäköinen suunta. Joissakin tutkimuksissa, joissa otos on suuri ja käytettävissä on paljon aineistoa, virhelähteitä voidaan ehkä kontrolloida mallintamalla keskeyttämisprosessia. Tästä on esimerkki tarkenteessa 12 (katso esimerkki 2).

## **Seuranta**

Yhtä tärkeää kuin tietojen keräämien intervention alussa ja lopussa vaikutusten selvittämiseksi, on yleensä ainakin yhden seuranta-arvioinnin toteuttaminen. Eräät vaikutukset voivat ilmetä vasta pitemmän ajan kuluessa. Esimerkiksi varhaisen puuttumisen ohjelmilla on osoitettu olevan merkittäviä

vaikutuksia lasten sosiaaliseen- ja koulumenestymiseen useita vuosia sen jälkeen, kun interventio on päättynyt. Muiden vaikutusten odotetaan ilmevän välittömämmin, mutta on silti tärkeää tarkistaa, kestääkö jokin myönteinen vaikutus pidemmän aikaa.

Seurantatutkimuksen ajankohta on aina hieman mielivaltainen, ja se määritellään usein käytännöllisillä perusteilla, kuten ihmisten tavoittamisesta aiheutuvat kustannukset pitkän ajan kuluttua. Mutta tulokset voivat olla hyvin herkkiä ajoitukselle (katso esim. pohdintaa synnytyksen jälkeisen tukiohjelman arvioinnista RCT-asetelmaa käsittelevässä kohdassa). Ajoitus pitäisi, aina kun mahdollista, perustaa aikaisempiin tutkimuksiin. On esimerkiksi osoitettu, että ihmisillä, jotka eivät ole syyllistyneet uusiin rikoksiin kahteen vuoteen tai sitä pitempään aikaan, on paljon pienempi todennäköisyys päätyä rikoksen uusijoiksi myöhemminkään. Siksi interventioissa, joiden tavoitteena on vähentää rikoksen uusimista, ihanteellista olisi seurata vaikutuksia vähintään kaksi vuotta.





## 9 Tulosten analysointi ja tulkinta

### Tulosten analysointi

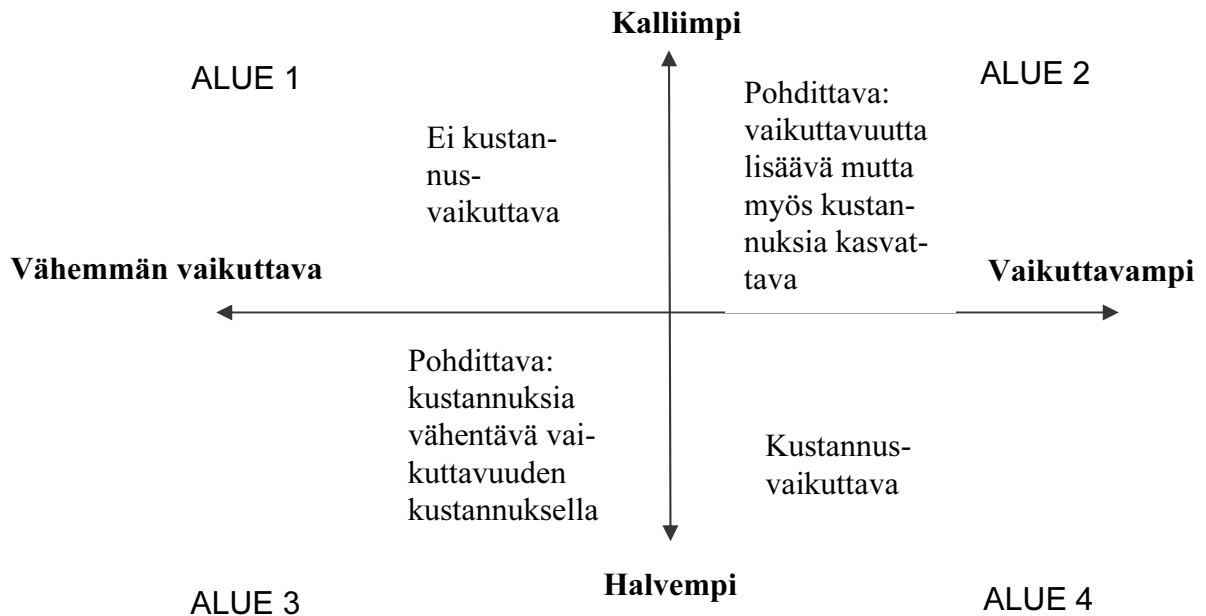
#### Kustannusten ja tulosten yhdistäminen

Kun kahta tai useampaa interventiota, joilla on tavoiteltu samaa vaikutusta, halutaan verrata keskenään, on hankkeiden kustannukset ja vaikutukset muokattava vertailua varten kustannus-vaikuttavuutta kuvaaviksi tunnusluvuiksi.

Jos vaikutuksia on enemmän kuin yksi, kustannus-vaikuttavuutta ei voi määrittää, ellei jokin interventio tuota parempaa tai vähintään yhtä hyvää tulosta kaikkien tarkasteltavien vaikutusten osalta eikä tämä ole muita kalliimpi. Jos näin ei ole, vaihtoehtojen paremmuusjärjestyksen ratkaiseminen (suhteellisesti ottaen kustannustehokkaimman nimeäminen) jää päätöksentekijöiden tehtäväksi. Heidän on ratkaistava, kuinka paljon painoa eri vaikutuksille annetaan. Arviointi voi tukea päätöksentekoa esittämällä tulokset selkeästi ja objektiivisesti.

Kustannus-vaikuttavuus –analyysi ei ole mahdollinen myöskään silloin, kun vertailtavien interventioiden vaikuttavuusmittarit eroavat toisistaan. Jos arvioidaan samaa vaikutusta, mutta vaikutus on vertailtavissa interventioissa mitattu eri tavoin, voi olla mahdollista verrata vaikutusten suhteellisia suuruuksia keskenään. (Tämä voidaan tehdä jakamalla verrattavat nettomuutokset standardipoikkeamillaan, jolloin muutos suhteutuu kunkin skaalan mukaiseksi). Näin tehtäessä on kuitenkin oltava varovainen. Lähestymistapaa käytettiin esimerkiksi arvioitaessa karkeasti kahden varhaista lukutaitoa kehittävän hankkeen kustannus-vaikuttavuutta. Hankkeet ja niiden arvioinnit oli toteutettu erikseen ja lukutaidon kehittyminen oli testattu erilaisella menetelmällä. Tutkijat päättelivät, että ryhmämenetelmään perustuva opetus oli kustannus-vaikuttavampaa kuin yksilöopetus, sillä vaikutus oli jokseenkin saman suuruinen ja kustannukset olivat likimain kymmenen kertaa pienemmät (Nicolson ym., 1999).

**Kuvio 9.** Kustannusvaikuttavuuden arvioiminen



Muodollinen kustannus-vaikuttavuus -analyysi ei ole välttämätön, jos jokin vaihtoehto on selvästi muita parempi tai heikompi; tällainen tilanne on, jos hanke on sekä vaikuttavampi että halvempi kuin muut (alue 4, kuvio 9), tai vähemmän vaikuttava ja kalliimpi (alue 1). Jos hankkeesta saatavat hyödyt ovat muita suuremmat, mutta hanke on myös kalliimpi, tai sekä hyödyt että kustannukset ovat pienemmät, on kustannus-vaikuttavuutta tarkasteltava yksityiskohtaisemmin laskemalla lisähyötyjen kustannus-vaikuttavuus. (Drummond ym., 1997).

Tämä tapahtuu laskemalla lisähyötyjen ja lisäkustannusten välinen suhde (toisin sanoen, kuinka paljon enemmän paremman tuloksen saavuttaminen maksaa kuin toisessa interventiossa). Kun vaikutus on mitattu samalla tavalla, eri interventioille laskettua kustannus-vaikuttavuus -suhdetta voidaan verrata, ja niin löydetään pienimmällä kustannuksella eniten lisähyötyä tuottava hanke. (Byford ym., tulossa).

### **Kustannusten ja vaikutusten vaihtelu**

Kustannukset ja vaikutukset voivat vaihdella merkittävästi eri yksilöiden välillä, erityisesti joustavissa palveluissa; esimerkiksi palveluissa, joissa käytetään palveluohjausta (case management) tai jotka perustuvat asiakkaan valintaan. Esimerkiksi ”Maudsley Daily Living” -ohjelman ensimmäisessä vaiheessa, tarkastellaan kustannuksia, jotka aiheutuvat, kun asiakas hakeutuu tilanteensa kriisiytyessä psykiatriseen sairaalaan. 12–20 kuukauden kulluttua tutkimuksen alkamisesta kustannusten ero yhteisöhoitoon perustuvassa (community-based) vaihtoehdossa oli 27-kertainen vähiten ja eniten maksavien potilaiden välillä (Knapp ym., 1994). Onkin harhaanjohtavaa puhua hoidon kustannuksista yleensä, jos vaihtelu on huomattavan yksilöllistä.

Kun aineisto sallii, vaihtelun tutkiminen on tärkeää. Osaryhmien analyysissä tarkastellaan kustannuksia ja tuloksia asiakas- tai käyttäjäryhmittäin. Kiinnostuksen kohteena on, onko interventio vaikuttavampi tai kustannusvaikuttavampi jollekin tietylle ryhmälle kuin muille ryhmille. Hanke, joka ei ole kustannus-vaikuttava keski-ikäisille, saattaa osoittautua kustannustehokkaaksi jollekin tietylle väestöryhmälle. Esimerkiksi yksi synnytyksen jälkeisen tukiohjelman arviointia kohtaan esitetty kritiikki kohdistui tähän asiaan (Esimerkki 1. tarkenne 8). Arviointia ei ollut suunniteltu siten, että ohjelman vaikuttavuus olisi voitu arvioida erikseen vaikeimmassa asemassa olleille äideille. Näiden olisi voinut olettaa hyötyvän lisätuesta eniten.

Kustannus- tai tuotantofunktiot ovat hienostuneempia tapoja tutkittaessa yksilöiden, palveluiden tai alueiden välistä kustannusten tai vaikutusten vaihtelua. Nämä käyttävät regressiotekniikoita (multiple regressions) ja etsivät tilastollista yhteisvaihtelua kustannusten, tulosten ja palvelun käyttäjän tai tarjoajan ominaisuuksien väliltä. Ne ovat hyvä tapa paljastaa tekijöitä, joilla on merkittävä vaikutus kustannuksiin tai tuloksiin, erityisesti, jos sekoittavia tekijöitä on paljon. Tämä on erityisen tärkeää, jos tapausten valinta ei ole tehty kontrolloidulla satunnaisotosmenetelmällä, sillä tällöin satunnaisuus ei poista eroja ryhmien välillä. Kun kontrolloidaan asiakkaiden ominaisuuksia, arvioitavan intervention nettovaikutus on mahdollista saada näkyville.

Kustannus- ja tuotantofunktiot ovat hyödyllisiä myös muissa tarkoituksissa (Knapp, 1998), esimerkiksi tutkittaessa, kohdentuuko joku ohjelma halutulla tavalla (toisin sanoen, liittyvätkö yksilökohtaiset kustannukset sellaisiin yksilöiden ominaisuuksiin, jotka osoittavat palvelutarvetta). Funktioiden avulla voidaan myös ennustaa kustannuksia tai vaikutuksia, jos käytettävissä on perustietoja potentiaalisesta asiakaskunnasta. Tämä on hyödyllistä palveluja suunnitteleville.

