

# Onnistunut ja vaikuttava arviointi



Koulutukseen liittyvä arviointi on lisääntynyt viime vuosina kansallisesti ja kansainvälisesti. Puheenvuoro tuo esille kasvatusalan kokeneiden arvioijien näkemyksiä siitä, millainen on onnistunut ja vaikuttava arviointi.

Hyvää, onnistunutta, tuloksekasta tai vaikuttavaa arviointia on kuvattu arviointikirjallisuudessa ehtimiseen. Sitä varten on kirjoitettu useita eettisiä ohjeita ja standardeja (AEA 2004; JCSEE 2016). Silti ihanteiden tavoittelu on käytännössä vaativaa ihmistieteellisillä aloilla, joilla syy-seuraus -suhteet eivät ole helposti hallittavissa. Jokainen arviointitoimeksianto pitäisi kuitenkin tehdä olosuhteisiin ja voimavaroihin nähden niin hyvin kuin mahdollista, että saavutettu tieto olisi pätevää ja oikeuttaisi harkittuja, ehkä mittaviakin seuraamuksia.

Mitä monisyisemmäksi länsimainen yhteiskunta on kehittynyt, sitä vahvemmaksi on tullut usko, että arvioinnilla voidaan saada paitsi erilaistumista kuvaavaa

niin tarvittaessa myös sen kontrolloimista edistävää tietoa. Arviointien määrä on lisääntynyt kansallisesti ja kansainvälisesti merkittävästi niin, että puhutaan jopa arviointimaniasta tai -ähkystä. Lisääntyminen voi johtaa sekä arviointiosaamisen vahvistumiseen että heikentymiseen riippuen siitä, onko koulutettua, eettisesti kestävin tietotaidoin toimivaa työvoimaa riittävästi (Jacob & Boisvert 2010).

Tässä puheenvuorossa annetaan ääni kasvatusalaa tunteville, kokeneille arvioijille siitä, millainen on onnistunut tai vaikuttava arviointi. Optimistinen näkökulma on tärkeä aiheessa, johon liittyy usein kielteisiä mielikuvia. Artikkelin perustuu haastatteluaineistoon,

josta on aineistolähtöisesti nostettu pää- ja alateemoja. Tutkitut arvioijat ovat toimineet aikuiskasvatustieteen kentällä esimerkiksi vapaan sivistystyön, aikuis-koulutusorganisaatioiden toimivuuden, EU- ja ESR-hankkeiden, aikuisten työkykyisyyden tai kansallisten koulutuspoliittisten ratkaisujen puntaroina.

## AINEISTO JA SEN ANALYSOINTI

Haastateltavien valinnan ensimmäisessä vaiheessa kartoitin laajasti tarkoituksiini sopivia ihmisiä hyödyntäen paitsi omaa työkokemustani myös suomenkielisiä arviointijulkaisuja, internet-sivustoja sekä keskusteluja kollegojen, arvioijien ja arviointiorganisaatioiden esimiesten kanssa. Pohdinnan jälkeen minulla oli 25 henkilön nimiluettelo, josta seuraavaksi suljin pois kymmenen henkilöä, jotta jäljelle jäi mahdollisimman heterogeeninen kohdejoukko erilaisissa organisaatioissa ja erikokoisissa arviointitoimeksiantoissa toimineita henkilöitä.

Kaikki valintahaaviini jääneet yksitoista naista ja neljä miestä vastasivat myönteisesti haastattelupyyntöni. Heillä oli arviointityökokemusta 4–30 vuotta. Osalla oli toimenkuvassaan muutakin työtä kuin arvioinnit. Kaikilla oli kasvatusalan, mutta monilla myös esimerkiksi nuorisotyön ja työvoimahallinnon arviointikokemusta; mukana oli myös kansainvälisiin arviointeihin osallistuneita. Paria poikkeusta lukuun ottamatta haastatteluja hallitsivat aikuisten tai aikuis-koulutuksen, -organisaatioiden ja -koulutuspolitiikan (mukaan lukien vapaa sivistystyö ja ammatillinen koulutus) arviointikysymykset. Oppimistulosten arviointeja ei käsitelty, vaan kohdejoukko profiloitui tila-, teema- ja ohjelma-arviointien tekijöinä.

Haastattelurunko käsitti seitsemän pääteemaa, joista tämä puheenvuoro perustuu ensisijaisesti teemaan ”myönteiset kokemukset”, mutta onnistuneita arviointeja koskevia kuvauksia löytyi myös monesta muusta haastatteluteemasta. Lähetin muutamin ranskalaisin viivoin avatun teemu luettelon haastateltaville etukäteen. Tein itse haastattelut useilla paikkakunnilla, tyypillisesti haastateltavan työpaikalla, ja ne kestivät 52–106 minuuttia.

Ensiksi tein koko haastattelulitteraatioon useita tekstihakua, jonka jälkeen lähiluvin huolellisesti koko

aineiston kahdesti ja poimin asiasisältöjen perusteella lisää tekstiotteita manuaalisesti. Näin rakentui 11 528 sanan tekstitiedosto, jonka luin useaan kertaan onnistumisten ja vaikuttavuuden hahmottamiseksi. Pelkistin ensimmäisessä vaiheessa 176 pienteemaa pyrkien säilyttämään näkemysvarianssin. Seuraavaksi integroin arviointia koskevan lukeneisuuteni perusteella pienteemat teemaryppäiksi. Onnistumista koskevia lausumia oli 137 ja vaikuttavuus-teemaisia 39.

