

Heikki K. Lyytinen &
Anu Räisänen (toim.)



Kehittämissuuntaa arvioinnista

Koulutuksen arviointineuvoston julkaisuja 6

Kehittämissuuntaa arvioinnista

Toimittaneet

Heikki K. Lyytinen & Anu Räisänen



Koulutuksen
arviointineuvosto

JULKAISUN MYYNTI:
Koulutuksen arviointisihteeristö
Maritta Leinonen
PL 35
40014 Jyväskylän yliopisto
puh. (014) 260 4685
maritta.leinonen@eval.jyu.fi
www.edev.fi

© Koulutuksen arviointineuvosto ja kirjoittajat

Kansi ja ulkoasu: Martti Minkkinen/Koulutuksen tutkimuslaitos
Taitto: Kaija Mannström/Koulutuksen tutkimuslaitos

ISSN 1795-0155 (painettu)
ISSN 1795-0163 (verkkójulkaisu, pdf)

ISBN 951-39-2172-7 (painettu)
ISBN 978-951-39-3242-8 (verkkójulkaisu, pdf)

Jyväskylän yliopistopaino
Jyväskylä 2005

Sisältö

Saatteeksi

Kauko Hämäläinen 3

Johdanto

Heikki K. Lyytinen & Anu Räisänen 7

I UUDEN SUUNNAN ASKELEITA

Kansainvälinen arviointiyhteistyö hyödyttämään kansallista arviointia

Pirjo Linnakylä & Jouni Välijärvi 13

Arviointitoiminta verkostoihin

Heikki Silvennoinen 27

Arviointi järjestelmälliseksi

Jouko Mehtäläinen 39

Med siktet på samverkan i utvärdering

Gunnel Knubb-Manninen & Sven-Erik Hansén 49

Arviointitiedot hyötykäyttöön

Harri Rönholm 59

Kunta koulutuksen arvioijana

Maisa Lovio & Juha Sihvonen 77

II ARVIOINTI KEHITTÄVÄKSI

Näytöt kehittämisen perustaksi	
Anu Räisänen	91
Kehittävään arviointiin	
Anu Räisänen	109
Kehittämiskoivaa arvioinnista	
Heikki K. Lyytinen & Pentti Nikkanen	129
Arviointi opetuksen ja oppimisen tueksi	
Päivi Atjonen	143
Utvärdering till stöd för lokalt beslutfattande	
Renata Svedlin	153
Ennakoivaan arviointiin	
Olli Hietanen & Jari Kaivo-oja	163

III UUTEEN OPPIMISTULOSTEN ARVIOINTIIN

Perusopetuksen oppimistulokset ohjausjärjestelmän osaksi	
Esko Korkeakoski	175
Oppimistulosten arviointi laadunvarmistuksena. Ammatillinen koulutus	
Anu Räisänen & Mari Rökköläinen	193
Principles and processes in evaluation systems: an outsider looks in.....	
Kathryn Eccleston	213
Mitataan oppilaita, mutta päätellään koulusta	
Jorma Kuusela & Jarkko Hautamäki	229
Kirjoittajaluettelo	243

Saatteeksi

Oppilaitoskohtainen, valtakunnallinen ja kansainvälinen arviointi ovat lisääntyneet joka puolella Eurooppaa viimeisten 15 vuoden aikana. Tämä koskee kaikkien kouluasteiden arviointia. Esimerkiksi 1980-luvun lopulla oli Euroopassa vain neljässä maassa korkeakoulujen arvioinnista vastaava kansallinen elin, mutta nyt niitä on jo kymmeniä, joissakin maissa jopa useampia, kuten Saksassa. Muiden asteiden koulutuksen arviointi on laajentunut samaan tahtiin.

Arviointitoiminnan laajentumisen perusteluina on ollut yleensä se, että julkisin rahoin tuotettujen palvelujen täytyy osoittaa omat tuloksensa. Koulutustoiminnan pitäisi olla läpinäkyvää niin, että voitaisiin seurata, mitä oppilaitokset ja koulujärjestelmä saavat aikaan. Monessa maassa haluttaisiin myös kontrolloida toimintaa. Usein tuodaan esille myös tarve kehittää koululaitosta arvioinnin avulla. Parhaimmillaan arvioinnit tuottavat tietoa sekä valtakunnallisen päätöksenteon pohjaksi että oppilaitosten oman työn kehittämiseksi.

Arvioinnin avulla ei ole saatu aikaan haluttuja tuloksia. Esimerkiksi kansainvälisissä oppimissaavutusten vertailuissa eivät suinkaan ole kärjessä ne maat, joiden arviointijärjestelmät ovat kehittyneet parhaiten. Sama koskee myös korkeakouluja. Tilanne voi olla jopa päinvastainen. Arviointitoiminta voi jopa haitata oppilaitosten toimintaa maissa, joiden arviointikäytänteet ovat massiivisia. Ulkopuolinen kontrolli voi laskea työmoraalia sekä kuluttaa paljon aikaa ja energiaa.

Itse asiassa Suomi näyttää olevan edelläkävijä arviointitiedon hyödyntämisessä. Meillä on jo paljon näyttöä siitä, että saatua arviointitietoa on käytetty tehokkaasti hyväksi esimerkiksi valtakunnallisen opetussuunnitelman ja erilaisten säädösten kehittämiseksi. Samoin erilaista kehittämisrahaa on suunnattu arviointien esille tuomiin kriittisiin kohtiin.

Ulkoisen arvioinnin hyödyllisyys ei siis ole itsestään selvää. Runsaskaan arviointitieto ei sinällään tuota muutosta, vaan tietoa täytyy yrittää hyödyntää tehokkaasti esi-

merkiksi opetussuunnitelmatyössä tai täydennyskoulutuksessa ja opetuksessa. Valtakunnan tasolla taas arviointien pitäisi johtaa tarvittaessa säädösten tai vaikkapa rahanjaon muutoksiin. Mitä enemmän arviointitietoa tuotetaan, sitä vaikeammalta päätöksentekijöiden näyttää olevan käyttää tietoa hyväksi. Suomessakin eri tahojen tuottamaa tietoa on jo niin runsaasti, että juuri kukaan ei ehdi paneutua siihen kunnolla. Harvalla päätöksentekijällä on aikaa seurata suurta julkaisumäärää ja yrittää vetää siitä järkeviä johtopäätöksiä.

Kansainväliset vertailut vaikuttavat tällä hetkellä hyödyllisimmiltä. Viimeisimmät PISA-vertailut ovat aktivoineet todella vilkasta keskustelua monissa maissa, joissa tulokset ovat olleet odotuksia heikompia. Joissakin maissa tulokset ovat herättäneet keskustelua valtakunnallisen opetussuunnitelman roolista, opettajakoulutuksesta, päätäntävällän hajauttamisesta, erityisopetuksen järjestämisestä ja yhtenäiskoulun merkityksestä. Hyvin menestyneissä maissa on keskusteltu puolestaan siitä, mistä menestys johtuu ja mitä avaintoimintoja pitäisi ylläpitää ja kehittää. Kansainvälinen yhteistyö on myös olennaisesti lisääntynyt. Viime vuosina on haluttu oppia muiden maiden hyvistä käytänteistä.

Suomessa on paljon hyvää arviointiosaamista. Monissa yliopistoissa ja keskusviirastoissa on kansainvälisestikin arvioiden ammattitaitoa, joka on jo tuottanut paljon arviointitietoa ja jolla on osaamista tuottaa uutta relevanttia tietoa. Tämä teos antaa hyvän kuvan tästä osaamisesta ja arviointitiedon runsaudesta. Koulutuksen arviointineuvoston yhtenä keskeisenä tehtävänä onkin ollut luoda arviointiosaajien verkosto, jota voidaan käyttää hyväksi erilaisten arviointien suunnittelussa ja toteuttamisessa.

Keskeisenä ongelmanamme on se, että meiltä puuttuu toimiva valtakunnallinen arviointisuunnitelma. Sekä oppilaitosten että niiden ylläpitäjien kannalta ongelmallista on se, että ne ovat lukuisten osittain päällekkäistenkin arviointien kohteena ja valtakunnalliselta arvioinnilta puuttuu suunnitelmallisuutta ja koordinoitua. Kaikkien osapuolten etu olisi, että voitaisiin ennakoita muutamaksi vuodeksi etukäteen, mitä arviointitietoa tarvitaan ja milloin tietoa kerätään. Tällöin oppilaitokset ja kunnat voisivat suunnitella omat arviointikäytänteensä valtakunnallisen tiedontarpeen mukaisesti.