### **Muita huomioita**

Taloustieteilijöitä syytetään joskus huomion kiinnittämisestä yksinomaan tehokkuuteen muitten tekijöiden, kuten oikeudenmukaisuuden kustannuksella (esim. House, 2000). Tutkimuksissa raportoidaan tavallisesti kokonais- tai keskimääräisiä vaikutuksia pohtimatta, kuka hyötyy (tai kärsii haittaa) eniten interventiosta. Oikeudenmukaisuuskysymykset ovat kuitenkin todennäköisesti tärkeitä hyvinvointisektorilla, ja arviointitutkimusten pitäisi vähintäänkin tutkia vaikutusten jakautumista merkittäville osaryhmille (esimerkiksi eri tuloryhmille, eri ikäisille tai vamman asteen mukaan). On silti väärin sanoa, että taloustieteilijät ovat laiminlyöneet julkisen politiikan jakautumisvaikutukset kokonaan, sillä monet tutkimukset ovat kohdistuneet erityisesti tähän kysymykseen (esim. Blow ja Crawford, 1997; Evans ym., 1994; Hills, 1989). Mutta, huolimatta voimakkaista periaatteellisista suosituksista (esim. Gold ym., 1996), oikeudenmukaisuuskysymykset eivät yleensä sisälly olennaisena osana taloudelliseen arviointiin.

Kehittyneempien analyysien pitäisi myös kiinnittää huomiota palvelujen saatavuuteen tai käyttöön suhteessa palvelutarpeeseen, taikka siihen, jae-

taanko palvelut oikeudenmukaisesti ja koetaanko tämä sellaiseksi. Tällainen analyysi menee tulosten ryhmäkohtaista esittämistä pidemmälle ja saattaa tarvita enemmän laadullista analyysiä. Laadullista otetta voidaan tarvita esimerkiksi silloin, kun halutaan ymmärtää, miten erilaiset ihmiset pääsevät palvelun käyttäjiksi tai miten päätetään siitä, mitä palvelua kullekin tarjotaan.

Jotkut taloustieteilijät ovat yrittäneet lisätä oikeudenmukaisuusnäkökulman taloudelliseen arviointiin pyytämällä ihmisiä kertomaan, mikä heidän mielestään on oikeudenmukaista. (Onko esim. oikeudenmukaista, että köyhemmät kotitaloudet saavat etua, mutta koko väestön tilanne heikkenee hieman). Näihin menetelmiin liittyy kuitenkin merkittäviä, joskaan ei ylipääsemättömiä ongelmia, ja suositellaan, että hyötyjen jakautumista koskeva informaatio esitettäisiin omana kokonaisuutenaan (Sassi ym., 2001).

Muut huomioon otettavat kysymykset, kuten vaikutusmahdollisuuksien lisääminen, kohtuullisuus ja oikeus palveluun, liittyvät tapaan, jolla palvelua jaetaan. Näillä prosessiin liittyvillä kysymyksillä on oma merkittävyytensä riippumatta siitä, miten ne vaikuttavat projektin tavoitteiden saavuttamiseen. Samoja asioita voidaan tavoitella keinoilla, jotka joko valtaistavat (empowerment) tai eivät valtaista asianosaisia riippuen siitä, miten ”määräilevä” ohjelma on ja onko se ”käyttäjakeskeinen”. Vaikutusmahdollisuuksien lisääminen ja muut arvot pitäisi sisällyttää arviointikehikkoon silloin, kun ne ovat arvioitavan toiminnan kannalta merkityksellisiä. Ne voivat olla joko toiminnan erillisiä tavoitteita tai ne voivat rajoittaa arvioitavien tavoitteiden valintaa. Ei ole mitään syytä, miksi näitä arvoja ei voi sisällyttää taloudelliseen arviointiin ja esimerkkejä niiden sisällyttämisestä onkin olemassa (esim. Kendall ja Knapp, 2000).

## **Miten ja miksi ohjelmat toimivat**

Suurin osa taloudellisesta arvioinnista kohdistuu kustannuksiin ja vaikutuksiin. Tyypillisesti paljon vähemmän huomiota kiinnitetään siihen, miksi interventiot toimivat ja miksi ne saavat aikaan tiettyjä kustannuksia ja vaikutuksia. Eräät taloustieteilijät pitävät näitä erillisinä - ehkäpä toissijaisina - asioina, joita ei tarvita arvioitaessa vaikuttavuutta tai kustannusvaikuttavuutta. Monipuolisemmin toteutetut tutkimukset saattavat joskus tutkia kustannusten ja vaikutusten vaihtelua eri ryhmien välillä. Ne saattavat jopa yrittää selittää tätä, esimerkiksi tarkastelemalla tilastollista yhteisvaihtelua vaikutusten ja tiettyjen ohjelman ominaisuuksien välillä. Mutta, koska otoskoko on harvoin riittävän suuri edellä mainittuihin tilastollisiin analyysihin, tulokseksi saadaan tavallisesti vain alustavia hypoteeseja vaihtelun systä.

On kuitenkin hyviä perusteita sille, miksi on tärkeää ymmärtää paremmin ohjelmien toimintaa, erityisesti monimutkaisten hyvinvointiin liittyvien interventioiden osalta. Ensinnäkin, parempi ymmärrys auttaa parantamaan tutkimuksen validisuutta; jos arvioijat kykenevät osoittamaan linkin (miksi ja miten) intervention ja haluttujen vaikutusten välillä, he voivat varmem-

min pitää vaikutuksia juuri intervention aiheuttamina. Tällä on vähemmän merkitystä, jos tutkimusasetelma on vankka, sillä hyvin toteutettu, satunnaisotokseen perustuva koeasetelma takaa luotettavan kustannusten ja vaikutusten arvioinnin. Satunnaisotokseen perustuvat koeasetelmat eivät ole kuitenkaan aina käyttökelpoisia sosiaalisen hyvinvoinnin alueella. Vähemmän täsmällisissä koeasetelmissa ei voida aina automaattisesti olettaa, että erot tutkimusryhmän ja vertailuryhmän välillä ovat juuri tutkittavan intervention aiheuttamia. Tällaisissa tilanteissa arviointi on vakuuttavampi, jos se pystyy tarjoamaan myös uskottavan selityksen sille, miten ohjelma aiheutti vaikutukset tutkimusryhmälle. Tarkenteessa 18 kuvataan ”muutoksen teoriaa”. Se on yksi lähestymistavoista, joissa kysymys intervention toiminnasta on keskeinen osa arviointia (esimerkki 1).

Toiseksi, intervention toiminnan ymmärtäminen (miten ja miksi) voi auttaa yleistämään arvioinnin tuloksia, erityisesti, jos vaikutusten oletetaan olevan olosuhdesidonniaisia. Yksi kritiikki kokeellisia arviointiasetelmia kohtaan onki, että samanlaisten interventioiden arvioinneissa saadaan usein erilaisia tuloksia. (Pawson ja Tilley, 1997). Tämä ei ole yllättävää, kun otetaan huomioon, että ihmiset reagoivat erilaisissa olosuhteissa eri tavoin samaan interventioon, tai että itse interventiot saattavat hieman erota eri alueilla toisistaan annetun koulutuksen, toimijoiden kompetenssin tai innostuksen suhteen. Tämä on yksi syy siihen, miksi onnistuneilta näyttäneet ohjelmat usein epäonnistuvat, kun ne lainataan muualle. Mutta ilman selitystä sille, miksi ohjelma vaikuttaa eri tavoin eri paikoissa ja eri väestöryhmille, päätöksentekijä ei ole yhtään viisaampi. Toisaalta, jos arviointi sisältää analyysin siitä, miten ja miksi ohjelma toimii, on helpompi arvioida, olisiko sama interventio vaikuttava tai kustannus-vaikuttava toisessa ympäristössä. Saattaa olla mahdollista esimerkiksi erotella institutionaalisen ympäristön ominaisuuksia, jotka vaikuttavat myönteisesti interventioon. Vaihtoehtoisesti, saattaa olla tiettyjä ominaisuuksia, jotka tekevät osan käyttäjistä muita vastaanottavemmiksi interventiolle. Realistinen arviointi menee vielä askelta pidemmälle väittämässään, että arvioinnissa tulisikin keskittyä tunnistamaan vaikuttavia toimintamekanismeja (policy mechanisms) mieluummin kuin mittaamaan vaikutuksia (esimerkki 2 tarkenteessa 18).

Osaryhmittäiset analyysit ovat siis hyödyllisiä ja tutkimukset pitäisi suunnitella tämä huomioon ottaen, mikä tarkoittaa suurempia otoskokoja. Taloustieteilijöiden pitäisi olla myös valmiita käyttämään muita lähestymistapoja tarkastellessaan tämän tyyppisiä kysymyksiä osana taloudellista arviointia. Vaikka muut lähestymistavat on usein esitetty vaihtoehdoiksi kokeellisille lähestymistavoille, ei ole mitään periaatteellista estettä yhdistää lähestymistapoja. Kokeellisilla tutkimuksilla hankitaan täsmällisiä arvioita vaikutuksista, ja muilla lähestymistavoilla, kuten realistisella arvioinnilla, arvioidaan missä määrin tulokset ovat sovellettavissa muissa ympäristöissä. Saattaa olla myös tilanteita, joissa teoriaperusteiset lähestymistavat voivat korvata systemaattisempia kustannus- ja vaikutusarviointeja (esimerkki tarkenteessa 19).

**Esimerkki 1: Muutosteoria (theory of change)**

Tämä teoriakeskeinen arvioinnin lähestymistapa kehitettiin laajojen ja monimutkaisten yhteisötason hankkeiden arviointiin, johon kokeellisen lähestymistavan väitetään soveltuvan huonosti (Connell, 1995). Lähestymistapaa voidaan kuitenkin käyttää myös muunlaisissa sosiaalialaa koskevissa hankkeissa. Muutosteorian (theory of change, TOC) taustalla on oletus, että ohjelmat perustuvat explisiittisiin tai implisiittisiin teorioihin siitä, miten halutut tavoitteet saavutetaan. Arvioinnin ensimmäisessä vaiheessa hankkeen keskeisiä toimijoita autetaan määrittelemään hankkeen tavoitteet sekä milloin ja miten ne aiotaan saavuttaa. Toisin sanoen, arvioija aloittaa määrittelemällä lopulliset tavoitteet ja etenee taaksepäin toimenpiteisiin, joita tavoitteiden saavuttamiseen tarvitaan. Tämän jälkeen määritellään välietapit tai välitavoitteet, joihin toimenpiteet liittyvät, siten, että ohjelmaa voidaan arvioida ja, jos tarpeen, muokata välietappien kohdalla.

Kun varsinaiset vaikutusmittarit liittyvät lopullisiin tavoitteisiin, tuotokset tai prosessivaikutukset ovat välietappeja. Taloustieteilijät suhtautuvat usein kriittisesti prosessia kuvaaviin mittareihin, koska ne eivät mittaa suoraan tavoitteiden saavuttamista. Kuitenkin, kun tukeudutaan taustateoriaan, kuten TOC-lähestymistavassa, prosessimittarit voivat olla tehokas tapa arvioida hankkeen tavoitteiden saavutettavuutta (Hughes ja Traynor, 2000). Prosessimittarit auttavat myös rakentamaan todistusketju, jonka avulla interventio voidaan yhdistää lopullisiin vaikutuksiin.