Vaikka haastateltavat muodostavat harkinnanvaraisen, yleistämiseen oikeuttamattoman näytteen, pidän esiteltäviä tuloksia yleismerkityksisinä. Huomioiden heidän taustakoulutuksensa monipuolisuuden, työvuosiensa määrän, toimeksiantojensa lukuisuuden sekä laajan toimintasektorinsa pidän tuloksia uskottavina, vaikka laadullisen aineistoanalyysin tekemiseen liittyy aina tulkinnallisuutta.

## TULOKSET

### Prosessitekijöiden toimivuus ja palkitsevuus

Yksi klassisista arviointilähestymistapojen erotteiluista koskee produkti- ja prosessiarviointia. Tuotoksien eli produktien ensisijaistamista koskeva arviointikulttuuri on niistä selvästi vanhempi ja nojaa erityisesti tavoitteiden ja tulosten vertailuun. Sen kapea-alaisuuden vuoksi on jo pitkään kiinnitetty huomiota mekanismeihin ja prosesseihin, joiden ansiosta tuloksiin päästään tai niistä jäädään esimerkiksi erilaisten hyvinvointi- ja koulutusohjelmien toteuttamisessa (esim. Forss ym. 2002; Pawson & Manzano-Santaella 2012).

Esimerkiksi reaaliaikaisen arvioinnin tarkoituksena on tuottaa kehittämisprosessin aikana tietoa, jota toimijat voivat käyttää toimintansa kehittämiseen. Päättöarviointihan ei hanketoimijoita enää auta. Toisaalta on Virtasen (2005, 63) tai Paasion (2006, 101) tavoin aiheellista suhtautua kriittisesti liialliseen prosessi-innostukseen vaikuttavuuden kustannuksella.

Haastateltavat puhuivat usealta kannalta (26 sitaattia) prosessien merkityksestä pohtiessaan, miksi jokin arviointi oli onnistunut. He analysoivat omaa rooliaan prosessien seuraajina ja fasilitoijina (H3, H5) ja valottivat hankkeen toiminnan aikana tuote-

tun palautetiedon merkittävyyttä (mm. H9, H11). Heidän mukaansa *”järkevintä ja parasta on tehdä tuettuja itsearviointia prosessinarviointien ja projektien toteutuksen aikana (H3)”* ja *”policy learning ei ole sitä, että arvioitsija suorittaa ja sit opitaan, vaan siinä on tää prosessihyöty isossa roolissa (H5)”*.

Onnistunut tai vaikuttava arviointi ei ollut välttämättä tulosta ongelmattomista työprosesseista, vaan usein niihin kuului vaikeita vaiheita, palaamista asiassa taaksepäin ja hankalien tilanteiden kääntämistä voitoksi. Kun arviointi mahdollisti kohteen kulttuurin näkemisen kunnolla, ristiriidat valottuivat kirkaasti (H6). Oivalluksia syntyi, jos rohkaistui pois mukavuusalueeltaan (H7). Vaikeasti tavoitettavien ja avautuvien arviointitiedon antajien luottamuksen saavuttaminen saattoi vaatia paljon työtä mutta lopulta palkita (H12).

Positiivinen palaute onnistuneista ratkaisuista sekä siivitti toimijoita parempiin tuloksiin että voimaannutti arvioijia (esim. H6, H7). Myös arviointiprosessien yllätyksellisyyteen kiinnitettiin huomiota. Vaikuttavuutta saattoi syntyä tekijöistä tai olosuhteista tavalla, jota ”ei oltu osattu ennakoita”, tai arvioija oli oivaltanut esittää työskentelyprosesseja ohjatessaan yllättäviä kysymyksiä (esim. H2, H14). Juuri tilanneherkkyyttä tunnistaa tavoitteista poikkeavia tuloksia – *”thinking outside evaluation boxes (Patton 2010b, 53)”* – on arvostettu kehittävässä arvioinnissa.

### **Arvioijan osaamisen merkitys arvioinnin onnistumiselle**

Arviointikirjallisuus (esim. Ryan & Cousins 2009) esittelee laajasti arvioijana toimivaan henkilöön kohdistuvia yksilöllisiä vaatimuksia. Arvioijan käsissä on usein koko prosessi kysymyksenasettelusta raportointiin, joista viimeksi mainittu on erityinen osoitus hänen vallastaan. Arvioijan henkilökohtainen osaaminen ja taidot voivat olla koko arvioinnin onnistumisen avain. Esimerkiksi H2, H11, H12, H14 ja H15 kuvasivat substanssin hallintaa koskevia omakohtaisia havaintojaan: aiheen tuntemus tehostaa osaamista, mikä motivoi arvioinnin eri osapuolia.

Kaikilla edellä mainituilla henkilöillä oli kokemuksia myös siitä, että oli joutunut arvioimaan koh-

teita tai ilmiöitä, joita ei eri syistä tuntenut omakseen. Yksi heistä totesi painokkaasti, että

*”ei voi aatella, et se on ihan samaa, et on tavallaan semmonen arvioitsija johon heitetään vaan, et tää on tämmönen tuotantokoneisto. Et ikään kuin liukuhihnalta, et heittääkää data sisään, ni täältä pulauttaa arviointisuositukset ulos.” (H2)*

Haastateltavien myönteisiin kokemuksiin (20 sitaattia) kuului tunnistaa oman ammattitaidon riittävyys tai tuntee sen paranemisen tuomaa tyytyväisyyttä, mikä vahvisti itseluottamusta selvität uusista toimeksiannoista. Arviointiosaamisen kehittäminen (evaluation capacity/capability) on kiinnostanut paljon myös arvioinnin tutkijoita (esim. McDonald ym. 2002; Preskill & Boyle 2008; Seppänen-Järvelä & Vataja 2009).

