

# Soveltavaa arviointitutkimusta ja kehityskumppanuutta

## Tekesin ohjelma-arviointien kehitysvaiheet ja aseointi tilaajan näkökulmasta

*Pekka Pesonen*

### ABSTRACT

The article studies the commissioner viewpoint of the evaluation processes in the field of research, development and innovation programmes. It is based on the experience gained in Tekes, the Finnish agency for technology and innovation. Tekes has commissioned more than one hundred evaluations during the last twenty years. Different evaluation approaches and procedures are described and the evolution phases of Tekes programmes and evaluation are presented. The article contains two case examples. The author recommends the need for both applied research and consultancy in evaluations. The role of strategic development and utilization is emphasized.

### 1. JOHDANTO

Tässä artikkelissa kuvataan tilaajan ja suunnittelijan näkökulma arviointitoimintaan. Erityisesti keskitytään tutkimus-, kehitys- ja innovaatiotoiminnan arviointiin ja siinä vielä tarkemmin ottaen Tekesin teknologiaohjelmiin. Pyrin kuvaamaan arviointien toteutuksessa olleita eri kehitysvaiheita ja asemoimaan Tekesin arviointeja muihin arviointeihin. Myös eri arviointikoulukuntien ja tavoitteiden välistä jännitettä tarkastellaan.

Artikkeli on laadittu käytännön toimijan näkökulmasta, ja siinä pyritään kriittiseen itsereflektioon.<sup>1</sup> Ensisijainen tavoite ei ole tieteellinen kontribuutio. Analyysi pohjautuu dokumenttiaineistoon ja osallistuvaan havainnointiin. Aineistona käytetään arviointikirjallisuuden lisäksi Tekesin arviointiraportteja ja niihin viittaavia julkaisuja. Erillisiä haastatteluja tai kyselyjä ei ole tätä artikkelia varten käytetty lähteinä.

Artikkeliin valittu tilaamisosaamisen näkökulma on jäänyt varsin vähälle huomiolle. Arviointien tilaamisen tueksi on muutama perusteos (Vuorela 1997, Virtanen 2007). Lisäksi on runsaasti eri alojen ohjeistuksia, joista muutamia käsitellään tässä artikkelissa. Tilaamista sivuavia lehtiartikkeleja on ilmestynyt Hallinnon tutkimus-lehden aiemmissa arviointiliitteissä.

Alan oppikirjoissa, lehdissä ja konferenssipapereissa on lähes poikkeuksetta muu painotus kuin tilaaminen. Yleensä arviointitutkimusjulkaisuissa kiinnostus kohdistuu tuloksiin, joita verrataan eri arviointien kesken. Arviointimenetelmiä on myös arvioitu ja niiden soveltuvuutta eri kohteisiin analysoitu. Arviointiprosessissa on jonkin verran huomiota kiinnitetty hyödyntämiseen, erityisesti arviointien oppikirjat ovat ottaneet tähän kantaa.

Seuraavassa luvussa tehdään katsaus eri arviointisuuntiin ja toteutustapoihin. Luvussa 3 käsitellään Tekesiä ja Tekesin ohjelma- ja arviointitoiminnan kehitystä. Luvussa 4 on lyhyt yhteenveto ja johtopäätöksiä.

## 2. ARVIOINTIEN MONIPUOLISUUS

### 2.1. Arvioinnin näkökulmia

Arviointia kuvastaa eri koulukuntien laajuus ja monipuolisuus. Sitä on vaikea kuvata omana tieteenalanaan, jolla olisi vakiintunut asema (evaluation as science), pikemminkin voidaan puhua arviointikäytännöistä (evaluation as practice). Osaamisen kannalta tämä tarkoittaa sitä, että menetelmiin ja teorioihin perehdytään esimerkiksi arviointikursseilla ja sen jälkeen oppii tekemällä ja osallistumalla arviointiprosessiin. Eri koulukuntien soveltaminen riippuu arviointien kohteista ja tavoitteista. Tässä luvussa esitellään laajasti eri arviointiteorioita; näkökulmana on eri koulukuntien väliset erot ja niiden vaikutukset tilaajan kannalta.

Arviointia tulkitaan usein tavoitteiden pohjalta. Chelimsky jakaa arvioinnin kolmeen osaan: tilivelvollisuus, kehittäminen ja tiedon tuottaminen. Tilivelvollisuusarvioinnilla tarkoitetaan toiminnan tuloksellisuuden arviointia, jolla on julkisella sektorilla tiivis yhteys tulosohjaukseen. Kehittävä arviointi tukee organisaatioiden kehittämistoimintaa. Tiedontuotantoarviointi tarjoaa uutta tietoa ja käsitteitä päätöksenteon tueksi. Varsin samantapainen on myös Ripin jaottelu. Arvioinnin rooleina ovat tilivelvollisuus (accountability), päätöksenteon tuki (decision support) ja strateginen muutos (strategic change). Tilivelvollisuus viittaa suorituksen mittaukseen, mitä rahalla on saatu aikaan. Päätöksenteon tuella haetaan arvioinnista apua valintaan. Tästä hyvä esimerkki peer review -menettely hakemusten asiantuntijaneelina. Hän korostaa arvioinnin strategista roolia kehittämisessä. Mark et al. määrittelevät arvioinnille seuraavat tavoitteet: politiikan/ohjelman arvon määrittäminen, ohjelman/organisaation kehittäminen, laillisuuden varmistus ja tiedon aikaansaanti. (Chelimsky 1997, Mark et al. 2000, Rip 2003)

Arvioinnin hyödyntämistä ja käyttöä tutkinut Patton erittelee tiedon käyttötavat seuraavasti: yhteisen ymmärryksen vahvistaminen, osallistujien sitoutumisen vahvistaminen, ohjelman tukeminen, ohjelman tai organisaation kehittäminen. Tiedon eri käyttäjien näkökulmat tulee ottaa huomioon prosessissa siten, että tiedon hyödyntäminen helpottuu. Arviointiprosessiin osallistumisella on tärkeä sija oppimisessa. Weiss nostaa esille tulosten hyödyntämisen konseptuaalisesta (conceptual) näkökulmasta. Vaikka tiedolla ei ole yhteyttä välittömään päätöksentekoon, sen avulla ymmärretään

ohjelmaa paremmin. Saadaan uusia ideoita ja näkemyksiä tulevien päätösten pohjalle. (Patton 1997, Weiss 1998).

Pitkälle vietyä osallistavuus lähenee voimaannuttamisen (empowerment) käsitettä. Tällöin organisaation tai ohjelman jäsenet vastaavat itse myös arvioinnin toteuttamisesta. Arvioijan rooli on asiantuntijatukena ja ohjaajana. Fetterman on tunnettu voimaannuttamiskoulukunnan keskeinen hahmo. Arviointitoiminnassa itsearviointi on luonteeltaan varsin lähellä voimaannuttamista. Hildén puhuu demokraattisesta arviointiprosessista, jossa kaikki asianosaiset arvioijaa myöten ovat tasavertaisia osallistujia. Tämä voi liittyä muun muassa ympäristövaikutusten arviointiin. (Fetterman 2001, Vataja 2005, Hildén 2007)

Arvioinnin hyödyntämistä paljon tutkinut Vedung on nostanut esille valtapyrkimykset arviointien taustalla. Legitimointi on usein arvioinnin tavoitteena ja se toimii siten suostuttelun välineenä. Usein myös kyse on ollut itseisarvoisesta toiminnasta, jolla arvioinnin tuloksista ja oppimisesta ei olla kiinnostuneita. Vedung käyttää myös käsitteitä: diffuusiostrategia, tuotantokeskeinen strategia ja metaevaluoinnit. Diffuusiostrategiassa kyse on tiedon tehokkaasta levittämisestä, esimerkiksi raporttien laadunparantamisella. Kuitenkaan tämä ei yksin riitä. Tuotantokeskeisellä strategialla pyritään vaikuttamaan arviointiprosessiin siten, että hyödyntäminen paranisi. Vedungin metaevaluoinnilla tarkoitetaan itse asiassa kolmea erityyppistä toimintaa: toisen arvioinnin arviointia, syntetisoivaa arviointia usean samaa tai samankaltaisia ohjelmia käsittelevän arvioinnin pohjalta sekä tarkastustyyppistä arviointitoiminnan arviointia. (Vedung 1997)

Ohjelmien arvioinnissa keskeinen asia on ohjelmateoria tai vaikuttavuusmalli: mitä asioita tavoitellaan. Clarken ja Dawsonin ohjelmateoria kuvaa sitä, miten ohjelman oletetaan toimivan ja mitä kausaalisuhteita syntyy toimien ja tulosten välillä. He jakavat mallin kahteen eri tarkastelutapaan. Deskriptiivisessä mallissa lähdetään kontekstista ja edetään mekanismien kautta tuloksiin. Pyritään siis ymmärtämään sitä, miltä yhteyttä ohjelman toimilla on tuloksiin. Preskriptiivisessä mallissa puolestaan lähdetään tavoiteltavista tuloksista, ja suunnitellaan toteutustapa ja ohjelman rakenne niitä palvelemaan. Georghiou ja Roessner korostavat ohjelmateorian ymmärtämisen merkitystä koko arvioinnin toteuttamisen kannalta. Heidän mukaansa arvioinnin rakenne, menetelmä ja analyysit riippuvat siitä, millaista muutosta ohjelmalla tavoitellaan. Chen

on esitellyt teorialähtöisen vaikuttavuusarvioinnin. Se on tulevaisuuteen suuntautunut ja painottaa muutosteorioita. (Clarke & Dawson 1999, Georghiou & Roessner 2000, Chen 2005)

Pawson ja Tilly ovat lanseeranneet realistisen arvioinnin, jota käytetään aktiivisesti erityisesti sosiaali- ja terveysalalla. Sitä kohtaan esiintyy myös Suomessa voimakasta mielenkiintoa: esimerkkinä tästä on sen ottaminen Suomen arviointiyhdistyksen vuoden 2009 vuosiseminaarin teemaksi. Realistisessa arvioinnissa selvitetään, mitkä vaikutusmekanismit sisältyvät interventioon, miten ne toimivat sekä millaisissa olosuhteissa vaikutukset ilmenevät. Realistisessa arvioinnissa korostetaan toteutuneiden ohjelmien ja politiikkojen mekanismien selvittämistä. (Pawson & Tilly 1997, ks. myös Paasio 2006)

Valtiontalouden tarkastusvirasto on käyttänyt tilivelvollisuudesta vastuullistamisen käsitettä. Vastuullistaminen on kiinteässä yhteydessä niin sanottuun uuteen julkishallintoon, jota Suomessa toteutetaan tulohajauksena. Tulohajaus on otettu käyttöön Suomessa valtionhallinnossa 1980-luvun lopulla. VM:n mukaan tavoite on ollut parantaa tehokkuutta ja laatua, mutta myös tilivelvollisuus on ollut tärkeällä sijalla. Alkuperäisinä kriteereinä olivat tuottavuus, taloudellisuus, vaikuttavuus ja tuloksellisuus. Vaikuttavuus on ollut näistä haasteellisista osista toteuttaa. Pitkän kehityksen tuloksena on päädytty tulosprismaan, joka otettiin käyttöön 2005. Tulohajauksessa organisaation yhteiskunnalliset vaikuttavuustavoitteet puretaan toiminnallisiin tulostavoitteisiin, joiden toteutumisesta ja aikaansaamasta vaikuttavuudesta raportoidaan osana tilivelvollisuutta. Pietilän mukaan toteutustavoissa on vielä kuitenkin suuria eroja. Lääkkeiksi esitetään tulohajauksen idean kirkastamista, mitareiden selkeää esitystä ja tulostavoitteiden konkreettisuutta. Oletettavaa onkin, että paineet vastuullistavan arvioinnin käyttöön tulohajauksessa eivät tule ainakaan vähenemään. (VTV 2008, VM 2008, Pietilä 2008)

Toisaalta myös muunlaisia kehitystrendejä on havaittavissa. Kuhlmann toteaa, että suoritusta mitaavan arvioinnin rajat tulevat vastan, kun monimutkaisuus lisääntyy. Kiinnostus kohdistuukin enemmän oppimissuuntautuneeseen arviointiin. Arviointiprosessista tehdään osallistava ja arvioinnille tulee välittäjärooli eri toimijoiden kesken. Kuhlmann tunnistaa neljä trendiä: a) toimien legitimoinnista ja vastuullistamisesta kohti parempaa ilmiön ymmärtämistä ja tulevien toimien tietopohjaa; b)

tehokkuuden ja taloudellisuuden mittauksesta toimenpiteiden soveltavuuden strategiseen kehittämiseen; c) riippumattomasta näytön tuottamisesta yhdessä tekemiseen ja oppimiseen; d) erillisistä arviointiraporteista yhdessä muiden tietolähteiden kanssa (benchmarking, ennakoinnit ja teknologian vaikutusten arvioinnit) niin sanotun strategisen älykkyyden (strategic intelligence) tuottamiseen. (Kuhlmann 2003)

Myös Guyn mukaan arvioinnin tarve on muuttunut legitimoinnista ymmärryksen parantamiseen ja tulevien toimien ohjaukseen. Arvioinnin kohdistuminen on laajentunut ja arvioijien rooli on muuttunut ulkopuolisesta tiedontuottamisesta osallistavaan suuntaan ja suositusten antoon samalla, kun käsitys innovaatioympäristön monimutkaisuudesta on lisääntynyt. (Guy 2003)

Valovirta ja Hjelt ovat todenneet, että innovaatioiden syntyamisprosessien monimutkaisuus ja systeemisyyden aikaansaa sen, että arvioinnille asetetaan uusia haasteita: legitimoivan palautetiedon tuottamisen sijasta tavoitteeksi asetetaan strategisen tiedon tuottamisen päämäärä. (Valovirta & Hjelt 2005)