Toisessa esimerkissä, Terveen elämän keskukset –hankkeen (Healthy Living Centres) arvioinnissa, tutkitaan teorioita siitä, millaisia muutoksia löytyy paikallisen aloitteellisuuden taustalta. Jotkut teorit perustuvat taloudellisten ongelmien vähentämiseen, jolloin vähenevät myös taloudellisten ongelmien kielteiset vaikutukset terveyteen. Lyhyen tähtäimen tavoitteena voi olla luotto-osuuskunnan (Credit Union) perustaminen ja mahdollisimman monen rohkaiseminen liittymään kassan jäseneksi. Seuraavassa – välivaiheessa - tämän pitäisi rohkaista useampia ihmisiä säästämään ja samalla se auttaisi yhteisöä kehittämään kykyään huolehtia omista tarpeistaan. Pitkällä tähtäimellä järjestely auttaisi ihmisiä pääsemään velkakierteestä ja siihen liittyvästä huolesta ja lisäisi luottamusta ja solidaarisuutta paikallisyhteisöön. Toivon mukaan lopulta seurauksena olisi vähäosaisimpien terveyden paraneminen, joka on Terveen elämän keskukset –hankkeen varsinaisen tavoite.

## Esimerkki 2. Realistinen arviointi

Realistisen arvioinnin kannattajat korostavat, että arvioinnissa ei pitäisi keskittyä mittaamaan vain ohjelman aikaansaamaa nettovaikutusta, koska se todennäköisesti kertoo melko vähän ohjelman vaikuttavuudesta erilaisissa olosuhteissa (Pawson ja Tilly, 1997). Sen sijaan arvioinnissa pitäisi pyrkiä löytämään tavoiteltavia vaikutuksia tuottavia mekanismeja. Esimerkiksi yksi tai useampi seuraavan luettelon mekanismeista vaikuttaa valvontakameran (CCTV, Closed Circuit Television) käyttökelpoisuuteen parkkipaikkojen valvonnassa:

- Sinut on huomattu –mekanismi: saa rikosta suunnitellun luopumaan aikeestaan lisäämällä kiinnijäämisriskiä
- Vaikuttava sijoittuminen –mekanismi: parantaa vartioiden ja poliisien mahdollisuutta siirtyä sinne, missä esiintyy epäilyttävää käyttäytymistä
- Muistin virkistys –mekanismi: muistuttaa autoilijoita huolellisuudesta (esim. autojen lukitseminen ja varashälyttimien päälle kytkeminen)
- Vetoa varovaisiin –mekanismi: turvapaikka turvallisuustietoisille autoilijoille.

Koska monien interventioiden onnistuminen riippuu olosuhteista, realistisessa arvioinnissa pitäisi myös yksilöidä konteksti eli ne olosuhteet, joissa nämä mekanismit todennäköisesti toimivat. Edellä olevassa esimerkissä Vaikuttavan sijoittumisen –mekanismi on uskottavampi kaupunkien parkkipaikoilla, missä turvallisuushenkilöstöä on enemmän saapuvilla kuin alueilla, joilla parkkipaikat ovat sijoittuneet kauemmas toisistaan. Sitä vastoin, muistin virkistys –mekanismin onnistunut toiminta riippuu varashälyttimien tasosta, esimerkiksi niiden yleisyydestä, vaikuttavuudesta ja helppokäyttöisyydestä.

Arvioija kehittää edellä kuvatun olosuhteet-mekanismi-vaikutukset –mallin (CMO; context-mechanism-outcome) käyttämällä apunaan hankkeen keskeisiä toimijoita, kuten suunnittelijoita ja käytännön toteuttajia. Liikkeelle lähdetään teorioista ja hypoteeseista, jotka testataan saatavilla olevan aineiston avulla. Esimerkiksi poliisia ja vartiointiyrityksiä haastatteleamalla voidaan selvittää, onko valvontakameran asentaminen muuttanut heidän tapansa käsitellä autoihin kohdistuvaa rikollisuutta. Poliisin tilastoja voidaan käyttää sen arvioimiseen, onko rikosten selvittämisaste parantunut. Autoilijoiden haastattelujen avulla voidaan selvittää valvontakameran vaikutusta heidän käyttäytymiseensä, esimerkiksi onko valvontakameralla vaikutusta parkkipaikan valintaan.

Realistinen arviointi ei edellytä mitään tiettyä tutkimusmenetelmää. Se paremminkin rohkaisee käyttämään samanaikaisesti sekä kvalitatiivisia että kvantitatiivisia menetelmiä. Tavoitteena on määritellä vaikuttava mekanismi, konteksti eli olosuhteet, joissa mekanismi käynnistyy, ja odotetut vaikutukset näissä



olosuhteissa. Tällaisen informaation väitetään olevan siirrettävämpää ja siksi päätöksentekijöiden näkökulmasta käyttökelpoisempaa kuin kokeellisten arviointiasetelmien tuottamat joko-tai –tulokset. Jos edellä esitellyssä esimerkkitapauksessa Vaukuttavan sijoittumisen –mekanismi olisi todettu hallitsevaksi mekanismiksi, päätöksentekijät saisivat myös tiedon siitä, että valvontakameran käyttö maaseutumaisien alueiden parkkipai-koilla ei olisi vaikuttavaa ilman riittävää poliisien ja vartioiden saatavilla oloa. Päättäjät saisivat myös tietää, että kaupunkimai-sillakin alueilla valvontakameran vaikuttavuus edellyttää koor-dinointia paikallisen poliisin ja vartiointiyritysten kanssa, jotta turvallisuushenkilöstöä olisi saatavilla.

### **Tarkenne 19. Esimerkki teoriaperusteisesta taloudellisesta arvioinnista**

#### **Yleislääkäreiden määrärahavastuukokeilun arviointi (Eva-luation of GP fundholding)**

Yleislääkäreiden määrärahavastuukokeilun tavoitteena oli lisätä sairaalapalvelujen tehokkuutta antamalla yleislääkäripraktiikalle välitön vastuu tiettyjen sairaalapalvelujen ostamisesta potilail-leen (praktiikoille annettiin myös määräraha). Ostettavia palve-luja ovat muun muassa laboratoriopalvelut ja halvemmat leik-kaukset, joista näin ollen muodostuu julkisen terveydenhoitojär-jestelmän (NHS; National Health Service) sisäiset markkinat. Tätä kokeilua olisi voitu arvioida vertaamalla määrärahasvas-tuussa olevien ja ei-määrärahavastuullisten lääkäreiden potilai-den terveydenhuoltokustannuksia ja terveysvaikutuksia, mutta tämä olisi tullut hyvin kalliiksi ja arviointitulosten saaminen olisi kestänyt kauan. Tämän sijasta arviointi perustui joukkoon yksityiskohtaisia arviointikysymyksiä. Kysymysten taustalla oli mikrotaloustieteellinen teoria sekä alun perin järjestelmälle ase-tettuja odotuksia ja sitä kohtaan esitettyjä huolia. Esimerkiksi

- *Parantaako lääkärin määrärahavastuu potilaan saamia palveluja:* mikrotalousteorian mukaan yleislääkäreille annetun tilaajavallan pitäisi rohkaista lähekkäin sijait-sevia sairaaloita kilpailemaan keskenään, minkä edel-leen pitäisi johtaa parempiin ja/tai halvempiin palvelui-hin.
- *Muuttuuko lääkäreiden lähetemenettely:* vanhassa sys-teemissä yleislääkärillä ei ollut taloudellisia kannusti-mia pohtia rationaalisesti läheteitään, koska kolmas osapuoli maksoi kustannukset. Päätöksentekijät toivoi-vat, että koska määrärahavastuullisella lääkärillä oli nyt kannustin kontrolloida lähetekäyttäytymistään, tämä vähentäisi läheteiden määrää ja johtaisi kustannusvai-kuttavampaan hoitoon.
- *Pysyvätkö määrärahavastuussa olevat lääkärit budjet-tinsa puitteissa:* yksi hallituksen huolenaiheista oli se, että rahat loppuisivat ennen vuoden loppua ja hallituk-sen olisi otettava vastuu, jotta potilaat eivät jäisi ilman hoitoa.
- *Yrittävätkö määrärahavastuulääkärit valikoida potilai-taan:* hallitus oli huolestunut myös siitä, että määrära-

havastuulääkärit saattaisivat yrittää ”kerman kuorintaa”. Käyttämällä hyväksi potilasrekistereitä he voisivat valita halvempihoitoiset potilaat ja pyrkiä välttämään raskashoitoisimpia ja näin ansaita enemmän voittoa praktikoillaan.

Jokaista kysymystä tarkasteltiin useiden tutkimusmenetelmien avulla. Muun muassa haastattelemalla yleislääkäreitä selvitettiin lääkäreiden motivaatiota määrärahavastuukokeiluun ja kokeilun vaikutusta praktiikkaan. Tietoja kustannuksista käytettiin budjetin toteutusvertailuun ja lähetetietoja määrärahavastuulääkäreiden ja ei-määrärahavastuussa olevien lähetekäytäntöjen vertaamiseen. Potilasrekisteriotosta analysoitiin sen selvittämiseksi, voiko rekisteritietojen avulla tunnistaa todennäköisesti kalliiksi tulevia potilaita.

Tutkijat päättelivät tulostensa perusteella, että yleislääkäreiden määrärahavastuun tehokkuutta lisäävä vaikutus oli merkittävä, mutta ei käänteentekevä ja että hallituksen huoli budjetin riittämättömyydestä oli aiheeton. Vaikka tutkimukset osoittivat, että järjestelmässä on kannustimia ”kerman kuorintaan”, siitä ei kuitenkaan löytynyt todisteita. Edellisen perusteella tutkijat esittivät, että määrärahavastuun edut voittavat todennäköisesti siitä aiheutuvat hallinnolliset lisäkustannukset.

(Lähde: Glennerster ym., 1994)

## **Tulosten tulkinta ja esittäminen**

Taloudellisten arviointien tulokset on usein esitetty huonosti tai harhaan johtavasti. Vaikka kvantitatiiviset (määrälliset, numeeriset) tulokset ovat todistusvoimaisia, niitä on myös helppo tulkita väärin. Niissä tutkimuksissa, joita kävimme läpi (ks. luku 3), oli taipumuksena liioitella tutkittavan intervention vaikutusta. Joissakin tapauksissa tutkimuksen lopussa oleva pohdinta oli kritiikitön ja kohdistui vain ohjelman myönteisiin puoliin; vedetyt johtopäätökset ylittivät sen, mitä tutkimustulokset näyttivät osoittaneen. Ristiriitaiset (monitulkintaiset) tulokset esitettiin joskus siten, kuin ne olisivat tukeneet ohjelmaa, ja riittämättömän informaation avulla oli tehty johtopäätöksiä tulosten luotettavuudesta. Yksi tärkeimmistä puutteista oli se, että huomiota ei ollut kiinnitetty riittävästi tulosten vaihtoehtoihin selityksiin.

