

# Evaluatio-oppia edistyville

*Pertti Ahonen*

## ABSTRACT

Evaluation for the advancing learner

Evaluation is a matter to learn, and not only in a linear way, and it is also a matter of unlearn on occasion. The complexities of evaluation concern as well the evaluators, those who manage them, those who subscribe evaluations and those who are being evaluated. Unfortunately, evaluation is deeply divided by its philosophical, theoretical and disciplinary origins, methods, techniques, uses and fields of application. All this prevents attainment of a unitary view. Books are necessary in evaluation to attain basic conceptual and methods wisdom, but still evaluation cannot be learned only from books. There are risks in defining benchmarks and best practices of evaluation. This threatens to consign to silence deeper knowledge on evaluation, which is received where the things do not go too well on the one hand, but on the other the best existential assurance is received that the evaluator has not trodden the path of the least resistance in the evaluation efforts. A particular risk is to believe that practical experience is a panacea to good evaluation; on the contrary, the implicit theories of that practitioners possess but deny to rely upon may lead evaluation astray. Recent efforts to put theoretical approaches to evaluation in perspective, such as that of Pawson and Tilley, fail to provide a conclusive classification despite the help they do render. Even more so, the same concerns conceptualisations that have their origin in the practice of evaluation; these are quite inconclusive. All the above has its deep message to the principal figure in evaluation, namely the evaluator. He or she does not have to remain in the sole role of an evaluation worker, but it is possible to proceed towards the role of the "reflective practitioner" of evaluation. What is more, the evaluator may evolve as a figure who

raises above everyday evaluation with a deep understanding of what evaluation is about and what is at stake in evaluation.

**Key words:** learning, unlearning, metaevaluation, evaluation approaches, evaluation roles

"On vain yksi, perspektiivinen, tapa nähdä, vain yksi perspektiivinen 'tietäminen'. Mitä enemmän affekteja saamme liikkeelle ne sanoiksi pukien jonkin asian merkeissä, sitä useampia silmiä, erilaisia silmiä, kykenemme kiinnittämään tuohon samaiseen asiaan, ja sitä täydellisemmäksi tulee käsittävä otteemme tuosta asiasta, 'objektiviteettimme'."

Nietzsche, Moraalin alkuperästä, kohta 12, s. 861

## 1 ASKELTAEN EVALUAATIOVIISAUTEEN

### 1.1 *Sorvissa näkökulma evaluatioon*

Kohteelle soveltuu nimeksi "evaluatio", jotta äänteellinen kosketus säilyy useimpiin eurooppalaisiin kieliin. Siten voi välttää sanan "arviointi" muodikasta viittaamista mitä moninaisimpiin ilmiöihin (vrt. Collins 2000, Birnbaum 2000). Evaluatioon keskeisin "näkökulma" ja "tekijä" on evaluaattori, joka on ainoa artikkelissa korostuva rooli. Evaluatiolle ovat kuitenkin tärkeitä myös teettäjät eli evaluoittajat, joihin kuuluu kuka tahansa evaluaattoriin nähden johtajan asemassa oleva. Evaluatioon roolipelin pelaajia ovat samoin tilaajat, evaluoittajat, sekä evaluoitavat. Niin yksinkertainen ei evaluatioon kenttä ole, että kukin neljästä roolista eriytyisi siististi saati että yksittäiset henkilöt sijoittuisivat täsmällisesti vain yhteen rooleista. Evaluoitavat saattavat hyvinkin evaluoida jopa itseään, kunhan kysymys ei ole jonkun muun ehdoin laadituista kuulustelun omaisista "itsearviointeista" tai siitä, että joku pääsee eva-

luaattorina ikään kuin tuomariksi omassa asiasaan.

Evaluaaation kohdetta koskeva oppiminen kuuluu evaluaaation vakiintuneisiin tarkoituksiin. Evaluuattori, evaluoittaja, evaluoituttaja ja evaluoitava saattavat oppia myös kustakin evaluuatioprosessista, -episodista ja -hankkeesta ja vähitellen myös evaluuatiosta yleensä. Yhtä paljon kuin lineaarisesta, "positiivisesta" oppimisesta, evaluuatiosta on kysymys epälineaarista oppimisesta, haasteista oikaista evaluuatio-oppimisen vinoumia sekä tarpeista taakaksi käyneen evaluuatiotietämyksen ja -taidon poisoppimiseen.

Evaluuatiota luonnehtii jakautuneisuus. Eri tahot pyrkivät paaluttamaan evaluuatiota kenttää, jotta kullekin niistä muodostuisi monopolietuja antava läänitys ajatellen sitä, mitä evaluuatioon on legitimiä lukea. Olipa kysymys evaluuatiosta sovelletuista lähestymistavoista ja metodeista tai niistä tarkoituksista, joissa evaluuatiota harjoitetaan, paalutuksia nousee paalutusten jälkeen ja monopoli- ja reviiiritaisteluja puhkeaa, raivoa mutta lopulta laantuu.

### 1.2 *Evaluuatiota hahmottaa sekin mitä evaluuatio ei ole*

Evaluuatio ei yhtene tarkastuksena, tilintarkastuksena tai revisiona tunnetun toiminnan kanssa. Suomalaisessa käytännössä toiminnan tarkastus saattaa laajeta evaluuatiota piiriin, mutta sen ala vastaa lähinnä englannin kielen ilmausta performance auditing (Pollitt ym. 1999). Tarkastus ja evaluuatio saattavat käyttää samoja perustietoja ja tietokantoja, mutta se ei tee evaluuatiosta tarkastusta eikä päinvastoin. Vielä vähemmän tarkastuksesta tulee evaluuatiota vain sillä perusteella, jos tarkastusalan organisaatiot harjoittavat myös evaluuatiota ja samat henkilöt sekä tarkastus- että evaluuatiotehtäviä.

Evaluuatio ei yhtene policyanalyysin kanssa. Poliittikan tutkimuksen otteena, sen piirteitä soveltavina käytäntöinä tai kansantaloustieteen sovelluksina policyanalyysi on kiinnostunut siitä, vaikuttaako julkinen politiikka, policy, itsenäisesti tavoitteiden suuntaan. Ellei näin käy, se mitä tapahtuu aiheutuu pelkistä olosuhdetekijöistä. Policyanalyysi antaa kuitenkin evaluuatiolle käsitteellisiä, metodisia ja teknisiä virikkeitä. (Esim. Pal 1992.)

Evaluuatio ei yhtene kehittämistyönä tunnettuun toimintaan. Huoli, että kytkennättä kehittämiseen evaluuatio olisi hyödytöntä, on yleensä liioiteltu. Sikäli kuin evaluuatiosta selvitetään, kuinka asiat ovat, kehittämistyön näkökulma saattaa olla hätköityä kenties kapeuttaessaan evaluuatiota alaa. Entä jos alaltaan kapeuttamaton evaluuatio toisikin esiin, että kehittämistyön kentän rajaus evaluuatiota lähtökohdaksi on liian suppea tai muuten sopimaton?

Evaluuatiota intellektuaalinen ydin on pidettävä erossa evaluuatiota sisäisistä käytännöistä. Ensimmäinenkään evaluuatio ei toteudu ilman suunnittelua, ohjelmointia, ohjeistamista, projektointia ja budjetointia, tarjouskilpailuja ennen laajoja evaluuatioita, raportointia tai tiimityötä. Ne eivät kuitenkaan ole evaluuatiota ydin.

Evaluuatiota identiteettiä hämää ja sen ytimen löytämistä vaikeuttaa, että evaluuatio on eriytynyt erikoisaloittain. Ulkopuoliselle on ammatillinen itsemurha mennä neuvomaan, kuinka on evaluoitava terveydenhuoltoa (Rimpelä 1992, Rimpelä ym. 1999, Rissanen 1996), sosiaalipalveluja (Cheetham 1992, Mäntysaari 1997), opetus-, koulutus- ja tutkimustoimintaa (Raivola 1999, Shinkfield & Stufflebeam 1995) tai aluekehitystä (MEANS 1999, Kelleher ym. 1999).

## 2. EVALUAATION OPPIMINEN, VÄÄRIN OPPIMINEN JA POISOPPIMINEN

### 2.1 *Evaluuatio: ilosanoma vai eksistentiaalinen haaste?*

Karl Marx (1974/1867, 47) luonnehti kapitalismia kohteena, jonka tunnusomaisin piirre on tarkastelijalle ilmenevä "valtava tavarajoukko". Myös evaluuatio ilmenee valtavana joukkona asioita, kirjoja, artikkeleita, ohjeistoja ja raportteja sekä käytännön ja tutkimuksen evaluuatiotöitä. Tutkijoiden tai muiden kirjoittamaa opastavaa evaluuatiokirjallisuutta on runsaasti (esim. Rossi ym. 1999, Clarke & Dawson 1999, Vedung 1998), joskin perusteokset puuttuvat suomeksi. Evaluuatiosta on kirjoitettu paljon, mutta ongelmana on evaluuatiota koskevan tiedon tulkinta ja soveltaminen (Pawson & Tilley 1997, 25).

Evaluuatiota koskeva teoreettinen, tekninen ja käytännöllinen oppiminen on monisyinen, -tasoinen ja -ulotteinen prosessi. Se tuskin etenee lineaarisesti, koska useimpien evaluuatiota oppivien

mieli on kaukana tabula rasasta, "tyhjistä taulusta". Prosessin tekee monisyiseksi myös inhimillinen unohtaminen.