Arviointien tilaamista on sivuttu useassa Hallinnon tutkimus-lehden artikkelissa. Temmes hahmottaa hyvää julkishallinnon arviointikäytäntöä ja -kulttuuria ja esittelee arvioinnin protokollan. Protokolla edustaa hyvää arviointikäytäntöä, jossa määritetään tilaajan ja arvioitsijan roolit ja vastuut. Temmes korostaa arvioinnin itsenäisyyttä ja riippumattomuutta arvioitavan toiminnan kehittäjistä kuten myös Kaakinen samassa lehdessä. Aivan keskeisenä tärkeänä on toimeksiantajan kysymysten selkeys ja arvioitsijan antamien vastausten kirkkaus sekä tulkinta. (Temmes 2004, Kaakinen 2004)

Arvioinnin tilaajana voi olla se organisaatio, virasto, jonka toimintaa ja toimenpiteitä arviointi koskee, tai tilaajana voi olla ylempi organisaatiotaso tai poliittinen päätöksentekijä. Eri tilaajilla voi olla toisistaan hyvinkin poikkeavat pyrkimykset sen suhteen, mitä ja miten ja milloin arvioidaan. Myös erilaisilla sidosryhmillä on omat ja erilaiset arviointiin kohdistuvat pyrkimyksensä. Tilivastuun toteuttaminen liittyy keskushallinnon ohjauksen kehittämiseen. Arviointitietämisen ja oppimisen välineinä ovat toimeenpanevien virastojen pyrkimyksiä. Valtionhallinnon yleisistä, konsernitason kehittämistarpeista, lähtevä arviointi on ylhäältä ohjattua erotuksena viraston omaan kehittämistarpeeseen perustuvasta arvioinnista. (Kauppi 2004)

Arviointikentän viime vuosien kehityksen leimallinen piirre on ollut sekä tilaajien että toteutta-

jiien ammattimaisuuden vahvistuminen. Tilaajien osalta tämä vahvistuminen on liittynyt ennen kaikkea arviointiprosessin administratiiviseen hallintaan: arviointitarjoukset, kilpailutus, sopimukset ja arviointityön toteutuksen seuranta sujuvat pääsääntöisesti muodollisesti moitteettomasti. Arvioitsijoiden osalta ammattimaisuus on tarkoittanut arviointiosaamisen vahvistumista, mihin liittyy monipuolista sekä teoreettista että menetelmällistä osaamista. (Kaakinen 2004) Kansainvälisten asiantuntijoiden vahvuutena on yleensä heidän ulkopuolisuutensa ja siten itsenäisyytensä suhteessa kohde-maahan. Myös heidän joukossaan on kuitenkin professioon liittyviä sidonnaisuuksia, jotka tulee tiedostaa ja tarvittaessa tasapainottaa käyttämällä erilaisista professionaalista lähtökohdista tarkastelevia arvioitsijoita. (Temmes 2004)

Tilaaminen on tietenkin hyvin pitkälle kytkeytyntä toteutukseen ja hyödyntämiseen. Toteutus vastaa tarjontaa ja hyödyntäminen kysyntää, jotka molemmat tilaajan on syytä tuntea ja pyrkiä tasapainottamaan: mitä tarvitaan ja mitä voidaan toimittaa. Georghiou on kuvannut tätä rajapintaa termein *delivery gap* ja *customer gap*. *Delivery gap* kuvaa arviointitulosten tulkintaa eri näkökulmista – tiedon hyödyntäjän ja arvioijan. Päätöksentekijät tarvitsevat ajantasaista tietoa resurssien allokointiin, selvää kytköstä tehtyjen investointien ja vaikutusten kesken, riippumattonta näyttöä tutkimuksen tasosta, ja indikaattoreita seurannan ja vertailun tueksi. Arvioijien näkemyksen mukaan tutkimuksen vaikutukset voivat näkyä vasta vuosien kuluessa, tutkimuksen lineaarimalli on harvinainen ja lisäarvo on monimutkaista selvittää, kollegat puolustavat toisiaan ja alaansa, ja mittarit ovat manipuloitavissa. *Customer gap* puolestaan kuvaa arviointiprosessia. Arvioijat haluavat ohjelmalle selkeästi määriteltäviä ja hierarkkisia tavoitteita, riippumattomuutta ja läpinäkyvyyttä, runsaasti aikaa ja resursseja, ja vapaan pääsy kaiken tiedon alkulähteille. Päätöksentekijöiden näkökulmasta taasen ohjelmat ovat kompromisseja, niillä monia ja ristiriitaisia tavoitteita ja intressejä, arviointisuositusten tulee olla realistisia, arviointitulokset tulee saada tiukassa ajassa pienin resurssein, ja toimijat ovat kiireisiä. (Georghiou 2001).

## 2.2. Arviointien toteutustapoja

Seuraavassa on katsaus eräisiin eri alojen arviointikäytänteisiin. Näkökulmana tässä luvussa on se,

miten eräät keskeiset arviointien tilaajat ovat käytännössä ratkaisseet ja ohjeistaneet toimintansa.

Euroopan unioni on koko toiminta huomioiden todennäköisesti maailman suurin arviointien toteuttaja. Komissiolla on systemaattinen ohjeistus vaikutusten arvioinnin tekoon (EC Impact assessment guidelines). Se on tarkoitettu komission henkilökunnan käyttöön, kun tulee tarve tehdä vaikutusten arviointi. Kyseessä voivat olla erilaiset aloitteet, lainsäädäntöhankkeet, tukiohjelmat ja muut selvitykset. Ohjeistus määrittelee yhtenäisen rakenteen ja laajuuden kaikille raporteille. Kaikille hankkeille asetetaan ohjausryhmä. Impact assessment board toimii laatu- ja kontrollielimenä ja hyväksyy kaikki raportit ennen kuin ne lähtevät komission sisäiseen konsultaatioon ja edelleen päätöksentekokäsittelyyn. Vaikutusten arviointi pyrkii ennakoimaan lainsäädännön ja muiden toimien tulevaisuudessa tuottamia vaikutuksia. (EU 2009)

Ohjelmien ja muiden toimien ja interventioiden arviointien teolle on oma ohjeistuksena (EC Evaluating EU Activities A practical guide for the commission services). Sitä käytetään sekä käytetään sekä hakemusten *ex ante* että ohjelmien *ex post* ja väliarviointeissa. EU:n käyttämässä mallissa tutkimusalojen asiantuntijat muodostavat paneelin, joka käyttää erilaista selvitys-, arviointi- ja haastatteluaineistoa. Arviointikonsulteilla on tätä tukeva rooli. Komissiossa arviointitehtävät on hajautettu pääosastoittain. Arvioinnin päätavoitteet ovat vaikutusta interventioiden suunnitteluun, avustaa resurssien tehokkaassa allokoinnissa, parantaa intervention laatua ja raportoida interventiolla saavutetuista tuloksista. Arviointi, seuranta ja tarkastus täydentävät toisiaan, mutta niillä on kuitenkin selvästi omat roolinsa. Hyväksi käytännöksi esitetään, että arvioinnin laajuus on noin 0,5 % arvioitavan ohjelman budjetista. (EU 2004)

EU:n tutkimus- ja kehitystoiminnan arvioinnin suurin ponnistus kohdistuu puiteohjelmiin. EU:n toteuttamat tutkimus- ja kehitystoiminnan arviointit ovat muuttuneet puiteohjelmien tavoitteiden mukaisesti. Ensimmäisissä puiteohjelmissä oli kyse korkeatasoisen tutkimuksen rahoittamisesta ilman laajempia vaikuttavuustavoitteita. Näin arvioinnissa kohteena oli tieteellinen laatu ja taso. (Guy 2003)

Keväällä 2009 valmistui kuudetta tutkimuksen puiteohjelmaa (2002–06) koskeva arviointi. Sitä toteuttamassa oli DG Researchin arviointiyksikön tilauksesta 12-henkinen asiantuntijaryhmä eri maista. Työtä tuki arviointikonsultti Technopolis. Arviointi oli pääosin meta-arviointia, jossa käytettiin teh-

tyjä selvityksiä ja arviointeja pohjalla. Näitä olivat puiteohjelmaosallistujille tehty kysely, verkostoanalyysit, selvitykset uusista instrumenteista ja bibliometrinen analyysi. Lisäksi työssä hyödynnettiin kansallisia arviointeja, joita oli tehty eri maissa, muun muassa Suomessa (ks. Kuitunen & Haila 2008). Tehty arviointia hyödynnetään uuden seitsemännen puiteohjelman arvioinnin suunnittelussa. (Fisch 2009)

Edelliseen arviointiin verrattuna oli tehty muutamia eroja. Arviointi kohdistui nyt koko puiteohjelmaan, eikä viisivuotiseen tutkimusjaksoon. Aineiston määrä ja laatu oli huomattavasti parempi. Arviointipaneelin työ kehittyi kohti näyttöperusteista arviointia. (Arnold 2009)

EU:n arviointijärjestelmää voi pitää sekä laajana että järjestelmällisenä. Parhaimmillaan arviointiraportit ovat erinomaisen kattavia tutkimuksia, ja niihin on koottu paras asiantuntemus eri osa-alueilta. Hyvä esimerkki tästä on vaikuttavuusarviointiraportti *Assessing the Socio-economic Impacts of the Framework Programme* (ks. Georghiu et al. 2002).

Suomen Akatemia on tieteellisen tutkimuksen rahoittajana vahvasti arviointiorganisaatio. Lähes kaikessa Akatemian toiminnassa on arviointi sisäänrakennettuna, ja ulkomaisia asiantuntijoita käytetään paneeleissa hakemusten, ohjelmien, tieteenalojen ja muiden arvioinnissa. Tutkimusohjelman arviointi nähdään tutkimuksen kehittämisen ja tiedepolitiikan välineenä. Arviointi tuottaa tietoa ohjelman tavoitteiden toteutumisesta, onnistumisesta uuden tiedon tuottamisessa sekä ohjelman aikaan saamista lisäarvosta. Arvioinnilla pyritään myös saamaan palautetta ohjelmajärjestelmästä ja koordinaatiosta sekä muuta tietoa tiedepoliittisen suunnittelun ja päätöksenteon pohjaksi: Lisäksi kukin arviointi edesauttaa arviointiprosessin edelleen kehittämistä ja toimii arviointiin osallistuvien oppimisprosessista. Keskeisiä käsitteitä ovat ohjelman relevanssi, vaikuttavuus, hyöty ja tehokkuus. Akatemian ohjelma-arvioinnin suunnittelusta ja läpiviennistä vastaa ohjelmajärjestelmä, johon kuuluu tieteellisten toimikuntien tai hallituksen jäseniä, muiden rahoittajien ja hyödyntäjien edustajia, muita asiantuntijoita sekä ohjelmajärjestelmän johtajia. Ohjelman arvioi asiantuntijajaneeli. Arvioinnin jälkeen toimikunnat laativat jälkihoitosuunnitelman arviointisuositusten toteuttamisesta. (Suomen Akatemia 2003). Akatemia tuottaa myös ohjelmatoimintaa laajempia analyyseja, kuten Suomen tieteen tila ja taso. (Suomen Akatemia 2009)

Ulkoasianministeriö hallinnoi kehitysyhteistyövärojen käyttöä. Arvioinneilla selvitetään mitä hankkeilla on saatu aikaan ja kuinka tehokkaasti resurssit on onnistuttu käyttämään. Arvioinnille on asetettu kansainvälisesti sovittuja periaatteita, joista tärkeimmät ovat OECD:n Kehityskomitean (DAC) tekemiä. Arviointiryhmät valitaan tarjouskilpailun perusteella. Asiantuntemusta tulee olla arvioinnista ja teemasta. Lisäksi ryhmässä tulee olla edustus maista, jota arviointi koskee ja sukupuolisen tasa-arvon pitää toteutua. Arviointikysymykset on pitkälti muotoiltuja. Ne koskevat relevanssia, tehokkuutta, vaikuttavuutta, pysyvyyttä, koherenssia ja lisäarvoa. Ohjeistus antaa myös tarkistuslistat eri työvaiheisiin. Ne koskevat sekä arvioinnin tilaajaa, vastuuhenkilöä, arviointiryhmää että arvioinnin kohdetta. Palautteenanto ministeriön yksiköiltä on otettu omaksi työvälineeksi. (Ulkoasianministeriö, 2007)

Yksi esimerkki arviointien ohjeistuksesta on kuntaliitolta. Tässä tapauksessa ohjeistus ei ole niin sitova vaan enemmän suositusluontoinen ja tukea antava. Kuntaliiton ohjeistus huomioi eri hallinnonalojen arvioinnit, kuten sosiaali- ja terveydenhuolto tai opetuspalvelut. Näissä on eroja muun muassa sen mukaan, mikä on keskitettyä tai kansallisen tason arviointia, missä määrin kukin kunta tekee arviointia itse. (Kuntaliitto 2006)

Edellä on kuvattu arviointiohjeistusten avulla eräiden merkittävien arvioinnin tilaajien toimintatapoja. Arviointiohjeistuksella on kehitetty käytäntöjä, mutta mikään näistä ei vaikuta suoranaisesti pohjautuvan tiettyyn teoriaan tai koulukuntaan, vaan eri aineksia on yhdistetty soveltavasti.<sup>2</sup>

### 2.3. Arvioinnit tilaamisosaamisvaatimusten kannalta

Edellä esitettyjen arviointikoulukuntien ja käytännön toteutustapojen pohjalta voikin vetää yhteen, että arvioinnin tilaamisessa on tarpeen ymmärtää ainakin alan toimintapolitiikkaa kuten ympäristö- tai innovaatiopolitiikkaa ja sen analysointia, arvioinnin teoriaa, käytäntöä ja lainsäädännöllistä pohjaa sekä sitä, kuinka arviointia sovelletaan ja mikä on arvioinnin kohde ja rajoitukset. Tilaamisessa tarvitaan analysointikykyä, verkostoitumiskykyä ja projektin hallinnointikykyä. Myös arviointiprosessia, tiedonlähteitä ja arviointimenetelmiä on syytä tuntea.