Terveystaloudellisessa arvioinnissa on melko selkeitä ohjeita arvioinnin tulosten esittämisestä. Ohjeilla pyritään varmistamaan se, että kaikki tutkimukset saavuttavat tietyn tason ja että keskeinen tieto arviointiasetelmasta kerrotaan selkeästi (esim. Eisenberg, 1989; Siegel ym., 1997; Task Force on Principles for Economic Analysis of Health Care Technology, 1995). Tämä johtaa siihen, että tuloksia aletaan esittää yhdenmukaisemmin, vaikka se saattaa myös tukahduttaa innovatiivisuutta.

Ei liene tarkoituksenmukaista soveltaa samoja kriteereitä koko sosiaalisen hyvinvoinnin alueelle. Terveystaloustieteessä kehitetyt ohjeet ovat liian

jäykkiä useimmille muille hyvinvoinnin alueille. Saattaa olla parempi pohtia yhteisiä periaatteita, joita tutkimusten tulisi pyrkiä noudattamaan. Hyvä toimintatapa voisi koostua seuraavista:

- Tulokset esitetään suhteessa aikaisempiin, samanlaisia interventioita koskeneisiin tutkimuksiin: tuloksia verrataan, korostetaan eroja ja etsitään niille selityksiä.
- Tunnustetaan mahdolliset tutkimusasetelman heikkoudet ja tunnustetaan näiden aiheuttamat mahdolliset virhelähteet. Tutkijoiden pitäisi pyrkiä arvioimaan virheiden suunta ja merkittävyys. (Tämän voi tehdä esim. tutkimalla, ja jos mahdollista, kontrolloimalla intervention ja vertailtavien ryhmien erilaisia ominaisuuksia). Arviointi joudutaan tekemään mahdollisesti laadullisesti, erityisesti, jos otoskoko on pieni.
- Tutkitaan ja suljetaan mahdollisuuksien mukaan pois tulosten vaihtoehtoiset selitykset.
- Testataan, ovatko tulokset yhdenmukaisia keskeisten odotusten kanssa ja yritetään selittää epätavalliset tai odottamattomat tulokset.
- Pohditaan tulosten merkitsevyyttä ja epävarmuuksia (esim. arvioimalla luottamusvälejä ja esittämällä keskeisille muuttujille tehdyn herkkyysanalyysin tulokset).
- Kuvataan avoimesti aineiston laatu ja määrä ja miten nämä saattavat vaikuttaa tulosten luotettavuuteen (esim. vaikutusmittareiden validius).
- Kerrotaan selkeästi, mitä sisältyy ja ei sisälly lopullisiin analyysiin: esim. mitä hyötyjä on laskettu rahaksi ja sisällytetty kustannus-hyöty –analyysiin. Lisäksi pohditaan, mitä merkitsee, jos jotain on jätetty pois.
- Esitetään selkeästi ja perustellaan analyysien taustalla olevat oletukset.

Täydellisempi lista hyvän taloudellisen arviointitutkimuksen kriteereistä on tämän kirja seuraavassa osassa (Byford ym., 2003).

Laadullisten tulosten oikea esittäminen on myös tärkeää. Valikoimalla laadullista aineistoa siten, että se tukee tiettyä johtopäätöstä, on luultavasti vielä helpompi kuin kvantitatiivisessa tutkimuksessa antaa vääristynyt kuva. Kuitenkin, myös laadullista aineistoa voidaan analysoida systemaattisesti (Pope ym., 2000). Laadullisen analyysin tuloksia on vaikeampi myös esittää ytimekkäästi, mistä on sekä haittaa että hyötyä. Yhtäältä, tästä seuraa, että laadullisilla tuloksilla on yleensä vähemmän välittömiä vaikutuksia lukijaan, mutta toisaalta, mahdollisuudet esittää monimutkaisia ja hienovaraisia asioi-

ta ovat paremmat. Kun on kysymys taloudellisesta arvioinnista, on todennäköistä, että suurin osa laadullisesta analyysistä käytetään kvantitatiivisten tulosten tukemiseen tai laadun parantamiseen. Ideaalitapauksessa näiden pitäisikin yhdistyä loppuraportissa.



## 10 Arviointiprosessin hallinta

Olosuhteet, joissa arviointi toteutetaan, voivat olla hyvin merkityksellisiä sille, mikä lähestymistapa valitaan. Tarkastelemamme tapaustutkimukset olivat erityisen hyödyllisiä tämän kysymyksen valottamisessa.

### Keskeisten tahojen osallistuminen

Yksi asia, jonka taloudellinen arviointi voi oppia muilta arvioinnin lähestymistavoilta on se, että on hyödyllistä, joskus jopa välttämätöntä, ottaa arvioitavan hankkeen kaikki toimijat tiiviimmin mukaan arviointiprosessiin. Näihin toimijoihin kuuluvat arvioitavan ohjelman rahoittajat, suunnittelijat ja hallinnoijat sekä arvioitavan palvelun käyttäjät.

Eräissä taloudellisen arvioinnin tavoissa, muun muassa kustannus-hyöty ja kustannus-utiliteetti –analyysissä, hankkeen toimijat ovat erityisen tärkeitä selvitetessä, miten ihmiset arvottavat toiminnan vaikutuksia. On ilmeistä, että tällainen vaikutusten arviointi on parasta hankkeen toimijoiden näkökulman huomioon ottamista: silloin vaikutusten arvottaminen perustuu niiden ihmisten mieltymyksiin, joihin toiminta vaikuttaa. Tekniikoita, joita taloustieteilijät käyttävät vaikutusten arvottamiseen, kuvattiin kohdassa ”Vaikutusten mittaaminen” (luvussa 7). Jotkut taloustieteilijät väittävät, että vaikutukset pitäisikin aina arvioida käyttäjien ja yleisön näkökulmasta, sillä nämä ovat hyötyjiä ja kantavat lopulta myös kustannukset, esimerkiksi veronmaksajina.

Tästä huolimatta taloustieteilijät ottavat harvoin hankkeen toimijoita mukaan arviointiin. He ovat haluttomia päästämään osalliseksi tahoja, joilla on hankkeeseen liittyviä omia intressejä (muut kuin käyttäjät/asiakkaat). He pelkäävät näiden tahojen mukaantulon heikentävän arvioinnin puolueettomuutta. Sitä vastoin on tyypillistä, että arvioijat, jotka ovat taustaltaan sosiaalipoliitikkoja, ottavat mukaan arviointiprosessiin laajan joukon toimijoita arvioitavan hankkeen eri vaiheista:

- auttamaan arvioinnin tavoitteiden selkiyttämistä
- selkiyttämään arvioitavan toiminnan tavoitteita
- tunnistamaan mahdollisia sivuvaikutuksia (vaikutuksia, joita ei odotettu, sekä lumipallovaikutuksia)
- auttamaan tärkeimpien vaikutusten ja kustannusten määrittelyä
- ymmärtämään paremmin, miten arvioitavan hankkeen oli tarkoitus toimia ja miten se todellisuudessa toimi

Hankkeen toimijat voivat siis osallistua arviointiin tavoilla, joita monet taloustieteilijät eivät tavallisesti tulisi ajatelleeksi. Joissakin tapauksissa panos on erittäin tärkeä. Esimerkiksi, jos hankkeen tavoitteita ei ole selkeästi määriteltä tai hanke on monimutkainen, hankkeen toimijoiden osallistuminen tavoitteiden selkiyttämiseen on tärkeää.

Joillakin sosiaalisen hyvinvoinnin sektoreilla (esim. yhteisöjen kehittämisessä; community development) toimijoita rohkaistaan osallistumaan arviointiprosessiin. Yksi yhteisöjen kehittämishankkeiden perimmäisistä tavoitteista on 'valtaistaa' (empower) paikallisia ihmisiä kehittämällä heidän tietojaan ja taitojaan. Oikealla tavalla toteutettu arviointi voi toimia osana oppimisprosessia: se antaa ihmisille mahdollisuuden reflektoida ja koota kokemuksia (Barr ja Hashagen, 2000). Kun annetaan hankkeiden toimijoille enemmän 'omistusoikeutta' arviointiin saatetaan arviointiprosessia myös helpottaa (esim. helpottaa tiedonsaantia) ja lisätä todennäköisyyttä, että arvioinnin suositukset otetaan huomioon ja niiden mukaan toimitaan (Patton, 1996).

Toimijoiden osallistumiseen arviointiprosessiin liittyy myös ongelmia. Jotkut saattavat olla haluttomia yhteistyöhön, jos he kokevat, että heitä määrällään tai he ovat kiireisiä. Erityisesti taloudellinen arviointi voidaan kokea osaksi kustannusten leikkausoperaatioita, ja ihmiset saattavat kokea tämän uhkaksi. Siksi onkin tärkeää selittää arvioinnin tarkoitus ja toimijoiden osallistumisen tärkeys. Vaarana on myös se, että arvioija saattaa omaksua liian yksipuolisesti jonkin toimijan näkökulman, joka ei edusta yleistä näkemystä.

## **Yhteistyö muiden tieteenalojen kanssa**

Taloustieteilijät työskentelevät usein samanaikaisesti muiden arvioijien kanssa toimimatta kuitenkaan monitieteellisenä ryhmänä. He ovat ottaneet tietyn roolin, useimmiten kustannusarvioinnin, ja hoitavat tehtävän enemmän tai vähemmän itsenäisesti suhteessa muuhun arviointiryhmään. Monesta syystä on kuitenkin toivottavaa, oikeastaan välttämätöntä, rakentaa siltaa taloustieteilijöiden ja muita tieteenaloja edustavien arvioijien välille. Sillan rakentuminen:

- auttaa hälventämään harhaluuloja, esimerkiksi sitä, mitä taloudellinen arviointi oikeastaan on
- muodostaa yhteistä ymmärrystä käsitteistä ja voittaa esteitä, joita aiheutuu teknisen slangin käyttämisestä
- luo mahdollisuuksia oppia toisilta tieteenaloilta: esim. taloustieteilijöillä on paljon opittavaa olosuhteiden (context) ja prosessin merkityksestä
- rohkaisee yhtenäisempään arviointikäsitteeseen (esim. arvioimaan kustannuksia ja vaikutuksia samanaikaisesti)
- antaa taloustieteilijöille realistisemmän kuvan siitä, mitä voidaan saavuttaa kun ymmärtää paremmin hankkeiden monimutkaisuus
- auttaa muita arvioijia ymmärtämään, mitä annettavaa taloustieteilijöillä on arviointiprosessin eri vaiheissa.

Kokemukset tapaustutkimuksista ja muista aikaisemmista arviointitutkimuksista tuovat kuitenkin esille monia esteitä, joita on taloustieteilijöiden ja muiden arviointitutkijoiden tehokkaalle yhteistyölle. Esteiden parempi ym-

märtäminen voi helpottaa eri tieteenalojen tutkijoita voittamaan vaikeudet arvioinnin alussa.