Evaluataation oppimisessa on virhe keskittyä yksipuolisesti vain best practice'ihin eli "parhaisiin käytäntöihin" ja benchmarkeihin. Ne eivät paljasta, onko evaluaattorin kohtaama vastus ollut kunnioitettava vai heikko. Onnistumisten lisäksi evaluaatiota opettavat epäonnistumiset, pettymykset, virheet, fiaskot ja huijaukset. Martin Heideggerin (1927/2001) ajatus sinänsä ikävien ääni-ilmiöiden opettavuudesta on tärkeä. Eigentlichkeit, "varsinaisuus", alkaa silloin, kun asioiden joustavan ja kitkattoman sujumisen Uneigentlichkeit, "epävarsinaisuus", rakoilee. Meille saattaa jopa tulla mieleen, ettemme ammatillisesti emmekä muutenkaan elä ikuisesti. Siksi ei ole samantekevää, mitä teemme ja jätämme tekemättä, mihin suostumme, mikä meille kelpaa ja päädymmekö olemaan evaluaattorinakin vain korvattavissa oleva "kuka tahansa", Heideggerin *das Man*.

Käytännöllisestä näkökulmasta Heideggerin ajatuksia on sovellettu informaatio- ja viestintätekniikan alueella (Winograd & Flores 1995, Dreyfus 1991). Siellä asioiden olemus paljastuu nimen omaan silloin, kun asiat lakkaavat sujumasta: ohjelmoinnista paljastuu virhe, ohjelmat eivät toimi ja järjestelmät lakkoilevat. Evaluaatioon yhtään syvemmin perehtyvä havaitsee, ettei sekään aina suju kitkattomasti paitsi silloin kun evaluaatiohaasteet jäävät niin helppoiksi, että evaluaatiosta oppiminen lakkaa. Evaluataation opettamista benchmarkeja ja parhaita käytäntöjä esittämällä on aiheellista täydentää evaluaation rajatilanteiden tarkastelulla ja analysoinnilla sekä koulutettavien koulumisella varautumaan niihin.

## 2.2 *Evaluatioviiisuuden tyypittelyä*

Evaluatiossa teoreettisella kirjaviisaudella on rajansa, mutta ilman sitä evaluaattori jää käsitteellisesti ja älyllisesti sokeaksi. Sama koskee jokaista, joka koettaa omaksua viisautta ulkoluvulla. Evaluaattori tarvitsee myös metodi- ja tekniikkaviisautta. Sitäkään ei voi noin vain painaa ulkomuistiin ja etsiä sitten sieltä tarpeen tullen. Evaluatiossa on laadittu ohjekirjoja, ohjeistoja ja eettisiä koodeja (PCM 1993, vrt. Ahonen 1999, CEC 1997, JCSEE 1994, MFA 1998, MEANS 1999, Kuntaliitto ym. 2001). Tuosta normatiivi-

sesta aineistosta ei tee evaluatioviiisautta edes aineiston laatijan arvovaltainen asema. Kenties aineisto antaa vain kielen, jota evaluaattorin on määrätyissä tilanteissa puhuttava henkensä pitimiksi. Toteutuneiden evaluaatioiden raportit, yhden tai useamman evaluataation uudelleenarvioinnit eli metaevaluaatiot (ks. esim. Uusikylä & Virtanen 2000) sekä evaluaatiota koskevat tutkimukset muodostavat nekin laajan aineiston. Parhaat metaevaluaatiot tuovat esiin, mitkä evaluaatioista ovat paradigmaattisia kelpaten malliesimerkkeinä evaluatioviiisuuden osiksi, olivatpa evaluaattoreiden kokemukset olleet auvoisia tai kovia. Osa evaluatioista kelpaa vain varottaviksi esimerkiksi, kun taas useimmat evaluaatiot eivät herätä huomiota puoleen eikä toiseen.

Evaluatioviiisautta saattaa sisältyä myös kokeiden evaluaattoreiden "tekijän tietoon". Osa siitä ei ole soveliasta panna paperille evaluaattoreiden, evaluoittajien, evaluoituttajien ja evaluoitujen nimiä mainiten. Muuten kritiikin kohteina esiintyvät myös ne, jotka ovat joutuneet vaihtuneissa e evaluatiosuhteissa ikävään "väläkkeeseen". Myös nimisuojan antamalla evaluaatiota tarkasteleva voi tuoda esiin aivan riittävästi virheitä, osaamattomuutta, inhimillistä alhaisuutta ja juonia.

Evaluataattorille tuskallisin tekijän tiedosta koskee hänen omia puutteitaan ja virheitään sekä hänen maksamiaan materiaalisia ja mentaalisia oppirahoja. Tuskin kovinkaan moni evaluaattori haluaa kirjata tuollaista tietoa muiden näkyviin. Osa kokemustiedosta ei lainkaan voi kirjata ylös vaikka halua ja uskallusta olisi. Sitä voi olla vain sillä, jolla on kokemusta, vaikka kokemus ei olekaan kaikki kaikessa.

Käytännön elämän piirissä esiintyy niitä, jotka halveksivat älyllisen suuntautumisen nimissä ylläpidetyissä instituutioissa kuten yliopistoissa työskenteleviä riippumattakin jälkimmäisten tosiasiallisesta suuntautumisesta. Tuohon halveksuntaan liittyy eräs harhaluulo. Sen mukaan "käytännön ihminen" suoritusi evaluaattorinakin käden käänteessä haasteista, joiden parissa älykkö vain mutkistaa asioita. Teoreettisia ja käsitteellisiä pohdintoja kaihtavilla käytännön ihmisillä on kuitenkin raudanlujat "käytännön teoriasa". Pahimmassa tapauksessa he eivät ole niistä tietoisia väittäen käsitystensä perustuvan vapaina ennakoedellytyksistä pelkkiin tosiasioihin ja omakohtaiseen kokemukseen. Yhtään tuollaista ihmistä eivät nykyiset tieteenfilosofit ja

tietoteoreetikot kuitenkin ole kohdanneet. He tuntevat vain meidät, joille jokainen tosiasia on mitä teoria-, käsitys- ja käsitepitoisin, olipa tuo pitoisuus meille tuttu tai ei.

Kokemuksesta saattaa tislautua viisautta, mutta siitä saattaa myös oppia väärin. On mahdollista, että oppii seuraamaan huonoja esimerkkejä olivatpa ne evaluaattoreita tai evaluaatioita, uskomaan hatariin auktoriteetteihin tai ryhtymään itse auktoriteetiksi perusteettomasti eli pelkän muodollisen aseman, fyysisen iän, virkaiän, suhteiden, ideologisten näkemysten tai oman loistavan menneisyyden voimalla. Voi olla, että päätyy toimimaan kaavamaisesti ja toistamaan itseään, jättämään uudistumatta, hutiloimaan milloin evaluaatiotehtävien antama vastus on alhainen ja kiire painaa päälle tai jättämään omaksumatta vankat evaluaation laatuksiteerit.

### 3 EVALUAATION YTIMESSÄ: TEORIA-LÄHTÖISTEN LÄHESTYMISTAPOJEN NÄKÖKULMAT

#### 3.1 Kuinka kiinni teoreettisimpaan evaluaatioon

Tutkimuksella on monia lähestymistapoja evaluaatioon. Niitä on kehitelty lähtien liikkelle useasta kohdin tiedonmuodostuksen prosessia, jota voi hahmottaa vaiheittain: filosofiset lähtökohdat → teoria → metodi → tiedonhankinnan tekniikat → tiedonkäsittelyn tekniikat → tiedon käyttö → palaute seuraaviin evaluaatioihin.

Evaluaatio on kuin elefanti tarinassa, jossa kolme sokeaa selvitteli sormituntumalla, millaisesta olenosta on kysymys. Jalkaa kaviopuolelta tunnustellut arveli elefantin olevan kivikuntaa, hännän pääosaa kopeloitunut kasvikuntaa ja kärsää käänneilyt matelevaa eläinkuntaa. Filosofit löytää evaluaation lähestymistavoista empirismää kokeellisen otteen sovelluksissa, filosofista pragmatismia käsitysten toimivuuden korostukseen sekä asioihin ymmärtävästi suhtautuvan filosofian piirteitä ja lopulta erilaisia filosofisia synteesejä. Teoreettisiin suuntauksiin keskittyvä havaitsee empiriisen luonnontieteen, taloustieteiden, politiikan tutkimuksen, sosiologian ja muiden tieteenalojen otteita. Metodikko päätyy pohtimaan kokeellisten, kvasikokeellisten ja ei-kokeellisten astelmimien sekä kokonaistutkimusten ja tapaustutkimusten (Stake 1995) asemaa evaluaatiossa. Teknikko saattaa etsiä evaluaation

ydintä tavoista, joilla tieto hankitaan ja käsitellään: onko se ekonometriää, tilastoanalyysia, surveyanalyysia, lähdekriittistä tekstianalyysia, matemaattista mallintamista, diskurssianalyysia vai teknisesti jotakin perin juurin muuta? Evaluaation lähestymistapoja voi tarkastella myös ikään kuin takaperin lähtien tarkoituksista, joita evaluaatiolle asetetaan (Owen & Rogers 1999). Lopulta voidaan tarkastella evaluaatioiden kytkeytymistä toisiinsa aika-akselilla alkaen kerta-luonteisista evaluaatioista muutaman evaluaation sarjoihin, laajempiin määräaikaisiin evaluaatio-ohjelmiin ja pysyväksi institutioituun evaluaatioon.

Tarkastelen Pawsonin ja Tilley'n (1997) jaottelea, jonka lähtökohdat ovat filosofiset. Nelijako on esimerkki, ei lopullinen filosofinen totuus evaluaatiosta (Taulukot 1 ja 2). Toisaalta juuri sen avulla luon oman silmäykseni osaan evaluaation ydintä.

#### 3.2 Kvasikokeellinen ja pragmatistinen evaluaatio

Kvasikokeellinen evaluaatio on evaluaation tutkimuslähtöinen peruslähestymistapa. Se on myös neljästä lähestymistavasta vanhin. (Campbell 1969, Campbell & Stanley 1963, Campbell & Russo 1998.) Lähtökohtaisesti kvasikokeellinen evaluaatio edustaa "kovaa" metodiikkaa, joskin sen periaatteita voidaan soveltaa myös, kun aineisto on laadullista.