Eri kouluasteiden arviointia olisi syytä koordinoita entistä paremmin. Erityisesti mielenkiintoista olisi saada tietoa eri kouluasteiden taitekohdista ja myös trenditietoja esimerkiksi kielten tai matematiikan oppimisesta eri ikävaiheissa. Eri asteiden arviointiyhteistyön lisäämistä on pyritty esimerkiksi Tanskassa tukemaan yhdistämällä eri asteiden arviointitoiminta samaan organisaatioon.

Suomessakin kannattaisi harkita sitä, että kaikkien asteiden kaikki arviointi kootaisiin yhteen organisaatioon. Jos sillä olisi oma budjetti ja päätösvalta arvioinnin kohteista ja toteutuksesta, voitaisiin monia päällekkäisten arviointien ja lyhyen aikavälin suunnittelun ongelmia voittaa. Etuna voisi olla myös entistä parempi arviointiosaaminen, kun nyt monessa paikassa työskentelevien ammattitaito tulisi yhteiseen käyttöön. Suurin ongelma eri koulutustasojen arvioinnin yhtyeenliittämisestä näyttää Tanskassa olevan se, että korkeakoulujen arvioinnin osuus koko arviointitoiminnassa

on jäämässä entistä vähemmälle huomiolle, koska niitä on lukumääräisesti vähän suhteessa muiden asteiden oppilaitoksiin.

Suomessa valtakunnallista arviointitoimintaa on monella tavoin hyvä kehittää, kuten edellä on todettu. Arviointiosaamisemme on monipuolista ja tasokasta, sitä osataan hyödyntää eri tavoin ja jopa eduskunta on osoittanut mielenkiintoa arvioinnin kehittämiseen. Lähivuosina valtakunnallista arviointia pitäisi suunnitella pitkäjänteisemmin. Arviointien päällekkäiset organisaatiot pitäisi purkaa ja arviointiosaamista hyödyntää. Lisäksi pitäisi pystyä kehittämään sitä, miten arviointituloksia hyödynnetään edelleen.

Johdanto

Koulutuksen arviointipoliittisissa linjauksissa on lähdetty siitä, että arviointia kehitetään turvaamaan koulutusta koskevan lainsäädännön tarkoituksen toteuttamista, tukemaan koulutuksen kehittämistä ja parantamaan oppimisen edellytyksiä. Arviointitietoa tarvitaan niin valtakunnallisen koulutuspoliittisen päätöksenteon kuin koulutuksen paikallisen kehittämisen pohjaksi. Arviointia kehitetään ja ylläpidetään oppivaa yksilöä ja tämän huoltajia, oppilaitoksia, koulutuksen järjestäjiä ja kansallisia päätöksentekijöitä varten. Arvioinnin tulee palvella sekä tilivelvollisuutta että kehittämistä. Näiden kahden näkökulman yhteensovittaminen on monimuotoinen haaste. Tilivelvollisuus viestii valvontaa, kehittäminen taas dynaamista muutosta ja siirtymistä uudelle tuntemattomalle alueelle.

Uuden vuosituhannen alkumatkan arvioinnin tienviitoina ovat voimassa oleva arvioinnin lainsäädäntö sekä koulutuksen ja tutkimuksen kehittämissuunnitelman linjaukset. Tulevaisuus näyttää sen, muodostavatko linjaukset tukevan kivijalan arvioinnin tavoitteiden saavuttamiselle. Tämän päivän arviointi edustaa aina myös historiallista jatkuvuutta, missä arvioinnin aiempi toimintakulttuuri säätelee sitä miten uusien tienviittojen suunnassa edetään. Arvioinnin pitkä historia liittyy oppilasarvosteluun ja sittemmin kansallisiin ja kansainvälisiin koulusaavutuskokeisiin. Arviointi on ollut aiemmin paljolti summatiivista tilivelvollisuusarviointia. Suomalainen tarkastustoiminta koulutusjärjestelmän arvioinnin muotona oli myös tilivelvollisuuskorosteista, joskin siihen liittyi myös kehittävä ulottuvuutensa. Koulutustutkimuksen piirissä on niin ikään arvioitu koko ajan kansallisen koulutuksen toimivuutta uudistusten toimeenpanovaiheessa.

Uusien viittojen pontimina on ollut tyytymättömyys olemassa olevaan tilaan. Tavoitteeksi on asetettu arvioinnin vaikuttavuuden lisääminen ja arvioinnin uskottavuus ja läpinäkyvyys. Tavoitteeksi asetettu verkostomainen ja kehittämistä aktiivisesti tukeva arviointi on ottamassa vasta ensi askeleitaan. Samanaikaisesti uuden suunnan

adaptaatiovaiheessa visioidaan jo uusia viittoja, vaikka kokemuspohja uudesta on varsin ohut ja aika näyttöjen saamiseen ja kulttuurisiin muutoksiin lyhyt.

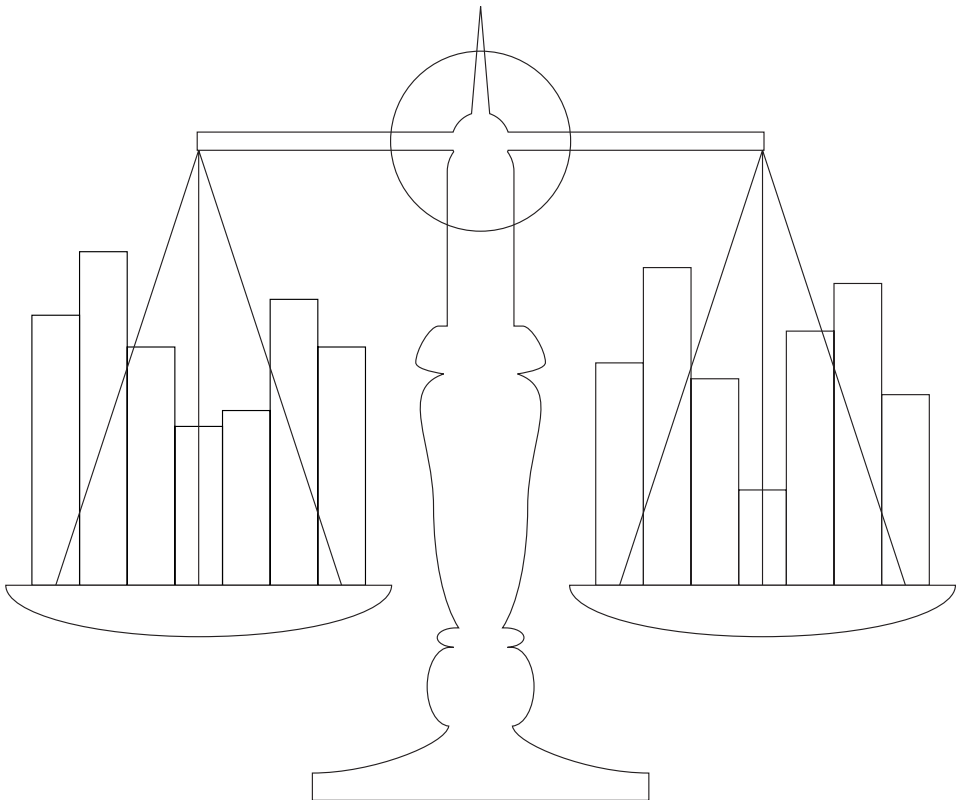
Jälkimodernin yhteiskunnan myötä syntynyt verkostomainen, toimijoita aktiivisesti osallistava ja eri toimijoiden rooleja muuttava arviointikulttuuri hämmentää uutuut-taan. Kulttuurisessa murroksessa arviointitoiminnan ehdoton selkeys on vaatimukse-na kova ja se on saavutettavissa vasta ajan myötä. Arvioinnin kehittämisessä tarvi-taan tulevaisuudenkuvaa – mieluiten kirkasta ja selkeää, jotta sitä kohti voi edetä. Muutoin kehittämistä voi häiritä enemmän vision kuin resurssien puute.