Erilaisista viitekehyksistä tähän artikkeliin mielenkiintoisin konteksti on arviointien jakaminen ta-

voitteiden mukaan: tilivelvollisuus; tiedon tuottaminen; kehittäminen. Näiden välillä näyttää olevan runsaasti jännitteitä. (VTV 2008, Kuhlmann 2003, Valovirta & Hjelt 2005, Kauppi 2004, Veuglers et al 2009) Tämä viitekehys tarjoaa myös hyvän pohjan keskittyä tilaajan rooliin ja rajata tarkastelua ohjelmatasolle. On toki huomattava, että jako ideaalittyyhin arviointeihin on teoreettinen, sillä

usein arvioinneissa pyritään huomioimaan eri näkökulmia. Tämä näkyy mm. edellisessä luvussa läpi käydyissä arviointiohjeistuksissa.

Eri tavoitteilla on eroja sen suhteen, mitä ne edellyttävät tilaajalta ja arviointiprosessilta. Taulukossa 1 esitetään eräitä eroja ja yhtäläisyyksiä. Vastaavanlaista tarkastelua on tehty mm. Kelan vaikutavuuden arvioinnista. (vrt. Rajavaara 2006)

**Taulukko 1. Arviointiprosessin typologia tavoitteiden ja tilaamisen näkökulmasta**

tavoite	tilivelvollisuus	tiedontuotanto	kehittävä
tilaaja	organisaation ulkopuolinen, esim. ministeriö tai Akatemia suhteessa yliopiston tutkimusryhmään	aloitteen tekijänä useimmiten ulkopuolinen, esim. tutkimusryhmä	organisaatio itse, esim. ohjelmasta vastuussa oleva organisaatio
toteuttaja	asiantuntijapaneeli, peer review	tutkimusryhmä	arviointikonsultti
ohjaus	tilaajan edustajia, seurantarooli, hands-off	aiheesta kiinnostuneiden ideointia, ohjauksen rooli pieni	tilaajan edustajia, tiivis vuorovaikutus, hands-on
aineistot	kvantitatiivista, indikaattorit, bibliometriikka, tilastoaineisto	sekä kvantitatiivista että kvalitatiivista, tutkimuksen menetelmät	kvalitatiivista, haastattelut, kyselyt, dokumenttiaineisto
hyödyntäminen	arviointiraportti, laaja julkisuus, välillinen vaikuttaminen	tutkimukselliset tulokset, tutkimuksen oppimisprosessi, välillinen vaikuttaminen	kehittävä oppimisprosessi, julkiset tulokset, välitön vaikuttaminen

Tilivelvollisuuden arviointi soveltuu parhaiten kokonaan ulkopuolisten tekemäksi, jotta saavutetaan selvä riippumattomuus. Tällöin sekä tilaaja että arvioinnin toteuttaja ovat vailla kytkentöjä arviotaan kohteeseen. Aineistojenkin tulisi olla pikemmin indikaattoripohjaisia, kuten kustannus/hyötyanalyysi tai viittausindeksi kuin esimerkiksi haastatteluihin perustuvia. Vastaavasti arvioinnin toteuttajat ovat luonteesta alan substanssiasiantuntijoita (peer review) tai ekonomisteja. Tilaajan tulisi panostaa aineistojen luotettavuuteen ja työn käynnistyttyä osallistua ohjaukseen lähinnä seurantana.

Tiedontuottamisarviointi voisi olla luonteeltaan esimerkiksi tutkimus, joka lähtee ilmiön kiinnostavuudesta tutkijan näkökulmasta. Tilaajan rooli olisi lähinnä rahoittaja, haastateltava ja tutkimuksen seuraaja. Aineistot voivat olla hyvinkin laajat, mutta ennen kaikkea tutkijan valitsemat.

Kehittävä arviointi on tilaajanäkökulmasta varsin intensiivinen. Tilaaja on organisaatio itse ja arvioijat ovat ulkopuolisia arvioinnin asiantuntijoita.

Tilaajan rooliin kuuluu eri tavoin hyödyntämisen huomiointi ja mahdollisimman tiivis vuorovaikutus arviointityön aikana. Aineistot ovat useimmiten kvalitatiivisia ja haastattelu/kyselypohjaisia.

Tekesin toteuttamissa arvioinneissa on eri aikoina ollut kaikkia näitä piirteitä. Tämänhetkinen suuntaus painottaa kehittävä näkökulmaa. Seuraavassa luvussa perehdytään tarkemmin Tekesin ja sen ohjelma-arviointitoiminnan kehittämiseen.

### 3. TEKES, OHJELMAT JA ARVIOINNIT

#### 3.1. Tekesin ja ohjelmien 25 vuoden kehityskaari

Teknologian ja innovaatioiden kehittämisskeskus Tekes on yritysten, yliopistojen, korkeakoulujen ja tutkimuslaitosten haastavien tutkimus- ja kehitysprojektien ja innovaatiotoiminnan rahoittaja ja aktivoija. Rahoitusta, asiantuntemusta ja verkos-

toja voivat saada yritykset, tutkimuslaitokset sekä julkisen ja kolmannen sektorin toimijat. Kaiken kokoiset yritykset voivat saada rahoitusta haasteellisiin ja uutuusarvoa sisältäviin kehittämissuomiin. Rahoitusta voi saada myös julkiseen tutkimukseen sekä julkisen ja kolmannen sektorin toimijoiden palveluiden kehittämiseen tietyissä Tekesin ohjelmissa. Tekesin ohjelmat ovat yrityksille ja julkisille tutkimusyksiköille tarkoitettuja rahoitus- ja asiantuntijapalveluiden kokonaisuuksia. Niitä käynnistetään aihealueille, joiden vauhdittaminen on Suomen elinkeinoelämän ja yhteiskunnan kannalta tärkeintä. Vuonna 2008 Tekes sijoitti 516 miljoonaa euroa yritysten, yliopistojen, korkeakoulujen ja tutkimuslaitosten tutkimus- ja kehitysprojekteihin ja innovaatiotoimintaan. Rungas puolet rahoituksesta kohdistui Tekesin ohjelmiin ja strategisen huippuosaamisen keskittymiin. Vuoden 2008 lopussa oli käynnissä 30 Tekesin ohjelmaa ja viisi strategisen huippuosaamisen keskittymien ohjelmaa. ([www.tekes.fi](http://www.tekes.fi), luettu 31.7.2009).

Tekes perustettiin vuonna 1983. Tätä ennen yritysten ja tutkimuslaitosten teknologian kehittämistä oli rahoitettu Kauppa- ja teollisuusministeriön ja Sitran toimin 60-luvun loppupuolelta alkaen. Heti alusta alkaen myös ohjelmatoiminta tuli mukaan Tekesiin. Näiden ensimmäisten ohjelmien päättyessä aloitettiin myös arviointitoiminta. (Murto et al 2007) Tekes on julkaissut yli 100 arviointiraporttia viimeisten 20 vuoden aikana.

Tekesin toimintakenttä on muuttunut ja laajentunut kuluneina vuosina. 1980-luvulla ja 1990-luvun alussa pääosin oli kyse teknologian kehittämisestä, jolloin asiakkaina olivat valmistavan teollisuuden t&k-yksiköt, VTT ja teknilliset korkeakoulut. Vähitellen toiminta laajeni innovaatioiden kehittämiseen, joka tarkoitti myös palvelusektorin tuloa asiakaskuntaan ja muutakin kuin teknologista tutkimusta. (Pesonen 1998, Gustafsson et al. 2003) 2000-luvulla ensin liiketoiminnan ja sitten työelämän kehittäminen ovat nousseet Tekesin tavoitteiksi. Yritysten ohella myös julkisen sektorin ja kolmannen sektorin kehityshankkeiden rahoittaminen on nykyisin mahdollisia. Tekesin toiminta on laajentunut myös sosiaali- ja terveystieteiden innovaatiotoimintaan sekä innovatiivisten julkisten hankintojen tukemiseen.

Teknologiaohjelmien muutostekijöitä ovat olleet seuraavat ohjelmatoiminnan kehityspiirteet:

- ohjelmien koko on kasvanut ja määrä vähentynyt
- kansainvälistymisen fokus on vahvistunut

- teknologiatyöntöisyydestä on siirrytty kohti markkinavetoisuutta
- poikkiteknologisuus ja teknologia-alojen rajat ylittävät ohjelmat ovat lisääntyneet – teknisen ongelman ratkaisuun pyrkivistä ohjelmista on siirrytty kohti kokonaisvaltaisempia ja monimutkaisempia ohjelmakokonaisuuksia
- ohjelmallisuuden lisäarvoon ja tulosten hyödyntämiseen on alettu kiinnittää enemmän huomiota
- ohjelmien rooli on muuttunut operatiivisesta rahoituksen allokoinnin välineestä strategisemmaksi – asiakastarpeiden lisäksi Tekesin sisältölinjaukset ovat tulleet vahvemmin mukaan ohjelmien lähtökohtiin
- ohjelmajärjestelmät ovat muuttuneet formaalimmiksi. (Valtakari et al 2007)

Teknologiaohjelmatoiminnan uudistaminen on tapahtunut Tekesissä varovaisesti pienten muutosten kautta. Selkeitä taitekohtia ohjelmatoiminnan kehityksessä ovat olleet vuodet 1999 ja 2006, jolloin Tekesin organisaatiota on muutettu. Organisaatiomuutokset ovat myös ilmentymä Tekesin strategiamuutoksista. Ohjelmatoiminnan kehityksen kannalta erityisen merkittävänä voidaan pitää vuotta 1999, jolloin ohjelmien johtamista alettiin kehittää määrätietoisesti ja mm. ensimmäinen ohjelmien kehittämisspällikkö aloitti tehtävässään. Ohjelmatoiminnan strategisuuden kasvua Tekesissä kuvaa se, että organisaatiouudistuksessa vuoden 2006 alussa teknologiaohjelmat on nostettu uudessa organisaatiossa yhdeksi ydinprosessiksi ja prosessi on saanut oman johtajansa. (Valtakari et al 2007)

Tekes seuraa asiakkaidensa hankkeiden etenevistä ja onnistumista ja teknologiaohjelmien vaikuttavuutta useilla keinoilla. Näitä ovat

- asiakkaille ja ulkopuolisille tehdyt kyselyt
- ulkomaisten asiantuntijoiden tekemät laadulliset arvioinnit
- ulkopuolisten tekemät, asiakaskyselyihin perustuvat arvioinnit
- makroekonomiset vertailututkimukset
- toimijoiden arvioinnit
- kansainvälisen tason indikaattorivertailut. (Tekes 2008)

Nämä muodostavat kokonaisuuden, joka raportoidaan osana tulossuunnitelmasta Työ- ja elinkeinoministeriölle. Vaikuttavuusarviointi kuuluu ministeriön ja Tekesin väliseen tulossopimukseen. Virallisen ra-

portoinnin lisäksi vaikuttavuusarviointi tuottaa suuren osan Tekesin julkaisuista. Tekes arvioituttaa kaikki ohjelmansa, mutta tätä ei ole lailla tai asetuksella säännelty. Tarkempi ohjeistus on Tekesin ja erityisesti arvioinnista vastuullisen tehtävä.

Nykyisin Tekesin ohjelmat arvioidaan ohjelmien päättyessä, yleensä myös puolivälissä ja usein myös muutama vuosi päättymisen jälkeen. Arvioinnin tavoitteena on antaa palautetta ohjelman tavoitteiden toteutumisesta ja sen vaikuttavuudesta, sekä tuottaa tietoa ohjelmatoiminnan ja Tekesin toiminnan strategisen kehittämisen tueksi. Arvioinnin avulla saadaan tietoa ja ymmärrystä tutkimus- ja kehitystoiminnan dynamiikasta sekä sen menestymiseen ja epäonnistumiseen vaikuttavista tekijöistä. Tekesin vaikuttavuusarviointiyksikkö vastaa ohjelmien loppu- ja jälkiarvioinneista. Arvioinnit tehdään ulkoisella arviointiasiantuntijaryhmällä monipuolisen ja riippumattoman vaikuttavuustiedon kokoamista varten. Samassa arvioinnissa voidaan tarkastella myös useita ohjelmia, kun ne kuuluvat samaan teknologia-alueeseen tai klusteriin tai kun niillä on samanlaiset tavoitteet tai jokin muu yhteinen nimittäjä. Yhteisrahoitteisissa ohjelmissa tehdään myös arvioinnissa yhteistyötä muiden rahoitajien kanssa. Väliarvioinnin suunnittelusta, toteutuksesta ja budjetoinnista vastaa ohjelma. Väliarviointi tehdään itsearviointina, mitä joissain tapauksissa voidaan täydentää ulkopuolisen arviointiryhmän työllä. Väliarvioinnin tueksi on kehitetty väliarviointityökalu verkkokyselyn toteutukseen ja raportointiin.

Tekesin loppu- ja jälkiarvioinnit julkaistaan Tekesin ohjelmaraporttien sarjassa (<http://www.tekes.fi/fi/community/Arviointiraportteja/669/Arviointiraportteja/1598>) ja tuloksia esitellään vuosittaisissa arviointiseminaareissa ja ohjelmien loppuseminaareissa. Väliarvioinneista esimerkkinä voi todeta Finnano-ohjelman väliarviointiraportin (<http://akseli.tekes.fi/opencms/opencms/OhjelmaPortaali/ohjelmat/NANO/fi/julkaisutjalinkit.html>).

Seuraavassa kuvataan arviointitoiminnan kehitystä kolmella vuosikymmenellä.