Kvasikokeellisella evaluaatiolla on vankat kosketuskohdat tieteellisessä arvojärjestyksessä korkealle sijoittuviin yhteiskuntatieteisiin kuten kansantaloustieteeseen ja tilastotieteeseen. Sillä on myös yhteyksiä osiin lääketieteitä, osiin psykologiaa sekä osiin kasvatustieteitä. Sellaisia evaluaatiokielen tavallisia sanoja kuin tehokkuus tai vaikuttavuus olisi vaikea käyttää, ellei kvasikokeellinen lähestymistapa antaisi jäsenystään. Sekä tehokkuuden että vaikuttavuuden käsitteet edellyttävät kausaaliasetelmaa ja jopa tiettyjä analogioita koeasetelmalle. Panoksia on uhrattava ennen kuin saadaan tuotoksia, tuotosten määrän tulisi muuttua samaan suuntaan kuin panosten määrän eikä muutos saisi aiheutua muusta kuin panosten määrän muutoksesta. Muuten ei voida päätellä tehokkuuden vallineen panosten ja tuotosten suhteessa. Vaikuttavuus edellyttää, että on ennalta asetettu vaikutuksia

**Taulukko 1. Kvasikokeellinen evaluaatio ja pragmatistinen evaluaatio (Pawsonin ja Tilleyn 1997 mukaan, täydennettynä).**

	Kvasikokeellinen evaluaatio	Pragmatistinen evaluaatio
Tieto-opillinen l. epistemologinen tausta	Lähestymistavan soveltamisen loogisena alkupisteenä on huomion kiinnittäminen kausaalisuhteisiin eli syy- ja seuraus –suhteisiin.	Päteväksi katsotaan pragmaattisesti hyväksyttävä tieto. Tuollainen tieto sovituu yhteen sen lähtökohtakehikon kanssa, jonka ehdoin politiikka, ohjelma tai hanke on muotoiltu.
Metodi	Kehitetään kokeellisia tai kvasikokeellisia (koeasetelman periaatteita mukauttavia) asetelmia mittauksineen ennen ja jälkeen toimenpidettä sekä kohde- että kontrolliryhmissä.	Tutkimuksen tehtävänä ”valistaa” tuottamalla informaatiota politiikasta, ohjelmasta tai hankkeesta eri tekniikoilla, jotka evaluaattori hyvänä käsityöläisenä hallitsee.
Käsitys siitä, mistä evaluaation kohteessa on kysymys l. ontologia	Pyrkimys on selvittää yhteiskunnassa vallitsevasta monimutkaisuudesta rakentamalla evaluaatio kontrollifunktioksi, jolla voidaan tutkia oletettujen kausaalisuhteiden esiintymistä, kun ajatus tuosta esiintymisestä on perusta poliitiikkatoimille.	Kohde ja evaluaatio ymmärretään valtapeleiksi. Evaluaatiot ja evaluaattorit joutuvat siksi tunkeutumaan vaikutusvallan käytön ketjuihin, jotka joko estävät tai edistävät muutosta.
Liittyminen käytännön toiminnan ohjelmointiin	Evaluaatio liittyy politiikkoihin, ohjelmiin ja hankkeisiin, joiden tarkoituksena on vaikuttaa kohteisiin saamalla aikaan kohteisiin halutulla tavalla kausaalisesti vaikuttavia seikkoja.	Politiikat, ohjelmat ja hankkeet ymmärretään omiksi institutionaalisisiksi todellisuuskseen ”palveluina”, ”urina”, ”elinkeinoina”, ”päämäärinä”, ”tingintäprosesseina”, ”asiakaina”, ”vastustajina” ja niin edelleen.
Käsitys edistyksestä	Koeasetelman sovellukset tuovat esiin ne politiikat, ohjelmat ja hankkeet, jotka toimivat tuottaen pysyvästi haluttuja tuloksia ja vaikutuksia.	Politiikkojen, ohjelmien ja hankkeiden kehittäminen liittyy poliittisen toteutavuuden ( <i>feasibility</i> ), kustannusvaikuttavuuden ja pienten askelten ”inkrementalistisen” politiikan näkökulmiin.
Käsitys evaluaation hyödystä	Politiikan laatu paranee, kun evaluaatio tuottaa kovaa dataa, jonka avulla voidaan valita, mitkä toimenpiteet tuottavat parhaat tulokset ja vaikutukset.	Lähestymistavan soveltaminen loogisena alkupisteenä ovat pohdinnat, millaisen kohdetta koskevan teorian päätöksentekijät hyväksyisivät heidän tavoitteensa huomioon ottaen.
Mahdollisia ongelmia	Tarvittavaa dataa on työlästä hankkia ja muodostaa. Lisäksi virhemahdollisuudet ovat suuret. Olojen muutos mukaan lukien poliittisen kiinnostuksen ehtyminen voi estää vaikutusten selvittämisen. Lähestymistapa ei sovellu kaikkiin kohteisiin.	Se, mikä on evaluaatiota, ei erotu siitä, mikä on politiikkaa. ”Evaluaatiolla politikointiin” tarjoutuu tietyt mahdollisuudet, josta saattaa aiheutua ammattievaluaattorille riskejä. Pragmaattinen evaluaatio on toisaalta konservatiivista ollessaan valtasuhteista riippuvaista.

**Taulukko 2. Konstruktivistinen evaluaatio ja pragmatistinen evaluaatio (Pawsonin ja Tilley'n 1997 mukaan, täydennettynä).**

	Konstruktivistinen evaluaatio	Realistinen evaluaatio
Tieto-opillinen l. epistemologinen tausta	Tällainen evaluaatio on relativistista uskomuksineen, että totuus riippuu näkökulmasta sen sijaan, että sen lähde olisi ulkoinen suhteessa minkään ryhmän uskomuksiin.	Lähestymistavan soveltamisen loogisena alkupisteenä on jokin kausaalinen selitysteoria. Se perustuu käsitykselle siitä, miten politiikat, ohjelmat ja hankkeet rakentuvat.
Metodi	Asianosaisten asioille antamia merkityksiä tarkastellaan laadullisesti, jotta niitä voidaan sovitella keskenään konsensukseen tähtäävissä neuvotteluprosesseissa.	Teorioita testataan sen selvittämiseksi, miten ohjelmien vaikutukset syntyvät eri mekanismien vaikutuksesta eri yhteyksissä. Tämä vaatii ohjelmien välistä vertailua.
Käsitys siitä mistä evaluaation kohteessa on kysymys l. ontologia	Lähestymistavan soveltamisen loogisena alkupisteenä on asettamus, että sosiaalinen maailmamme perustuu siihen liittyviin arkipäiväisiin merkityksiin, jotka liittyvät kaikessa sosiaalisessa vuorovaikutuksessa vaikuttaviin järjkeilyprosesseihin.	Säännönmukaisuudet yhteiskunnassa tapahtuvien toimintojen rakenteistuksessa syntyvät, kun vaikuttavat mekanismit, joissa on kysymys ihmisten järjkeilystä ja heidän voimavaroistaan kussakin toimintayhteydessä.
Liittyminen käytännön toiminnan ohjelmointiin	Lähestymistapa katsoo ohjelmat löysiksi koosteiksi useiden asianosaisten (päättäjät, toteuttajat, kohteet ym.) asioista omaamista ajatuskonstruktioista.	Ohjelmointi mukaan lukien politiikkojen muotoilu ja projektointi on yritys rakentaa evaluaatioihin tietoa, joka tuo esiin, mikä toimii kenenkin kannalta ja miten missäkin yhteyksissä.
Käsitys edistyksestä	Tuloksia ja vaikutuksia ei pidetä "löydöksinä" eikä "selityksinä" vaan "konstruktioina", jotka voivat tulla kohteiksi asianosaisten jatkoneuvotteluille siitä, mistä on kysymys.	Tietoa kertyy toisiaan seuraavilta politiikka-, ohjelma- ja hankekiirroksilta ja muista lähteistä. Päättäjät saavat sitä kautta tietoa menestykseen johtavista politiikkojen, ohjelmien ja hankkeiden lajeista.
Käsitys evaluaation hyödyistä	Avoimena päämääränä on laajentaa yhteistyön prosesseja, jotta asianosaisia voitaisiin vallallistaa ja opettaa.	Opetus- ja oppimisprosesseissa evaluaattori auttaa nostamaan esiin asianosaisille palasittain kertyneen tiedon.
Mahdollisia ongelmia	Lähestymistapa saattaa olla epämääräinen, kenties ajelehtiva ja riippuvainen osapuolten halusta yhteistyöhön ja neuvotteluihin. Se ei ilmeisesti sovellu läheskään kaikkien politiikkojen, ohjelmien ja hankkeiden arviointiin.	"Realistinen evaluaatio" on profiililtaan jäsentymätön synteesi kolmesta muusta lähestymistavasta. "Realistinen evaluaatio" on tuskin mikään yleispätevä ratkaisu.

koskevia tavoitteita, että toteutuneet vaikutukset ovat samansuuntaisia kuin tavoitteet edellyttävät eivätkä vaikutukset ole aiheutuneet muista syistä kuin niistä toimenpiteistä, jotka on toteutettu asetettujen tavoitteiden saavuttamiseksi. Kummassakin tapauksessa "muiden syiden" eliminointi on se seikka, joka johtaa koeasetelman analogeihin.

Luonteeltaan lähelle kvasikokeellista evaluatiota sijoittuu ei-kokeellinen evaluatio, jonka käsitteistö on osin peräisin kokeellisen tutkimuksen piiristä (klassikkona ks. Blalock 1964). Eikokeellisessa evaluatiossa evaluattori saattaa kyllä vertailla toisiinsa tilannetta ennen toimenpidettä ja toimenpiteen jälkeen, mutta käytävissä ei ole kvasikokeellista asetelmaa saati koeasetelmaa. Sen sijaan evaluattori käyttää niitä alkutilannetta koskevia tietoja, joita tilastot tai muut lähteet tuovat saataviin.