Tällä hetkellä suomalaista koulutuksen arviointia leimaa rakenteellinen moninapai-suus ja siitä seuraavat työnjaolliset selkiyttämisyrittäykset. Moninapaisuus johtaa helposti myös arviointien päällekkäisyyksiin koordinaatiivisen ohjauksen etsiessä uudessa tilanteessa vielä muotojaan. Taloudelliset ehdot säädellävät jatkossa arvioin-nin sellaista keskittämistä, jossa pystytään voimavarojen tarkoituksenmukaiseen käyt-töön säilyttämällä kuitenkin asiantuntijuuden synergiaetujen käyttö verkostoajattelun arvioinnin riippumattomuuden pohjalta.

Tämä Koulutuksen arviointineuvoston määrittämää koulutuksen arvioinnin uutta suuntaa konkretisoiva julkaisu on näyttö arviointiasiantuntijaverkoston toiminnasta. Julkaisun tavoitteena on avata valitun arvioinnin suunnan sisältöä ja edistää uuden arviointikulttuurin juurtumista. Julkaisu on tarkoitettu kaikille koulutuksen arviointi-toiminnasta kiinnostuneille, erityisesti arviointia aktiivisesti sekä kansallisella että paikallisella tasolla kehittäville ja tekeville.

Julkaisu muodostuu kolmesta osasta. Lukija voi syventyä kirjan ensimmäisessä osassa arviointijärjestelmän toimintaan niin kansainvälisestä, kansallisesta ja paikka-lisesta näkökulmasta. Toisessa osassa keskitytään arvioinnin kehittävään tehtävään, jotka tarkastellaan niin lähestymistapojen ja arviointimetodien näkökulmista. Kolmas osa rajautuu oppimistuloksiin, joita käsitellään sekä yleissivistävän että ammatillisen koulutuksen näkökulmista.

UUDEN SUUNNAN ASKELEITA



Koulutuksen arviointitoiminnassa otetaan askeleita lainsäädännössä ja Koulutuksen arviointineuvoston hyväksymässä arviointiohjelmassa määritettyyn uuteen suuntaan. Keskeisenä lähtökohtana on arviointitoiminnan vakiinnuttaminen ja sen vaikuttavuuden lisääminen. Määritetty suunnan muutos asettaa haasteita arvioinnin osuvuudelle, tehokkuudelle, taloudellisuudelle ja vaikuttavuudelle. Sen lisäksi, että arviointikohteiden täsmentämisvaatimukset ovat kasvaneet, muutos haastaa kehittämään myös arviointimetoja ja arviointien organisointitapoja.

Arvioinnin uusi suunta haastaa arvioijat kohdistamaan arviointia aiemmin arvioinneissa painottuneen tehokkuuden rinnalla myös vaikuttavuuteen ja taloudellisuuteen. Vaikuttavuuden arvioinnin ytimessä ovat aina oppimistulokset. Yksittäisten poikkileikkausarviointien vaihtoehtona ovat pitkittäisarvioinnit, joissa päästäisiin entistä enemmän analysoimaan myös oppimistapahtuman, oppimisympäristöön ja muihin kontekstitekijöiden yhteyksiä oppimistuloksiin.

Arviointijärjestelmä koostuu kansallisesta, paikallisesta ja kansainvälisestä arvioinnista. Järjestelmän toimivuutta ja tehokkuutta voidaan parantaa muun muassa ottamalla näiden tasojen tarpeet huomioon jo suunnittelussa ja sovittamalla nämä järjestelmät yhteen. Arvioinnin tehokkuuden ja vaikuttavuuden lisäämiseen pyritään muun muassa parantamalla paikallisen, kansallisen ja kansainvälisen arvioinnin toiminnallista yhteyttä, karsimalla päällekkäisyyksiä, panostamalla arviointitulosten levittämiseen ja kehittämällä toimiva lainsäädännön edellyttämä palautejärjestelmä. Vaikuttavuutta voidaan niin ikään lisätä sillä, että kansallisen tason arvioinneissa hyödynnetään arvioinnin osuvuutta lisäävää paikallista arviointia.

Uusi suunta merkitsee siirtymistä verkostomaiseen arviointikulttuuriin. Arviointitoiminta organisoidaan tieteen, opetushallinnon, käytännön opetustyön ja sidosryhmien ja arviointiasiantuntijoiden verkostona. Verkostomaisen toiminnan vaatimukset kohdistuvat yksittäisiin arviointihankkeisiin ja muuhun arviointitoimintaan kuten arviointimenetelmien kehittämiseen. Painopisteet ovat arviointiverkoston identifioinnissa, yhteisten toimintaperiaatteiden ja yhtenäisen verkostokulttuurin kehittämisessä, vastuiden selkeyttämisessä ja verkostoon kuuluvien asiantuntijoiden sitoutumisessa uuteen toimintakulttuuriin.

Pirjo Linnakylä ja **Jouni Välijärvi** tarkastelevat artikkelissaan **Kansainvälinen arviointiyhteistyö hyödyttämään kansallista arviointia** kansainvälistyvää arviointiympäristöä. Kirjoittajat tähdentävät sitä, että kansainväliset arvioinnit voivat osoittaa kansallista terävämmin oman koulutusjärjestelmän vahvuudet ja kehittämistarpeet. Kirjoittajat pitävät erityisen tärkeänä ja jopa välttämättömänä kansallisen ja kansainvälisen arviointitoiminnan koordinoitua arviointien vaikuttavuuden vahvistamiseksi ja kiinnostuksen säilyttämiseksi arviointiin.

Heikki Silvennoisen artikkelissa **Arviointitoiminta verkostoihin** tarkastellaan verkostotoiminnan lähtökohtia, uuden verkostomaisesti toimivan arviointijärjestelmän toimintaperiaatteita ja rajoituksia sekä kehittämistä. Artikkelin antaa myös eväitä verkostomaisen toiminnan käsitteellistämiseen ja mallintamiseen. Verkostomainen arviointitoiminta sidostetaan artikkelissa erityisesti kehittävän arvioinnin paradigmaan.

Jouko Mehtäläinen hahmottaa artikkelissaan **Arviointi järjestelmälliseksi** järjestelmällistä kokonaisuutta. Kokonaisuuden jäsentämisellä tavoitellaan arvioitavien ilmiöiden syvempää ymmärrystä ja ennen pitkää arvioinnin nykyistä parempaa osuvuutta ja taloudellisuutta. Kirjoittaja korostaa kokonaisuudesta lähtevän arviointitiedon kumuloitumisen merkitystä.

Gunnel Knubb-Manninen ja **Sven-Erik Hansén** tarkastelevat artikkelissaan **Med sik-tet på samverkan i utvärdering** koulutuksen arviointia ruotsinkielisen koulutuksen näkökulmasta. Uuden kansallisen arviointiverkoston tehtäväksi nähdään koulutuksen kehittäminen palvelemaan elävää kaksikielistä yhteiskuntaa, jossa ruotsinkielisen koulutuksen erityistarpeet otetaan huomioon.

Harri Rönholm käsittelee artikkelissaan **Arviointitiedot hyötykäyttöön** arvioinnin hyödynnettävyyttä. Hyödynnettävyyttä tarkastellaan artikkelissa toisaalta säädösten asettamien muodollisten vaatimusten, toisaalta arviointiprosessin tuottamien hyötyjen kannalta. Kirjoittajan mielestä on keskeistä ymmärtää, että arviointiprosessi itsessään ja siihen liittyvät menetelmälliset valinnat vaikuttavat arvioinnin hyödynnettävyyteen. Hyödyntämistä voidaan edistää luovilla ja joustavasti toteutetuilla arvioinneilla, osallistamalla käyttäjät ja sidosryhmät arvioinnin eri vaiheisiin, liittämällä arviointi johtamiseen ja strategiseen suunnitteluun ja tekemällä arviointituloksiin perustuvia päätöksiä.

Maisa Lovio ja **Juha Sihvonen** tarkastelevat artikkelissaan **Kunta koulutuksen arvi-oijana** kunnan koulutuksen arviointia kokonaisuutena, johon sisältyy myös koulutuksen järjestäjän arviointi sekä koulujen ja oppilaitosten itsearviointi. Tärkeimmäksi kirjoittajat nostavat kunnan valtuuston asettamien tavoitteiden saavuttamisen arvioinnin ja kunnan oman vastuun oman toimintansa arvioinnissa. Keskeiseksi haasteeksi nostetaan kansallisen ja paikallisen arviointitoiminnan toisiaan täydentävyys ja eri tahojen välisen vuoropuhelu ja arviointitoiminnan avoimuus. Artikkelissa esitetään myös esimerkki EFQM:n mukaisesta oppilaitoksen itsearvioinnista.