### 3.2. Tasonmittausarvioinnit 80- ja 90-luvulla

Tekesin teknologiaohjelmien arviointi on aloitettu ensimmäisten ohjelmien päättyessä vuonna 1988 (Lemola & Siivonen 1988). Mainittakoon, että arviointitoiminta yleistyi samaan aikaan myös muualla, mm. VTT:llä tämän artikkelin kirjoittaja toimi ar-

viointiryhmien teknillisinä sihteerinä. Vallitsevana käytäntönä oli joko peer review perinteisessä mielessä, jolloin alan tunnetut professorit muodostivat raadin, tai henkilökeskeisempi malli, jolloin kutsuttiin tutkimusyksikön johtaja tai yrityksen teknologiajohtaja toimimaan arvioijana. Tuohon aikaan oli t&k-arvioinneissa ammattimaisten arviointikonsulttiyritysten käyttö vielä hyvin harvinaista. Arviointitoiminnan alkuvuosina palautteen anto ohjelmalle oli keskeisin tavoite. Ohjelmat toimivat tilaajina.

Arviointitoiminnan vakiinnuttua noin vuoteen 1994 mennessä oli havaittavissa kahdenlaiset kiinnostuksen kohteet. Toinen kohdistui ohjelmien toteutukseen, hallintoitiin ja tuloksiin teollisesta näkökulmasta, toinen taas arvioi tieteellistä ja teknologista laatua kansainvälisestä näkökulmasta. Tekesin arviointia ohjasi Tekesin toiminta-ajatus Suomen elinkeinoelämän kehittämisestä teknologisin keinoin. Yksittäisten ohjelmien arvioinnin ohella 90-luvun lopulla arviointia laajennettiin joissakin tapauksissa koskemaan useampaa, toisilleen läheistä ohjelmaa. Myös muita Tekesin palveluita on arvioitu, näitä olivat teknologiaklinikat ja teknologiastrategiakonsultointipalvelut. (Oksanen 2000)

Tekesin toteuttamissa arvioinneissa korostui vahvasti tasonmittaus. Ulkomaiset professorit tunnistivat tieteen ja teknologia-aloja, jotka olivat hyvää kansainvälistä tasoa. Vastaavasti esille nostettiin myös niitä projekteja, joita ei pidetty riittävän tasokkaina ja haasteellisina tieteellisin perustein. Joissakin tapauksissa professorit kommentoivat myös suotuisasti hallinnollista toteutusta. Lähes poikkeuksetta kannustettiin jatkamaan panostuksia. Usein kehoitettiin tutkimuslaitehankintoihin ja väitelyiden alentamiseen, mitkä olivat viestejä opetusministeriölle ja yliopistoille, eivät niinkään Tekesille.

Luvussa 2 esitelty tilivelvollisuus oli kiinteä osa tasonmittausarviointia. Tällöin poimittiin yksittäisiä projekteja, jotka ulkomaiset professorit arvioivat. Raportteihin saattoi tulla voimakasta yhtä tutkimusryhmää koskevaa kritiikkiä. Tästä löytyy useita esimerkkejä arviointiraporttien yhteenvedosta (ks. Oksanen 2000).

Oksanen mukaan arvioinneissa oli kaksi merkittävää heikkoutta: käytettävissä olleen ajan niukkuus ja arviointisijoiden heikko olosuhtetuntemus. Kehittämiskohteita nähtiin siinä, että paneuduttiin liikaa tiedon tuottamiseen menneistä toimista ja päätöksistä eikä katsottu eteenpäin. Myös tiedon kulkua arviointisijoiden sekä tiedon hyödyntäjien ja muiden osapuolten välillä tulisi parantaa. (Oksanen 2000) Tekesin selvityksessä tehdyissä haastat-



teluissa useat totesivat, että monet arviointiraportit jäivät vain hyllyyn, ja niiden käyttö tietolähteenä oli puutteellista. (Pihlajamaa et al 1998).

Vaikka Tekesissä arviointien hyödyntäminen olikin vähäistä, niin avoimeksi kysymykseksi jää, oliko juuri ennen vuosien 1990–93 lamaa aloitetulla tutkimus- ja kehitystoiminnan arvioinnilla jotain vaikutusta poliittiseen päätöksentekoon. Suomessa panostettiin myös laman olosuhteissa kasvavasti rahoitusta t&k-toimintaan, vaikka muuten julkinen talous oli vaikeuksissa. Tämä viittaa joko päättäjien vahvaan näkemykseen t&k-toiminnan hyödyllisyydestä tai saatuun informaatioon esimerkiksi arviointitoiminnasta. Tämä olisi mielenkiintoinen jatkotutkimuksen kohde.

Tasonmittausarviointien voi sinänsä katsoa sopineen varsin hyvin ajankohtaan. Suomalainen tutkimus- ja kehitystoiminta oli vielä suhteellisen pientä, mutta vahvassa kasvussa. Oli tärkeää saada näkemystä niistä aloista, joissa on päästy hyvää kehitysvauhtiin. Oli myös aika rakentaa näille kansainvälisiä yhteyksiä parhaisiin tutkimusryhmiin muissa maissa. Se taas kantoi hedelmää, kun Suomi liittyi ensin EU:n tutkimusohjelmiin ja sitten EU:n varsinaiseksi jäseneksi.

Kuten muuallakin (ks. luku 2.2.) eräs kehitysvaihe oli myös Tekesissä ohjeistuksen laadinta. Teknologiahjelmien arviointi ja ohjaus -raportti on lopputulos Tampereen teknillisellä korkeakoululla tehdystä monivuotisesta tutkimushankkeesta. Se sisältää ohjelman arvioinnin työvaiheet ja niitä tukevat työkalut lähtien liikkeelle arviointistrategian laadinnasta ja päätyen hyödyntämiseen päätöksenteossa. Painopiste on ohjelman päällikön ja johtoryhmän työssä ja sitä tukevan palautemekanismin toteutuksessa. (Berg & Lindberg 1997) Kyseinen raportti palveli siten ennen kaikkea osallistujien oppimista. Sitä käytettiin useita vuosia Tekesissä tilaajan oppaana erityisesti, kun arviointi oli vielä 90-luvun lopulla ohjelmien vastuulla.

### 3.3 Kokeilevaa arviointitutkimusta

Tekesiä koskeva arviointi toteutettiin 1995 osana kauppa- ja teollisuusministeriön alaisten organisaatioiden arviointiohjelmaa. Arvioinnissa nostettiin kehityskohteina esille seurannan ja arvioinnin prosessien kehittäminen ja kytkeminen tavoitteisiin ja toimenpiteisiin. (KTM 1995)

90-luvun lopulla vaikuttavuusarviointi koettiin laajalti valtionhallinnossa ja myös Tekesissä entistä

tärkeämmäksi, mikä osaltaan johtui tulosohjauksen tuomasta paineesta. Tekesissä organisoitiin arvioinnille oma yksikkönsä. Sen nimenä oli alkuun Vaikuttavuus ja laatu, vuodesta 2006 Vaikuttavuusarviointi.

Yksikön toiminta-alue määriteltiin varsin laajaksi. Arviointitoiminta kytkettiin innovaatiotutkimukseen ja innovaatiopolitiikkaan sekä ennakkointiin. Lisäksi yksiköllä oli osavastuu Tekesin johtamisessa tarvittavan tiedon tuottamisessa ja vaikuttavuutta kuvaavan materiaalin tuottamisessa. Arviointiyksikkö laati vaikuttavuuden arvioinnin strategian ja organisoi arviointimenetelmien ja tietokantojen kehityshankkeita. (Tekes 2000) Laaja ja monivuotinen ponnistus on ollut innovaatiotoiminnan vaikuttavuusmallin kehittäminen: panokset – tulokset – välittömät vaikutukset – vaikutukset kansantalouteen ja yhteiskuntaan (ks. Tekes 2008)

Ohjelmien arviointien tilausvastuu tuli keskitetyksi ja arviointiasiantuntijan rooli muuttui osapäiväisestä tukihenkilöstä päätoimiseksi vastuuhenkilöksi. Osaamista alkoi kumuloitua uuteen yksikköön, kun aiemmin arviointien osaaminen oli hajallaan.

Tekes teetti muutaman selvityksen ohjelmaarviointien hyödyntämisestä. Hyödyntämisessä todettiin kehitettävää. Tekesin toimintakenttä oli laajenemassa teknologian kehittämisestä laajemman innovaatiotoiminnan suuntaan. Teknologiahjelmat kärsivät inflaatiosta, enimillään oli samanlaisesti käynnissä 78 ohjelmaa. Näin suurta joukkoa ei pystytty hallitsemaan enää perinteisin keinoin. Edellisessä luvussa kuvattiin arviointeihin kohdistunutta kritiikkiä. Tekesissä pyrittiin laajentamaan näkökulmaa hyödyntämisen kehittämiseksi. (Pesonen 1998, Pihlajamaa et al 1998, Oksanen 2000).

Arvioinnin kasvava kysyntä merkitsi arviointiin erikoistuneiden organisaatioiden syntyä ja kasvua ja arvioinnin ammattimaistumista. Tekes kehitti yhteistyössä näiden kanssa arviointimenetelmiä, erityisesti arviointien tutkimuksellista otetta tiedontuottajana. Tätä kuvataan tarkemmin seuraavassa case-esimerkissä, joka käsittelee Tekesin elektroniikan ja tietoliikenteen ohjelmia.

#### 3.3.1 Case-esimerkki: Elektroniikkaa ja tietoliikennettä

Vuonna 1996 Suomen hallitus päätti ohjata valtionyhtiöiden myynnistä saatuja varoja noin 500 miljoonaa € t&k-toimintaan, pääosin Tekesin kaut-

ta. Tekes valmisteli elektroniikka-alalle ETX- ohjelman ja tietoliikennealalle TLX-ohjelman. INWITE oli pohjoismainen yhteistyöohjelma. Alan perustutkimusta varten Suomen Akatemia käynnisti Telectronics-ohjelman.

ETX- ja TLX-ohjelmat olivat siihen aikaan Tekesin laajimpia. Kun keskimääräinen ohjelman koko vuonna 1997 oli 15 miljoonaa €, niin ETX:n ja TLX:n volyyymi oli kymmenkertainen, kummankin yli 150 miljoonaa €. Akatemian Telectronics oli kooltaan huomattavasti pienempi, noin 5 miljoonaa € ja INWITE vielä pienempi (Suomen osalta 2 miljoonaa €).

Ohjelmien poikkeuksellisen laajan koon takia arviointiin kiinnitettiin tavanomaista enemmän huomiota. Myös kohdistuminen vahvasti kasvussa olleelle teollisuudenalalle puolsi arvioinnin tärkeyttä. Väliarviointi ohjelmien puolivälissä oli vielä tuohon aikaan suhteellisen vähäistä, ja sitä oli käytetty kuudessa ohjelmassa. ETX:n ja TLX:n arviointisuunnitelmaan otettiin mukaan ulkopuolisten tekemänä sekä väli- että loppuarviointi.

Väliarvioinnissa vuonna 1999 lähdettiin vahvasti arviointitutkimuksen suuntaan kokeilemaan uudenlaista lähestymistapaa. Lähtökohtana pidettiin edellä mainittua Tekesin ohjeistusta (Berg & Lindberg 1997) ja arvioinnin kysymykset tarkennettiin arvioinnin ohjausryhmässä. Arvioinnin ohjausryhmään kutsuttiin Tekesin vaikuttavuusarvioijan lisäksi Tekesin ohjelmien vastuuhenkilöt sekä ohjelmapäälliköt. Arviointikysymykset tarkennettiin koskemaan tavoitteiden toteutumista, strategista relevanssia, johtamista, sisältöä ja vaikutuksia.

Arviointitutkimushankkeen suunnitelman laati ja arvioinnista vastasi TKK:n systeemanalyysin laboratorio. Arviointi aloitettiin toimialakartoituksella, jonka jälkeen tehtiin 50 haastattelua ja kymmen-sivuinen kysely kaikille projekteille. Kyselyn vastausprosentiksi tuli 57 %, joka oli tuolloin kohtuullinen, mutta tänä päivänä sitä voisi pitää erinomaisena. Suurimmat projektit (28) valittiin peer review-tarkastelun kohteeksi. Yhteensä kymmenen yhdysvaltalaisasiantuntijan voimin projektit reitattiin ja niistä laadittiin projektikohtainen lausunto. Asiantuntijat perusivat arvionsa kirjalliseen materiaaliin, erityistä vierailua ei organisoitu. Lisäksi käytettiin japanilaista teollisuusasiantuntijoista koostunutta neuvottelukuntaa, joka antoi oman lausuntonsa ohjelmista. Pääosin arviointi oli positiivinen, mutta myös kehittämiskohteita löydettiin. Perustutkimuksen osuus jäi suhteellisen vaatimattomaksi ja sitä toivottiin lisättäväksi. Laajat ohjelmat oli hallinnol-

lisista syistä jaettu aihealuerhmiksi, mutta niiden toiminta ei ollut tyydyttävää. (ks. Salo et al 2000)

Loppuarvioinnissa 2001 arviointikysymykset koskivat ohjelmien strategiaa ja projektisalkkua, vaikutusta alalla, ohjelman hallinnointia ja yhteistyötä ohjelmien kesken. Haluttiin välttää liiallista päällekkäisyyttä väliarvioinnin kanssa, koska se oli toteutettu vain kaksi vuotta aiemmin, mutta osittain kysymykset olivat toki samoja, kun haluttiin nähdä, miten väliarvioinnin suositukset ovat toteutuneet. Arviointi tilattiin Technopolikselta, joka oli tehnyt useita arviointeja Tekesille 90-luvun lopulla, muun muassa laajan energiatutkimusohjelmien arvioinnin.