Käytännössä moni evaluatio on epäilemättä sitä, mitä on nytemmin kutsuttu pragmatistiseksi evaluatioksi. Pawson ja Tilley (1997) pitävät Pattonin (esim. 1978) hyödyntämiskeistä evaluatiota (utilization-focused evaluation) sekä tunnetussa evaluatio-ohjeistossa heijastuvia periaatteita (JCSEE 1994) malliesimerkkeinä pragmatistisesta evaluatiosta. Tuollainen evaluatio on Pawsonin ja Tilleyn mukaan tavallista, mikäli evaluatation kohde on poliittisesti merkittävä ja jopa poliittisesti tulehtunut. Silloin evaluointi ja politikointi eivät selkeästi erotu toisistaan eikä tämä ole ehkä tavoitteenakaan. Myös "politikointi evaluoimalla" saattaa tulla kysymykseen.

Pragmatistinen evaluatio ei määriy tietä-opillisesti eikä metodisesti kiteytyneesti kuten kvasikokeellinen evaluatio. Sen sijaan pragmatistinen evaluatio kehkeytyy kytkeytyneenä poliittisiin käytäntöihin sen mukaan, mikä on noiden käytäntöjen asema evaluatation ennako-oletusten joukossa ja millainen on evaluatiotulosten ennakointi käyttö.

### 3.3 *Konstruktivistinen ja realistinen evaluatio*

Konstruktivistisen evaluatation (Guba & Lincoln 1989, vrt. Fischer & Forester 1993) taustalla voi erottaa määrättyjä yhteiskuntatieteellisiä lähtökohtia. Etnometodologiassa (Heritage 1986) omaksuttiin jo vuosia sitten fenomenologisesta filosofiasta ajatus, että sosiaalinen todellisuus

konstruoituu pienissä ihmisten välisissä vuorovaikutusepisodeissa. Niissä muodostuu käytäntöjä, joista on vaikea poiketa. Muuten määrittäyttyä oudoksi, osaamattomaksi, erilaiseksi tai huonosti käyttäytyväksi. Professori ja opiskelija yliopistossa, evaluattori neuvotellessaan mahdollisen toimeksiantajan kanssa tai haastattellessaan evaluatation kohteen edustajia, palveluhenkilö ja asiakas keskinäisessä kontaktissaan tai lapsi suhteessa vanhempiin on vahvojen mutta kirjoittamattomien sääntöjen alainen. Vankkumatoman tiukkoja nuo säännöt eivät kuitenkaan ole. Siitä todistaa niiden syntytyapa: ne kehkeytyvät julkilausumattomissa "neuvotteluissa" eri osapuolten kesken siten, että vähitellen muodostuu vakiintunut käsitys siitä, mikä on asianmukaista.

Toisin kuin Pawson ja Tilley (1997) esittävät, heidän realistiseksi kutsumansa evaluatio (Pawson & Tilley 1997) näyttäytyy synteessä kolmesta muusta lähestymistavasta. "Realistinen" evaluatio vaikuttaa itse asiassa hyvin pragmatistiselta, vaikka se on saanut oman nimityksensä. Mitä ilmeisemmin Pawson ja Tilley ovat Anthony Giddensin (alkaan 1979, ks. myös esim. Giddens 2000) lukeneet. Se, mitä Giddens on halunnut olla Tony Blairin "kolmannelle tielle", sitä Pawsonin ja Tilleyin evaluattorit haluavat olla evaluattorin rooleissaan: opettajia, auttajia ja asioiden palasten kokoajia.

### 3.4 *Pawsonin ja Tilleyin jaottelun arviointia*

Pawsonin ja Tilleyin jaottelua voi pitää yhtäältä kirjoittajien tahtomattaan pirtämänä karikatyyriä, toisaalta sepitteenä, joka korostaa kirjoittajien parhaana pitämän lähestymistavan vahvuuksia. Missään tapauksessa nuo kaksi kirjoittajaa eivät esitä koko totuutta evaluatiosta ottaen huomioon kaikki filosofiset lähestymistavat.

Pawsonin ja Tilleyin tarkastelu tekee miten kuten oikeutta kvasikokeelliselle evaluatiolle, joskin he esittävät sen pelkistäen. Pragmatistinen evaluatio jää Pawsonilla ja Tilleyllä sisällyttäen vaille vankkaa tarkennusta ja rajoitetaan epämääräiseksi. Yhtä paljon kuin evaluatation lähestymistavasta, pragmatistisessa evaluatiossa näyttäisi kysymys olevan mistä tahansa evaluatation lähestymistavasta silloin, kun sitä sovelletaan hyvin politisoituneissa olosuhteissa. Lähestymistavan määrittämisessä on yhtä paljon kysymys evaluatiota ehdollistavista olosuhde-

tekijöistä kuin evaluaation varsinaisen intellektuaalisen ytimen mistään osasta.

Sellaisenaan pitää paikkansa, että Guban ja Lincolnin laajalti huomattu tarkastelu "neljännen sukupolven evaluaatiosta" rakentuu konstruktivismina kutsutulle ajattelulle. Toisaalta konstruktivismi kattaa vain rajatun osan ymmärtävien otteiden sovelluskentästä yhteiskuntatieteissä. Guban ja Lincolnin teoksen ilmestymisen jälkeen kuluneina 12 vuotena tuo sovelluskenttä on laajentunut, jäsentynyt ja muutenkin muuttunut. Eri-laiset "laadulliset" otteet ovat kiteytyneet paljon pidemmälle kuin vielä 1980-luvun lopussa. Tämä on oleellista myös evaluaation kentällä.

Pawsonin ja Tilley'n tarkastelua voisi tarkentaa täsmentämällä, millaisissa yhteyksissä mikin evaluaation neljästä filosofialähtöisestä otteesta toimii parhaiten ja millaisissa yhteyksissä niitä puolestaan tulisi välttää. Kvasikokeelliselle asetelmalle löytyy edelleen soveltamisyhteyksiä. Pawson ja Tilley tuskin saavat terveydenhuoltoa tai koulutus-alaa lakkauttamaan kvasikokeellista tai edes kokeellista evaluaatiotoimintaansa. Moni muukin yhteiskunta- ja talouspolitiikan ala on sekä niin pitkälle rakenteistunut että siinä määrin suljettu, että edellytykset kvasikokeille vallitsevat. Myös konstruktivistisella evaluaatiolla on vahvuutensa määräyhteyksissä. Niin sanotussa ihmistyössä ja siihen liittyvissä intensiivisissä palvelutilanteissa, -episodeissa ja -prosesseissa "realistinen evaluaatio" saattaa osoittautua mitä epärealistisimmaksi verrattuna konstruktivistiseen evaluaatioon tai jonkin muun ymmärtävän ja laadullisen evaluaatio-otteen sovellukseen.

"Realistinen evaluaatio" sijoittuu Blairin Britanniaassa hallitsevan evaluaatioideologian ytimeen. "Pragmatistinen evaluaatio" ei kuitenkaan ole menettänyt merkitystään Euroopan maissa, joilta puuttuu kaksipuoluejärjestelmä ja siitä juontuva yhden hallitsevan puolueen hegemonia, kunhan se säilyttää parlamenttienemmistönsä. "Pragmatistinen evaluaatio" on osin artifakta: määrätyt olot saavat tietyn lähestymistavan näyttämään ylivertaiselta, vaikka tuon yliveraisuuden pätevyys ehdot olisivat itse asiassa kapeat. Niin oli myös thatcherismin aikana vallatessa läntistä maailmaa laajemmalti kuin pelkkä taloudellinen stagnaatio selitti. Tuollaisten artifaktojen vaikutusta voimistaa englannin kielen yhä voimistuva ylivalta.

## 4 EVALUAATION YTIMESSÄ: EVALUAATION KÄYTÄNNÖN TEORIOIDEN NÄKÖKULMA

### 4.1 Skemaattinen tarkastelu

Mikä tahansa jäsentely, malli, skeema tai "sabluna" on kaksiteräinen miekka. Tämä koskee myös evaluaation "käytännön teorioita". Yhtäältä kukin tuollainen jaottelu leikkaa kohdetta, jota on tarkoituskin leikata, mutta toisaalta jaottelu kapeuttaa tarkastelun alaa.

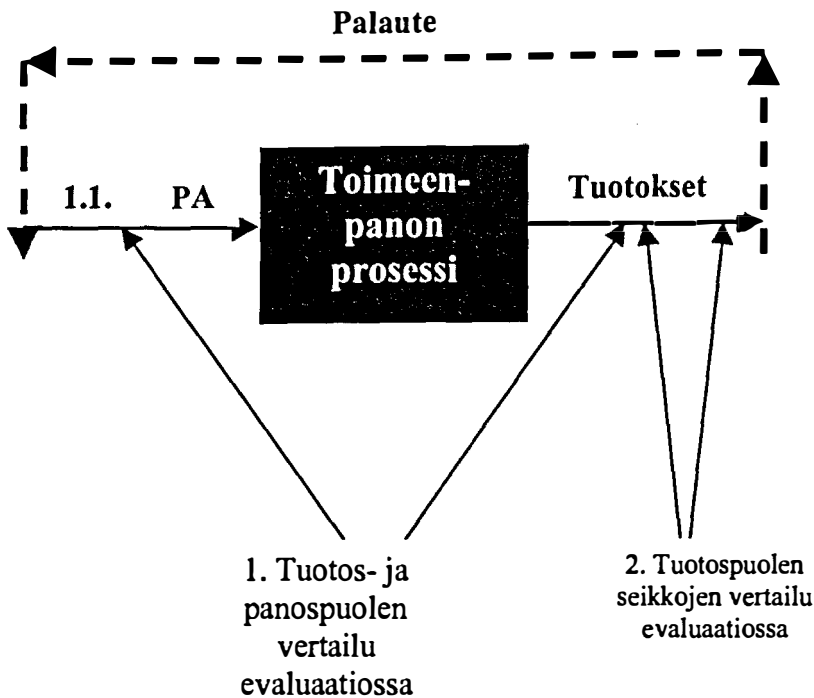
Esitän skeeman (Kuvio 1), joka on pelkistetty malli tietystä "järjestelmästä". Kuten vastaavissa skeemoissa yleensä, malliin sisältyy (1) "musta laatikko", jonka sisällöstä ei oleteta tiedettävän mitään. "Musta laatikko" on toimeenpanoprosessi sen sisäisiltä, tarkastelussa näkymättömiin jääviltä osiltaan. Päähuomio suuntautuu (2) syötteisiin eli panoksiin, (3) tuotoksiin sekä (4) palautteeseen tuotoksista tuleviin panoksiin. Skeemaan sisältyy myös viittaus kahteen tyyppiin (5) vertailuja, jotka ilmaisevat skeeman hahmotetun evaluaation ytimen. Yhtäältä vertaillaan toisiinsa (5.1) tiettyjä panos- ja tuotospuolen seikkoja, toisaalta (5.2) tiettyjä tuotospuolen seikkoja.