- *Taloudellisen arvioinnin tarkoituksen väärinymmärtäminen:* Taloustieteilijöiden mukana olo on usein nähty osaksi kustannusten leikkaustoimenpiteitä. Sitä vastoin taloudellinen arviointi on kiinnostunut resurssien tehokkaasta käytöstä ja sen vuoksi sitä voidaan yhtä hyvin käyttää perusteltaessa suurempaa varojen käyttöä kustannusvaikuttavampaan toimintaan.
- *Tutkimusasetelmaa koskevat erimielisyydet:* Ristiriita taloustieteilijöiden suosimien suureen satunnaisotokseen perustuvien kontrolloitujen kokeiden ja sosiaalitieteilijöiden suosiman syvällisen laadullisen tutkimuksen välillä saattaa aiheuttaa merkittäviä ongelmia monitieteellisessä tutkimuksessa. Monissa tapauksissa ryhmämme taloustieteilijöitä on neuvottu jättäytymään pois tutkimuksista, joissa tutkimusryhmiä ei ole muodostettu satunnaisotoksen avulla. Jos neuvo olisi otettu vastaan, sosiaalisen hyvinvoinnin tuottamisen taloudellinen arviointi ei olisi kehittynyt. Taloustieteilijöiden pitäisi suhtautua ymmärtäväisemmin suuriin vaikeuksiin käyttäen satunnaisotokseen perustuvia koeasetelmia sosiaalisen hyvinvoinnin alueella. Samoin laadullisia menetelmiä käyttäneiden tutkijoiden pitäisi ymmärtää paremmin kvantitatiivisen tutkimuksen hyötyjä. Pitäisi keskustella eri lähestymistapojen toisilleen tarjoamista eduista ja pitäisi yrittää löytää paras mahdollinen kompromissi. Keskustelujen perusteella on monissa tutkimuksissa päädytty asetelmaan, jossa ensin tehdään suuri otos kvantitatiivista tutkimusta varten, ja otoksesta valitaan pienempi joukko laadulliseen tutkimukseen.
- *Vakiintuneiden mittaristojen käyttöä koskeva erimielisyys:* Joillakin tieteenaloilla ollaan epäileviä standardisoitujen mittaristojen käyttöä kohtaan vaikutusten arvioinnissa, koska mittarit eivät näytä mittaavan tarkalleen sitä, mitä tutkija haluaa tutkia. Sen sijaan halutaan suunnitella uusia mittareita, jotka tarkemmin mittaavat juuri arvioitavan hankkeen tavoitteiden toteutumista. Vaikka kritiikki vakiintuneita mittaristoja kohtaan on perusteltua, standardimittaristojen etuja ei pitäisi jättää hyödyntämättä kevyin perustein. Jos halutaan luotettava arvio kustannus-vaikuttavuudesta ja laajempaa vertailtavuutta muihin hankkeisiin, vakiintuneet mittaristot ovat arvokkaita välineitä. Edelleen, tutkijoiden pitäisi pohtia standardimittaristojen ja tapausta varten suunniteltavien mittareiden hyötyjä ja haittoja ja pyrkiä yhdistämään mittaamistapoja.
- *Taloustieteilijät eivät ymmärrä laadullisia menetelmiä:* Laadullinen tutkimus on tärkeä väline, jolla on keskeinen rooli hankittaessa näyttöjä ja kehitettäessä toimintapolitiikkaa monella sosiaalisen hyvinvoinnin alueella. Taloustieteilijät ovat kuitenkin usein tietämättömiä laadullisista menetelmistä ja laadullisen tutkimuksen käytöstä eikä monikaan tiedä, miten laadullinen analyysi voi tukea heidän omaa



tutkimustaan. Taloustieteilijöiden pitäisi olla avoimia muiden menetelmien käytölle ja ottaa apu vastaan laadullisen tutkimuksen tekijöiltä.

- *Kiista resurssien kohdentamisesta taloudellisen tiedon keräämiseen:* Taloudellinen arviointi nähdään tavallisesti 'lisukkeena' tärkeimmille arviointituloksille. Taloustieteilijälle ei ole epätavallista huomata olevansa tilanteessa, jossa hänen odotetaan tekevän työnsä vaikutusarvioijien määräämin ehdoin. Hänellä voi olla käytössään suhteellisen pieni osuus arviointiin varatuista resursseista, niin rahasta, ajasta kuin kyselyssä varatusta tilastakin. Muitten arvioijien on tärkeää ymmärtää, että merkittävät taloudellisen arvioinnin rajoitukset voivat heikentää koko tutkimuksen käyttökelpoisuutta.
- *Halutaan pitää tutkimuksen taloudelliset komponentit irrallisina:* Taloustieteilijän odotetaan usein työskentelevän yksin ja pelkästään kustannusarvioinnin parissa. Taloudellista arviointia pidetään suppeampana, mitä se on, sillä taloudellinen arviointi vaatii sekä kustannus- että vaikutustietoja (sopivassa muodossa). Tällainen heikentää myös mahdollisia monitieteellisestä tutkimuksesta saatavia hyötyjä.

Jos nykyistä tilannetta halutaan parantaa, on välttämätöntä muodostaa aidosti monitieteisiä ryhmiä kokoamalla samaan ryhmään taloustieteilijöitä, sosiolitieteilijöitä ja muita asiantuntijoita (esim. psykiatreja, jos arvioidaan mielenterveyteen liittyvää hanketta). Muiden arvioijien pitäisi ottaa taloustieteilijät mukaan arviointiprosessiin aikaisemmassa vaiheessa, ja taloustieteilijöiden pitäisi olla halukkaampia 'likaamaan kätensä' esimerkiksi vieraillemalla projekteissa tai osallistumalla haastatteluihin. Lisäksi kaikkien pitäisi olla halukkaampia käyttämään monenlaisia tutkimusmenetelmiä. Taloustieteilijöille olisi hyödyllistä opetella laadullisia tutkimusmenetelmiä tai edes seurata, miten näitä menetelmiä käytetään. Mutta kaikkein eniten tarvitaan sen aitoa ymmärtämistä ja arvostamista, mitä eri asiantuntijoilla on annettavaa arvioinnille.

## **Rajoittavia tekijöitä arvioinnissa**

Arviointitilanteisiin liittyy aina erilaisia rajoittavia tekijöitä, joilla saattaa olla merkittävä vaikutus arvioinnin tutkimusasetelmaan. Yleisin rajoite aiheutuu arviointiin käytettävistä resursseista. Vaikka suurin osa julkisista ja vapaaehtoisista toimijoista nykyisin hyväksyy, että heidän hankkeitaan arvioidaan, aina ei varata riittävästi resursseja, että arviointi voitaisiin tehdä kunnolla. Esimerkiksi hankkeessa 'Turvapaikka kaupungissa', arviointiin ei ollut alun perin varattu lainkaan resursseja, ja arvioinnin rahoittamisessa jouduttiin turvautumaan pääosin ulkopuolisiin varoihin.

Resurssien puute voi rajoittaa otoskokoa, kerättävän tiedon määrää ja käytettävää aikaa. Käytännössä on vaikea saada rahoitusta pitkäkestoisiin tutkimuksiin, vaikka joidenkin hankkeiden odotetaan vaikuttavan vain pitämällä viiveellä. Joissakin tapauksissa, kun otetaan huomioon arvioitavan

hankkeen koko riittäviä resursseja kunnollisen arvioinnin tekemiseen on vaikea perustella. Tällöin arvioinnin tavoitteista on tingittävä vastaavasti tai hankkeen rahoittajien ja poliitikkojen pitäisi harkita lisärahoitusta arvioinnille. Ei pitäisi korostaa perusteellisen arvioinnin tärkeyttä, ellei tarvittavaa rahoitusta ole varattu hankkeen budjettiin. Budjetointi vuorostaan edellyttää hyvää käsitystä hyvälaatuisen taloudellisen arviointitutkimuksen kustannuksista, jotka vaihtelevat melkoisesti. Taulukkoon 3 on koottu keskeiset arviointikustannuksiin vaikuttavat tekijät. Kun resursseja on rajoitetusti, joudutaan tekemään valintaa kerätyn tiedon kokonaisvaltaisuuden ja tiedon keräyksen vaatimien kustannusten ja ajan välillä. Haastatteluihin perustuvat tiedonkeräystavat ovat useimmiten parhaita kerätyn tiedon laadun kannalta, mutta tulevat verraten kalliiksi. Hallinnollinen tieto on edullista hankkia, mutta ei välttämättä ole parasta mahdollista arviointikäyttöön.

Muita rajoituksia arvioijalle aiheutuu arvioitavan hankkeen suunnitelmasta ja toteutustavasta, ja arvioijalla on yleensä vähän mahdollisuuksia vaikuttaa näihin. Tällaisia ovat:

- Päätökset siitä, ketkä osallistuvat hankkeeseen: klassinen kokeellinen tutkimus edellyttää, että jotkut potentiaaliset asiakkaat on satunnaisesti määrätty kontrolliryhmäksi, mutta ohjelman toteuttajat eivät ehkä halua estää ihmisten osallistumista tai haluavat käyttää muita kriteereitä valitessaan kohteita.
- Olennaiset muutokset arvioitavassa toiminnassa tai luonnollinen, ajan myötä tapahtuva kehittyminen haittaavat arviointia. Yleinen ongelma on, että vertaisryhmälle tarjotut palvelut alkavat jäljitellä varsinaisen arvioitavan hankkeen toimintatapoja, erityisesti jos hanke on ”muodikas”.
- Jos hanke on kovin pieni, se voi rajoittaa mahdollista otoskokoa ja heikentää kvantitatiivisen tutkimuksen tilastollista merkitsevyyttä. Se rajoittaa todennäköisesti myös arviointiin käytettävissä olevaa rahoitusta.

Lisäksi eräät rahoittajien asettamat rajoitukset vaikuttavat arviointiin. Rahoittajat hyötyvät tuloksista ja saattavat asemansa avulla pyrkiä ohjaamaan tutkimusta. Rahoittajat saattavat asettaa arvioinnille myös keinotekoisia aikarajoja, ehkä koska he haluavat nopeita tuloksia poliittiseen käyttöön. Akateemisilla tutkijoilla on omat paineensa tuottaa merkityksellisiä tai odottamattomia tuloksia, koska ne saadaan paremmin julkaistua akateemisissa julkaisuissa (esim. Easterbrook ym., 1991).

Jotkut arvioinnin rajoitteet yhdistyvät tyypillisesti taloustieteilijöihin ja liittyvät heidän rooliinsa arviointitiimissä. Taloustieteilijä saatetaan ottaa mukaan arviointiin vasta myöhäisessä vaiheessa, ja vaikka taloustieteilijä on mukana arviointiryhmässä, hänen roolinsa voidaan rajata vain tiettyihin arvioinnin näkökulmiin. Esimerkki tästä on tuore tutkimus lapsuusajan käytöshäiriöiden liittyvistä kustannuksista (Scott ym., 2001). Tässä tutkimuk-

sessä taloudellista analyysiä ei tehty rinnan sosiaalisen ja kliinisen arvioinnin kanssa, eivätkä taloustieteilijät olleet mukana suunnittelemassa tietojenkeräyslomaketta, johon kustannusten arvioinnin kuitenkin täytyi perustua. Kun tehtiin taloudellista analyysiä jälkeen päin, monia kustannuksiin ja resurssien käyttöön liittyviä tietoja ei ollut enää saatavilla. Tällaisia olivat esimerkiksi tiedot sosiaalipalvelujen käytöstä, vapaaehtoissektorin kustannuksista ja huoltajan kustannuksista. Tämän johdosta tutkijat kykenivät tuottamaan vain vaillinaisen (ja varovaisen) arvion pitkän ajan kustannuksista.

Lopuksi voidaan todeta, että taloudellinen arviointi on vaikeampaa, jos samantyyppistä hanketta ei ole arvioitu aikaisemmin. Voi olla, että standardisoituja mittaristoja vaikutusten arviointiin ei ole käytettävissä, eikä valmiita yksikkökustannuksia voi 'lainata' mistään. Tarvittava otoskoko voi myös olla vaikea määrittää eikä tuloksiakaan ole aina helppo tulkita ilman, että niitä voidaan verrata muihin tutkimuksiin.