Haastateltavat kiittivät vertaispalautteen saamista (H2, H3) ja mahdollisuutta oppia rikkaasta arviointiaiheiden ja -kohteiden kirjosta (H8, H11) sisällöllisesti ja menetelmällisesti. Iloa tuotti huomata, että monipuolinen arviointityökokemus taipui uusien toimeksiantojen joskus varsin kryptisten ongelmien ratkomiseen (H14).

### **Yhteistyön voima arvioinnin vaikutusten turvaamisessa**

Konstruktivististen arviointiparadigmojen vallattua alaa arvioijan asemaa on horjutettu ja korostettu erityisesti sidosryhmien oikeutta osallistua arvioinnin kaikkiin vaiheisiin, tulostulkinnat mukaan lukien (Kushner 2011, 309–301). Nykyarvioija nähdään yhdeksi toimijaksi monien arviointiosallisten joukossa, kehittämiskumppaniksi ulkopuolisen valankäyttäjän sijasta. Taitavakaan yksilö ei voi hallita monimutkaisia arvioitavia ilmiöitä ja arviointiosa puolten vuorovaikutussuhteita, vaan arvioinnissa tarvitaan osaamistaustaltaan monialaista asiantuntijaryhmää.

Yhteistyö tuli selvästi (25 sitaattia) esille myös tekemissäni haastatteluissa. Osa arvioijista mainitsi yhteistyön tai yhteistyökumppaneiden merkityksen erittelemättä sitä tarkemmin (esim. H6, H7, H10, H15), mutta osa syventyi yksilöimään erityisesti

# HAASTATELTAVAT KOROSTIVAT YHTEISTYÖN JA VUOROVAIKUTUKSEN MERKITYSTÄ.

vuorovaikutuksen merkitystä. He tarkastelivat keskustelujen, yhteisen pohdinnan sekä eri osapuolten kohtaamisen hyötyjä esimerkiksi näin:

*"Nythän on paljon enemmän menty siihen, että sitä vuoropuhelua käydään koko ajan, mikä mun mielestä on sinänsä ihan ehdottomasti hyvä siinä mielessä, et silloin harvoin tulee sellasia tilanteita, että siinä vaiheessa, kun me toimitetaan vaikka loppuraportti... niin et silloin tulis ihan yllätyksenä, että toimeksiantaja on odottanu esimerkiks ihan jotain muuta." (H2)*

*"Vuorovaikutteisuus lisääntyy koko ajan, ja asiakkaat osaa ja haluaa, osaa vaatia. Ja ne kyllä sitä selvästi haluaa, ja se sitten vaatii taas meiltä myösi tiettyä semmosta [vuorovaikutusta]. Se on meillä se lähtökohta, et me tehdään asiakaslähtöisesti ja vuorovaikutuksellisesti, jos asiakkaat sitä haluaa, ja itse koen sen kyllä varsin mielekkääksi." (H11)*

Monilla oli kokemuksia arviointi- tai ohjausryhmän käytöstä arviointitoimeksiantojen toteuttamisessa. Ryhmästä tuli arvioitaville ilmiöille uusia tulkintoja (H2), käytettävissä oleva asiantuntemus monipuolistui (H4), ratkaisuvaihtoehdot lisääntyivät useamman miettijän ansiosta (H6) ja vaikeiden asioiden läpivienti helpottui (H2).

## Muutoksien aikaansaaminen

Yksi arvioinnin ydintehtävistä on lisätä tietoa. Joskus tieto on uutta, toisinaan arviointi vahvistaa käytännössä jo havaittuja vahvuuksia tai heikkouksia. Perusolettama on, että ellei nykytoimintoja katsota kriittisesti, *status quo* säilyy. Muutos voi tarkoittaa sekä positiivista että negatiivista siirtymää. Se voi olla toimijoille suuri, vaikka näyttää ulkopuoliselle pieneltä.

Liikkeelle sysäävä tekijäkin voi olla melkein mitä tahansa. Eduksi on ainakin se, että ihmiset ovat itse voineet osallistua tiedon tuottamiseen: silloin halu sitoutua tuloksiin ja toimijuuden tunne vahvistuvat (esim. Atjonen 2015a; Cousins ym. 2012). Näin päästään pelkän "tilivelvollisuuden" ja tiedontuotannon palvelemisesta kehittävään arviointiin.

Haastattelemini arvioijien puheesta kuulin usein (25 sitaattia) kuvauksia arvioinnin ansiosta aikaansaaduista muutoksista, joiden laajuus vaihteli arviointitoimeksiannon koon mukaan. Arvioinnista oli ollut seurauksia jopa kansalliseen päätöksentekoon (H4), mutta yhtä lailla tyytyväisyyttä tuotti pienissä hankkeissa selkeästi todettavissa ollut hyöty (H1). Muutoksista oli tullut konkreettista palautetta arvioijille jopa vuosien jälkeen (H3), ja vaikka arvioinnilla oli ollut tuntuvia seurauksia (esim. irtisanomiset), silti päätelmät tuntuivat edelleen oikeilta (H5).

Useat arvioijat (H2, H6, H11, H13, H15) olivat kokemuksensa perusteella vakuuttuneita siitä, että nimenomaan kehittämiseen tähtäävä arviointi oli johtanut muutoksiin. Esimerkiksi H13 kuvasi, että

*"kehittävässä arvioinnissakin on minun mielestä semmonen viisaus, että kun otetaan hetkeksi siitä omasta työympäristöstään ja normaalista työympäristöstään irralleen ja hetki punniskellaan semmosia asioita, mitä yleensä ei oo aikaa katsoa, että miten me muuten itse asiassa tehdäänkään tätä työtä, ni jo se interventio itsessään on kehittävää. Se, että he istuivat siinä, ni on jo ensimmäinen askel siinä [muutoksessa]."*

Pattonin käsitys (2011, 38–41) kehittymisestä on vaativa, kun hän vertailee parantamista ja kehittämistä (*improvement vs. development*). Hänelle muutos on syvälinen transformaatio (emt., 174–175), jonka tulisi johtaa arvioivan kulttuurin kehittymiseen organisaation sisälle. Tämänäyttöisiä kokemuksia haastateltavillani ei oikeastaan ollut, mikä selittynee ainakin osan kohdalla toimeksiantojen lyhyydellä.