### 4.2 Panokset ja tuotokset

Panospuolta tarkasteltaessa otetaan usein huomioon (a) tarpeet ja muut olosuhteet, joihin vaikuttamiseksi koko prosessi kehkeytyy. Toisinaan on tarpeiden ohella tai asemasta on otettu esiin (b) vaateet (demands), joilla yksilöt, ryhmät tai muut tahot lausuvat julki mitä haluavat. Tämä tapahtuu vaaleissa, julkisen mielipiteen ilmaisuihin, parlamenttidebateihin tai hallituspuolueiden aktiivisuutena. Se tapahtuu myös esimerkiksi palvelujen piiriin hakeutuvien määrrien tai palvelujen käyttäjälukujen muutoksina. Julkinenkin palvelu, jolta ehtyvät käyttäjät, on vaikeuksissa jo ennen kuin sen tarjonnan poliittinen tuki ehtyy. Panoksena esiintyvät erittäin usein (c) tavoitteet (engl. objectives). Samoin esiintyvät tarkoitukseen osoitetut (d) voimavarat eli resurssit.

(e) Toimeenpano jää "mustaksi laatikoksi". Toimeenpanosta kumpuavat tuotokset. Välittömimminkin niitä ovat (f) ensi vaiheen tuotokset, outputs. Niitä seuraavat vaikutukset, jotka saatetaan jakaa (g) tuloksiin (results) ja (h) varsinaisiin vaikutuk-





**Kuvio 1. Skemaattinen esitys eräistä käytännön evaluaatioviitekehyksistä.**

siin (esim. impacts). Esimerkiksi koulun tuotoksena saattaa olla opettajien opetusvelvollisuuden täyttyminen ja tietyn määrän oppilaita läsnäolo eri oppiaineiden tunneilla sekä heidän osallistumisensa määrättyihin kokeisiin. Koulun tuloksia saattavat olla päästötodistuksiin saakka suoritettut oppikurssit, menestys valtakunnallisissa kokeissa sekä todistuksia saaneiden pääsy seuraavien koulutusvaiheiden laitoksiin. Vaikutuksia saattaisivat olla esimerkiksi ne kansantaloudelliset lisähyödyt, joita yleisen koulutustason nousu vähitellen saa aikaan.

Lisäksi vaikutuksista on usein kiinnitetty huomiota (i) negatiivisiin sivuvaikutuksiin ja (j) positiivisiin sivuvaikutuksiin. Negatiivisia sivuvaikutuksia ovat esimerkiksi kasvi- ja eläinpopulaatioiden ja maisemien tuhoutuminen tie- ja vesirakennustöiden vuoksi, julkisten toimenpiteiden haitalliset kannustinvaiikutukset, yllättävät vallansiirtymät hallinnonuudistuksissa tai toimenpiteiden vaikutukset sukupuolten asemiin työelämässä. Positiivisista sivuvaikutuksista voidaan mainita oppiminen ja innovatiiviset ratkaisut, jotka

sittenkin kenties halutaan vakiinnuttaa pysyviksi osiksi toimintaa. Oma, erityinen vaikutuksensa, joka saatetaan ottaa huomioon, on (k) palaute (feedback) vaikutuksista seuraavan toimintakierroksen panoksiin. Ahkeroina työntekijä voi luottaa laiskottelevaa paremmin siihen, että saa säilyttää työpaikkansa. Vaikuttavuudeltaan, tuloksiltaan tai tuotoksiltaan esimerkillinen organisaatio tai organisaation tulosyksikkö johtoineen voinee olettaa, että rahoitusta saadaan jatkossakin. Maata tai kuntaa menestyksellisesti luotsannut poliittinen taho voi olettaa kannatuksensa jatkuvan myös seuraavissa vaaleissa suurempana kuin epäonnistumisesta toiseen rämpinyt vastineensa.

#### 4.3 *Evaluaatiokriteerit tehokkuus ja vaikuttavuus*

Evaluaatioissa arvioidaan tarkasteltuja kohteita erityisesti määritellyillä kriteereillä. (A) Tehokkuus (engl. efficiency, ruotsiksi jostakin syystä produktivitet) esiintyy kriteerinä eri nimillä ja eri

muodoissa, kuten tuottavuutena, säästäväisyytenä (economy) eli kustannusten minimointina. Julkisellakin sektorilla se saattaa joissakin tapauksissa esiintyä liiketaloudellisena kannattavuutena sen jossakin muodossa kuten voittona tai sijoitetun pääoman tuottoasteena. Tehokkuus merkitsee ensi vaiheen tuotosten suhteuttamista niihin voimavaroihin, joilla nuo panokset saatiin aikaan. Kaavana ilmaisten on kysymys osamäärästä

$$(I) A=f/d$$

(B) Vaikuttavuus (engl. effectiveness, ruots. effektivitet) on useimmiten esitetty joko tulosten tai vaikutusten suhteena siihen, mitä on edellytetty tavoitteita asetettaessa eli jompana kumpana suhteista

$$(II) B(1)=g/c$$

$$(III) B(2)=h/c$$

Edellisten lisäksi tai asemasta vaikuttavuus on saatettu esittää jonakin erilaisten tuotospuolen tekijöiden välisistä suhteista tarkastellen vastavasti, mitä vaikutuksia on tuloksilla tai tuotoksilla tai mitä tuloksia on tuotoksista eli

$$(IV) B(3)=h/g$$

$$(V) B(4)=h/f$$

$$(VI) B(5)=g/f$$

On myös mahdollista, että varsinaisia vaikutuksia (impacts) on korjattu vähentämällä niistä tavalla tai toisella negatiiviset sivuvaikutukset mutta lisäämällä niihin positiiviset sivuvaikutukset. On toisin sanoen sovellettu korjauskaavaa

$$(VII) h' = h-i+j,$$

Korjauskaavaa soveltamalla saadaan kolme tarkennettua kaavaa kolmelle vaikuttavuuskriteeristä:

$$(VIII) B(6)=h'/c$$

$$(IX) B(7)=h'/g$$

$$(X) B(8)=h'/f$$

On syytä tarkkuuteen, kun on puhe vaikuttavuudesta, koska se osoittautuu monikäsitteiseksi. Joissakin evaluaation kielimurteissa tulee kysymykseen tuloksellisuuden eriyttäminen vai-

kuttavuudesta. Tuolloin viitataan lähinnä siihen, mitä on edellä ilmaistu kaavoilla (II) ja (VI). Tietty tuloksellisuuskriteeri on myös kustannusvaikuttavuus eli lähinnä:

$$(XI) B(9)=h/d$$

$$(XII) B(10)=h'/d$$

#### 4.4 Muita evaluaatiokriteereitä

Evaluaatiossa, joka koskee (C) relevanssia, kysymys on jommasta kummasta vertailuista

$$(XIII) C(1)=h/a$$

$$(XIV) C(2)=h'/a$$

Kysymys on saavutettujen viime käden vaikutusten tai viime käden nettovaikutusten vertailusta tarpeisiin, joita tulisi tyydyttää. Riittävyys on eräs mahdollinen synonyymi relevanssille silloin, kun sekä vaikutus h tai h' sekä tarpeet a on ilmaistu määrällisinä suureina. Lisäksi riittävyttä toisaalta voidaan hahmottaa myös erällä kaavoista, joilla luonnehditaan alla myös utiliteettia, kuten kaavoilla

$$(XV) C(3)=h/b$$

$$(XVI) C(4)=h'/b$$

(D) Kestävyys (sustainability) tarkoittaa toimenpiteellä aikaan saatujen vaikutusten pysyvyyttä pidempään kuin vain jonkin määrätyn, yleensä lyhyen ajan. Tuo lyhyt aika saattaa olla jakso, jonka kuluessa tarkoitukseen on kohdennettu ja käytetty tilapäisesti saatavissa olleita lisävoimavaroja. Tuolloin kestävyys on yhtä kuin tuon kohdennuksen ehdyttyä kenties esiintyvä "pysyvä" parannus jonkin tavoitellun suureen arvossa. Toisaalta kestävyys saattaa tarkoittaa fiskaalista eli jonkin julkisen talouden kestävyyttä. Kysymys on siitä, tapahtuuko julkisen talouden tilassa parannuksia, pysyykö tila ennallaan tai pysyvätkö huononukset siedettävissä rajoissa. Joissakin hankkeissa tai toimintaohjelmissa, kuten tie-, kaavoitus- tai energiahankkeissa, saattaa korostua ekologinen kestävyys. Juuri tuosta aihepiiristä sana "kestävyys" on alun perin peräisin tarkasteltaessa uusiutuvien luonnonvarojen varantojen järkipäistä, kestäväää käyttöä.