Kansainvälinen arviointiyhteistyö hyödyttämään kansallista arviointia

Artikkelissa tarkastellaan kansainvälistyvää arviointiympäristöä ja korostetaan kansallisen ja kansainvälisen arviointitoiminnan koordinoimnin välttämättömyyttä. Ennen kaikkea tähdennetään sitä, että kansainväliset arvioinnit voivat osoittaa kansallista terävämmin oman koulutusjärjestelmän vahvuudet ja kehittämistarpeet. Kansallisen ja kansainvälisen arviointitiedon vertailu tuottaa myös syvällisempää ja jalostuneempaa ymmärrystä koulutuksen ja opetuksen tuloksista. Koordinointi mahdollistaa näin monipuolisen, syntetisoivan ja laadullisesti korkeatasoisen tiedon välittymisen päättäjille ja julkiseen keskusteluun. Yksittäisten arviointien tulokset ovat usein kapea-alaisia. Niitä tulisikin tarkastella erilaisissa konteksteissa, jos niiden avulla halutaan rakentaa pitkäjänteisesti objektiivista ja eri näkökulmista tarkentuvaa kuvaa koulutuksestamme ja sen kehittymisestä. Koordinoimattomuus korostaa yksittäisten arviointien tulosten riskitietäisyyksiä, hämmentää ja heikentää vaikuttavuutta ja johtaa pahimmillaan kiinnostuksen ja uskon sammumiseen koko arviointitoimintaan. Kansainvälisen arviointiyhteistyön hyötynä kansalliselle arvioinnille painotetaan metodologisen asiantuntemuksen jakamista ja eettistä kestävyyttä. Vaikka kansallisen ja kansainvälisen arviointitoiminnan yhteensovittaminen ei ole ongelmaton, koordinointi tuottaisi nykyistä laajemman kehysten koulutuspolitiikan, opetussuunnitelmien ja oppimisympäristöjen kriittiseen tarkasteluun, etenkin jos arviointitoiminta integroidaan läheisesti myös koulutuksen tutkimukseen ja tutkijakoulutukseen.

Asiasanat: kansallinen arviointi, kansainvälinen arviointi, oppimistulokset, koulutusjärjestelmä



Kansainvälistyvä arviointiympäristö

Koulutusjärjestelmien ja etenkin niiden tuottamien oppimistulosten kansainvälinen arviointi on laajentunut talouden globalisaation, tieto- ja viestintäteknologian kehityksen sekä kulttuurin ja koulutuksen kansainvälistymisen myötä. Osaamisesta ja opiskelun taidoista on tullut keskeisin yksilöiden, yhteisöjen ja kansakuntien voimavara ja kilpailukeino tietoyhteiskunnassa. Tämän seurauksena työvoiman ja opiskelijoiden liikkuvuus sekä koulutusjärjestelmän ja tutkintojen läpinäkyvyys korostuu entisestään tulevaisuudessa (Koulutuksen arvioinnin uusi suunta 2004, 8).

Talouden globalisoitumiseen ja koulutuksen kansainvälistymiseen liittyy myös uhkatekijöitä. Näitä ovat taloudellisen kilpailun koveneminen, alueiden eriarvoistuminen sekä kansakuntien ja yksilöiden jakautuminen menestyjiin ja syrjäytyjiin. Koulutusjärjestelmää ei voida arvioida vain taloudellisen tai sivistyksellisen pääoman tuottajana, vaan myös sosiaalisen oikeudenmukaisuuden ja yksilöiden elämänlaadun näkökulmasta. Monet järjestelmätason arvioinnit ovatkin osoittaneet, että ne kansakunnat, jotka huolehtivat hyvin heikoimmasta väestöosastaan, menestyvät parhaiten niin sivistyksellisesti kuin taloudellisesti (OECD 2000; OECD 2001).

Avoimempi koulutuspolitiikka, päätösvallan hajauttaminen, informaatio- ja tulosohtaus, julkisten menojen leikkaus, oppilaitosten ja opiskelijoiden itseohjautuvuus sekä pedagogisten ratkaisujen valinnaisuus ovat osaltaan lisänneet tarvetta koulutuksen arviointiin. Koulutusjärjestelmän arviointi ja kehittäminen ei kuitenkaan enää onnistu yksinomaan kansallisesti. Euroopan ja maailman yhdentyessä koulutuksen, opetuksen ja oppimisen ongelmat ovat paljolti yhteisiä, ja niitä on myös ratkaistava yhdessä sekä seurattava ratkaisujen toimivuutta ja tuloksellisuutta kansallisen koulutusjärjestelmän kehittämiseksi (Linnakylä 2002).

Koordinoinnin välttämättömyys

Parhaimmillaan kansainvälinen arviointitoiminta täydentää kansallisia arviointeja. Ennen kaikkea kansainväliset arvioinnit voivat osoittaa kansallista arviointia terävämmin oman koulutusjärjestelmän vahvuudet ja kehittämistarpeet. Läheltä on joskus vaikea nähdä, mikä kansallisessa koulutusjärjestelmässä ja oppimiskulttuurissa on omaperäisintä ja onnistuneinta ja missä ovat pahimmat puutteet ja ongelmat, joihin on syytä kansallisesti puuttua. Tässä auttaa eurooppalainen tai maailmanlaajuinen vertailukehys. Kansalliset arvioinnit taas voivat kohdentua kansainvälisiä tarkemmin ja syvällisemmin ongelma-kohtiin ja keskittyä opetussuunnitelman toteutumisen seurantaan.

Kansallisen ja kansainvälisen arviointitiedon tarkastelu rinnakkain voi myös tuottaa entistä syvällisempää ja jalostuneempaa tietoa sekä ymmärrystä koulutuksen ja opetuksen tuloksista ja taustatekijöistä. Koordinointi mahdollistaa myös monipuoli-

sen, syntetisoivan ja laadullisesti korkeatasoisen tiedon välittymisen päättäjille ja julkiseen keskusteluun. Yksittäisten arviointien tulokset ovat usein kapea-alaisia. Niitä tulisikin tarkastella erilaisissa konteksteissa, jos niiden avulla halutaan rakentaa pitkäjänteisesti objektiivista ja eri näkökulmista tarkentuvaa kuvaa koulutuksestamme ja sen kehittymisestä. Koordinoimattomuus korostaa helposti yksittäisten arviointien tulosten ristiriitaisuuksia, hämmentää ja heikentää arvioinnin vaikuttavuutta ja johtaa pahimmillaan kiinnostuksen ja uskon sammumiseen koko arviointitoimintaan.

Kansallisen ja kansainvälisen arviointitoiminnan koordinointi lisää myös taloudellisuutta ja tehokkuutta. Koulutuksen arviointiin suunnataan huomattavia summia julkisia varoja, jotka voitaisiin käyttää tehokkaammin, jos kansainvälisistä arvioinneista saatavaa monipuolista tietoa hyödynnettäisiin kansallisesti nykyistä monipuolisemmin. Tiheästi ja samansisältöisinä toistuvien oppimistulosten arviointien rinnalle ja sijaan tarvittaisiin opetuksen ongelmakohtiin syventyviä arviointeja, joissa voitaisiin puretua perusteellisemmin ongelmien syihin ja ratkaisukeinoihin. Jos esimerkiksi nuorten lukutaitoa ja matematiikan taitoja arvioidaan kansainvälisesti monipuolisilla mittareilla kolmen vuoden välein, näitä alueita tuskin kannattaa mitata kansallisesti joka toinen vuosi. Huomiota voisi kiinnittää välillä esimerkiksi matematiikan ja äidinkielen opiskeluasenteiden ja motivaation tai heikoimmin menestyneiden oppilaiden tukitoimien arviointiin.