## Arvioinnin aiheen ja uuden sisältötiedon merkitys

Tutkimusten mukaan väärän tai vähämerkityksisen asian arvioiminen (mahdollisesti vielä vääränä

ajankohtana) on tyypillinen syy mitätöidä arviointi ja sen tulokset (esim. Stame 2010b). Arvioinnin aiheen koettu merkittävyys selittää ymmärrettävästi myös tuloksellisuutta koskevia kokemuksia. Arviointitiedon käytöstä sosiaalisissa organisaatioissa on tutkimusten perusteella selvää oikeastaan vain se, että sitä selittävät monet, keskenään vuorovai-  
kutteiset seikat.

Arviointitulosten käyttäjät vaikuttavat merkittävästi tulosten hyödyntämiseen, eli arvioijaosaamisen lisäksi tarvitaan myös käyttäjäosaamista. Forssin ym. (2002) peräämä ”moraalinen buusti” vaikuttaa tulosten toimeenpanoon organisaatiossa, samoin käyttäjien polarisoituminen hyödyn ja työmäärän vertailemisen perusteella (Contandriopoulos & Brousselle 2010).

Tiedonkäsitykselläkin on sijansa arvioinnin kohdeaihetta koskevan tietämyksen käytössä: varhempi absoluuttinen totuuskäsitys on ainakin osittain korvautunut näkemyksellä sosiaalisesti rakentuvasta tiedosta, mikä on omiaan rohkaisemaan osallistaviin arviointiprosesseihin (Cousins ym. 2012).

Haastateltavani kuvasivat arvioinnin tuottaman tiedon merkitystä (22 sitaattia) usealta kannalta. Onnistunut ”arviointi vastasi todelliseen tarpeeseen” (H3, H6, H7) tai se oli ”tuottanut uutta tietoa” (H2, H4, H7, H8):

*”Siin oli selkee tavoite, tarkoitus, ja se ennen kaikkea sopi tähän kunnalliseen palvelujärjestelmään. Sille oli tilausta, ja se myös onnistuttiin markkinoidaan niin, et kuntapäätäjät ymmärsi sen.” (H7)*

*”Mehän arvioidaan pääosin asioita, joihin kohdistetaan paljon julkista rahoitusta.. Et nää on kuitenkin isoja rahoja varsinkin nykyisessä taloustilanteessa. Ja kun käytiin niitä tuloksia läpi, niin tulihan siinä sellanen [olo, että arviointi oli tärkeä]. Et pysty tuomaan sinne sellasta tietoa, mitä heillä ei selvästi ollut.” (H2)*

Arvioija saattoi kokea olevansa auttaja tai puolestapuhuja tuodessaan arviointituloksia esille (esim. H8). Yksi haastateltava kuvasi tällaista merkittävyyttä näin:

*”... on ollu asiakkaita, jotka on lähtökohtasesti vaikeita haastatella. Et oli... asunnottomia, moniongelmaisia ihmisiä, jotka tuli sinne haastatteluun aikamoisessa kunnossakin. Mut tavallaan, et mä sain heistä irti. He halus tulla sinne, ja sain heistä irti asioita... Et mä oon tavottanu sellasia ihmisiä, et ne on saanu heidän äänensä kuuluviin... vaikeesti työllistyviä ihmisiä, et siellä on ollu ihan vammautuneita... Tulee sitten sillä tavalla välitettyä se viesti myös sille hankkeelle ja niille, mitkä on sit jossain ohjausryhmässä tai muussa.” (H11)*

Housen (1980) korostamien arvolähtökohtien ja MacDonalidin (1976) aloittaman demokraattista arviointia käsittelevän keskustelun ansiosta arvioinnissa on opittu voimistamaan marginaalisten ryhmien ääntä, eikä tyydytty arviointitulosten toimeenpanoon vain enemmistön ehdoilla.

### Arviointimenetelmien toimivuus

Väittely arviointimenetelmistä näkyy kirjallisuudessa erilaisten tiedonkeruuvälineiden (kysely, haastattelu, havainnointi) hyvyyden vertailuna, laajempien konseptien (esim. EU-impaktiarviointi, Common Assessment Framework- eli CAF-malli, 10-asteleinen voimaannuttava arviointi) tuloksellisuuden pohdintana sekä kvalitatiivisen tai kvantitatiivisen lähestymistavan kyseenalaistamisena. (Bamberger 2012.) Menetelmän pätevyuden kysymykseen on lopulta sama yksinkertainen vastaus kuin tutkimuksessa: arviointikysymys ratkaisee sopivan menetelmän.

Haastattelemani arvioijien menetelmiä koskeva puhe (19 sitaattia) keskittyi kahteen pääteemaan: menetelmälliset kokeilut ja itsearviointin tuloksellisuus. ”Uusien analyysimenetelmien kokeilu” (H1) tai ”usean tiedonkeruumenetelmän yhdistelmä” oli nostanut arviointikohteista pohdittaviksi ennakoimattomia näkökohtia (H5, H6). Alun perin kvantitatiivisiin menetelmiin viehtynyt (H15) henkilö oli huomannut kvalitatiivisen lähestymistavan voiman, tai laadullisten menetelmien vankka kannattaja (H9) oli koonnut valaisevia kvantitatiivisia aineistoja. Kokonaisen arviointijärjestelmän luominen tiettyyn

organisaatioon oli tuottanut tyydyttävän lopputuloksen (H5).