Organisaation sisäisestä näkökulmasta kestävyys määrittyy jossakin määrin toisin kuin edellä. Missä suhteessa organisaatiolle sen ulkopuolelta käsin asetettavat tavoitteet ovat organisaation toiminnan toteutuneisiin vaikutuksiin tai tuloksiin? Missä suhteessa ulkopuolelta saatavat voimavarapanokset ovat organisaation tuotoksiin, tuloksiin tai vaikutuksiin? Tällaisissa tapauksissa organisaatio arvioi, missä määrin sitä palkitaan sen saavutuksista ja laskelmoi, kuinka kannattaa toimia. Sisäisen rationaalisuutensa kannalta organisaation on perusteltua olla piittaamatta toimintansa vaikutuksista, mikäli sitä palkitaan voimavaroin vain tuotoksista tai tuloksista. Tulosten määrän kasvattaminen ei ole organisaatiolle rationaalista, mikäli organisaatio silloin saa yhä vaikeammin saavutettavia tulostavoitteita. Tämän ongelman tunsijokainen tehtaanjohtaja suunnitelmataloudessa, ja sen tuntee jokainen suomalaisen nyky-yliopiston rehtori, dekaani, laitosjohtaja ja oppiainevastaava professori (vrt. Milgrom & Roberts 1992). Vastaavia kaavoja ovat esimerkiksi:

$$(XVII) D(1)=c/h$$

$$(XVIII) D(2)=c/g$$

$$(XIX) D(3)=d/h$$

$$(XX) D(4)=d/h'$$

$$(XXI) D(5)=d/g$$

$$(XXII) D(6)=d/f$$

(E) Utiliteetin eli "hyödyn" kriteeri esiintyy myös määrätyissä evaluaatioissa. Kaavoina ilmaisten kysymys on jostakin vertailuista

$$(XXIII) E(1)=h/a$$

$$(XXIV) E(2)=h/b$$

$$(XXV) E(3)=h'/a$$

$$(XXVI) E(4)=h'/b$$

Esimerkiksi soveltuu terveydenhuoltoalan QALY-mittaus (quality adjusted life years; ks. esim. Rissanen ym., 1995). Oletettakoon, että hoito-, palvelu- ja prosessi-innovaatioiden seurauksena kansalaisten keskimääräinen elinlänodote kasvaa. Haasteeksi jää evaluoida,

vastaako tätä "objektiivisin" toimintakykyisyyden mittarein mitattavissa oleva tai asianomaisten itsensä arvioima elämisen laadun pysyminen siedettävänä keskimääräisen elämän pituuteen voitettujen lisävuosien aikana. Kuvaavasti joskin epätarkasti asia on ilmaistu kysymyksellä, tuleeko "elämään lisää vuosia vai vuosiin lisää elämää". Käytännössä QALY-mittaukset ovat toisaalta usein enemmänkin kustannus-vaikuttavuus-analyysia (Rissanen ym., em.), jolloin menettelytapaa hahmottavat kaavat (XI) ja (XII).

Arviointikriteerit eivät tyhjene tähän. Esitetyt kriteerit eivät tyhjentävästi kata esimerkiksi niitä seikkoja, joita (F) palvelujen saatavuuskriteerit pyrkivät tavoittamaan. Vastaavien kaavojen osoittajaan kuuluu sellaisia suureita kuin vaikutukset, tulokset tai tuotokset ja nimittäjään suureita, joissa on kysymys vaikutuksen, tuloksen tai tuotoksen jakaantumisesta esimerkiksi alueellisesti, ikäryhmittäin tai väestön sosioekonomisten ryhmien mukaan. Merkittäköön jakaumasuuretta yleisesti kirjaimella m, jolloin saadaan

$$(XXVII) F(1)=h/m$$

$$(XXVIII) F(2)=h'm$$

$$(XXIX) F(2)=g/m$$

$$(XXX) F(3)=f/m$$

#### 4.5 Temaattisen evaluaation kriteerejä

Sateenvarjonimitystä temaattiseksi evaluaatio, thematic evaluation, käytetään viittaamaan huomattavaan joukkoon evaluaatiotapoja, jotka perustuvat niille ominaisiin kriteereihin. (G) Sukupuolten tasa-arvo on usein käytetty temaattisen evaluaation kriteeri. Kriteerin kiperyys käy esiin, kun pitäisi arvioida tasa-arvon toteutumista yhtäältä miesvaltaisilla aloilla kuten diplomi-insinöörikkoulutuksessa tai naisvaltaisilla aloilla kuten sairaanhoitajakoulutuksessa. Olisiko otettava käyttöön sukupuolikiintiöt vai peräti muutettava koulutussisältöjä tasa-arvon nimissä? Terveydenhuollossa on puolestaan kysytty, saavatko "miessairaudet" kohtuutonta huomiota verrattuna "naissairauksiin"? Toisaalta kuinka arvioida sairaalapalvelujen tasa-arvoa vallitsevassa tilanteessa, jossa keskeinen käyttäjäryhmä koostuu iäkkäistä naisista, joiden useimmat

miespuoliset ikätoverit ovat jo manan majoilla? Formaalisissa mielessä tasa-arvokriteerit vastaavat esitettyjä saatavuuskriteerejä, joita koskevat kaavat (XXVII)-(XXX).

Toimenpiteiden vaikutukset sukupuolten tasa-arvoon saattavat olla tärkeän, silmällä pidettävän sivuvaikutuksen asemassa silloinkin, kun tasa-arvon edistäminen ei ole varsinaisena tavoitteena. Tuolloin ajatustapana on se, että negatiiviset vaikutukset sukupuolten tasa-arvoon pitäisi tavallaan vähentää saavutetusta vaikutuksesta ja positiiviset tasa-arvoaikutukset lisätä varsinaiseen vaikutukseen korjauskaavan (VII) mukaisesti. Sukupuolten tasa-arvokriteerejä muistuttavat kriteerit, jotka koskevat sukupuoliselta suuntautumiseltaan, etnisiteitiltään, äidinkieltään tai elämänvaiheeltaan erilaisia henkilöitä.

(H) Temaattisen evaluaation kriteeri on myös vallallistaminen tai "valtaistaminen", englanniksi empowerment. Perusajatuksena on, että huonossa asemassa olevia autetaan ottamaan vähitellen yhä lisää vastuuta olojensa ja asemansa parantamisesta. Vallallistamisessa tavoitteena voi olla parantaa esimerkiksi kehitysmaiden tai kehittyneiden maiden syrjäseutujen tai syrjäytyneiden väestönosien asemaa. Mikään ei estä rakentamasta vallallistamista ja esimerkiksi sukupuolten tasa-arvon edistämistä toistensa lomaan ja sitten evaluoimasta huono-osaisemman sukupuolen ja sen jäsenten vallallistamista. (Ks. Fetterman & Kaftarian 1996.)

Monet temaattisen evaluaation kriteerit saavat sisältönsä erityisten käyttöyhteytensä mukaan. Tuollainen kriteeri on EU-yhteyksissä mutta myös kansainvälisessä kehitysyhteistyössä käytetty (I) kumppanuus (Kelleher ym. 1999), (J) EU-yhteyksissä sovellettu subsidiariteettiperiaate eli läheisyysperiaate sekä kunnallisen demokratian laajentamishankkeissa esiintyvä (K) osallisuuden periaate, jolla on tarkka vastine vain ruotsinkielessä, delaktighet (Salmikangas ym. 1999). Hieman naiivisti on tähdennettävä, että kumppanuuskriteerin ja subsidiariteettiperiaateen sisältö niiden soveltamisyhteydessä määrittäyty tarkoin sen mukaan, mitä EU-lainsäädäntö ja muu EU-norminanto määrää. EU-kumppannus merkitsee vain sitä, että toimintaohjelmien hallintoelimiissä on jäsenenä määrättyjen intressitahojen jäseniä eikä sitä, että kumppanuuden tarvitsisi koskea myös ohjelmien kohteena olevia tahoja. Subsidiariteettikaan ei merkitse muuta kuin sitä, että

EU-toimielimet ovat periaatteessa toimivaltaisia silloin ja vain silloin, kun jäsenvaltioiden hallitukset eivät kykene hoitamaan jotakin asiaa.

Määrätyissä evaluaation yhteyksissä esiintyy vielä muita evaluaatiokriteereitä. Niitä ovat esimerkiksi kehitysyhteistyön evaluoinneissa (L) koherenssi eli eri politiikkatoimien keskinäinen sopivuus, (M) kongruenssi eli toimintaohjelman sovittavuus yhteen poliittisten periaatteiden ja ohjelmien kanssa sekä (N) hyvä hallinta. Viimeinen kriteereistä on englannin kieliseltä nimitykseltään good governance. Sen virallinen EU-käännös "hyvä hallinto" on virheellinen, koska kysymys on monimutkaisesta yhdistelmästä kansanvallan ja oikeusvaltiota luonnehtivan hallinnon piirteitä. Governancen hyvä käännös "hallinnaksi" esiintyi politiikan tutkimuksessa ja hallintotieteissä jo 1980-luvulla mutta unohtui vailettavasti 1990-luvun virallisesta EU-suomesta.

#### 4.6 *Evaluaation käytännön teorian huokuva talo*

Evaluaatiota koskevan jäsenyyksen kohteet menettävät selkeytensä, kun asiaan paneudutaan syvemmin. Olen jo tuonut esille tehokkuuden moniselitteisyyden. Taloustieteelliset tarkastelutavat ovat koettaneet ratkaista vastaavat ongelmat sulkeistamalla tarkastelunsa ulkopuolelle sinänsä tärkeitä seikkoja turvataksaan formaalien mallien laadinnan ja eksaktit empiiriset tarkastelut. Pahimmassa tapauksessa seurauksena on pyrkimys tehdä maailmasta sellainen kuin maailma on mallien mukaan. Valitettavasti tästä on laajaa näyttöä, ei vain talouspolitiikassa ja muillakin politiikkalohkoilla. Maailmaa pakotetaan malliin eikä mallia soviteta maailmaan, olipa mallissa kysymyksessä monetaristisesta rahapolitiikasta, verotuksen ja julkisten siirtomenojen oletetuista yhdistetyistä kannustinvaikutuksista tai yrityksen laskentatoimen sovelluksista julkisen sektorin järjestämis- ja rahoitusvastuulla olevassa palvelutoiminnassa. Uhkana on fundamentalismi, se, että periaatteet ovat tärkeämpiä kuin se empiria, jota esimerkiksi evaluaatiot tuovat esiin.