Arviointi koostui haastatteluista, joita tehtiin eri tahojen kanssa yhteensä 62, verkostanalyysista, itsearviointikyselystä, peer reviewsta ja vertailusta muissa maissa toteutettuihin ohjelmiin. Peer review tehtiin neljän asiantuntijan voimin, jotka saivat ennakolta runsaasti kirjallista materiaalia ja joille projektit esiteltiin Suomessa. Kokeellisena uudistuksena olivat projektipäälliköiltä pyydyttävät itsearviointiraportit. Itsearviointiraporttien avulla oli tarkoitus tehostaa triangulaatiota ja niiden suunnitteluun ja hankkimiseen käytettiin paljon aikaa. Tässä ei kuitenkaan onnistuttu suunnitellusti. Lopulta raportit saatiin vain pienestä osaa hankkeita, ja niiden taso oli liian kirjava, jotta asiantuntijat olisivat voineet niitä käyttää. Itsearviointiraporttien käytöstä tässä tarkoituksessa luovuttiinkin tämän kokelun jälkeen.

Arvioinnin mukaan ohjelmia pidettiin pääosin onnistuneena. Fokus olisi voinut olla pidemmän aikavälin kysymyksissä, mutta näinkin tieteellistä ja liiketoiminnallista vaikuttavuutta pidettiin hyvänä. Tekesin ja Akatemian yhteistyö sujui hyvin. Kehityskohteiksi nostettiin aihealuerhmän toiminnan kehittäminen, rahoituspäästösten suurempi avoimuus ja kansainvälistymisen tukeminen, jotta EU:n ohjelmista saadaan parempi hyöty. (ks. Arnold et al 2002)

Tämä arviointi otettiin myös valtionalouden tarkastusviraston tarkastuksen kohteeksi case-esimerkinä. Tarkastuskertomus suhtautuu varsin kriittisesti arviointiin ja erityisesti sen hyödyntämiseen. (ks. VTV 2008)

ETX-ohjelmaan palattiin vielä jälkiarvioinnin merkeissä 2006. Siinä arvioitiin päättymässä ollut Elektroniikan miniatyrisointi -ohjelma (ELMO) ja tehtiin jälkiarviointi paitsi ETX-ohjelmasta, niin myös sitä edeltäneistä ohjelmista alkaen vuodesta 1982. Nämä olivat Puolijohdeteknologia, Anturitekniikka,

Mikroelektroniikka ja Elektroniikan suunnittelu- ja valmistustekniikat. Tämä arviointi tilattiin kilpailutettuna arviointihankintana. Tilaajan näkökulmasta pyrittiin kartoittamaan pitkää kehityskaarta ja Tekesin roolia siinä. (ks. Raivio et al, 2006)

### 3.4 Arviointihankinnat 2000-luvulla

Vuoden 2006 paikkeilla tapahtui eräitä muutoksia, joilla oli vaikutusta Tekesin arviointitoiminnan kehittämiseen. Tekesin tuorein iso organisaatiomuutos tehtiin vuoden 2006 alussa. Tällöin Tekesin ydinprosessien (Asiakkuudet, Ohjelmat, Rahoitus) asemaa vahvistettiin. Ohjelma-arvioinnit kytkettiin entistä tiiviimmin ohjelmanprosessiin siten, että ohjelmanprosessi voitiin nähdä arvioinnin sisäisenä asiakkaana. Tällä oli vaikutusta arvioinnin suunnitteluun, kun yksi selkeä kysyntä- ja hyödyntäjätaho voitiin tunnistaa.

Toinen arviointien tilaamiseen vaikuttava muutos koski julkisten hankintojen lainsäädännön kehitystä. Se painottaa työnjaon eriytymistä ja tilaamisosaamisen tarvetta. Hankintaohjeistus ohjaa asiantuntijapalvelun ostoon eikä yksittäisten asiantuntijoiden käsin poimintaan. Toimintamallissa laaditaan tarjouspyyntö ja sen osana palvelukuvaus, jossa kuvataan se, mitä hankinnalla tavoitellaan. Valintakriteerien tulee olla läpinäkyviä ja tarjousvertailu tehdään vain tarjouspyynnössä ilmoitettujen kriteerien perusteella. Tekesin käyttämät pääkriteerit ovat hinta, arviointiryhmän osaaminen ja arviointisuunnitelma.

Kolmas trendinomainen muutos liittyy tarjonnan kehittymiseen. Suomeen kehittyi useita ammattimaista arviointipalvelua tarjoavia organisaatioita. Arviointiosaamista oli riittävästi, jotta hankintamenettelyn edellyttämä aito kilpailuttaminen tuli mahdolliseksi.

Tekesissä arviointitoimintaa alettiin kehittää osallistavaan ja toiminnan strategisen kehittämisen suuntaan. Seuraavassa esimerkissä kuvataan Tekesin arviointia Tekesille uudella alueella, terveydenhuollon innovaatiotoiminnassa.

#### 3.4.1. Case-esimerkki Tekes terveydenhuollon sektorilla

Tekes lähti vaiheittain mukaan sosiaali- ja terveydenhuollon innovaatiotoimintaan. Ensimmäinen teknologiaohjelma alalle oli Terveydenhuollon digitaalinen media (1996–99). Seuraavaksi aloitettiin

iWell-ohjelma (Hyvinvointi ja terveys, 2000–2003), joka kohdistui yksilön kotona, vapaa-aikana, työssä tai matkalla tarvitsemiin ja käyttämiin terveyden- ja hyvinvoinnin teknologia- ja palvelutuotteisiin. Erittäin painopisteenä oli uuden liiketoiminnan luominen hyvinvointia lisäävien tuotteiden ja palvelujen kehittämisen yhteydessä. Välillisenä tavoitteena oli panostaa ikääntyneiden hyvinvointiin ja ehkäistä täten varhaisen laitoshoidon tarvetta, jonka kustannukset sosiaali- ja terveyspalvelujärjestelmälle ovat huomattavia. iWellin laajuus oli 30 miljoonaa €.

iWell-ohjelman arviointi päätettiin toteuttaa yhdessä kolmen muun ohjelman kanssa temaattisena arviointina. Tällä pyrittiin tutkimukselliseen otteeseen ja ilmiön tunnistamiseen. Ohjelmia yhdistävänä tekijänä oli rakenteiden uudistaminen, jota tarkasteltiin kiinteistösektorilla (Rembrand-ohjelma), Puuenergia-ohjelmassa ja teollisuuden palvelutoiminnassa (UTT-ohjelma) sekä iWell-ohjelmassa. Arviointi painottui tiedon tuottamiseen ohjelmien roolista rakenteiden muutoksessa. Arviointi hankittiin vuorovaikutteisen suunnittelun jälkeen Gaia Consultingilta.

Rakenteiden uudistamisen arvioinnissa todettiin, että ohjelmatoiminnan laajeneminen on nähtävissä ohjelmille asetetuissa tavoitteissa ja Tekesin strategisen toiminnan tasolla. Arvioituissa ohjelmissa muutos ei ole vielä kuitenkaan täysin näkyvissä kaikessa ohjelmien toiminnassa tai muutos ei ole ollut selkeää Tekesin asiakkaille. iWell-ohjelman yksi tärkeä saavutus oli Tekesin toiminnan laajentaminen selkeämmin hyvinvointi- ja palvelukentälle. (ks. Hjelt et al. 2004)

Tekesin toiminta-ajatusta ja strategiaa muutettiin merkittävästi vuonna 2005. Tekesin toiminta laajeni teknologiasta innovaatiotoimintaan ja asiakaskunta valmistavasta teollisuudesta palvelualoille. Uusissa ohjelmissa muutos näkyi siten, että Tekes voi tukea uudentyyppisiä projekteja liiketoiminnan ja palvelujen kehittämiseksi. Kun terveydenhuollon sektorille valmisteltiin uutta ohjelmaa, niin Tekesin muutokset pyrittiin ottamaan huomioon.

FinnWell-ohjelman (2004–2009) tavoitteena oli parantaa kansallisen terveydenhuollon laatua ja tuottavuutta sekä edistää alan yritystoimintaa ja kansainvälistymistä. Ohjelman perusajatus oli se, että teknologian avulla voidaan parantaa terveyspalvelujen laatua ja tuottavuutta vain jos samanaikaisesti kehitetään uusia toimintatapoja. Ohjelman visiona oli, että FinnWell -ohjelmaan osallistuvat yritykset, palveluntarjoajat sekä tutkimuslaitokset

loisivat terveydenhoitoalan tuotteita, palveluja ja toimintatapoja, jotka olisivat aiempaa tehokkaampia, avoimempia ja erityisesti asiakaslähtöisempiä. FinnWell oli laajuudeltaan 177 miljoonaa €, siis noin viisi kertaa edeltäjänsä suurempi.

FinnWell -ohjelman väliarviointi vuonna 2007 oli merkittävä ponnistus ja laajuudeltaan suurempi kuin yksittäisen ohjelman loppuarviointi. Arviointi toteutettiin kolmessa osassa. Ensin ohjelmaan osallistuvilta kerättiin palautetietoa ohjelmatoiminnasta. Toinen arviointiosa kohdistettiin erityisesti julkisten toimijoiden innovaatiohankkeisiin ja niiden kokonaistaloudellisen merkityksen arviointiin. Kolmannessa arviointiosassa tarkasteltiin terveydenhuollon innovaatiostategiaa ja Tekesiä ja FinnWell-teknologiaohjelmaa suhteessa siihen. Tämä poliittikatason arviointi kohdistui terveys-, elinkeino- ja innovaatiopolitiikan rajapinnoille.

Arviointi toteutettiin tarjouskilpailun pohjalta hankintaohjeistuksen mukaisesti ja tilattiin osina laajalta konsortiolta, johon kuuluivat VTT, Lillrank Consulting, Gaia Consulting, Net Effect ja PricewaterhouseCoopers. Tekesin puolelta väliarvioinnin ohjausryhmään saatiin poikkeuksellisen voimakas sitoutuminen, muun muassa Tekesin pääjohtaja Saarnivaara oli aktiivisesti mukana. Tämä osoitti arvioinnin saamaa painoarvoa, jolla tuli myöhemmin olemaan vaikutusta suositusten toteuttamisessa. Väliarvioinnissa hyödynnettiin laajaa dokumenttiaineistoa ja selvityksiä, tehtiin kysely (saanto 31 %), 17 haastattelua ja pidettiin työpaja.

Arviointi julkaistiin kahtena eri raporttina. Raporteissa todetaan, että FinnWell-ohjelman toimeenpanon kautta on vahvistettu teknologiaymmärtämystä ja ymmärtämystä siitä, että terveydenhuollon uudistamisessa teknologia ja innovaatiot ovat yksi keskeinen moottori ja väline. Kuitenkin FinnWellin ja Tekesin toiminnasta riippumaton tekijä on ainakin se, että terveys-, elinkeino- ja innovaatiopolitiikat ovat logiikaltaan erilaisia. Merkittävää tehokkuushyötyä terveydenhuollossa saadaan aikaan vasta sitten, kun palvelu/hoitoprosesseja ja toimintatapoja uudistetaan ja vanhoista luovutaan. Suosituksissa korostetaan strategis-operatiivista kumppanuutta ja kumppanuusverkoston rakentamista. (Tekes 2007, Kuitunen & Haila 2007)

FinnWell-ohjelman väliarvioinnilla voi katsoa oleen merkittävä vaikutus. FinnWellin rakennetta muutettiin siten, että rahoitus kohdistui hyvinvointia tukeviin hankkeisiin. Ohjelmassa ei enää panostettu alan rakenteellisiin seikkoihin. FinnWellin rinnalle ja jatkoksi alettiin valmistella uutta ohjelmaa,

joka sittemmin sai nimen Innovaatiot sosiaali- ja terveyspalvelujärjestelmässä. Sitä rakennettiin luonteeltaan voimakkaammin verkosto-ohjelman luonteiseksi.

Innovaatiot sosiaali- ja terveyspalvelujärjestelmässä -ohjelma (SOTE, 2008–2015) on kumppanuusohjelma, jossa Tekes toteuttaa hallitusohjelman sekä sosiaali- ja terveysministeriön linjauksia ja tavoitteita sosiaali- ja terveyspalvelujärjestelmän kehittämiseksi. Ohjelman tavoitteena on uudistaa julkisen sektorin järjestämisvastuulla olevaa sosiaali- ja terveyspalvelutuotantoa, edistää palvelujen saatavuutta, laatua ja vaikuttavuutta sekä edistää yritysten liiketoimintamahdollisuuksia alalla. Tavoitteisiin pyritään pääsemään kehittämällä palvelujärjestelmää uudistavia, asiakaslähtöisiä yhteistyö- ja toimintamalleja. Uudistaminen tapahtuu julkisen, yksityisen ja kolmannen sektorin yhteistyönä.

Kun FinnWell oli päättymässä ja sen loppuarviointia alettiin suunnitella, oli väliarviointi syytä ottaa painokkaasti huomioon. Tekesin asemoitumista ja strategiaa koskevat seikat oli tyhjentävästi käyty läpi väliarvioinnin yhteydessä, ja niihin tiedettiin jo reagoidun, osin myös uuden SOTE-ohjelman toimin. Näin loppuarvioinnissa suuntauduttiin katsomaan tuloksia. Oman mielenkiintonsa kohteeksi tuli iWell-ohjelma, joka toimi erilaisessa ympäristössä ja eri pelisäännöin kuin FinnWell.

Verrattuna aiempiin arviointihankintoihin entistä enemmän painoa asetettiin johtopäätöksille ja suosituksille. Arvioinnin tarjouspyyntö rakennettiin niiden ympärille, millä pyrittiin korostamaan tulosten hyödyntämisen näkökulmaa. Rakenne tehtiin samankaltaiseksi kuin muissa vuoden 2009 aikana toteutettavissa arvioinneissa. Tähän syynä oli tarve löytää parempaa vertailukelpoisuutta arviointien kesken, mikä helposti jää toteutumatta, jos arvioinnin tavoitteet muuttuvat jatkuvasti.