Kansainvälisten arviointien lisäarvo kansalliselle arviointitoiminnalle

Kansainväliset arvioinnit tuovat monenlaista lisäarvoa kansalliselle arviointitoiminnalle (Leimu 2001; Linnakylä 2001; Plomp 2001). Ennen kaikkea ne tarjoavat yhtä koulutusjärjestelmää laajemman kehyksen, jossa tarkastella myös koulutuspolitiikkaa ja sen tuottamia tuloksia. Kansainväliset vertailukohteet voidaan valita kulttuurisesti lähellä olevista naapurimaista, EU:n piiristä tai maailmanlaajuisesti. Maamerkeiksi voidaan ottaa esimerkiksi vertailumaiden keskimääräinen laatutaso, parhaiten menestyneiden maiden taso tai oman maan aiempi tai kansallisten arviointien tuottamat tulokset (Linnakylä 2002).

Kansainväliset arvioinnit antavat myös mahdollisuuden tutustua pintaa syvällisemmin erilaisiin koulutusjärjestelmiin ja niiden tapoihin ratkaista samoja ongelmia. Ne antavat myös mahdollisuuden arvioida erilaisten ratkaisujen tuloksellisuutta. Tätä ei pelkän kansallisen aineiston varassa voi tehdä. Useimpia maita esimerkiksi kiinnostaa oppimistulosten tason ohella tulosten ja oppimismahdollisuuksien vaihtelu koulutuksen tasa-arvoperiaatteen näkökulmasta (Lie ym. 2003). Tähän liittyen on erityisen kiinnostava selvittää, miten koulutusjärjestelmän rakenne sekä hajautettu tai keskitetty koulutusjärjestelmä liittyy oppimisen laatuun ja oppimismahdollisuuksien tasa-arvoon ja miten suurien oppilaiden, koulujen ja alueiden väliset erot ovat. Vaikka

erot saattavat näyttäytyä kansallisesti selvinä, ne voivat kansainvälisesti verrattuna olla kuitenkin vähäisiä. Tämä tieto voi olla tärkeä etenkin silloin, kun pohditaan kehittämiskorjausratkaisuja, jotka ovat vertailumaissa tuottaneet huomattavasti suurempia eroja.

Kansainväliset arvioinnit syventävät ja laajentavat ymmärrystämme opiskeluun, kouluympäristöön, opetussuunnitelmiin, opettajankoulutukseen ja kulttuuriin liittyvistä tekijöistä ja niiden yhdysvaikutuksista, jotka vaikuttavat oppimistulosten taustalla ja ovat yhteydessä niihin. Miten sukupuolten välinen tasa-arvo toteutuu eri koulutusjärjestelmissä ja oppiaineissa? Miten vahva selittäjä kodin tai kouluympäristön sosio-ekonominen tausta on, ja miten nämä tekijät liittyvät koulutusjärjestelmään, koulujen välisiin eroihin, koulun ulkopuolisiin opiskelumahdollisuuksiin ja jatko-opiskeluhalukkuuteen? Miten erilaisia oppijoita, erityisesti heikosti menestyviä, ohjataan ja mitkä ovat heidän mahdollisuutensa oppia erilaisissa koulutusjärjestelmissä? Mikä on maahanmuuttajien osuus ja kielitausta, ja mitkä järjestelmätason ja pedagogiset ratkaisut tukevat parhaiten maahanmuuttajien opiskelua? Näihin kysymyksiin vastaamisessa kansainväliset arvioinnit antavat usein kansallista laajemman näkökulman ja tietoa, joka suuntaa ongelmien ennalta ehkäisyä. Esimerkiksi samoilta alueilta kotoisin olevien maahanmuuttajien oppimistulokset vaihtelevat huomattavasti eri koulutusjärjestelmissä ja erilaisten opiskeluympäristöjen tuloksena (esim. Hvistendahl & Roe 2004; Linnakylä ym. 2004).

Selitys- ja ennustemallien kautta arvioinnit tuottavat myös uusia käsitteellisiä kehyksiä sekä entistä pätevämpiä ja luotettavampia mittavälineitä. Arvioinnit tuottavat myös laajoja empiirisiä tiedostoja, joilla on jatkuvaa tutkimusmerkitystä ja jotka voivat toimia kansallisten ja kansainvälisten jatkotutkimusten ja tarkempien analyysien perustana (esim. Döbert ym. 2004; Linnakylä ym. 2004; Välijärvi & Malin 2003). Kansainväliset arviointiaineistot ovat kaikkien tutkijoiden käytettävissä heti ensi tulosten julkistamisen jälkeen.

Kansainväliset arviointitutkimukset tarjoavat tekijöilleen monipuolista tutkijakoulutusta ja metodologista asiantuntemusta, jota on vaikeaa ja kallista hankkia kansallisesti. Hankkeita koordinoivat tunnettujen tutkimuslaitosten konsortiot ja niitä johtavat tieteellisesti arvostetut asiantuntijat. Hankkeissa työskentelee tavoitteellinen, kansainvälinen yhteistyöverkosto, joka vastaa eri maissa arvioinnin toteutuksesta vaativin laadullisin kriteerein. Arviointeihin liittyy nykyisin myös metodisia koulutusjaksoja. Esimerkiksi PISA-hankkeen yhteydessä on järjestetty useita viikkoja kestävä metodikoulutus eri maiden tutkijoille koskien sitä, miten hyödynnetään osioanalyysseja ja monitasomalleja, kun analysoidaan hierarkkisesti rakentuneita kouluaineistoja. Kansainvälisten arviointihankkeiden yhteydessä virinnyt yhteistyö on usein johtanut syventäviin tulosanalyysihin tai joidenkin ongelmien laaja-alaiseen tutkimukseen. Esimerkiksi IEA:n matematiikan ja luonnontieteiden arviointeihin on liittynyt opetussuunnitelmien ja autenttisten opetustilanteiden vertailua ja analysointia videoaineistoja hyödyntäen (Stigler & Hiebert 1999).

Valmiin yhteistyöverkoston avulla on myös mahdollista vahvistaa kansainvälistä kehittämis-, tutkimus- ja koulutusyhteistyötä maailmanlaajuisen koulutusvastuun edistämiseksi. Monet kehitystyöjärjestöt liittävätkin rahoitukseen vaatimuksen koulutuksen kehittämisestä ja sen arvioinnista. Esimerkiksi UNESCO:n World Education Indicators -toimintaa tuetaan OECD:n koulutusindikaattoriyhtiöllä. Myös PISA-hankkeeseen on liitetty erillinen PISA Plus -hanke, joka mahdollistaa OECD:n ulkopuolisten maiden osallistumisen ja jonka avulla on mahdollista vahvistaa erityisesti Balkanin ja Itä-Euroopan maiden demokratisoitumiskehitystä (OECD & UNESCO 2003).

Kansainväliset arvioinnit antavat erityisesti pienille maille tilaisuuden tehdä tunnetuksi omaa kansallista koulutus- ja opetuskuulttuuriaan. Erityisesti silloin kun maan oppimistulokset ovat erinomaiset, muut maat kiinnostuvat myös tulosten taustalla olevista kulttuuri- ja järjestelmätekijöistä.

Kansainvälinen arviointitoiminta laajenee, tieteellistyy ja politisoituu

Oppimistulosten sekä kansallinen että kansainvälinen arviointi on saanut viime vuosikymmen aikana entistä virallisemmän aseman Suomessa. Kansainvälisessä arviointitoiminnassa oppimistulosten vertailevasta arvioinnista vastasi 1950-luvun lopulta vuosituhanen vaihteeseen ensisijaisesti kansallisten tutkimuslaitosten yhteistyöjärjestö IEA (The International Association for the Evaluation of Educational Achievement). Suomessa IEA:n kansallisena keskuksena on toiminut Koulutuksen tutkimuslaitos (aiemmin Kasvatustieteiden tutkimuslaitos). Toiminta oli useita vuosikymmeniä pääosin yksittäisten tutkijoiden ja tutkijaryhmien aloitteiden ja intressien varassa, joskin opetusministeriö tuki toimintaa taloudellisesti. Suomi on osallistunut aktiivisesti IEA-järjestön puitteissa 15 kansainvälisen arviointihankkeen toteuttamiseen lähinnä lukutaidon, matematiikan ja luonnontieteiden, yhteiskuntatiedon sekä tietoteknologian opetuksen alueilla (Leimu 2004).