Vaikuttavuus osoittautui vieläkin monikäsitteisemmäksi kuin tehokkuus. Monikäsitteinen on myös utiliteetti: uskotaanko objektiivisena esiintyvän mittaamisen olevan oikeutettua vai hyväksyttävääkö vain se, että asiaa kysytään asianomaisilta itseltään. Tämä seikka kuuluu yhteisöetiikan, kansanvallan ja hyvinvointivaltion peruskysymyk-

siin. Sen enempää kestävyys ei vaikuta yksiselitteiseltä.

Epäselvyyttä lisää edelleen se, että tavoitteita voidaan johdannaiseen tapaan asettaa paitsi vaikutuksille, tuloksille ja ensi vaiheen tuotoksille, myös tuottavuudelle, kannattavuudelle, yksikkökustannuksille, siedettäväksi katsottaville kielteisille sivuvaikutuksille, myönteisille sivuvaikutuksille, kestävyydelle, relevanssille, utiliteetille, saatavuudelle ja sukupuolten, sukupuolisten suuntausten, etnisiteettien ja lingvististen ryhmien tasa-arvolle, vallallistamiselle ja muillekin seikoille. Sitä paitsi tavoitteita saatetaan asettaa paitsi optimoinnin mielessä pyrkien parhaaseen mahdolliseen lopputulokseen, lähtien myös jostakin vaatimattomammasta tavoitetasosta.

Toimeenpanon pysyttely useissa evaluaatiomalleissa tuntemattomana "mustana laatikkona" kiinnittää huomiota. Eikö "laatikkoa" avaten voisi evaluoida toimeenpanon sisäistä tehokkutta, vaikuttavuutta, relevanssia, kestävyttä ja utiliteettia sekä tasa-arvoa ja vallallistamista toimeenpanon yhteyksissä? Paljon vastaava onkin tehty (esim. Pressman & Wildavsky 1984, Teittinen 1985, Kettunen 1994, King ym. 1998). Toisinaan laatikko on sellaisenaan evaluaation kohteena ollessaan kehittämistyön piirissä tai ollessaan tulossa kehittämisvaiheeseen. Evaluaation "käytännön teorioita" tämä ei kuitenkaan yksinkertaista vaan päin vastoin: nuo teoriat monitukaistuvat jälleen yhden monikerran lisää.

## 5 EVALUAATION PLATONIA MUKAILEVA ROOLIPELI

Evaluaation aihepiiriin kuuluu laajassa mielessä eettisen ja moraalisen ajattelun alueeseen. Tuolla alueella ihmiskunnan ajattelu on saattanut suorastaan taantua viimeksi kuluneiden 2 500 vuoden aikana. Evaluaatioviisauden tilan tutkinta tuskin todistaa muusta. Platonin (Valtio, 439e-442a) esittämä jaottelu mukautuu hyvin evaluaattoriroolien pelin tarkasteluun.

### 5.1 *Ahnas käsityöläisevaluaattori*

Suuri osa evaluaation kentästä käytäntöineen ja instituutioineen on luonteeltaan teknistä. Tällöin evaluaattori toimii aivo-, käsi- ja puhetyöläisenä henkensä pitimiksi. Hänen on hallittava

kohdealueidensa lainsäädäntöä ja kirjoittamattomia käytäntöjä, oltava perillä kunkin kohdealueen henkilögalleriasta, nokkimisjärjestyksestä ja muista valtasuhteista sekä siitä, minkä seikkojen hallinta on osoitettava, jottei leimautuisi epäajan kohtaiseksi ja taidottomaksi.

Evaluaattorin on tunnettava evaluaation teorioita, alan käsitepuolta sekä sen metodeja, tekniikoita ja aineistoja. Hänen on hallittava evaluaation tekniikka tarjouksiin vastaamisesta evaluaati-osuunnitelmien tekemiseen, tiedon kokoamiseen ja analysointiin sekä raportointiin saakka. Muuten ei evaluaatiomarkkinoiden "näkyvätön käsi" uusinna evaluaattorin ammattillista mainetta siten, että hänestä tulisi pysyvä toimija ainakaan noilla markkinoilla.

Evaluaatiotyöläisen taidot ovat standardisointia sekä korkeakoulujen evaluaatio-opetuksessa että "käytännön ihmisille" kohdistetussa evaluaatiokoulutuksessa. Toisaalta evaluaattorille saattaa jo "käsityöläisen" roolissa muodostua pelkkää teknistä syvällisempää ja monipuolisempaa ymmärrystä ja hallintaa siitä, mitä hän tekee. Hän saattaa kyetä muodostamaan alustavaa evaluaationäkemyistä antoisien tai karujen evaluaatiokokemustensa pohjalta, kehittämään tilanneherkkyyttä sekä kehittämään monimutkaisuuden ja ajoittaisen kaaoksen sietokykyä. Muuten hän jää "nälkäiseksi vatsaksi" ilman "päätä": pelkkään välineelliseen rooliin sekä itselleen että muille.

### 5.2 *Urhea evaluaattorisoturi*

Syvämmälle etenevä evaluaattori saattaa kehittyä "ajattelevaksi praktikoksi" (vrt. Schön 1984). Kysymys ei ole pitkäaikaisen kentällä olon tuomasta "vakaumuksen rintaäänestä" vaan kyvystä käsitellä sitä huomattavaa, toistuvaa, alituista ja vakavaakin kaaosta ja ajoittaisia absurditeetin tuntoja, joita evaluaatioon saattaa liittyä. Tämä koskee sekä evaluaation kohdekenttää, sen prosessia, tulosten raportointia ja käyttöä että uudelleenevaluaatiota ja kokonaan uusia evaluaatioita. Vastaavan evaluaattorirooliin rinnastus soturin rooliin sodan melskeiden kaaoksessa ei ole kaukaa haettua.

Syvällinen kyky oppia omista virheistä eksistentiaalisesti paikoittuu evaluaattorin rooliin reflektioivana praktikkona. Rooli ei ole puoluepoliittinen eikä poliittinen muun puolueellisuuden mielessä.

Sen sijaan se on poliittinen sikäli, että ajattelevana praktikkona evaluaattori on oppinut ottamaan tilaisuuksista vaarin oppien sietämään ja käyttämään hyväkseen kontingenssejä. Kontingenssin filosofisessa ongelmassa on kysymys siitä (Rorty 1989, Palonen 1998), että jokin oleva voisi olla toisinkin tai kokonaan olematta, asia joka tapahtuu olisi voinut jäädä toteutumatta ja odotetut tapahtumat saattavat jäädä tapahtumatta. Aina eivät ennako-odotukset pidä paikkaansa; evaluaation asiakas ja kohde osoittautuvat toisenlaiseksi kuin odotti, jokainen odotettu evaluaatiohanke ei toteudu eikä kukaan eikä mikään toteuttaja voita kilpailua jokaisesta evaluaatiohankkeesta. Kontingenssi kuuluu niihin avaintermiin, joilla filosofit ovat hahmottaneet jälki- tai myöhäismodernia ajan henkeä.

Evaluaattorin on "ajattelevana praktikkona" kyettävä selviämään myös tilanteista, joissa evaluaatio ei tavallaan ole evaluaatiota lainkaan vaan ainoastaan jonkinlainen naamio, naamiaispuuku tai kulissi. Evaluaatio saattaa esimerkiksi olla pääasiassa valtapoliittista symboliikkaa, jossa itsearviointit ovat ylhäältä käsin pakotettua tunnustavaa itseruoskintaa. Valtapoliittisissa yhteyksissä arvovaltaisetkaan evaluaattorit eivät itse asiassa evaluoi vaan toimivat äänitorvena vahvistaen sitä "isännän ääntä", jonka takana on evaluaation maksaja. Silloin evaluaatioeuroja ei jaeta parhaille vaan sopivimmille, vahvimille, sopuisimmille ja vakiintuneimmille evaluaattoritahoille. On ikävää joutua itse pelkäksi kohteeksi symbolisille evaluaatioille, jotka toteutetaan huolimatta siitä, että ratkaisut on jo tehty ja evaluaation tehtävänä on kenties vain koettaa jälkikäteen oikeuttaa ratkaisuja, joiden takana valta seisoo. Tuolloin evaluaattori tuskin selviää omatunto puhtaana.

Ajattelevana praktikkona evaluaattorin on opittava luopumaan. Hänen on kyettävä poisoppimaan omista, rakkaudella vaalimistaan käsityksistä. Hänen on opittava olemaan havittelematta hankkeita, joista hänen olisi parasta pysyä erossa. Hänen on oltava valmis keskeyttämään hankkeet, jos se on lain, hyvien tapojen, oman ja muiden edes minimaalisen maineen sekä muiden tilannesidonnaisten seikkojen rajoissa mahdollista. Ellei evaluaattori tuota kaikkea opi, hän jää "teräväksi miekaksi", jota sopii heilutella ja jolla sopii iskeä, kunhan ei iske omaa itseään. Pahimmassa tapauksessa hän ei huomaa luopua ajoissa evaluaattorin roolista,

vaikka jokaisen aika tulee siihenkin jossakin vaiheessa.

### 5.3 Viisas filosofievaluaattori

Kun jo evaluaatiotyöläisen ja ajattelevan evaluaatiopraktikon roolit ovat vaikeita ja toisinaan ylivoimaisia, kaikkein vaativin on jaloön asioiden tarkasteluun kykenevän evaluaattorin rooli. Tähän saakka edennyt evaluaattori omaa kirkkaan käsitteellisen hallinnan evaluaation kysymyksistä, ilmenivätpä ne tutkimuksen, metodien ja tekniikoiden tai käytäntöjen haasteina. Kehittyneimmillään evaluaattori tosiaan tietää, "mistä on kysymys" ja "mitä mikäkin on".