**Taulukko 3.** Taloudellisen arvioinnin kustannuksiin vaikuttavia tekijöitä

<b>Tekijä</b>	<b>Selitys vaikutuksesta arvioinnin kustannuksiin</b>
Tutkimusasetelma	Satunnaisotokseen perustuvat kontrolloidut kokeet ovat usein (mutta eivät aina) kalliimpia kuin muut asetelmat, koska asetelma korostaa monia jäljempänä käsiteltäviä tekijöitä (esim. suurta otoskokoa, ajantasaista tiedon keräystä ja haastatteluja).
Otoskoko	Kustannukset riippuvat käytettävistä tietolähteistä (ks. jäljempänä). Tavallisesti kustannukset nousevat otoskoon kasvaessa, koska tarvitaan enemmän haastatteluja, tietoa on tallennettava enemmän ja aikaa tarvitaan enemmän.
Tietolähteet	Hallinnolliset tiedostot ovat halvimpia kun taas yksilölliset haastattelut ovat kalleimpia. Muut lähteet, esimerkiksi postikyselyt ja palvelujen tarjoajien rekisterit sijoittuvat näiden välille, vaikka jälkimmäisten analysointi saattaa olla aikaa vievää ja sen vuoksi kallista.
Seurantajakson pituus ja seurantojen tiheys	Useammat seurannat nostavat tiedonkeräyskustannuksia. Seurantajakson pituus nostaa myös kustannuksia, koska tutkijat on pidettävä palkkalistoilla kauemmin ja vastaajia saattaa olla hankala jäljittää.
Jälkikäteinen tai ajantasainen tietojen keräys	Vaikka jälkeinpäin tehtävä arviointi ei ole niin toivottava kuin parhailaan toteutettava, se on yleensä halvempi – voidaanhan jälkikäteisessä arvioinnissa käyttää jo olemassa olevaa tietoa.
Tavallisesta poikkeavien haastattelujen tai tiedonkeräysmenetelmien tarve	Vaikka uusien kyselyjen suunnittelu onkin usein välttämätöntä, se saattaa olla aikaa vievää erityisesti, jos kyselyn soveltuvuus myös testataan (pilotoidaan).
Taloudellisen arvioinnin tyyppi	Kustannus-hyöty –analyysi on yleensä kalliimpi, jos vaikutusten rahalliseen arvottamiseen tarvitaan erityinen kysely.
Eettinen lupa	Eettisen luvan hankkiminen arvioinnin tekemiseen on aikaa vievä ja hallinnollisesti kallis prosessi. Sitä ei tavallisesti vaadita terveydenhuollon ulkopuolella, mutta se saattaa tulla merkittävämmäksi asiaksi tulevaisuudessa.
Onko yksikkökustannukset arvioitava itse	Joissakin tapauksissa on mahdollista ja tarkoituksenmukaista käyttää valmiina saatavia yksikkökustannuksia, esim. valtakunnallisia keskiarvoja. Tällöin välttyään aikaa ja resursseja vievältä, arvioitavaa palvelua koskevien yksityiskohtaisten resurssitietojen keräämiseltä.
Tietojen tallentamisen tarve	Tietojen tallentaminen lisää merkittävästi arvioinnin kustannuksia, ellei olemassa olevia tietolähteitä voi hyödyntää, tai ellei palvelun tarjoaja tai käyttäjä itse tee tallennusta osana palveluprosessia.
Tiedonkeräyspisteiden määrä	Kustannukset ovat sitä suuremmat, mitä hajautetumpaa tietojen keräys on (esim. monista toimipaikoista, eri puolilta maata tai eri maista).
Tietojen analysointi	Satunnaisotokseen perustuvien kontrolloitujen kokeitten tuottama tieto on suhteellisen helppoa analysoida. Kvasi-kokeellisissa tutkimuksissa saatetaan sitä vastoin käyttää asiantuntija-apua vaativia monimutkaisia tilastollisia tekniikoita, mikä saattaa tulla kalliiksi.



# 11 Yhteenveto ja johtopäätökset

## Yhteenveto

Sosiaalisen hyvinvoinnin alueella tarvitaan enemmän ja parempaa taloudellista arviointia. Näitä tietoja, samoin kuin muita poliittisia, eettisiä ja käytännöllisiä perusteluja, voidaan hyödyntää päätettäessä rajallisten resurssien käyttämisestä. Tähän päivään mennessä muista kuin terveydenhuoltoon liittyvistä sosiaalisen hyvinvoinnin hankkeista on (Isossa Britanniassa) tehty vähän taloudellista arviointia. Suurin osa tehdyistä arvioinneista on tehty USA:ssa ja voittopuolisesti lähellä terveydenhuoltoa olevilta alueilta. Tutkimusten taso vaihtelee huomattavasti.

On monta hyvää syytä, miksi taloudellinen arviointi tällä alueella on kehittynyt niin hitaasti. Taloudellista arviointia ei ole vaadittu eikä suosittu, ja sitä on myös vastustettu muiden arvioijien ja käytännön työntekijöiden taholta. Lisäksi monet sosiaaliseen hyvinvointiin liittyvät hankkeet ovat hyvin monimutkaisia. Tyypillistä on, että käyttäjien panos on suuri, saman hankkeen sisällä voi olla vaihtelevuutta, vaikutukset ovat monimutkaisia ja pitkällä ajalla toteutuvia ja hankkeisiin osallistuu monia toimijoita. Tämä monimutkaisuus epäilemättä tekee arvioinnista haasteellisen, mutta sen ei pitäisi estää hyvätasoisien taloudellisten arvioinnin toteuttamista.

Taloudellisen arvioinnin pitäisi olla osa laajempaa arviointistrategiaa, jonka tavoitteena on määrittellä, hioa ja ymmärtää hanketta, ennen kuin pyritään tekemään systemaattisempi arvio hankkeen vaikuttavuudesta ja kustannusvaikuttavuudesta. Taloustieteilijöillä on tärkeää annettavaa kaikilla arvioinnin aloilla heidän erityisen näkökulmansa ansiosta. Aikainen taloudellisen näkökulman mukaan tulo auttaa luomaan pohjaa perusteelliselle taloudelliselle arvioinnille esimerkiksi varmistamalla, että tarpeellinen tieto tulee kerättyä. Rahoittajien ja päätöksentekijöiden pitäisi olla tietoisia hankkeen kehityksen eri vaiheista, koska hankkeen vaihe määrittelee arvioinnin tärkeimmät tehtävät ja sen, mitä arvioinnilta voidaan realistisesti odottaa.

Minkä tahansa arvioinnin ensimmäinen ja tärkein tehtävä on selkiyttää hankkeen tavoitteet. Tämä tulee todennäköisesti olemaan perustavanlaatuisen ja olennainen elementti sosiaaliseen hyvinvointiin liittyvien hankkeiden arvioinnissa. Asia on tärkeä, koska tavoitteet ovat usein moninaisia, huonosti määriteltäviä ja voivat vaihdella ajan ja toimijoiden mukaan. Merkityksellisten kustannusten ja vaikutusten määrittely- ja arviointiprosessi on myös arvioinnin kriittinen kohta. Erityisesti kustannuksiin pitäisi kiinnittää nykyistä enemmän huomiota ja niitä pitäisi käsitellä arvioinnin olennaisena osana. Nykyisin suurin osa taloudellisista arvioinneista ottaa kapean näkökulman kustannuksiin, jättää usein huomiotta kustannukset, joita ei voi mitata rahassa ja muiden palvelujen käyttöön liittyvät kerrannaisvaikutukset.

On sekä toivottavaa että mahdollista toteuttaa enemmän satunnaisotokseen perustuvia kontrolloituja koeasetelmia myös sosiaaliseen hyvinvointiin liit-

tyvissä hankkeissa. Jotkut tähän näkökulmaan liittyvät huolenaiheet ovat joko virheellisiä tai liioittelevia, ja ne voidaan ratkaista huolellisella suunnittelulla ja täydentävillä tutkimusmenetelmillä. Milloin satunnaisotokseen perustuva tapausten valinta ei ole käyttökelpoinen joko käytännöllisistä tai eettisistä syistä, tavoitteena pitäisi olla niin vankka arviointiasetelma kuin mahdollista. Satunnaisotoksen jäljittelyyn voidaan löytää luovia ratkaisuja. Tilastollisilla tekniikoilla voidaan kontrolloida tutkimusryhmän ja vertailuryhmän välisiä eroja ja vakiintuneita (validoituja) mittaristoja voidaan käyttää vaikutusten mittaamiseen. Lisäksi useampien vaikutusten seuraaminen ja pitkä seuranta-aika saattavat auttaa kokonaisvaltaisemman arvioinnin tekemisessä. Joissakin tapauksissa hanke voidaan suunnitella helpommin arvioitavaksi esimerkiksi suunnittelemalla pilotointia tai etenemällä toimeenpanossa asteittain, siten että vertailua voidaan tehdä hankkeen sisällä.

Sosiaalisen hyvinvoinnin alueella arviointiasetelmissa joudutaan usein poikkeamaan 'ideaalista'. Yleisesti voidaan todeta, että mitä monimutkaisempi arvioitava hanke on, sitä todennäköisemmin kokeellinen taloudellisen arvioinnin lähestymistapa on sellaisenaan riittämätön. Silloin suuremman roolin saavat erityisesti laadulliset tutkimusmenetelmät, koska ne auttavat käsittelemään kokeellisen lähestymistavan rajoituksia, niiden avulla voidaan esimerkiksi tarkistaa vaikutusmittareiden soveltuvuutta ja antaa vahvistusta kvantitatiivisen analyysin tuloksille.

Taloudellisen arvioinnin tulokset pitäisi esittää ja tulkita huolellisemmin kuin on tapana tehdä. Tutkimusasetelman heikkoudet pitäisi tunnustaa ja tutkia tästä johtuvat mahdolliset vinoumat. Kustannusten ja vaikutusten arvioinnin epävarmuustaso (vaihteluvälit) pitäisi esittää selkeästi. Taloustieteilijöiden olisi syytä kiinnittää enemmän huomiota myös siihen, miten ja miksi hankkeet toimivat, erityisesti kun on kysymys monimutkaisista hankkeista. Kun kyetään määrittämään toimenpiteiden ja vaikutusten väliset yhteydet, paranevat tutkimuksen luotettavuus. Kun ymmärretään, mitkä olosuhteet ovat vaikuttavuuden kannalta merkityksellisiä, lisääntyy tulosten yleistettävyyttä.

Nämä ehdotukset nykyisen taloudellisen arviointitavan muokkaamisesta ja kehittämisestä saattavat edellyttää, että taloustieteilijät ottavat erilaisen roolin kuin mihin he ovat tottuneet. Taloustieteilijät vartioivat mustasukkaisesti objektiivisuuttaan. Kuitenkaan sosiaaliseen hyvinvointiin liittyvän hankkeen arviointia voidaan harvoin toteuttaa täysin ilman, että hankkeen toimijat ovat arviointiprosessissa mukana, mukaan lukien hankkeen suunnittelijat, käytännön toteuttajat sekä palvelujen käyttäjät. Tämä voi olla välttämätöntä aikaisessa vaiheessa, koska se auttaa selkeyttämään hankkeen tavoitteita ja määrittelemään mahdollisia sivuvaikutuksia. Myöhemmässä vaiheessa se auttaa arvioimaan vaikutuksia ja ymmärtämään miten nämä vaikutukset syntyivät. Perinteisesti muut arvioijat ovat vastanneet useista edellä mainituista arviointitehtävistä, mutta mielestämme ei ole viisasta erottaa puhtaasti taloudellisia komponentteja muista arvioinnin näkökulmista. Tämä tarkoittaa, että eri tieteenalaja edustavien arvioijien on opittava työskentelemään tiiviimmin yhdessä ja kunnioittamaan ja oppimaan toistensa näkökulmista.