1990-luvulla myös OECD ryhtyi aktiivisesti hankkimaan arviointitietoa oppimistuloksiin liittyvien koulutusindikaattoriensa pohjaksi (Owen 2001). OECD:n aloitteesta toteutettiin 1994–1998 Aikuisten kansainvälinen lukutaitotutkimus (IALS/SIALS). Samoihin aikoihin käynnistettiin 15-vuotiaiden nuorten oppimistulosten arviointiohjelma PISA (Programme for International Student Assessment), johon kaikki OECD-maat ja joukko muita maita osallistui ensimmäisen kerran vuonna 2000. PISA-arviointi toteutetaan kolmen vuoden väliajoin, ja se pyrkii selvittämään, miten 15-vuotiaat nuoret hallitsevat tietoyhteiskunnassa keskeisiä taitoja erilaisissa koulutusjärjestelmissä ja opetuskuultuureissa. PISA on muuttanut olennaisesti arviointitoiminnan luonnetta ja kytkenyt sen entistä läheisemmin koulutuspoliittiseen päätöksentekoon. Ohjelmaa toteutetaan hallitusten välisenä yhteistyönä, ja siinä hyödynnetään monipuolisesti sekä koulutuspoliittista että tutkimuksellista asiantuntemusta (Owen ym. 2004).



Katse tulevaisuuden avaintaitoihin

Kansainvälisissä arviointiohjelmissa hankitaan säännöllisin väliajoin pätevää, luotettavaa ja kansainvälisen vertailun mahdollistavaa tietoa oppimistuloksista ja jatko-opiskelun edellytyksistä sekä niihin yhteydessä olevista yksilöllisistä, sosioekonomisista ja kulttuurisista tekijöistä. Eri ohjelmien tavoitteet ja tarkoitukset kuitenkin poikkeavat jossain määrin toisistaan. PISA-ohjelma poikkeaa IEA:n ja kansallisten arviointien tavoitteista etenkin siten, että se ei kohdistu suoranaisesti kansallisiin opetussuunnitelmiin. Sen sijaan arviointiperustaksi on yhdessä neuvoteltu tulevaisuuden jatko-opintojen, työelämän ja aktiivisen kansalaisuuden sekä elinikäisen oppimisen kannalta keskeiset avaintaidot (Owen ym. 2004). PISA:n edustamia avaintaitoja on määritelty ja käsitteellistetty tietoyhteiskunnan elinikäisen oppimisen strategian, tietotyön ja aktiivisen kansalaisuuden vaatimusten näkökulmasta (OECD 2001; Välijärvi & Linnakylä 2002).

Vaikka sekä IEA:n että OECD:n ohjelmissa arviointikohde vaihtelee, ja huomio kiinnitetään ensisijaisesti yhteen tai muutamaankin oppiaineeseen, yksittäisistä oppiainearvioinneista ollaan siirtymässä arviointeihin, jotka yhdistävät useampia aineita ja tiedonalueita sekä painottavat tiedon soveltamista. Erityisesti tähdennetään tiedon hakua, ymmärtämistä, tarkoituksenmukaista käyttöä, ongelmanratkaisua ja kriittistä pohdintaa. Arviointimenetelmissä suositaan elämänläheisiä, mahdollisimman autenttisia tehtäviä, joissa oppilas joutuu käyttämään monenkin oppiaineen tietoja ja taitoja, niin koulussa oppimistaan kuin koulun ulkopuolelta hankkimiaan. Lisäksi arvioidaan opiskeluun ja oppimiseen liittyviä asenteita ja metataitoja, esimerkiksi itse-tuntoa, motivaatiota, metakognitioita, oppimisen strategioita, oman opiskelun suunnittelua ja säätelyä, ryhmässä työskentelyn ja ongelmanratkaisun taitoja sekä tietoteknologian käyttöä. Taustatekijöistä selvitetään monipuolisesti oppilaan ja kouluympäristön kulttuurista, taloudellista ja sosiaalista kontekstia. Erityistä huomiota kiinnitetään koulutuspoliittisesti polttaviin kysymyksiin, kuten tasa-arvon toteutumiseen koulutusjärjestelmässä ja oppimisympäristöissä, koulujen resursseihin ja päätöksentekoon sekä vanhempien ja koulujen yhteistyöhön, opetus- ja opiskeluympäristön muutokseen ja teknologian hyödyntämiseen, oppilaiden asenteisiin ja uskomuksiin sekä tulevaisuuden jatko-opiskelusuunnitelmiin. Kansallisia koulutusjärjestelmään ja opetusratkaisuihin liittyviä tietoja kerätään myös arviointien yhteydessä. (OECD 2001; Peschar 2004; Klieme 2004; Välijärvi & Linnakylä 2002.)

Metodologiset laatuvaatimukset tiukentuvat

Kansainvälisten arviointien teoreettiset ja metodiset vaatimukset ovat viime vuosina entisestään kasvaneet, mikä vahvistaa arviointien tutkimuksellista otetta ja tämän myötä aineistojen luotettavuutta, pätevyyttä ja vertailukelpoisuutta. Kansainvälisissä

arvioinneissa hankittua metodologista asiantuntemusta voitaisiin hyödyntää entistä monipuolisemmin myös kansallisissa arvioinneissa.

Kansainvälisiä arviointiohjelmaa edeltää nykyisin vuosien kehittämistyö käsitteellisesti vahvojen arviointikehysten rakentamiseksi. Kehittämistyö on pohjana myös kun laaditaan koetehtäviä, tulkitaan tuloksia ja tehdään johtopäätöksiä. Esimerkiksi PISA-ohjelmaan vuonna 2003 sisällyttävä ongelmanratkaisun arviointia valmisteltiin kansainvälisenä yhteistyönä jo vuodesta 1997 (Klieme 2004). Arviointikohteen teoreettinen käsitteellistäminen toteutetaan useimmiten kansainvälisissä asiantuntijaryhmissä, joihin valitaan osallistujamaista alan parhaita asiantuntijoita. Tavallisesti teoreettinen kehys julkistetaan kaikkien osallistujamaiden yhteiseen keskusteluun ja kritiikkiin. Mittavälineet rakennetaan teoreettisen kehäyksen pohjalta kansainvälisissä tutkimuskeskuksissa siten, että kaikki mukana olevat maat voivat osallistua tehtävien tarjontaan ja kehittämiseen.

Mittavälineiden laatimiseen liittyy tavallisesti useita esikoevaiheita eri maissa. Näissä testataan sekä tehtävien kulttuurista sopivuutta, mittavälineiden pätevyyttä ja soveltuvuutta erilaisille oppilasryhmille että arviointiprosessin toimivuutta. Tehtävien kulttuurisen sopivuuden arvioinnissa käytetään yksilö- ja ryhmäkokeita, haastatteluja sekä erilaisia asiantuntija- ja oppilaspaneelita. Tehtäviä testataan eri puolilla maailmaa jo ennen esikoetta. Validius varmistetaan monipuolisesti arviointikohteen teoreettisesta määrittelystä ja arviointialueiden spesifioinnista mittavälineiden rakentamiseen ja arviointiprosessin monitorointiin. Myös tulkinnan pohjana olevien asteikojen rakentamisessa hyödynnetään uusimpia osioanalyysimenetelmiä (Item Response Model).

Mittavälineiden kehittämisessä huomio on erityisesti kiinnitetty autenttisten ja avoimien tehtävätyyppien laatimiseen sekä laadullista analyysia vaativien vastausten pätevään ja luotettavaan laadulliseen arviointiin sekä kansallisesti että kansainvälisesti (OECD 2001). Ongelmanratkaisutehtävissä seurataan myös ratkaisuprosessin etene mistä ja metakognitiivista sääätelyä. Uuden teknologian mahdollisuuksia pyritään myös hyödyntämään monipuolisesti sekä tiedon kulun että aineistojen keruun, analyysien tarkistuksen ja raportoinnin tehostamiseksi.

Arviointitutkimusten otannat suoritetaan useimmiten keskitetysti kansallisten rekisteriaineistojen pohjalta siten, että alueellinen, etninen ja kielellinen edustavuus taa taan. Jälkipainotuksilla korjataan otantaa ja vastaamattomuuden aiheuttamia vääristymiä (Chromy 2002). Katoanalyysit ovat tarkkoja, ja puuttuvia havaintoja estimoidaan erilaisin kertoimin. Myös systeemisten selitys- ja ennustemallien rakentamiseen kehitetään jatkuvasti uudenlaisia tilastollisia menetelmiä, jotka ottavat huomioon aineistojen rakenteisuuden. Monitasomallien hyödyntäminen tulosaineistojen analyysissa onkin tuonut uusia välineitä, joiden avulla voidaan tarkentaa erityisesti koulujen ja alueiden samoin kuin osa-alueiden välisiä eroja (Bosker & Snijders 2004; Raudenbush & Kim 2002). Uudet graafiset ja visuaaliset esitysmuodot helpottavat tulosten tulkin taa ja johtopäätösten tekoa sekä monipuolista tiedottamista erilaisille yleisöille.