Yksittäisistä arviointimenetelmistä eniten puhetta synnytti itsearviointi, positiivisessa, vaikutuksia tai muutoksia aikaansaavassa merkityksessä. Tässä suhteessa tutkittavat edustivat ”mittausorientaation” – Vedungin (2010) sanoin uusliberalistisen tai tietoperustaisen arviointiaallon – vastaista liikettä. Konsultaatiotyypillisessä työssä kohteet olivat halunneet tuettua itsearviointia (H5, H7, H11). Joidenkin mielestä kehittävä arviointi edellytti juuri itsearviointia (H11, H13).

Vaikka itsearvioinnin uskottavuuteen on perusteltuja syitä suhtautua kriittisesti (esim. Basson 2009, 63; Vanhoof & Petegem 2007, 109), siihen on kohdistunut paljon toiveikkuutta ja vaikutuksista on saatu hyviä kokemuksia (esim. Huusko 2010; Vataja 2012; Yliruka 2015). Itsearviointi saatetaan yhdistää ulkoiseen arviointiin, ellei sen yksinomaiseen muutosvoimaan rohjeta luottaa.

Eriytin haastattelusitaateista 39 sellaista katkelmaa, joissa käsiteltiin tarkennetusti arvioinnin vaikutuksia tai vaikuttavuutta, ei vain ”väljästi” myönteisyyttä, tuloksellisuutta tai onnistuneisuutta. Katkelmien valintakriteerinä oli vaikutus-, vaikuttavuus- tai seura(am)us-käsitteen käyttäminen – viimeksi mainittu tulkittuna asiayhteydestään niin, että haastateltava puhui hankkeen ”vakavista” tai ”varsinaisista” seurauksista tai seuraamuksista (ks. Kajamaa ym. 2008).

Sivuutan tässä ne 15 puheenpartta, jotka käyttiin mainitsemalla jokin konkreettinen arviointihanke nimeltä (esim. H4, H5), lueltiin vaikuttavuutta hämärtäneitä tekijöitä (esim. H3, H6, H7, H9) tai puhuttiin fokuusoimattomasti vaikutuksista (esim. H8, H10, H13). Niistä ei varsinaisesti löytynyt vaikuttavuusmekanismeihin liittyvää pohdintaa. Keskityn seuraavassa kahteen vaikuttavuuspuheen teemaan eli menetelmähaasteisiin (f = 12 sitaattia) ja vaikuttavuuden synnyttämiseen (f = 12 sitaattia).

#### **Vaikuttavuuden menetelmällisistä haasteista**

Vaikutusten ja vaikuttavuuden arviointi on vaativala. Vaikka periaatteessa tiedetään päteviä menetelmiä

niiden osoittamiseen (esim. satunnaistettu koe-kontrolli -asetelma ”kultaisena sääntönä” Stame 2010a, 32), ne eivät käytännössä tahdo taipua monimutkaisista vuorovaikutussuhteista muodostuviin opetus- ja kasvatustilanteisiin. Arviointimallit tuottavat joko hyvin kapea-alaisen kuvan vaikuttavuudesta, osoittavat interventiot tuloksettomiksi tai ohittavat merkittävät sivuvaikutukset. (Dahler-Larsen 2005.) Vaikuttavuuden todentamisen vaikeus oli hyvin tuttua haastatteleilleni henkilöille, joista erityisesti H13 virittyi kriittisiin mietteisiin:

*”Kyllähän jatkuvasti törmää esimerkiksi tutkimuksiin, missä vaikuttavuustutkimus-sana esiintyy otsikossa. Ja voisko sanoa pahimmillaan oon [nähty] semmosia, missä on muutama teemahaastattelu tehty... ja sit se oli kerrottu, et tämä on nyt vaikuttavuustutkimus. Ni vaikka en vaikuttavuustutkimusta ehkä pidä niin suurella arvossa kuin sitä keskimäärin pidetään, mut siitä huolimatta sillä tarkotetaan jotain. Et sitten ei voi ihan muuta väittää vaikuttavuustutkimukseksi.” (H13)*

H15 katsoi vaikuttavuuden syntyvän erityisesti ”hyödyllisestä tiedosta”. Hyödyllisellä, kehittämistä edistävällä tiedolla hän tarkoitti tietoa, joka ei ole liian yksityiskohtaista tai fokuusoimattomaa. Pelkkä tieto riittää harvoin, joten hyödyntämistä pitää ajatella koko arviointiprosessin ajan, arviointikysymyksen asetelusta raportin kirjoittamiseen asti (Forss ym. 2002; Patton 2012). Esimerkiksi ydintoimijoiden ja hyvien sidosryhmien monitahoisella osallistamisella on suuri merkitys hyödyntämiselle (Cousins ym. 2012); siihen ei välttämättä riitä osuvakaan faktatieto.

H6 uskoi arviointien hyödyn toteutuvan tehokkaimmin avainryhmien tai -henkilöiden kautta, koska ”keskeiset avainhenkilöt puskee sitä asiaa eteenpäin ja kääntää sen myöskin hyödyksi”. H7 toi esille koeasetelmille vaihtoehtoisia, laadullisempia tapoja (ks. Hind 2010) todentaa vaikutuksia kertomalla hyviä kokemuksia case-tarinoiden käytöstä ja erilaisten palvelujärjestelmien toimivien mallien ”ristiinpölytyksestä”, kuten hän asiaa kuvasi. Viimeksi mainituista voitaisiin puhua myös hyvien käytänteiden levittämisenä.