Taloudellinen arviointi on lähes aina toteutettava erilaisten rajoitusten vallitessa, jotka vaikuttavat siihen, mitä käytännössä voidaan tehdä. Tällaisia ovat rajalliset resurssit, aikarajoitteet, 'poliittiset' paineet, vähäiset mahdollisuudet vaikuttaa arvioitavan hankkeen toteuttamiseen ja soveltuvien tutkimusinstrumenttien puute. Rajallinen rahoitus on suurin este täsmällisemmille taloudellisten arvioinnin lähestymistavoille ja se on asia, johon taloudellisia arviointeja teettävien tulisi kiinnittää enemmän huomiota.

## **Johtopäätökset**

Taloustieteilijöillä on tärkeää annettavaa näyttöön perustuvalla politiikan teolle, ja vaatimukset hankkeitten vaikuttavuuden ja kustannusvaikuttavuuden täsmällisemmästä toteen näyttämisestä kasvavat. Hyvin tehty taloudellinen arviointi tarjoaa arvokasta informaatiota päätöksentekijöille, jotka joutuvat tekemään vaikeita niukkojen resurssien kohdentamista koskevia päätöksiä. Vaikka laadukas taloudellinen arviointi on harvinaista monella sosiaalisen hyvinvoinnin alueella, löytyy hyviä esimerkkejä, jotka osoittavat, että täsmällinen taloudellinen arviointi on mahdollista suuressa joukossa sosiaaliseen hyvinvointiin vaikuttavia hankkeita.

Useiden sosiaaliseen hyvinvointiin vaikuttavien hankkeiden monimutkaisuus tekee kuitenkin taloudellisen arvioinnin haastavammaksi. Terveystieteiden vakiintuneet taloudellisen arvioinnin lähestymistavat tarjoavat hyvän lähtökohdan, mutta niitä on muokattava siten, että ne soveltuvat monimutkaisempiin arviointikohteisiin. Tämä vaatii käytännöllistä lähestymistä ja sen hyväksymistä, että saatetaan joutua tekemään arviointiasetelmaan liittyviä kompromisseja ja että tarvitaan sopivassa määrin muita tutkimusmenetelmiä parantamaan tulosten pätevyyttä ja yleistettävyyttä. Taloustieteilijöillä on melkoisesti opittavaa muista arvioinnin näkökulmista, mutta taloustieteilijöillä on vastaavasti myös paljon tarjottavaa, edellyttäen että he ovat halukkaita omaksumaan joustavamman ja moninaisemman taloudellisen arvioinnin lähestymistavan.



## LÄHTEET

Bagley, C. ja Pritchard, C. (1998) The reduction of problem behaviours in at-risk youth: an experimental study of school social work with cost-benefit analyses. *Child and Family Social Work*, Vol. 4, No. 4, pp. 219-26.

Barr, A. ja Hashagen, P. (2000) *ABCD Handbook: A Framework for Evaluating Community Development*. London: Community Development Foundation.

Beck, E.J., Beecham, J., Mandalia, S., Griffith, R., Walters, M.D., Boulton, M. ja Miller, D.L. (1999) What is the cost of getting the price wrong? *Journal of Public Health Medicine*, Vol.21, No. 3, pp. 311-17.

Beecham, J. (2000) *Unit Costs not Exactly Child's Play: A Guide to Estimating Unit Costs for Children's Social Care*. London: Department of Health.

Beecham, J. ja Knapp, M. (1992) *Costing psychiatric interventions*. Teoksessa G. Thornicroft, C. Brewin ja J. Wing (toim.) *Measuring Mental Health Needs*. London: Gaskell.

Blow, L. ja Crawford, I. (1997) *The Distributional Effects of Taxes on Private Motoring*. London: Institute of Fiscal Studies.

Briggs, A. (2000) Economic evaluation and clinical trials: size matters. *British Medical Journal*, Vol. 321, pp. 1362-3.

Britten, N. (1995) Qualitative research: qualitative interviews in medical research. *British Medical Journal*, Vol. 311, pp. 251-3.

Buxton, M.J., Drummond, M., Van Hout, B., Prince, R., Sheldon, T., Szucs, T. ja Vray, M. (1997) Modelling in economic evaluation: an unavoidable fact of life. *Health Economics*, Vol. 6, pp. 217-27.

Byford, S., Harrington, R., Togerson, D., Kerfoot, M., Dyer, E., Harrington, V., Woodham, A., Gill, J. ja McNiven, R. (1999) Cost-effectiveness analysis of a home-based social work intervention for children and adolescents who have deliberately poisoned themselves: the results of a randomised controlled trial. *British Journal of Psychiatry*, Vol. 174, pp. 56-62 .

Byford, S., McDaid, D. ja Sefton, T. (2003) *Because it's worth it: A practical guide to conducting economic evaluations in the social welfare field*. York: Joseph Rowntree Foundation.

Centre for Management and Policy Studies (2001) *Modernising Policy Development: Nine Features of Modern Policy Making*. <http://www.cmps.gov.uk/> (Lokakuu 2001).

Challis, D., Darton, R., Johnson, L., Stone, M. ja Traske, K. (1995) Care Management and Health Care of Older People: The Darlington Community Care Project. Aldershot: Ashgate.

Clinch, P.J. ja Healy, J.D. (2001) Cost-benefit analysis of domestic energy efficiency. *Energy Policy*, Vol. 29, pp. 113-24.

Coast, J. (1999) The appropriate uses of qualitative methods in health economics. *Health Economics*, Vol. 8, pp. 345-53.

Connell, J.P. (toim.) (1995) *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*. Queenstown. MD: Aspen Institute.

Davies, B. ja Challis, D. (1986) *Matching Resources to Needs in Community Care: An Evaluated Demonstration of a Long-term Care Model*. Aldershot: Gower.

Department of the Environment, Transport and the Regions (2000) *Fuel Poverty: The New HEES*. London: The Stationery Office.

Department of Health (2000) *Framework for the Assessment of Children in Need and their Families: The Family Pack of Questionnaires and Scales*. London: The Stationery Office.

Deschenes, E. ja Greenwood, P. (1998) Alternative placements for juvenile offenders: results from the evaluation of the Nokomis Challenge Program. *Journal of Research in Crime and Delinquency*, Vol. 35, No. 3, pp. 267-94.

Drummond, M.R, O'Brien, B., Stoddart, G.L. ja Torrance, G.W. (1997) *Methods for the economic Evaluation of Health Care Programmes*. 2. painos. Oxford: Oxford Medical Publications.

Easterbrook, P., Berlin, J., Gopalan, R. ja Matthews, D. (1991) Publication bias in clinical research. *The Lancet*, Vol. 337, No. 8746, pp. 867-72.

Eisenberg, J. (1989) Clinical economics: a guide to the economic analysis of clinical practices. *Journal of the American Medical Association*, Vol. 262, No. 20, pp. 2879-86.

Ekblom, P., Law, H. ja Sutton, M. (1996) *Safer Cities and Domestic Burglary*. London: Home Office.

Elixhauser, A., Halpern, M., Schmier, J. ja Luce, B. (1993) Health care cost-benefit and cost-effectiveness analysis (CBA/CEA) from 1979 to 1990: a bibliography. *Medical Care*, Vol. 31, No. 7, Supplement.

Elixhauser, A., Halpern, M., Schmier, J. ja Luce, B. (1998) Health care CBA and CEA from 1991 to 1996: an updated bibliography. *Medical Care*, Vol. 36, No. 5, Supplement.

EuroQol Group (1990) EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy*, Vol. 16, pp. 199-208.

Evans, M., Piachaud, D. ja Sutherland, H. (1994) Designed for the Poor - Poorer by Design?: The Effects of the 1986 Social Security Act on Family Incomes. *Welfare State Programme Discussion Paper No. 105*. London: London School of Economics.

Farkas, G. (1998) Reading one-to-one: an intensive program for serving a great many students whilst still achieving large effects. Teoksessa J.Crane (toim.) *Social Programs that Work*. New York: Russell Sage Foundation.

Forrester, D., Frenz, S., O'Connell, M. ja Pease, K. (1990) *The Kirkholt Burglary Prevention Project: Phase II*. London: Home Office.

Gardiner, K. (1997) *Bridges from Benefit to Work: A Review*. York: YPS for the Joseph Rowntree.

Foundation Glennerster, H. ja Turner, T. (1993) *Estate Based Housing Management: An Evaluation*. Report for the Department of the Environment. London: HMSO.

Glennerster, H., Matasaganis, M., Owen, P. ja Hancock, S. (1994) *Implementing GP Fundholding: Wild Card or Winning Hand?* London/Bristol, PA: Open University Press.

Gold, M., Siegel, J., Russell, L. ja Weinstein, M. (1996) *Cost-effectiveness in Health and Medicine*. New York: Oxford University Press.

Greenwood, DJ. (1998) *Introduction to Action Research: Social Research for Social Change* Thousand Oaks, CA: Sage.

Cuba, E.G. ja Lincoln, Y.S. (1989) *Fourth Generation Evaluation*. London: Sage.

Hasluck, C., McKnight, A. ja Elias, P. (2000) *Evaluation of the New Deal for Lone Parents: Early Lessons from the Phase One Prototype - Cost-benefit and Econometric Analyses*. Department of Social Security Research Report No. 110. Huddersfield: The Charlesworth Group.

- Hausman, D. ja McPherson, M. (1994) Economics, rationality, and ethics. Teoksessa D. Hausman (toim.) *The Philosophy of Economics: An Anthology*, toinen painos. Cambridge: Cambridge University Press.
- Healey, A. (2001) An economic approach to assessing the impact of violent criminal offending on individual well being: results from exploratory survey of the general public. Presentation at conference organised by the Institute of Psychiatry. *Economics of mental health*, 12 October.
- Hills, J. (1989) *Distributional Effects of Housing Subsidies in the United Kingdom*. Welfare State Programme Discussion Paper No. 44. London: STICERD.
- HM Treasury (1997) *Appraisal and Evaluation in Central Government*. London: The Stationery Office.
- Holtermann, S. (1995) *Investing in Young Children: A Reassessment of the Cost of an Education and Day Care Service*. London: National Children's Bureau.
- Hokermann, S. (1998) *Weighing it up: Applying Economic Evaluations to Social Welfare Programmes*. York: YPS for the Joseph Rowntree Foundation.
- House, E.R. (2000) The limits of cost-benefit analysis. *Evaluation: The International Journal of Theory, Research and Practice*, Vol. 6, No. 1, pp. 79-86.
- Hughes, M. ja Traynor, T. (2000) Reconciling process and outcome in evaluating community initiatives. *Evaluation: The International Journal of Theory, Research and Practice*, Vol. 6, No. 1, pp. 37-49.
- Hutton, J. (1992) Cost-benefit analysis in health care expenditure decision-making. *Health Economics*, Vol. 1, pp. 213-16.
- Jan, S. (1998) A holistic approach to the economic evaluation of health programs using institutionalist methodology. *Social Science and Medicine*, Vol. 47, No. 10, pp. 1565-72.
- Johannesson, M. (1995) A note on the depreciation of the societal perspective in economic evaluation of health care. *Health Policy*, Vol. 33, pp. 59-66.
- Johannesson, M., Joensson, B. ja Karlsson, G. (1996) Outcome measurement in economic evaluation. *Health Economics*, Vol. 5, No. 4, pp. 279-96.
- Joseph Rowntree Foundation (2001a) *Current Priorities*. <http://www.jrf.org.uk/funding/priorities/> (Lokakuu 2001).