Eettisesti kestävä arviointi

Arviointien toteutukselle asetetaan nykyisin entistä mittavampia eettisiä ehtoja, joita on syytä pohtia sekä kansallisesti että kansainvälisesti (House 1997; Vuorenmaa 2001). Erityisesti arvioinnin motiivit ja tavoitteet sekä julkisuus ja sanktiot ovat herättäneet keskustelua. Onko arvioinnin tarkoitus kontrolli, tehokkuuden ja taloudellisuuden takaaminen vai pätevän ja luotettavan arviointitiedon hankinta kehittämistaroituksessa? Merkitseekö avoimuus tulosten julkisuutta ja ketä julkisuus palvelee? Miten tuloksia tulkitaan, kenen eduksi ja millä seurauksin? Palvelevatko arvioinnit koulutuspolitiikan tai hallinnon legitimointia vai tuovatko ne rehellisesti esille sekä ongelmia että vahvuuksia? Miten arvioitavien etuja ajetaan, mikä oikeus heillä on arviointitietoon ja palautteeseen? Tämäntapaisiin kysymyksiin arviointien toimeenpanijat ja tutkijat joutuvat vastaamaan.

Arviointien kohdentumisessa, mittavälineissä ja toteutuksessa ei saisi kansallisesti syrjiä pieniäkään kielellisiä, etnisiä, kulttuurisia tai sosiaalisia vähemmistöryhmiä eikä kansainvälisesti väheksyä pieniä kansoja, kehitysmaita, alueita tai kieliryhmiä. Tulokinnassa ei kuitenkaan saa sulkea silmiä ikäviltäkään tuloksilta. Tuloksia on uskallettava katsoa avoimesti ja kriittisesti, mahdollisesti vallitsevaa yhteiskunta- ja koulutuspolitiikkaa, yleistä mielipidettä, median välittämiä näkemyksiä tai jonkin eturyhmän intressejä vastaan (House 1997; Lindqvist 1999).

Joskus yhteinen etu voi olla ristiriidassa yksilön tai jonkin ryhmän edun kanssa. Esimerkiksi erilaisten vähemmistöryhmien oppimistulokset voivat poiketa merkittävästi edukseen tai haitakseen valtaväestön tuloksista. On vaikea kysymys, miten avoimesti tulokset voidaan välittää päätöksentekijöille ja julkisuuteen. Esimerkiksi Suomessa eri maahanmuuttajaryhmien oppimistulosten taso on varsin vaihteleva. Miten tulos voitaisiin välittää siten, että arviointitietoa käytetään näiden ryhmien koulutuksen kehittämiseen eikä ryhmien leimaamiseen tai jopa syrjinnän pahentamiseen? Tulosten kontekstointi ja kulttuuristen taustojen valottaminen on tärkeää, vaikka se lisääkin tulkinnanvaraisuutta ja joskus jopa tulosten vähättelyä (LeTendre 2002).

Kansainvälisissä arvioinneissa kansallinen ja kielellinen syrjintä nousee jatkuvasti kiivaidenkin keskustelujen kohteeksi. Etenkin afrikkalaiset ja aasialaiset kokevat angloamerikkalaisen valtakulttuurin ja kielen jyräävän muita. Ranskalaiset, italialaiset ja espanjalaiset valittavat taas omaa kielikulttuuriaan syrjittävän. Pohjoismaalaisista taas tuntuu siltä, että englantilaisen, ranskalaisen ja saksalaisenkin kulttuurin asema on selvästi vahvempi kuin pohjoismaisen. Kehitysmaat tuovat puolestaan esille resurssiensa vähäisyyden, joka tulisi ottaa huomioon myös tuloksia esitettäessä. Näitä ongelmia ei kohdata kansallisissa arvioinneissa, mutta ne herättävät pohtimaan kehitysalu-eiden, maahanmuuttajien ja kielivähemmistöjen asemaa arvioinnissa.

Arvioinnin laatukriteerien määrittely on kansainvälisessä arviointitoiminnassa erityisen ongelmallista, sillä arvioinnin laatukriteerit perustuvat koulutuksen laadun yleisiin kriteereihin (Lundgren 2001). Kansainvälisissä arvioinneissa koulutuksen kehittä-

tämisen intressit ja arvoperustat vaihtelevat tietysti enemmän kuin kansallisesti. Yhteisen arviointiperustan luominen onkin aina syvällisten ja kiivaittenkin tutkimus- ja koulutuspoliittisten keskustelujen kautta neuvoteltava niin, että kaikki mukana olevat maat voivat sen hyväksyä. Keskustelujen monitahoisuus paljastaa usein todellisia arvostiririitoja, jotka näkyvät kansallisissa koulutusjärjestelmissä erilaisina järjestelmätason ratkaisuina, koulutusrakenteina, opetussuunnitelmina ja hyvinkin vaihtelevina opetusjärjestelyinä, materiaali- ja menetelmäpainotuksina sekä arviointimenetelminä (Linnakylä 2001). Näiden keskustelujen kautta selkiytyy myös oman kansallisen koulutusjärjestelmän ja opetuskulttuurin arvopohja ja oppimistulosten laadun kriteeristö.

Kansallisten ja kansainvälisten arviointien yhteensovittamisen ongelmia

Kansallisten ja kansainvälisten arviointien yhteensovittaminen ei ole ongelmatonta. Kansallisesti polttavimmat koulutuspoliittiset ja pedagogiset ongelmat eivät välttämättä kiinnosta kansainvälisesti. Esimerkiksi kansallisesti kiinnostavat kysymykset oppimismahdollisuuksien alueellisista eroista, taloudellisen laman jälkimainingeista, hajauttamisen ja valinnaisuuden kasvun seurauksista, erityisopetuksen tarpeesta tai eri kieliryhmien oppimistulosten vertailusta eivät muita maita välttämättä kiinnosta, joten näiden kysymysten tarkempaan arviointiin ei kansainvälisesti keskitytä. Tällöin kansallista arviointia on syytä kohdentaa esimerkiksi näihin teemoihin.

Kansainväliset arvioinnit tarjoavat kuitenkin kansallista arviointia paremmat mahdollisuudet koulutuspolitiikan, opetussuunnitelmien ja oppimisympäristöjen kritiikkiin, kun erilaisten koulutuksellisten ratkaisujen tuloksellisuutta voidaan verrata. Tämä voi olla kansallisesti melko kiusallinenkin mahdollisuus, etenkin silloin kun kansainväliset vertailutulokset osoittavat muuta kuin kansalliset poliittiset ja hallinnolliset ohjelmat tai opettajajärjestöjen julkilausumat. Tulosten vertailussa ja kehittämistoimien suuntaamisessa syntyy joskus jännitteitä ja ristiriitoja. Jos esimerkiksi keskitetyt koulutusjärjestelmät ja tiukasti tulosstandardeihin perustuvat opetusjärjestelyt tuottavat matematiikassa parempia tuloksia kuin hajautetut ja väljemmin määritellyt opetussuunnitelmat, tuloksia on syytäkin pohtia monesta näkökulmasta. Ovatko hyvät tai heikot tulokset todella keskitetyn järjestelmän seurausta vai liittyvätkö tulokset pikemminkin pedagogisiin painotuksiin ja menetelmävalintoihin, kuten opettajankoulutukseen, oppilaiden käsityksiin ja motivaatioon, yksityiseen lisäopetukseen tai koulujen ja vanhempien odotuksiin? Jos taas kansalliset äidinkielen arviointitulokset osoittavat, että neljäsosalla peruskoulun päättävistä on vakavia ongelmia, mutta kansainväliset tulokset viittaavat erinomaiseen yleistason, tulosten ristiriitaisuus voi olla näennäistä. Kansallinen arviointi on ehkä painottanut kirjoittamista ja oman kielen tuntemusta, kun taas kansainvälinen arviointi on kohdistunut lukutaidon eri osa-

alueiden hallintaan. Arvioinnin kohde, kehys, mittavälineet, kriteerien ja tulosasteikkojen tuntemus on aina tarpeen tuloksia tulkittaessa.