## Vaikuttavuuden aikaansaamisesta

Haastateltavat pohtivat monipuolisesti vaikuttavuuden, kehittävyuden ja hyödyllisyyden edistämistä. Jonkun mielestä toimeksiantosopimuksen mukaisesti vain viestitään tulokset (H2) tai annetaan palautetta (H9), ja vastuu siitä eteenpäin on arvioinnin kohteilla. Kun oli onnistunut arvioinnin avulla ”rakentamaan organisaation sisälle organisaation (H5)”, se turvasi kehittämisen jatkossa ilman ulkopuolisia toimijoita. Tätä arviointikirjallisuus kutsuu evaluatiivisen kulttuurin juurruttamiseksi (Suarés-Herrera ym. 2009, 334). Samaan viittasi myös H10 puhuessaan johdon merkityksestä ja arviointihavaintojen kytkemisestä oppilaitoksensa vuoden kierron muihin laatuprosesseihin.

Jos aikaisemmin riitti, että arvioija kuvasi toimijoiden kokemuksia asioiden kulusta, nykyisin haluttiin selkeitä esityksiä erilaisten vaihtoehtojen vaikuttavuudesta (esim. H12). Tämä viritti toisen informantin kritiikkiin:

*”Se haaste tulee siinä, että arviointihan tapahtuu kuitenkin avoimessa systeemissä, jossa tavattoman moni eri asia vaikuttaa siihen kohteeseen. Ja ei semmosta arviointisijaa ole tullu vastaan, joka pystyisi sanomaan, että tämä vaikutus johtuu vain ja pelkästään tästä tietyistä asiasta. Eli kun avoimessa maailmassa toimitaan, ni me ei koskaan voida sanoa mitään satavarmasti.” (H7)*

Arviointi vaatii siitä vastaavalta henkilöltä monenlaisia yksilötaitoja esimerkiksi vuorovaikutuksen rohkaisemisessa tai kohteiden lähestymisessä kannustavasti silloinkin, kun pitää käsitellä ongelmia. Vaikka arvioijan tehtävä on ”to speak truth to power” (Patton 2010b, 50; Eliadis ym. 2011), ”totuuden” esittämispää kannatti harkita. Esimerkiksi H1 kuvaili onnistuneita kohtaamisia arvioitavien kanssa niin, että keskityttiin miettimään tapoja päästä tilanteesta eteenpäin, sen sijaan, että pysähdyttäisiin osoittelemaan sormella ongelmia.

H13 muistutti hyvien käytäntöjen identifioinnin merkityksestä kehittämiselle ja H15 nettovaikuttavuudesta, eli tarpeesta tarkastella positiivisten saavutusten rinnalla myös kielteisiä vaikutuksia tai uhkakuvia. Kehittävän arvioinnin kannattajaksi haastatteluissa

selvästi osoittautunut H15 kuitenkin totesi, ettei ongelmilla kannata ”retostella” ja jäädä niihin ”vellomaan”, koska hyvät vaikutukset tuottavat varmemmin muutoksia.

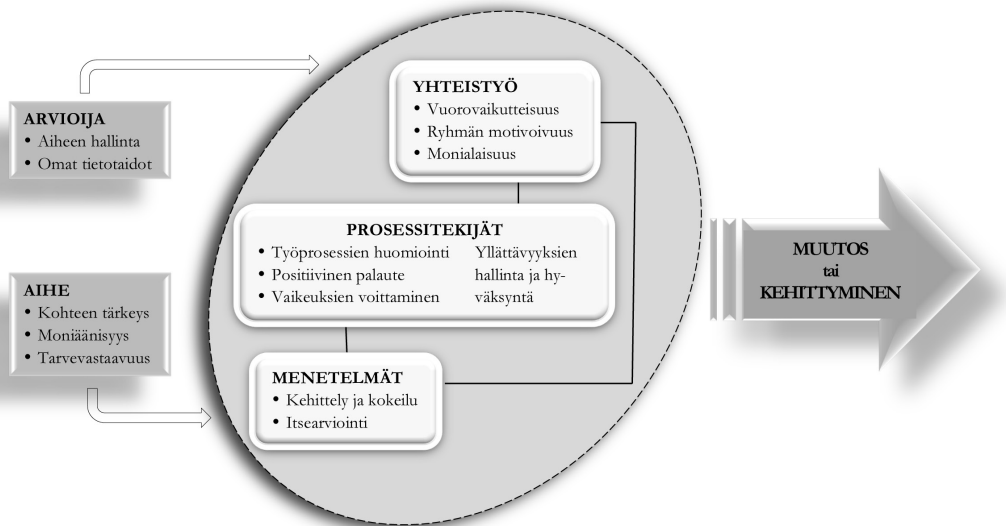
## LOPUKSI

Koostin haastattelun tuloksista **kuvion 1**. Kuviota voi käyttää varoen kuvaamaan onnistunutta ja vaikuttavaa arviointia. Arvioijan osaaminen sekä arvioinnin kohdallinen aihe vaikuttavat siihen, millainen koko arviointiprosessista tulee. Harmaassa soikiossa on arviointiprosessin aikana muunneltavissa olevia elementtejä, joista keskiössä ovat arviointiprosessia itseään koskevat tekijät. Viimeksi mainittuja voidaan ohjata arviointimenetelmiä varioimalla ja hyvällä yhteistyöllä. Hyvien prosessielementtien vallitessa ja ollessa sopivassa keskinäisessä vuorovaikutuksessa arvioinnin tuloksena voi olla muutoksia tai kehittämistä.

Viidentoista, vaikkakin kokeneen arvioijan tuotama kuva voi tietysti olla aukollinen. Toisaalta kokemuskirjo nosti esiin valtaosan esimerkiksi niistä tekijöistä, joita Morris (2008) tarkastelee arvioinnin eettisinä kipupisteinä. Mitä enemmän arviointeja tehdään, sitä suurempi tarve on eettisille arviointiperiaatteille (Atjonen 2013). Hyvä, onnistunut arviointi ei ole vain taitavaa tiedonkeruutekniikkaa ja sofistikoituneita analyysimetodeja, vaan enenevästi ”hyvän” ja ”oikean” pohtimista, mikä kohdistuu arviointiprosessin kaikkiin latentteihin tekijöihin, ihmiskäsitystä myöten.