Jaloon asioiden tarkasteluun kyetessään evaluaattorilla on kiteytynyt mutta joustava ja muuntumiskykyinen näkemys siitä, mitä evaluaatio on ja mikä on evaluaattorin oma osa. Evaluaattorilla on myös kehittynyt kyky arvioida joka hetki itseään, asemaansa ja toimintaansa evaluaattorina sekä evaluaatiotyötään sen kehkeytymisen ja käytön eri vaiheissa, mukaan lukien evaluaatiotulosten soveltamisessa evaluaation kohteissa.

Evaluaattori "ajattelevana praktikkona" kokee evaluoitavan maailman, evaluaation ja oman itsensä kontingenssit. Jaloön asioiden tarkasteluun kykenevä evaluaattori näkee evaluaatioon, sen kohteisiin ja evaluaattorina toimintaan sisältyvät mahdollisuudet niine valinnan haasteineen, joita evaluaattori kohtaa ja joihin hänen on vastattava. Vasta vaativimmassa roolissaan evaluaattori on vapaa, mutta hän ei silloinkaan ole vapaa valitsemaan muuta kuin hyvän ja pahan välillä. Hyvän ja pahan välillä valitseminen evaluaattorin tehtävässä on vaikeaa, eikä kukaan muu voi kertoa hänelle, mikä milloinkin on hyvää ja mikä pahaa.

### LÄHTEET

- Ahonen, Pertti: Learning by using the project cycle management method (PCM): general considerations, and user experiences in a small donor country-transitional country context. *Impact Assessment and Project Appraisal*, 17(1999): 2, s. 97-105.
- Birnbaum, Robert: Management fads in higher education: where they come from, what they do, why they fail. Jossey-Bass, San Francisco 2000.
- Blalock, Hubert M.: Causal inferences in nonexperimental research. University of North Carolina Press,

- Chapel Hill 1964.
- Campbell, Donald T.: Reforms as experiments. *American Psychologist*, 24(1969), s. 409-429.
- Campbell, Donald T. & M. Jean Russo: *Social experimentation*. Sage, London 1998.
- Campbell, Donald T. & Stanley, Julian C.: *Experimental and quasi-experimental designs for research*. Rand McNally, Chicago 1963.
- Cheetman, Juliet, Roger Fuller, Gill McIvor & Alison Petch: *Evaluating Social Work Effectiveness*. Open University, London 1992.
- CEC (Commission of the European Communities): *Evaluation of expenditure programmes*. CEC, Brussels 1997.
- Clarke, Alan & Ruth Dawson: *Evaluation research*. Sage, London 1999.
- Collins, David: *Management fads and buzzwords: critical-practical perspectives*. Routledge, London 2000.
- Dreyfus, Hubert L.: *Being-in-the-world: a commentary on Heidegger's Being and time, Division I*. MIT Press, Cambridge, MA 1991.
- Drummond, Michael & Alistair McGuire: *Economic evaluation in health care*. Oxford: Oxford University Press 2001.
- Fetterman, David M. & Shakeh J. Kaftarian, toim.: *Empowerment evaluation: knowledge and tools for self-assessment and accountability*. Sage, London 1996.
- Fischer, Frank & John Forester: *The argumentative turn in policy analysis and planning*. Duke University Press, Durham, NC 1993.
- Giddens, Anthony: *Central problems in social theory*. Macmillan, London 1979.
- Giddens, Anthony: *The third way and its critics*. Polity Press, Cambridge 2000.
- Guba, Egon G & Yvonna S. Lincoln: *Fourth generation evaluation*. Sage, London 1989.
- Heidegger, Martin: *Sein und Zeit*. 18. p. 1. p. 1927. Max Klostermann, Tübingen 2001. Suomeksi Oleminen ja aika. 2. p. Vastapaino, Tampere: 2001.
- Heritage, John: *Harold Garfinkel ja etnometodologia*. Gaudeamus, Helsinki 1986.
- JCSEE (Joint Committee on Standards for Educational Evaluation): *The program evaluation standards: how to assess evaluations of educational programs*. Sage, London 1994.
- Kelleher, John, Sarah Batterbury & Elliot Stern: *The thematic evaluation of the partnership principle: final synthesis report*. Tavistock Institute, London 1999.
- Kettunen, Pekka: *Implementation in a multi-organisational setting: local networks in environmental health policy*. Turun yliopisto, Turku 1994.
- King, Jean A., Lynn Lyons Morris & Carol Taylor Fitz-Gibbon: *How to assess program implementation*. Sage, London 1998.
- Kuntaliitto ym. Kohti julkisten palvelujen yleistä arviointia: keskustelualoita. Suomen kuntaliitto, sisäasiainministeriö, valtiovarainministeriö 2001.
- Marx, Karl: *Pääoma*. I osa. 1. p. 1867. Edistys, Moskova 1974.
- MEANS. *Evaluation of socio-economic programmes*. 6 volyyymia. Uudistettu p. European Commission, Brussels 1999.
- MFA (Finnish Ministry for Foreign Affairs): *Guidelines for project design, management and evaluation*. Helsinki: Finnish Ministry for Foreign Affairs. 2. p. MFA, Helsinki 1998.
- Milgrom, Paul & John Roberts: *Economics, organization and management*. Prentice-Hall, Englewood Cliffs, NJ 1992.
- Mäntysaari, Mikko *FinSoc edistää menetelmien arviointia*. *FinSoc News* 1/1997. [Http://www.stakes.fi/palvelut/finsoc/news0197%5F1.html](http://www.stakes.fi/palvelut/finsoc/news0197%5F1.html). Stakes, Helsinki 1997.
- Nietzsche, Friedrich: *Zur Genealogie der Moral*. 1. p. 1887. Teoksessa Friedrich Nietzsche. Werke in drei Bänden. Carl Hanser Verlag, München: 1994.
- Owen, John M. & Patricia J. Rogers: *Program evaluation: forms and approaches*. Sage, London 1999.
- Pal, Leslie: *Public policy analysis: an introduction*. Scarborough, ONT. Nelson Canada 1992.
- Palonen, Kari: *Das 'Webersche Moment': Zur Kontinenz des Politischen*. Westdeutscher Verlag, Opladen 1998.
- Patton, Michael: *Utilization-focused evaluation*. Sage, Beverly Hills, CA: 1978. Useita uusia p., mm. 3. p. 1997.
- Pawson, Ray & Nick Tilley: *Realistic Evaluation*. Sage, London 1997.
- PCM (Project Cycle Management): *Manual project cycle management. integrated approach and logical framework*. Commission of the European Communities, Brussels 1993.
- Platon: *Valtio*. Teokset. IV osa. Alkup. kreikk. Politeia 5. vuosis. e.Kr. Keuruu: Otava 1999. Kreikankielinen teksti esim. 2-osaisessa painoksessa Plato. Republic. Books 1-5. Books 6-10. Harvard University Press, Cambridge, MA 1999, 2000.
- Pollitt, Christopher, Xavier Girre, Jeremy Lonsdale, Robert Mul, Hilikka Summa & Marit Warness: *Performance or compliance?* Oxford University Press, Oxford 1999.
- Pressman, Jeffrey L. & Aaron B Wildavsky: *Implementation*. 3. p. University of California Press, Berkeley 1984.
- Raivola, Reijo: *Tehoa vai laatua koulutukseen*. WSOY, Porvoo 1999.
- Rimpelä, Arja: *A critical analysis of the Finnish Tobacco Act : implementation and legitimacy 1977-1989*. *Tobacco Control - an International Journal* (1992): 1, s. 285-293.
- Rimpelä, Arja, Matti Rimpelä & Andres Vikat: *Is prohibition of tobacco sales to minors effective? The case of Finland*. Teoksessa C. Tudor-Smith, toim.: *Working together for better health, Tackling tobacco*. Health Promotion Wales, Cardiff 1999.
- Rissanen, Pekka: *Effectiveness, costs and cost-effec-*

- tiveness of hip and knee replacements. Stakes, Helsinki 1996.
- Rissanen, Pekka, Harri Sintonen & Markku Pekurinen: 15D: terveyteen liittyvän elämänlaatumittarin, visuaalisen analogiamittarin ja koetun terveydentilan mittarin arvot aikuisväestössä. *Sosiaalilääketieteellinen aikakauslehti* 32(1995), s. 207-211.
- Rorty, Richard: *Contingency, irony and solidarity*. Cambridge University Press, Cambridge 1989.
- Rossi, Peter H., Howard E. Freeman & Mark W. Lipsey: *Evaluation: a systematic approach*. 6. p. Sage, London 1999.
- Salmikangas, Anna-Katriina, Kristiina Veräväinen & Terho Vuorela: Osallisuusprojektien monet kasvot – osallisuushankkeen arviointia. Sisäasiainministeriö, Helsinki.
- Schön, Donald A.: *The reflective practitioner: how professionals think in action*. Basic Books, New York 1984.
- Shinkfield, Anthony J. & Daniel L. Stufflebeam: *Teacher evaluation: guide to effective practice*. Kluwer, Dordrecht 1995.
- Stake, Robert E.: *The art of case study research*. Sage, London 1995.
- Teittinen, Tuomo: *Julkisten toimintaohjelmien ominaisuudet ja toimeenpano-organisaatio*. Väitösk.: Kuopion yliopisto, Kuopio 1985.
- Uusikylä, Petri & Petri Virtanen: *Meta-evaluation as a tool of learning: a case study of the European Structural Fund evaluation in Finland*. *Evaluation* 6(2000): 1, s. 50-65.
- Vedung, Evert: *Utvärdering i politik och förvaltning*. 2. uudist. p. Studentlitteratur, Lund 1998.
- Winograd, Jerry & Fernando Flores: *Understanding computers and cognition: a new foundation for design*. Uudistettu p.: Addison-Wesley, New York 1995.