Joseph Rowntree Foundation (2001b) Specific Requirements of the Committee. [http://www.jrf.org.uk / funding / priorities / scdc.asp](http://www.jrf.org.uk/funding/priorities/scdc.asp) (Lokakuu 2001).

Keen, J. ja Packwood, T. (1995) Qualitative research: case study evaluation. *British Medical Journal*, Vol. 311, pp. 411-46.

Kendall, J. ja Knapp, M. (2000) Measuring the performance of voluntary organisations. *Public Management*, Vol. 2, No. I, pp. 105-32.

Kitzinger, J. (1995) Qualitative research: introducing focus groups. *British Medical Journal*, Vol. 311, pp. 299-302.

Knapp, M. (1998) Making music out of noise - the cost function approach to evaluation. *British Journal of Psychiatry*, Vol. 173, Supplement 36, pp. 7-11.

Knapp, M. (1999) Economic evaluation and mental health: sparse past ... fertile future? *The Journal of Mental Health Policy and Economics*, Vol. 2, pp. 163-7.

Knapp, M., Cambridge, P., Thomason, C., Beecham, J., Alien, C. ja Darton, R. (1992) *Care in the Community: Challenge and Demonstration*. Aldershot: Ashgate.

Knapp, M., Beecham, J. ja Koutsogeorgopoulou, V. ym. (1994) Service use and costs of home-based versus hospital-based care for people with serious mental illness. *British Journal of Psychiatry*, Vol. 165, pp. 195-203.

Kushner, Saville (1996) The limits of constructivism. *Evaluation: The International Journal of Theory, Research and Practice*, Vol. 2, No. 2, pp. 189-200.

Long, C., Williams, M. ja Hollin, C. (1998) Treating alcohol problems: a study of programme effectiveness and cost effectiveness according to length and delivery of treatment. *Addiction*, Vol. 93, No. 4, pp. 561-71.

McDaid, D. (2001) Estimating the costs of informal care for people with Alzheimer's Disease: methodological and practical challenges. *International Journal of Geriatric Psychiatry*, Vol. 16, pp. 400-5.

McDaid, D., Sefton, T. ja Byford, S. (tulossa) *Economic Evaluations in Social Welfare: A Review of the Literature*. LSE Health and Social Care Discussion Paper.

MacDonald, G. (1996) Evaluating the effectiveness of social interventions. Teoksessa A.Oakley ja H. Roberts (toim.) *Evaluating Social Interventions: A Report of Two Workshops Funded by the Economic and Social Research Council*. Ilford: Barnardos.

- McKee, M., Britton, A., Black, N., McPherson, K., Sanderson, C. ja Bain, C. (1999) Interpreting the evidence: choosing between randomised and non-randomised studies. *British Medical Journal*, Vol. 319, pp. 312-15.
- Mannion, R. ja Small, N. (1999) Postmodern health economics. *Health Care Analysis*, Vol. 7, pp. 255-72.
- Martin, S. ja Sanderson, I. (2000) Evaluating public policy experiments. *Evaluation: The International Journal of Theory, Research and Practice*, Vol. 5, No. 3, pp. 245-58.
- Mays, N. ja Pope, C. (1995a) Rigour and qualitative research. *British Medical Journal*, Vol. 311, pp. 109-12.
- Mays, N. ja Pope, C. (1995b) Observational methods in health care settings. *British Medical Journal*, Vol. 311, pp. 182-4.
- Medical Research Council (2000) *A Framework for Development and Evaluation of RCTs for Complex Interventions to Improve Health*. London: Medical Research Council.
- Mitton, L., Sutherland, H. ja Weeks, M. (2000) *Microsimulation Modelling for Policy Analysis: Challenges and Innovations*. Cambridge University Department of Applied Economics Occasional Paper No. 65. New York: Cambridge University Press.
- Morrell, C.J., Spiby, H. ja Stewart, P., Walters, S. ja Morgan, A. (2000a) Costs and benefits of community postnatal support workers: a randomised controlled trial. *Health Technology Assessment*, Vol. 4, No. 6.
- Morrell, C.J., Spiby, H. ja Stewart, P., Walters, S. ja Morgan, A. (2000b) Costs and effectiveness of community postnatal support workers: randomised controlled trial. *British Medical Journal*, Vol. 321, pp. 593-8.
- Moser, C.A. ja Kalton, G (1993) *Survey Methods in Social Investigation*. Aldershot: Dartmouth Publishing Company.
- Mugford, M. (2001) National evaluation of school breakfast clubs. Presentation at workshop on Economic evaluation and social welfare: developing methods and learning from practical experience. University of Sheffield, 26 July.
- National Centre for Social Research (NCSR) (2000) In-depth interviewing course, five-day training course run by NCSR, June.
- Netten, A ja Curtis, L (2000) *Unit Costs of Health and Social Care*. Canterbury: Personal Social Services Research Unit, University of Kent.

Newburn, T. (2001) What do we mean by evaluation?. *Children and Society*, Vol. 15, pp. 5- 13.

NHS Centre for Reviews of Research and Dissemination (CRD) (2001) *Undertaking Systematic Reviews of Research on Effectiveness: CRD Guidelines for those Carrying out or Commissioning Reviews*, NHS CRD Report No. 4, 2. *painos*. York: NHS CRD.

NHS Executive (1994) *Capital Investment Manual: Business Case Guide*. London: HMSO.

Nicolson, R., Fawcett, A., Moss, H., Nicholson, M. ja Reason, R. (1999) Early reading intervention can be effective and cost-effective. *British Journal of Educational Psychology*, Vol. 69, pp. 47-62.

Nocon, A. ja Qureshi, H. (1996) *Outcomes of Community Care for Users and Carers: A Social Services Perspective*. Buckingham: Open University Press.

Oakley, A. ja Fullerton, D. (1996), *The lamp-post of research: support or illumination*. Teoksessa A. Oakley ja H. Roberts (toim.) (1996) *Evaluating Social Interventions: A Report of Two Workshops Funded by the Economic and Social Research Council*. Ilford: Barnardos.

Patton, M. (1996) *Utilization-focused Evaluation*. London: Sage.

Payne, J. (2000) *Evaluating Training Programmes for the Long-term Unemployed*. PSI Report No. 871. London: Policy Studies Institute.

Pawson, R. ja Tilley, N. (1997) *Realistic Evaluation*. London: Sage.

Pecora, PJ., Fraser, M., Nelson, K., McCroskey, J. ja Meezan, W. (1995) *Evaluating Family-based Services*. New York: Walter de Gruyter.

Peled, E. ja Spiro, S.E. (1998) Goal-focused evaluation. *Evaluation: The International Journal of Theory, Research and Practice*, Vol. 4, No. 4, pp. 455-68.

Pope, C. ja Mays, N. (1995) Researching the parts other methods cannot reach: an introduction to qualitative methods in health and social services research. *British Medical Journal*, Vol. 311, pp. 42-5.

Pope, C., Ziebland, S. ja Mays, N. (2000) Analysing qualitative data. *British Medical Journal*, Vol. 320, pp. 114-16.

Pratt, T. ja Maahs, J. (1999) Are private prisons more cost-effective than public prisons? A meta-analysis of evaluation research studies. *Crime and Delinquency*, Vol. 45, No. 3, pp. 358-71.

Revolving Doors Agency (2000) *Mental Health, Multiple Needs, and the Police*. London: Revolving Doors Agency.

Robertson, A., Grimes, P. ja Rogers, K. (2001) A short-run cost-benefit analysis of community-based interventions for juvenile offenders. *Crime and Delinquency*, Vol. 47, No. 2, pp. 265-85.

Ruhm, C. (2000) Parental leave and child health. *Journal of Health Economics*, Vol. 19, pp. 931-60.

Sassi, F. (2000) *Methods for the evaluation of health care interventions*, Teoksessa R. Cookson, A. Maynard, D. McDaid, F. Sassi ja T. Sheldon *Analysis of the Scientific and Technical Evaluation of Health Interventions*. Report to the European Commission, July.  
[http://www.lse.ac.uk/Depts/lsehsc/pdf\\_files/methods.pdf](http://www.lse.ac.uk/Depts/lsehsc/pdf_files/methods.pdf) (Helmikuu 2002).

Sassi, F., Archard, L. ja Le Grand, J. (2001) Equity and the economic evaluation of health care. *Health Technology Assessment*, Vol. 5, No. 3, pp. 1-138.

Scott, S., Knapp, M., Henderson, J. ja Maughan, B. (2001) Financial cost of social exclusion: follow-up study of antisocial children into adulthood. *British Medical Journal*, Vol. 323, pp. 1-5.

Sefton, T. (2000) *Getting Less for More: Economic Evaluation in the Social Welfare Field*. CASE Paper 44. London: Centre for the Analysis of Social Exclusion.

Shiell, A. ja Hawe, P. (1996) Health promotion, community development and the tyranny of individualism. *Health Economics*, Vol. 5, pp. 241-47.

Siegel, J., Torrance, G. ja Russell, L. ym. (1997) Guidelines for pharmacoeconomic studies: recommendations from the panel on cost effectiveness in health and medicine. *Pharmacoeconomics*, Vol. II, No. 2, pp. 159-68.

Stinnett, A., Murray, A., Mittleman, M., Weinstein, M., Kuntz, K., Cohen, D., Williams, L., Goidman, P., Staiger, D., Hunink, M., Tsevat, J., Tosteson, A. ja Goidman, L. (1996) *The cost-effectiveness of dietary and pharmacological therapy for cholesterol reduction in adults*. Teoksessa M. Gold, J. Siegel, L. Russell ja M. Weinstein *Cost-effectiveness in Health and Medicine*. New York: Oxford University Press.



Task Force on Principles for Economic Analysis of Health Care Technology (1995) Economic analysis of health care technology: a report on principles. *Annals of Internal Medicine*, Vol. 122, pp. 61-70.

Thompson, R.W., Ruma, P., Schuchmann, M. ja Burke, R. (1996) A cost-effectiveness evaluation of parent training. *Journal of Child and family Studies*, Vol. 5, No. 4, pp. 415-29.

Welsh, B.C. ja Farrington, D.P (1999) A review of the costs and benefits of situational crime prevention. *British Journal of Criminology*, Vol. 39, No. 3, pp. 345-68.

White, M. ja Lakey, J. (1992) *The Restart Effect: Evaluation of a Labour Market Programme for Unemployed People*. London: Policy Studies Institute.

Williams, A. (1992) Cost-effectiveness analysis: is it ethical?. *Journal of Medical Ethics*, Vol. 18, No. I, pp.7-11.