Arvioinnin tulevaisuudesta on yhtä monia ennusteita kuin on ennustajiakin, mutta monissa toistuu arviointiosaamisen parantamisen suuri tarve. Siihen on esimerkiksi Patton (2010a) kiinnittänyt huomionsa. Sitä tarvitsevat niin yksilöarvioijat kuin kokonaiset organisaatiot, jotta ne voisivat yltää omaehtoiseen evaluatiiviseen kulttuuriin. Ellei halua nihilistisesti kyseenalaistaa yhteiskunnallisen jatkuvan kasvun, edistyksen ja kehityksen tarpeellisuutta, kriittistä arvioivaa katsetta tarvitaan nykykäytäntöjen ja toivottujen tulevaisuuksien välisen kuilun tunnistamiseen ja kaventamiseen.

Arvoiltaan kovenevassa ajassa katse saattaa jyrketä ylikriittiseksi, sillä esimerkiksi OECD-maissa (esim. Eliadis ym. 2011; Vedung 2010) tietoon perustuva, uusliberalistinen arviointiote voimistuu,



Kuvio 1. Onnistuneen arvioinnin tekijät haastatteluaineiston teemojen mukaan

mihin liittyy vahvoja sanktioita (esim. koulutusyksiköiden sulkeminen). Näyttöjen puuttuminen mistä tahansa asiasta voi siis olla kohtalokasta millaiselle toiminnalle tahansa.

Tällaisessa ajassa on suuri tarve työstää arviointikäytänteitä, joita läpäisee Pattonin (2010a, 241) sanoin ”arvostava tarkastelu” (*appreciative inquiry*). Kehittämisen takia tulee olla kriittinen, mutta sitä tulee sävyttää samalla rohkeus tunnustaa osaamista ja vahvuuksia. Ulkopuolisen tekemillä arvioinneilla on oma tärkeä tehtävänsä, mutta toimijoiden osallistaminen ja sisäiset arvoinnit ovat välttämättömiä. Ne tunnistavat paikallista asiantuntijuutta, turvaavat toimijuuden tunnetta ja parantavat tuloksiin sitoutumista siksi, että mukaan osallistettuna on ulkopuolista helpompi löytää arviointituloksista aitoon kehittämiseen johtavia tulkintoja ja päätelmiä.

Tämän puheenvuoron tarkoitus on ollut (uudelleen)suunnata katsetta arviointiin kahdella tavalla. Ensinnäkin otan etäisyyttä oppivelvollisuusikäisten oppimistuloksiin, joiden mittaaminen hallitsee suomalaista kasvatusalan arviointikeskustelua. Vain

kansainvälinen aikuistutkimus PIAAC on onnistunut rikkomaan tätä ”koululaisrintamaa”.

Toiseksi puheenvuoro kiinnittää huomiota vähän tutkittuun aiheeseen: vaikka Suomessa on tehty runsaasti tila-, teema- ja ohjelma-arviointeja, niiden tekemistä ei ole sanottavasti valittu arviointitutkimuksen kohteeksi omien julkaisujeni (Atjonen 2014 ja 2015b) lisäksi sen paremmin kasvatus- kuin aikuiskasvatustieteessä. *Hallinnon tutkimus* -lehdessä arviointia käsitteellistetään tutkimusartikkeleilla säännöllisesti, mitä pidän tärkeänä arviointiaktiivisuuden yleisen lisääntymisen takia.

Aikuiskasvatustieteen keskiössä olevan koulutuksen ja työelämän kehittämisen yksi merkittävistä vipusimista neoliberalistisessa, markkinasuuntautuneeksi kutsutussa ajassa on juuri arviointi: kaikki sitä sanovat tekevänsä, mutta tietävätkö kaikki, mistä (hyvässä) arvioinnissa on kyse?

Päivi Atjonen  
 KT, professori  
 kasvatustieteiden ja psykologian osasto  
 Itä-Suomen yliopisto