Johtopäätöksiä ja suosituksia koskeva osuus jaettiin kolmeen osaan:

- Alan tutkimus-, kehitys-, ja innovaatiotoiminnalle: Miten alan tutkimus-, kehitys- ja innovaatiotoimintaa voitaisiin kehittää jatkossa? Mikä on tulevaisuudessa huippuosaamisen keskittymien (SHOK) ja osaamiskeskusten (OSKE) rooli? Mitkä kansainväliset ohjelmat ja aloitteet ovat alan kannalta merkityksellisimpiä ja kuinka niihin voitaisiin enemmän vaikuttaa?
- Tekesille: Mitä konkreettisia ja toimivia hyviä käytäntöjä on tunnistettavissa ohjelmajärjestelmien ja ohjelmatoiminnan kehittämiseen? Mitä suo-

situksia on Tekesin rahoitusinstrumenttien, sisältölinjausten ja strategian kehittämiseen?

- Innovaatiopolitiikalle: Miten kansallisen innovaatiostrategian suositukset tulisi huomioida? Mitä muita toimenpiteitä innovaatiopolitiikassa on tarpeen vaikuttavuuden parantamiseksi Tekesin toimien ohella?

Tämän lisäksi oli vielä tarkentavia kysymyksiä ohjelmien tuloksista ja odotettavissa olevista vaikutuksista.

Arviointihankkeesta järjestettiin tarjouskilpailu, johon tehdyistä tarjouksista valittiin tekijäksi PricewaterhouseCoopers. Arviointi rakentui tuttuun tapaan dokumenttianalyysin, haastattelujen, kyselyn ja työpajan pohjalle. Mainittakoon, että vastausprosentti kyselyssä on vain 24 %, mikä on kuitenkin tämän päivän mittapuulla kohtuullinen.

Raportissa todetaan, että FinnWell-ohjelma on arvioinnin perusteella ollut onnistunut strategisen tason interventio ja piristysruiske hyvinvoinnin ja terveydenhuollon toimijakenttään. Ohjelman tuloksista ja vaikutuksesta ei ole kuitenkaan vielä tunnistettavissa kansallisia läpimurtoja, joiden perusteella voitaisiin todentaa asiakaslähtöisyyden lisääntyneen merkittävästi terveydenhuollon järjestelmätasolla. FinnWell-ohjelman väliarviointi ja sen pohjalta tehdyt isot strategiset linjavedot ovat osoitus oikeanlaisesta ohjelmajohtamisesta sekä tähän liittyvästä strategisesta ketteryydestä. Ohjelmatoiminta ja sen painopisteet ovat muuttuneet iWell-ohjelman laitekeskeisyydestä FinnWell-ohjelman prosessinäkökulmaan ja siitä kohti SOTE-ohjelman systeemitasoa. (ks. Saarinen et al. 2009)

### 3.4.2 Arvioinnin kehittämistä innovaatiotutkimuksena

Arviointihankinnat ovat tuoneet mukanaan julkisten hankintojen kilpailutuksen edellyttämät piirteet arvioinnin suoritukseen. Tämä tarkoittaa muun muassa täsmällisiä suunnitelmia, arviointikokemuksen ja osaamisen tarkkaa kuvausta ja tarjousasiakirjojen julkisuutta.

Arvioinnin ostamiseen hankintoina suhtaudutaan positiivisesti. Valtiontalouden tarkastusviraston raportissa todetaan, että arviointihankkeiden kilpailuttamisella on ollut positiivisia vaikutuksia sekä arviointitoiminnan avoimuuteen ja läpinäkyvyyteen että arviointiosaamisen paranemiseen. Arviointien tilaamisen ja toteutuksen asiantuntemus näyttää tapausanalyysien valossa hyvältä. Arviointien

toteutuksen (tiedonhankinta ja analyysi) ja raportoinnin laadussa ei tapaustarkastelujen perusteella myöskään ole tiedon hyödynnettävyyttä heikentäviä merkittäviä ongelmia. (VTV 2008)

Toisessa tutkimuksessa todetaan kilpailutuksesta seuraavasti: Se, mitä kilpailutus varmistaa, on laatua markkinaehtoisessa mielessä joko siten, että tietty arviointien vakioalaatu saavutetaan minimikustannuksin tai siten, että tietyillä kustannuksilla on mahdollista tilata mahdollisimman laadukas arviointi. Kilpailutuksen laadunvarmistus riippuu keskeisesti tilaajatahon tilaamisosaamisesta ja sen lisäksi sen arviointiosaamisesta sisältökysymyksiä myöten. (Ahonen et al. 2009)

Kun arviointia tehdään toimeksiantona pitkälle suunniteltujen tavoitteiden mukaan ja määrämutoiseen suuntaan, niiden kokeilevuus ja tutkimuksellisuus vähenee. Tästä haittapuolena on se, että liikkumavaraa tutkimusartikkeleille ei arviointihankkeen puitteissa jää. Aiemmin toteutetut arviointitutkimukset olivat varsin tuotteliaita julkaisujen suhteen.

Myös Ahonen et al. kiinnittää asiaan huomioita. Heidän mukaansa olisikin suotavaa, että arvioijat käyttävät arvioinneissa kertynyttä aineistoa laatiakseen tieteellistä tutkimusta, jonka tulokset julkaistaan laadultaan varmennetuilla areenoilla eli refereemenettelyä käyttävissä tieteellisissä aikakausjulkaisuissa. (Ahonen et al. 2009). Arviointihankintojen rinnalle tarvittaisiin siten myös tutkimuksellista otetta arviointien kehittämiseen.

Tekesissä tämä on nähty innovaatiotutkimuksen ja arvioinnin välisenä vuorovaikutuksena. Tekes on rahoittanut teknologian ja innovaatioiden tutkimusta 1990-luvun alusta lähtien. Teknologian tutkimuksen ohjelma toteutettiin 1997–2001 ja sitä seurasi 2001–2005 toteutettu ProACT-ohjelma. Näiden avulla tutkittiin monitieteisesti teknologia- ja innovaatioprosesseja ja -politiikkaa. (Pesonen 2006)

Vuonna 2008 käynnistettiin haku hankkeille, jotka kohdistuvat innovaatiotoiminnan ilmiöiden tunnistamiseen. Kuten edellä kuvattiin, Suomessa on vähitellen siirrytty teknologiapolitiikasta laajalaiseen innovaatiopolitiikkaan, mikä on heijastunut myös Tekesin ohjelmiin. Yhdeksi haun osaluokaksi valittiin arvioinnin ja ennakoinnin kehittäminen. Sen tavoitteeksi asetettiin kehittää arviointitoimintaa ja ennakoitotoimintaa sekä teoreettisesti, menetelmällisesti että osana strategiaprosesseja. Laajentuvaa innovaatiokäsitetä ja innovaatiopolitiikan horisontaalisuutta haluttiin tutkia ar-

viointi- ja ennakoitotoiminnan kannalta ja huomioida innovaatiotoiminnan muutokset arviointi- ja ennakoitimenetelmien kehittämisessä.

Haku tuotti muun muassa kaksi hanketta, joissa tutkitaan strategisen huippuosaamisen keskitymien (SHOK) ohjausta sekä yhden ohjelmien arviointitutkimuksen: Kehittävä vaikuttavuusarviointi (Forias), joka on yksityiskohtaisemmin kuvattu tämän lehden toisessa artikkelissa. Foriaksessa on erityisesti kyse siitä, että tarvitaan hyviä eväitä uusien laajojen kumppanuusohjelman toteutuksessa, seurannassa ja arvioinnissa. Kohteina ovat yllä mainittu SOTE-ohjelma ja Rakennetun ympäristön ohjelma. SHOK-ohjauksen tutkimushankkeilla on taas yhteys arviointitoiminnan tulevaan haasteeseen eli miten keskittyviä tullaan jatkossa arvioimaan.

### 3.5. Meta-arviointia arviointiprosessista ja -toiminnasta

Arviointitoiminnan kehittämistä varten on tarpeen saada palautetta eri suunnista. Eräs palautemekanismi voidaan toteuttaa Vedungin esittelemällä tavalla meta-arvioinnin muodossa. Seuraavaan on koottu eräitä tutkimus- ja innovaatiotoiminnan arviointiin liittyviä meta-arviointeja.

Tekes, Akatemia ja muut tutkimus- ja kehitystoiminnan arvioinnissa mukana olleet tahot olivat valtiontalouden tarkastusviraston tarkastuksen kohteena 2007. Tarkastuksen näkökulma oli, millainen mahdollisuus päätöksentekijöillä on olemassa olevien arviointikäytäntöjen pohjalta vastuullistaa toimijoita näille asetettujen tavoitteiden toteutumisesta. (VTV 2008)

VTV:n mukaan julkisen hallinnon uudistaminen on voimakkaasti mukana arvioinnissa. Tilivelvollisuuden korostaminen on ajavana voimana, jolla lisätään tietoa siitä, onko rahat käytetty parhaalla mahdollisella tavalla. Tutkimus- ja kehitystyössä tiedon tuottaminen on tärkeää, koska lopputulokseen liittyy epävarmuutta, ja syyt ymmärtää ilmiötä paremmin, mikä onnistuu ja mistä syystä. Vastuullistavaa arviointia vierastetaan, koska vaikutukset ovat välillisiä ja toteutuvat pitkällä aikajänteellä ja huomio kiinnittyy virheiden välttämiseen. Arviointiraportit eivät yleensä ole kriittisiä, mikä heijastanee tutkimus- ja kehitystyön ymmärrystä. Kyse on riskipitoisesta toiminnasta, jossa tuloksia ei aina etukäteen voi taata. Toimialalla tyypillisen ajattelutavan mukaan arviointiin on suunnattava pikemmin-

kin uusia vaatimuksia yhteiskunnan kehityksen ja tulevan toiminnan ennakoimisesta ja kehittämistä kuin vastuita tähänastisesta toiminnasta. Tämä ilmenee myös vastuullistavan arviointitiedon kysynnän vähäisyytenä.

Raportti toteaa, että ohjelma-arviointien roolina ei niinkään pidetä ohjelmien vaikuttavuuden varmistamista kuin uusien näkökulmien ja kehittämisideoiden tuottamista ohjelmatoimintaan. Arviointituloksille on tietoisesti etsitty asiakasta ja sellainen on löytynyt erityisesti ohjelmatoiminnan lisäarvo- palvelujen kehittäjistä ja ohjelman osallistuneista. Nykyinen toimintamalli tuo mukanaan sen, että ohjelman toteutukseen osallistuvat arviointiin. Tämän katsotaan vähentävän riippumattomuutta ja vähentävän vastuullistamista. Positiivista palautetta tuli arviointitoiminnan kehittämisestä, monipuolistamisesta ja systematisoinnista, minkä todettiin saaneen myös kansainvälistä tunnustusta. Tekesin vaikuttavuusarviointi tekee jatkuvasti konseptien kehittämistyötä, mikä toisaalta voi parantaa arviointitoiminnan relevanssia joissakin valmisteluprosesseissa (kuten ennakointi), mutta samalla vähentää jatkuvuutta näkökulmien ja mallien käytössä ja tiedon kumuloitumista. Arviointikysymykset ja arviointiprosessi suunnitellaan niin, että se tuottaisi tietoja, ideoita ja ehdotuksia ohjelman toteuttajien käyttöön ohjelman päättyessä ja mietittäessä vielä mahdollisia jatkotoimia. Toisaalta ohjelmien hallintaa ja toimintatapoja koskevat ehdotukset voidaan hyödyntää myös tulevia ohjelmajärjestelmien suunniteltaessa. Lähtökohtana on kehittävä ja osallistuva arviointiprosessi. (VTV 2008)

Ahonen, Tala ja Hämäläinen ovat selvittäneet politiikkatoimien vaikuttavuusarvioinnin menetelmiä ja toimintatapoja tuoreessa tutkimuksessaan. Tässä tutkimuksessa on laajempi näkökulma, joka huomioi myös muun arvioinnin kuin vastuullistamisen.

Tässä raportissa todetaan Tekesistä seuraavaa:

- Tekes itsessään on ollut useiden, mukaan lukien kansainvälisten arviointien kohteena
- Tekes on itse kehittämistarkoituksessa arvioituttanut omat käytäntönsä ja niiden vaikuttavuuden
- Tekes on maamme tutkimuksen ja teknologisen kehittämisen politiikan EU-lähtöisten ja muiden kansainvälisten kytkentöjen johdosta monin tavoin kansainvälisesti kytkeytynyt myös arvioinneissaan
- Tekesin toteuttama sen harjoittamaa toiminta-

taa yleisesti tai kohdennetummin pohjustava arviointitoiminta on kaiken kaikkiaan erittäin laajaa ja erityyppistä

- Työ- ja elinkeinoministeriön vastuualueella teknologiapolitiikan politiikkatoimien vaikuttavuusarviointi on ilmeisen vankka Tekesin toiminnan ja osin myös VTT:llä olevan kapasiteetin ansiosta
- Tekes tulee kysymykseen vertailuarvioinnin perusyksikkönä eli benchmarkina ajatellen tutkimuksen tyyppisen toiminnan politiikkatoimien vaikuttavuusarvioinnin kehittämistä muuallakin kuin Tekesin omalla tai sen omimmilla vastuualueilla. (Ahonen et al. 2009)

Valovirta ja Hjelt ovat tarkastelleet ja analysoineet laajasti innovaatio- ja teknologiapolitiikan ennakointi- ja arviointitoimintaa. Heidän mukaansa suuri osa arviointityöstä ja konkreettisesta tiedon hyödyntämisestä tapahtuu Tekesin piirissä. Tekes tilaa ja organisoii Tekesin toimenpiteitä (mm. teknologia-ohjelmia) koskevia arvioiteja suhteellisen vakiintuneen toimintamallin mukaisesti. Tekes myös teettää, tekee ja kokoaa yhteen tietoa julkisen t&k-rahoituksen vaikuttavuustutkimuksista, millä pyritään suoranaisesti hyödyttämään teknologiapolitiikan strategista kehittämistä. Erityisesti teknologiapolitiikan toimenpiteiden painottuminen ohjelmallisiin rakenteisiin tarjoaa otollisen maaperän täsmälliselle arviointityöskentelylle. Ohjelmien ja niiden osana toimivien hankkeiden arviointi on arviointitoiminnalle ja -menetelmille toimiva ympäristö, missä toimenpiteiden vaikutuksiin päästään pureutumaan kohtuullisen tarkasti. Innovaatio- ja teknologiapolitiikan arvioiteja tehdään hyvin monipuolisesti eri toiminnan tasoilla:

- Policy-arviointi / strateginen arviointi, joka voi kohdistua esimerkiksi organisaatioihin tai koko innovaatiojärjestelmään, kansallisiin strategioihin tai hallinnonalan laajoihin ohjelmiin ja linjauksiin
- Ohjelma- ja toimenpidetason arviointi, jossa tarkastellaan yksittäistä ohjelmaa tai toimenpidetason kokonaisuutta
- Hanketason arviointi, jossa analysoidaan yksittäisen, tarkasti kohdennetun hankkeen tai yksittäisen toimenpiteen toimivuutta ja vaikutuksia
- Vaikuttavuusarviointi, joka kohdistuu kaikkien edellä mainittujen tasojen osalta täsmällisten vaikuttavuustavoitteiden saavuttamisen analysointiin. Vaikuttavuusarvioinnilla viitataan tässä niin tulosten, tuotosten, vaikutusten kuin vaikuttavuuden analysointiin. (Valovirta & Hjelt 2005)

### 3.6. Päätelmiä arviointitoiminnan ja tilaajaosaamisen kehityksestä Tekesissä

Edellä esitetyt kolme julkaisua kuvaavat hyvin Tekesin arviointitoiminnan asemoitumista. Tekesissä arviointia on rakennettu monella tasolla. Ohjelma- arviointien rinnalla on tehty hanketason ja vaikuttavuuden arviointia sekä osallistuttu tiiviisti innovaatiopolitiikan arviointeihin. Vaikuttavuusarviointiyksikön toimintakenttä on varsin laaja, sillä se sisältää myös innovaatiotutkimuksen, ennakoinnin ja strategisen tiedontuotannon. Toimintaa on voitu kehittää suhteellisen itsenäisesti, erityistä normiohjausta tai lainsäädäntöön pohjautuvaa sääntelyä ei ole ollut. Toisaalta ohjelmien arviointitoiminta on rajattu Tekesin rahoittamiin ohjelmiin, eikä sille ole annettu laajempaa mandaattia ns. konsernitasolla tai innovaatiopolitiikassa.

Ohjelma- arviointien kehitys on asemoitavissa taulukon 1 viitekehukseen. 1980-luvun lopulla ja 1990-luvun alussa ohjelma- arviointi oli konsultointia ohjelmajohtajille tai osa ohjelmaprosessia ja ohjelman palautemekanismia. Ohjelmat itse olivat vastuussa arvioijien kutsumisesta ja arvioinnin toteutuksesta. Pääpaino oli ohjelman sisäisessä operatiivisessa kehittämisessä.

Seuraavassa vaiheessa, noin vuodesta 1994 vuoteen 1999 painotus oli tiilivelvollisuudessa. Arviointien pääsisältönä oli kehitetyn teknologian laatu ja taso. Toteutustapana oli vertaisarviointi, jolloin toimialan tai teollisuuden ja tutkimuksen edustajat suorittivat arvioinnin. Arviointien tilaajana toimivat edelleen Tekesin ohjelmat, joten tiilivelvollisuusnäkökulma ei ollut puhdasoppinen.

Vuoden 2000 alussa Tekesiin perustettiin arviointiyksikkö. Arvioinnille nimettiin oma vastuuhenkilönsä, joka suunnittelee ja ohjaa kaikki ohjelma- arvioinnit toimialasta riippumatta. Hän toimi tapauksesta riippuen joko tilaajana tai tutkimusaloitteen tekijänä, jolloin toteutussuunnitelmasta vastasi tutkimusryhmä. Yksikön laajan toiminta-alueen myötä arviointiin tuli tutkimuksellista näkökulmaa ja arviointia pyrittiin kytkemään vaikuttavuusmalleihin ja ohjelmateorioihin. Arvioitavina asioina alkoivat korostua ohjelmatoiminnan logiikka ja vaikuttavuus ja huomiota painotettiin tiedon tuottamiseen. Painotus oli ilmiöiden tunnistamisessa ja oppimisprosessissa.

Vuodesta 2006 alkaen Tekesin organisaatiomuutokset ja julkisten hankintojen kehitys painottivat työnjaon eriytymistä. Tekesissä on suunnattu arviointitoimintaa entistä enemmän osallistuvaan

ja ohjelmatoiminnan ja -salkun strategisen kehittämisen suuntaan. Arvioinneille on haettu sekä sisäisiä että ulkoisia käyttäjiä. On pidetty tärkeänä saada tiedon käyttäjät mukaan seuraamaan ja osallistumaan tulosten käyttöönottoon. Arvioinnin tilaajana toimii edelleen arvioinnin vastuuhenkilö. Lisäksi rahoitetaan rinnalla joitakin arvioinnin tutkimusprojekteja, joiden aloite on lähtenyt tutkijoilta.

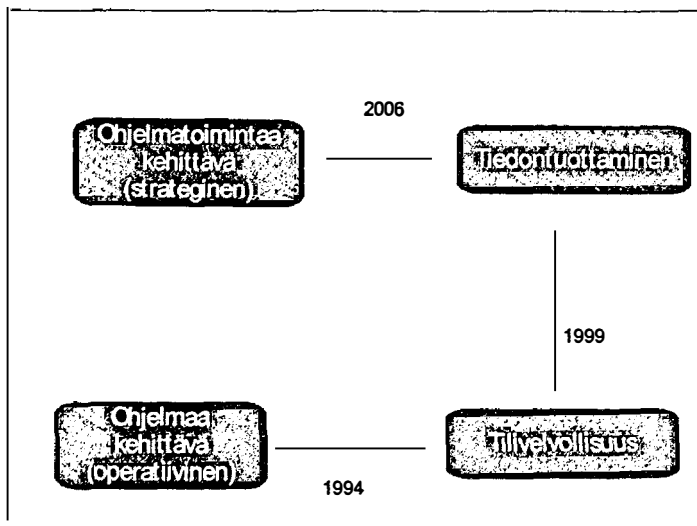
Tilaaminen on Tekesissä erityisesti hyödyntämisestä, mutta kehittävän arvioinnin luonteen mukaisesti sen huomiointi on oleellinen osa arviointihankintojen suunnittelua. Vaikka hyödyntäminen on rajattu tämän artikkelin ulkopuolelle, voi todeta, että arviointitiedon hyödyntämisestä on kasvavassa määrin pyritty edistämään. Arviointiraportit julkaistaan, ne esitellään erilaisissa seminaareissa ja kaikista kiteytetään keskeiset viestit sekä Tekesille ja ohjelmoproessille että laajemmin hyödynnettäväksi. Haasteellista kuitenkin on, miten yksittäisistä arviointihankkeista saadaan syntetisoitua tietoa instrumenttien toimivuudesta ja valintojen onnistumisesta. Ohjelma-arviointi on vain yksi tietolähde monien muiden joukossa, joten sen hyödyntämiseen liittyy usein ehkä liiankin suuria odotuksia.

Arviointien eri sukupolvet ovat edellyttäneet osittain erilaista tilaajaosaamista. Ensimmäinen ja toinen sukupolvi toteutettiin ohjelmien toimesta ja tilaamisosaaminen edellytti alan substanssiasiantuntijoiden tunnistamista. Arvioinnit hankittiin ohjelmien koordinoitubudjetista, niiden suunnittelu oli epämuodollisempaa ja siinä oli paljon liikkuma-

vara. Kolmannen sukupolven arvioinneissa oli vahva kytkentä innovaatiotutkimusosaajiin. Hankkeissa voitiin tehdä kokeiluja ja toteuttaa niitä tutkimuksellisempaan. Tilaajaosaaminen oli intuitiivista, ideointia ja inspirointia. Huonojakin puolia toki oli, ohjaus ja ohjelmassa mukana olleiden osallistuminen jäi yleensä vähäiseksi ja tutkimushankkeet saattoivat venyä ja lopputulos saattoi poiketa paljon ennalta odotetusta.

Tilaajaosaaminen neljännen sukupolven hankinnoissa on toisenlaista: pitää olla tarkkuutta, tasa-puolisuutta ja tehokkuutta. Hankinnan suunnittelussa on pyrittävä mahdollisimman pitkälle määrittelemään ne kysymykset, joihin halutaan vastaus. Itse asiassa arvioinnin tilaajana joskus tuntuu siltä, että työ annetaan pois, kun mielenkiintoinen osuus alkaisi. Roolien eriytyminen tarkoittaa sitä, että toteuttaja on vastuussa osaamisestaan ja menetelmien ja työkalujen kehittämisestä. Tilaaja voi tukea tätä suunnittelemalla haasteellisia arviointeja. Hankinnoissa valinta kohdistuu kokonaistaloudellisesti edullisempaan tarjoukseen, jossa huomioidaan hinnan lisäksi arviointiryhmän osaaminen ja suunnitelman laatu.

Mielenkiintoinen seikka on se, että Tekesin ohjelma-arviointitoiminnan murroskohdat osuvat samoihin aikoihin, kuin Tekesin ja ohjelmatoiminnan eli vuosiin 1999 ja 2006 (ks. Valtakari et al 2007). Näinä vuosina arviointitoiminnan painotus muuttui ensin tilivelvollisuudesta tiedon tuottamiseen ja sitten strategiseen kehittämiseen (kuva 1).



Kuva 1. Tekesin ohjelma-arviointitoiminnan sukupolvet



Tämän tarkastelun perusteella näyttää siltä, että muutokset toimintaympäristössä ovat aikaansaaneet arviointitoiminnan kriittisen uudelleentarkastelun tarpeen. Arviointitoiminta on ollut reaktiivista, ja murroskohdat ovat ratkaisevalla tavalla muokanneet sitä, vaikkakin muutoksen keskellä tätä on ollut vaikea huomata.

On vaikea sanoa, mikä on oikea tasapaino eri arviointimuotojen kesken. Kaikilla on omat vahvuutensa ja heikkoutensa ja niiden välillä on selvätkin vahvoja jännitteitä. Ehkä tilaajan roolin parempi ymmärtäminen auttaa hahmottamaan eroja. Ohjelman kehittäminen on luonteeltaan ohjelman itsensä tehtävää. Ohjelmien väliarvioinnit vastaavat tähän tarpeeseen. Ohjelmatoiminnan kehittäminen on viime aikoina ollut Tekesin ohjelma-arvioinnin fokuksessa, kun tiedon ensisijaisena käyttäjänä on ohjelmaproessi. Tiedontuottaminen on puolestaan eriytynyt tutkimusprojekteiksi, joita rahoitetaan vaikuttavuusarviointiyksikön innovaatiotutkimuksena. Sen sijaan Tekesin arvioinneissa ei ole nyt juurikaan tilivelvollisuusarvioinnin elementtejä. On jossain määrin kyseenalaista, kuinka uskottavaa olisi tilivelvollisuuden arviointi organisaation itsensä tuottamana. Sen tulisi olla täysin riippumaton. Voisiko sisäinen tai ulkoinen tarkastustoiminta vastata tilivelvollisuuden haasteeseen?

Avoimeksi kysymykseksi jää, olisiko muita Tekesin ulkopuolisia toimijoita, jotka voisivat tarttua tilivelvollisuuden arviointiin innovaatiotoiminnassa? Voisiko korkeakoulujen arviointineuvoston toimintaa kehittää tähän suuntaan, kuten innovaatioympäristön arviointi ehdottaa? (ks. Veugelers et al. 2009)

Tulevaisuudessa ohjelmatoiminta tulee kokemaan muutoksia, sillä uudet strategisen huippuosaamisen keskittymät ovat aloittaneet toimintansa. Keskittymiä on perustettu kuusi, ja niiden ensimmäiset tutkimusohjelmat on käynnistetty. On arvioitu, että noin kolmen – neljän vuoden kuluessa Tekesin ohjelmiin kohdistuu puolet siitä rahoituksesta, mitä tällä hetkellä ja ohjelmien lukumäärä putoaa kymmeneen nykyisestä 25:stä. Keskittymät ovat Tekesistä riippumattomia osakeyhtiöitä, jotka laativat oman tutkimusstrategiansa ja hankkivat eri tahoilta rahoituksen ohjelmiinsa. Keskittymien ohjelmat eivät siten ole Tekesin ohjauksessa ja arvioinnin kohteena samalla tavoin kuin Tekesin omat ohjelmat.

Keskittymätoiminnasta tehtiin syksyllä 2009 itsearviointikysely. Siinä pyydettiin SHOK-toiminnan johtoryhmältä ja keskittymien avainhenkilöiltä kan-

nanottoa sekä keskittymien ohjelmien, keskittymien itsensä että koko SHOK-toiminnan tulevasta arvioinnista. Arvioinnin suhteen saatiin runsaasti toisistaan poikkeavia ehdotuksia. Vastauksissa korostetaan kuitenkin sitä, että ohjelmilla täytyy olla työrauhaa ja aikaa saavuttaa tuloksia ennen arviointia. Ohjelmien osalta väliarviointia ehdotetaan, kun käynnistämisestä on kulunut 2–4 vuotta. Keskittymiä ehdotetaan arvioitavaksi 4–6 vuotta käynnistämisen jälkeen. Koko SHOK-toiminnan osalta arviointi olisi seuraavan kerran ajankohtaista, kun toiminta on saavuttanut täyden laajuutensa. (Pesonen 2009)

Pohdinnan arvoista on, mikä rooli Tekesin ohjelma-arvioinnilla on strategisen huippuosaamisen keskittymien arvioinnissa. Tämä on nähtävästi seuraava murroskohta, jolloin Tekesissä muotoutuu viidennen sukupolven arviointikäytäntö.

#### 4. YHTEENVETO JA JOHTOPÄÄTÖKSIÄ

Artikkelissa on tarkasteltu ja asemoitu tilaajan näkökulmasta Tekesin ohjelma-arviointitoimintaa 20 vuoden ajalta. Kehityksessä on havaittavissa murroskohtia, jolloin ohjelma-arvioinnin asema on muuttunut merkittävästi. Ensimmäinen sukupolvi oli ohjelman operatiivinen kehittäminen. Tätä seurasi tilivelvollisuus ja sitten tiedon tuottaminen. Neljäs sukupolvi on myös kehittävä, mutta ei yksittäisen ohjelman vaan ohjelmaproessin strategisella tasolla.

Ohjelmatoiminnan arvioinnissa tuleva viides sukupolvi on todennäköisimmin vaihe, jossa arvioinnin kohteena ovat strategisten huippuosaamisen keskittymien ohjelmat. Tämä edellyttää oletettavasti arviointitoiminnalta taas uudenlaista otetta.

Tilaamisosaaminen on avainasemassa arviointitoiminnan tulevan kehittymisen kannalta. On syytä pitää kirkkaana mielessä, mitä arvioinneilla kulloinkin tavoitellaan: tilivelvollisuutta, tiedontuottamista vai kehittämistä. Tämä määrittää sen, kenen rooliin arvioinnin tilaaminen kuuluu ja minkä tyyppinen osaaminen ja ote on tarpeen. Tarvitaanko esimerkiksi intuitiota, ideointia ja inspiointia vai tarkkuutta, tasapuolisuutta ja tehokkuutta? Ensin mainittu koskee soveltavaa arviointitutkimusta, kuten ilmiöiden tunnistaminen ja arviointimenetelmien kehittäminen, kun taas jälkimmäinen koskee arviointien tilaamista hankintana.

Tilajalla on selvästi oma poikkeava roolinsa ver-

rattuna arviointien toteuttajiin ja hyödyntäjiin. Arvioinnin toteuttajat ovat vastuussa arviointiteorioiden ja menetelmien hallinnasta ja hyödyntäjät puolestaan eri kehitysimpulssien käytöstä toiminnan kehittämiseen. Tilaajan roolina on parhaimmillaan luoda mahdollisimman hyvät edellytykset sekä toteuttajille että hyödyntäjille. Kehittävän arvioinnin puitteissa tätä voitaneen kutsua kehityskumppanuudeksi.

Tilivelvollisuuden arviointiin liittyy paljon vielä ratkaisemattomia kysymyksiä. Arviointitoiminnan tutkimukselle ja analysoinnille onkin selkeästi tarvetta. Myös arviointiosaamiseen kohdistuu vaatimuksia, mikä näyttäisi edellyttävän myös arviointikoulutuksen entistä määrätietoisempaa kehittämistä. Tätä hankaloittaa tosin arvioinnin monialaisuus ja käytännönläheisyys, jolloin ei hevillä löydy omaa tieteenalaa ja professuuria toteuttamaan esimerkiksi arvioinnin maisterikoulutusta tai tohtorikoulua.<sup>3</sup>

## VIITTEET

<sup>1</sup> Artikkelin kirjoittaja vastaa Tekesissä ohjelma-arviointien suunnittelusta, tilaamisesta ja toteuttamisesta. Esitettävät näkemykset ovat kirjoittajan eivätkä edusta Tekesin kantaa.

<sup>2</sup> Myös Tekesissä kehitettiin 90-luvulla arvioinnin ohjeistusta tilaajan tueksi. Tähän palataan tarkemmin luvussa 3.2.

<sup>3</sup> Kirjoittaja esittää kiitoksensa Ville Valovirrälle, Markus Koskenlinnalle ja kahdelle refereeelle kommentoinnista.

## LÄHTEET

- Ahonen P, Tala J, Hämäläinen K. Poliittikatoiminen vaikuttavuusarvioinnin menetelmien ja toimintatapojen kehittäminen: esiselvitys. Sektoritutkimuksen neuvottelukunta 8/2009. Helsinki 2009
- Arnold E. How meta-evaluation helps us understand the effects of the framework programmes. Esitys Eufordia Conference. Prague, 25.2.2009
- Arnold E, Luukkonen T, Joerg L, Oksanen J, Thuriaux B, Whitehouse S. Evaluation of Finnish R&D Programmes in the Field of Electronics and Telecommunications. Tekesin teknologiaohjelmaraaportteja 2/2002. Helsinki 2002
- Berg P, Lindberg R. Teknologiaohjelmien arviointi ja ohjaus. Tekesin ohjelmaraaportteja 15/97. Helsinki 1997
- Chelimsky E. The coming transformations in evaluation. Julkaisussa Chelimsky E, Shadish W (eds.). Evaluation for the 21st century. A handbook. Thousand Oaks: Sage Publications 1997
- Chen H. Practical Program Evaluation., Thousand Oaks, Sage Publications 2005.
- Clarke A, Dawson R. Evaluation research: An introduction to principles, methods and practice. London Sage 1999
- Euroopan Unioni. EC Impact assessment guidelines, SEC(2009)92). Brussels 2009
- Euroopan Unioni. Evaluating EU Activities A practical guide for the comission services. Brussels 2004
- Fetterman D. Foundations of Empowerment Evaluation, Sage Publications, Thousand Oaks 2001
- Fisch P. FP6 Ex-Post Evaluation. Esitys Eufordia Conference. Prague, 25.2.2009
- Georghiou L. Evaluation in the balance: Matching methods and policy needs. Esitys Evaluation conference, Brussels, 10.–12.11 2001
- Georghiou L, Rigby J, Cameron H (eds.). Assessing the Socio-Economic Impacts of the Framework Programme. Brussels 2002
- Georghiou L, Roessner D. Evaluating technology programs: tools and methods. Research Policy 29: 657–678. 2000
- Gustafsson R, Ahola E, Ilmola L, Kuusinen J, Pesonen P. Uuden sukupolven teknologiaohjelmia etsimässä, Tekes katsauksia 135/2003, Helsinki 2003.
- Guy, K. Assessing RTD program portfolios in the European Union. Julkaisussa Shapira P, Kuhlmann S (eds.) Learning from science and technology policy evaluation. Edgar Elgar 2003
- Hildén M. Laajojen kansallisten ohjelmien vaikutusten etukäteisarviointi – kokemuksia ympäristöarvioinneista. Hallinnon tutkimus 25(2007) 4, s.72–81
- Hjelt M, Vanhanen J, Raivio T. Rakenteiden uudistuminen – teknologiaohjelmat taustatukena ja muutoksen veturina. iWELL-, Puuenergia-, Rembrand- ja Uusi teollinen toimintatapa (UTT)-ohjelmien loppuarviointi. Arviointiraportti. Teknologiaohjelmaraaportti 12/2004. Helsinki 2004
- Kaakinen J. Arvioinnin paikka. Arvioitsija intressien ristipaineessa. Hallinnon tutkimus 23 (2004): 2, s. 94–99.
- Kauppa- ja teollisuusministeriö. The Technology Development Center of Finland, Tekes. International evaluation. Helsinki 1995
- Kauppi H. Kuka tarvitsee arviointia, mihin? – Arvioinnin ristiriitaiset pyrkimykset. Hallinnon tutkimus 23(2004): 2, s. 73–85.
- Kuhlmann S. Evaluation as a source of strategic intelligence. Julkaisussa Shapira P, Kuhlmann S (eds.) Learning from science and technology policy evaluation. Edgar Elgar 2003
- Kuitunen S, Haila K. Terveys-, elinkeino-, ja innovaatiopolitiikan rajapinnoilla. FinnWellin ja Tekesin asemoitumisen arviointi. Tekesin teknologiaohjelmaraaportteja 8/2007. Helsinki 2007
- Kuitunen S, Haila K, Kauppinen I, Syrjänen M, Vanha-

- nen J, Ahonen P, Tuomi I, Kettunen P, Paavola T. Finns in the EU 6th Framework Programme. Tekesin ohjelmaraportteja 6/2008. Helsinki 2008
- Kuntaliitto. Suositus arvioinnista kuntien valtuustoille. Helsinki 2006
- Lemola T, Siivonen T. Teknologiaohjelmien arviointi. Tietotekniikan kehittämisohjelman väliarviointi. Tekesin raportti 14/88. Helsinki 1988.
- Mark M, Henry G, Julnes G. Evaluation. An integrated framework for understanding, guiding, and improving public and nonprofit policies and programs. San Francisco. Jossey-Bass, Wiley 2000
- Murto E, Niemelä M, Laamanen T. Altavastaaajasta ykköskenttään – Suomen teknologia politiikan ja sen toimijaorganisaatioiden kehitysvaiheita 1960-luvulta nykypäivään. Edita Helsinki 2007
- Oksanen J. Research Evaluation in Finland. Practices and experiences, past and present. VTT working papers 51/00. Espoo 2000
- Paasio P. Yleinen ja erityinen viitekehys arvioinnista. Hallinnon tutkimus 24(2006): 3, s. 92–107.
- Patton M. Utilization-focused evaluation. The new century text. Third printing. Thousand Oaks. Sage Publications 1997
- Pawson R, Tilley N. Realistic Evaluation. London, Sage 1997
- Pesonen P. Teknologiaohjelmat 1998. Tekesin sisäinen raportti 1998
- Pesonen P. Uutta tietoa ja osaamista innovaatiopolitiikkaan. Teknologiaohjelmaraportti 5/2006. Helsinki 2006
- Pesonen P. Strategisen huippuosaamisen keskittymä (SHOK) -toiminnan itsearviointikysely 2009. Tekes 2009.
- Pietilä M, Tuloksia vai toiminta. Selvitys virastojen ja laitosten toimintakertomusten tuloksellisuuden kuvauksista. Valtiokonttori 2008.
- Pihlajamaa J, Nummi J, Berg P. Analyysi Tekesin teknologiaohjelmien arvioinneista. Sisäinen raportti 1998.
- Raivio T, Syrjänen M, Halonen M. Tekes elektroniikan moottorina. Tekesin elektroniikka-alan ohjelmatoiminnan arviointi. Teknologiaohjelmaraportti 14/2006. Helsinki 2006
- Rajavaara M. Yhteiskuntaan vaikuttava Kela. Sosiaali- ja terveyturvan katsauksia. Helsinki 2006
- Rip A. Societal Challenges for R&D evaluation. Julkaisus-  
Shapira P, Kuhlmann S (eds.). Learning from science and technology policy evaluation. Edgar Elgar 2003
- Saarinen J, Kiviniemi K, Pekkala H, Mäenpää M, Vuori J. Terveydenhuollon laitekehityksestä systeemien kehittämiseen – FinnWell- ja iWell-ohjelmien arviointi. Tekesin ohjelmaraportteja 6/2009. Helsinki 2009
- Salo A, Pahlavan K, Salmenkaita J. R&D Programmes in Electronics and Telecommunication ETX, TLX, INWITE and Teletronics. Mid-Term Evaluation. Tekesin teknologiaohjelmaraportteja 5/2000. Helsinki 2000
- Shapira P, Kuhlmann S (eds.). Learning from science and technology policy evaluation. Edgar Elgar 2003
- Suomen Akatemia. Tutkimusohjelmastrategia. Akatemian julkaisuja 2/03. Helsinki 2003
- Suomen Akatemia. Suomen tieteen tila ja taso. Akatemian julkaisuja 9/09. Helsinki 2009
- Tekes. Tekesin vaikuttavuusarvioinnin strategia. Hyväksytty Tekesin johtoryhmässä 27.3.2000. Tekes 2000
- Tekes. Innovaatioita julkiseen terveydenhuoltoon – FinnWell-ohjelman ensimmäinen väliarviointi. Arviointiraportti. Tekesin teknologiaohjelmaraportteja 3/2007. Helsinki 2007
- Tekes. Toivoo, tahtoo, tarmoo. Tekes 25 vuotta. Tekes 2008
- Tekes. Innovaatiotoiminnan vaikutukset. Tekes 2008
- Temmes M. Arviointityön autuus ja salahaadat arvioijan näkökulmasta. Hallinnon tutkimus. 23 (2004):2, s. 86–93.
- Ulkoasiainministeriö. Evaluation Guidelines. Between Past and Future. Helsinki 2007
- Valovirta V, Hjelt M. Onnistumista punniten, tulevaa luodaten. Arviointi ja ennakointi strategisen päätöksenteon tukena innovaatio- ja teknologia politiikan esimerkkien valossa. Hallinnon tutkimus 23(2005): 3, s. 95–111.
- Valtakari M, Rajahonka M, Tinnilä M, Kujala A. Teknologiaohjelmien strategiatyön johtamisen arviointi. Teknologiaohjelmaraportteja 6/2007. Helsinki 2007
- Valtiontalouden tarkastusvirasto. T&k-arviointitoiminta. Toiminnantarkastuskertomukset 157/2008. Helsinki 2008
- Valtiovarainministeriö. Monta tietä vaikuttavuuteen. Valtiovarainministeriön julkaisuja 29/2008. Helsinki 2008
- Vataja K. Arviointiprosessin monet roolit ja mahdollisuudet arviointiprosessin hyödyntämisessä. Hallinnon tutkimus 23(2005): 3, s. 123–130.
- Vedung E. Public policy and program evaluation. New Brunswick (N. J.): Transaction Publishers. 1997
- Veugelers, R et al. Evaluation of the Finnish National Innovation System – Full Report. Helsinki 2009.
- Virtanen P. Arviointi. Arviointitiedon luonne, tuottaminen ja hyödyntäminen. Edita 2007
- Vuorela T. Arvioinnin tilaajan opas. Valtiovarainministeriö. Helsinki 1997
- Weiss C. Evaluation. Methods for studying programs and policies. Second edition. New Jersey. Prentice Hall 1998