## LÄHTEET .....

- AEA (2004). *American evaluation association guiding principles for evaluators*. <http://www.eval.org/p/cm/ld/fid=51>. (Luettu 7.1.2016)
- Ajtonen, P. (2013). Arvioinnin eettiset periaatteet. Teoksessa G. Knubb-Manninen, H. Niemi & V. Pietiläinen (toim.) *Kansallinen arviointi kohti tulevaisuutta. Koulutuksen arviointineuvoston 10-vuotisjuhlajulkaisu*. Koulutuksen arviointineuvoston julkaisuja 63, 117–129.
- Ajtonen, P. (2014). Kehittävän arvioinnin periaatteiden ilmeneminen opetus- ja kasvatustieteen arviointiraporteissa vuosina 2005–2012. *Kasvatus* 45(3), 212–227.
- Ajtonen, P. (2015a). *Kehittävä arviointi kasvatustieteillä*. Joensuu: Kirjokansi.
- Ajtonen, P. (2015b). "Your career will be over" – Power and contradictions in the work of educational evaluators. *Studies in Educational Evaluation* 45, 37–45.
- Bamberger, M. (2012). *Introduction to mixed methods in impact evaluation. Impact Evaluation Notes* No. 3. <https://www.interaction.org/sites/default/files/Mixed%20Methods%20in%20Impact%20Evaluation%20%28English%29.pdf>. (Luettu 11.1.2016)
- Basson, R. (2009). Evaluation research in education – Developments in qualitative approaches. *Evaluation Research in Education* 10(1), 55–70.
- Contandriopoulos, D. & Brousselle, A. (2010). Evaluation models and evaluation use. *Evaluation* 18(1), 61–77.
- Cousins, B., Whitmore, E. & Shulha, L. (2012). Arguments for a common set of principles for collaborative inquiry in evaluation. *American Journal of Evaluation* 34(1), 7–22.
- Dahler-Larsen, P. (2005). *Vaikuttavuuden arviointi. Hyvät käytännöt -menetelmäkäsikirja*. [http://www.julkari.fi/bitstream/handle/10024/77071/vaikuttavuuden\\_arv.pdf?sequen](http://www.julkari.fi/bitstream/handle/10024/77071/vaikuttavuuden_arv.pdf?sequen). (Luettu 5.1.2016)
- Eliadis, P. Furubo, J-E. & Jabob; S. (eds.) (2011). *Evaluation. Seeking truth or power? Comparative policy education*. Volume 11. New Brunsvik & London: Transaction.
- Forss, K., Rebien, C. & Carlsson, J. (2002). Process use of evaluations. Types of use that precede lessons learned and feedback. *Evaluation* 8(1), 29–45.
- Hind, J. (2010). Additionality: a useful way to construct the counterfactual qualitatively? *Evaluation Journal of Australasia* 10(1), 28–35.
- House, E. (1980). *Evaluating with validity*. London: Sage.
- Huusko, M. (2010). Itsearviointi laitosten kehittämisen välineenä. *Yliopistopedagogiikka* 17(2), 6–14.
- Jacob, S. & Boisvert, Y. (2010). To be or not to be a profession. Pros, cons and challenges for evaluation. *Evaluation* 16(4), 349–369.
- JCSEE (2016). *Program evaluation standard statements. Joint committee on standards for educational evaluation*. <http://www.jcsee.org/program-evaluation-standards/program-evaluation-standards-statements>. (Luettu 7.1.2016)
- Kajamaa, A., Kerosuo, H. & Engeström, Y. (2008). Työelämän kehittämisprojektien seuraamusten jäljillä – uusi näkökulma arviointitutkimukseen. *Hallinnon tutkimus* 27(4), 61–79.
- Kushner, S. (2011). A return to quality. *Evaluation* 17(3), 309–312.
- MacDonald, B. (1976). Evaluation and the control of education. Teoksessa D. Tawney (toim.) *Curriculum evaluation today: Trends and implications*. London: MacMillan, 125–136.
- McDonald, B., Rogers, P. & Kefford, B. (2002). Teaching people to fish? Building the evaluation capability of public sector organizations. *Evaluation* 9(1), 5–25.
- Morris, M. (2008). Ethics and evaluation. Teoksessa M. Morris (toim.) *Evaluation ethics for best practice. Cases and commentaries*. New York: Guilford, 1–29.
- Paasio, P. (2006). Yleinen ja erityinen viitekehys arvioinnista. *Hallinnon tutkimus* 25(3), 92–107.
- Patton, M. (2010a). *Developmental evaluation. Applying complexity concepts to enhance innovation and use*. New York: Guilford press.
- Patton, M. (2010b). The future of evaluation in society. Top ten trends plus one. Teoksessa S. Donaldson (toim.) *The future of evaluation*. Charlotte: IAP – Information Age Publishing, 45–62.
- Patton, M. (2012). *Essentials of utilization-focused evaluation*. Los Angeles: Sage.
- Pawson, R. & Manzano-Santaella, A. (2012). A realist diagnostic workshop. *Evaluation* 18(2), 176–191.
- Preskill, H. & Boyle, S. (2008). A Multidisciplinary model of evaluation capacity building. *American Journal of Evaluation* 29(4), 443–459.
- Ryan, K. & Cousins, B. (toim.) (2009). *The SAGE international handbook of educational evaluation*. London: Sage.
- Seppänen-Järvelä, R. & Vataja, K. (2009). Kehittävän itsearvioinnin taidot, kyvykkyudet ja osaaminen. Empiirinen analyysi sosiaalitoimen työyhteisöistä. *Hallinnon tutkimus* 28(5), 60–73.
- Stame, N. (2010a). US sociology and evaluation: Issues in the relationship between methodology and theory. Teoksessa J. Waessen & F. Leeuw (eds) *Mind the gap. Perspectives on policy evaluation and the social sciences. Comparative policy evaluation* 16, 29–44.

- Stame, N. (2010b). What doesn't work? Three failures, many answers. *Evaluation* 16(4), 371–387.
- Suárez-Herrera, J., Springett, J. & Kagan, C. (2009). Critical connections between participatory evaluation, organizational learning and intentional change in pluralistic organisations. *Evaluation* 15(3), 321–342.
- Vanhoof, J. & Petegem, P. (2007). Matching internal and external evaluation in an era of accountability and school development: Lessons from a Flemish perspective. *Studies in Educational Evaluation* 33(2), 101–119.
- Vataja, K. (2012). *Kehittyvä työyhteisö – Itsearviointin hyödyntäminen työyhteisön kehittämisessä kunnallisessa sosiaalitoimessa*. Terveiden ja hyvinvoinnin laitos. Tutkimus 86.
- Vedung, E. (2010). Four waves of evaluation diffusion. *Evaluation* 16(3), 263–277.
- Virtanen, P. (2005). Neljännen sukupolven evaluaation jälkeinen arviointitutkimus? *Hallinnon tutkimus* 24(1), 62–76.
- Yliruka, L. (2015). *Itsearviointi reflektiivisenä rakenteena. Kuvastin-menetelmän toimivuus, käyttöönotto ja kehittäminen*. Heikki Waris -instituutin tutkimuksia 1.