Kansallisessa arvioinnissa tehtävät liittyvät yleensä läheisesti opetussuunnitelmien perusteisiin. Kansainvälisissä arvioinneissa ei voida sitoutua mihinkään kansalliseen opetussuunnitelmaan, vaan arviointi on suunnattava opetussuunnitelmien yhteisiin tekijöihin (IEA) tai tulevaisuuden yleisiin avaintaitoihin (PISA). Edellä mainitut ovat monesti kansainvälisesti neuvoteltuja kompromisseja eivätkä tietenkään vastaa minäkään yksittäisen maan opetuksen painotuksia (Owen ym. 2004). Kansallisesti arvioinnit sen sijaan fokusoidaan koulutuspolitiikan ja opetussuunnitelmien painoalueille. Kansallisissa arvioinneissa huomio kohdistetaan usein vaativiin, kehittämisen kohteena oleviin tietoihin ja taitoihin. Perusasioita ja -taitoja ei niinkään arvioida, koska niiden oletetaan jo olevan hallussa. Kansainvälisissä arvioinneissa myös perusosaaminen on usein arvioinnin kohteena, sillä mukana on muitakin kuin pitkälle kehittyneitä tietoyhteiskuntia. Kansainvälisissä arvioinneissa kehys ja tulosasteikko ovat näin ollen huomattavasti laajempi kuin kansallisissa. Laaja arviointikehys antaa samalla mahdollisuuden verrata kansallisesti toteutuneen opetussuunnitelman tuloksia muiden maiden tuloksiin.

Kun kansainvälisten arviointien aineistot on avattu tutkijoiden käyttöön, arviointikehysten, mittavälineiden ja asteikkojen rakentamisen periaatteiden tuntemus on tärkeää. Jokaisen käyttäjän tulisi tutustua myös mittauksen teoreettiseen ja empiiriseen validointiin, muuttujien luokitteluun sekä otantaratkaisuihin ja erilaisiin jälkipainotuksiin. Kansainväliset arviointiaineistot ovat asiantuntevasti käytettyinä arvokkaita lähteitä myös kansallisesti. Vastavuoroisesti kansallisten aineistojen avaaminen tutkijoille monipuolistaisi vertailuanalyysseja ja keventäisi aineistojen jatkuvaa keräämistä kouluissa.

Kansallisten ja kansainvälisten arviointien yhteistyöverkosto

Kansallisten ja kansainvälisten arviointien ei tulisi kilpailla keskenään vaan täydentää toisiaan siten, että päällekkäisyys vältetään ja kansallisesti tärkeimmät kehittämiskohdeet löydetään ja niitä tuetaan. Kansainväliset arvioinnit tarjoavat laajan vertailukontekstin, jota on kuitenkin syytä tarkentaa ja syventää kansallisesti ja kohdistaa arviointia ongelma-alueille, kuten opiskeluasenteisiin ja -strategioihin sekä kriittiseen, pohtivaan ja päättävään oppimiseen kaikkien oppiaineiden alueella. Nuorten vähäinen kiinnostus yhteiskunnalliseen osallistumiseen on myös tärkeä selvityksen kohde. Samoin taide- ja taitoaineiden sekä fyysisen kunnon ja terveystiedon alueita olisi syytä arvioida. Arvioinnin ja kehittämisen tarvetta on myös taidoissa ja asenteissa lähestyä muita kulttuureita ja maahanmuuttajia.

Myös innovatiivisuutta, luovaa ja rohkeaa ajattelua sekä aloitekykyä ja sosiaalisia taitoja olisi paikallaan arvioida, vaikka näiden alueiden käsitteellistäminen ja siten

myös arviointi on vaikeaa. Kansallisesti arvioinnissa voidaan kuitenkin olla innovatiivisia, sillä käsitteellistämisen tueksi voidaan liittää laadullista tutkimusta, eikä kymmenien kansakuntien yksimielisyyttä arviointikohteen käsitteellistämässä tarvitse saavuttaa. Arviointi-innovaatioista olisi puolestaan hyötyä myös kansainvälisesti, kun suunnitellaan uusia yhteisiä hankkeita.

Kansainväliset ja kansalliset arvioinnit tulisi nivoa yhteen niin, että ne täydentävät toisiaan ja muodostavat kokonaisvaltaisen arviointistrategian, jota tutkimus voi tukea. Vahvan arviointistrategian rakentamiseksi tarvitaan viisasta koordinaointia, jossa yhdistyy vahva näkemys sekä yhteiskunnan että koulutuksen kehittämisen suunnista ja joka näkee niin lähelle kuin kauas ja laajalle kansallisten rajojen ulkopuolelle. Kansainvälinen arviointitoiminta voi tukea erinomaisesti kansallista strategiaa, jos kansainvälisessä toiminnassa ollaan aloitteellisia ja valppaita. Kansallisia velvollisuuksia ei myöskään tulisi väistellä, kun toimitaan kansainvälisissä yhteisöissä. Jälkijoukoissa eteneminen voi tuntua turvalliselta, mutta matkan suuntaan voi vaikuttaa eniten, kun on rohkeutta astua koulutuksen arvioitsijoiden ja kehittäjien kansainväliseen kärkijoukkoon.

Kansallinen arviointistrategia vaatii tuekseen monitieteistä arviointitutkijoiden koulutusta, joka rakentuu näkemykselle, että arviointiasiantuntijuus on laaja-alaista, tutkimuksellisesti suuntautunutta ja yhteiskunnallisesti merkittäviä ongelmia ratkaisevaa sekä innovatiivista, eri tieteenaloilta ratkaisuja etsivää toimintaa. Pyrkimyksenä on edistää koulutuksen arvioinnin teoreettisten, eettisten, metodologisten samoin kuin kulttuuristen ja koulutuspoliittisten kysymysten syvällistä ymmärtämistä. Erityisesti tulisi painottaa kansainvälistä yhteistyötä sekä arviointitutkijoiden liikkuvuutta ja verkottumista. (Linnakylä 2001; Välijärvi 2001.)

Arviointi, tutkimus ja koulutuspolitiikka nivoutuvat väistämättä yhteen, etenkin silloin kun puhutaan kansallisesta tai kansainvälisestä arvioinnista. Arvioinnin juuret ovat yhteisesti sovitussa arvostuksessa ja niiden ohjaamisessa päämäärissä eli politiikassa; perusta taas on systemaattisessa, luotettavassa ja pätevässä tiedon keruussa ja analysoinnissa eli tutkimuksessa. Arviointitulosten merkitys on viisaissa ja käyttökelpoisissa johtopäätöksissä sekä arviointitiedon käytössä koulutuksen ja yhteiskunnan kehittämisen välineenä. Tämä onnistuu parhaiten kansallisen ja kansainvälisen arviointitutkimuksen yhteistyöllä.

Lähteet

- Bosker, R. & Snijder, T. 2004. Moving beyond bivariate indicators: the role of multi-level modeling in international comparisons. Teoksessa J. Moskowitz & M. Stephens (toim.) Comparing learning outcomes. International assessment and education policy. London: Routledge Falmer, 178–189.
- Chromy, J. 2002. Sampling issues in design, conduct, and interpretation of international comparative studies of school achievement. Teoksessa A. Porter & A. Gamoran (toim.) Methodological advances in cross-national surveys of educational achievement. National Research Council Wash-



- ington, DC: National Academy Press, 80–114.
- Döbert, H., Klieme, E. & Stroka, W. (toim.) Conditions of school performance in seven countries. A quest for understanding the international variation of PISA results. Munster: Waxmann.
- House, E. R. 1997. Ethics of evaluation studies. Teoksessa J.P. Keeves (toim.) Educational research methodology, and measurement: an international handbook. 2nd edition. Cambridge: Cambridge University Press, 257–261.
- Hvistendahl, R. & Roe, A. 2004. The literacy achievement of Norwegian minority students. *Scandinavian Journal of Educational Research* 48 (3), 307–324.
- Klieme, E. 2004. Assessment of cross-curricular problem-solving competencies. Teoksessa J. Moskowitz & M. Stephens (toim.) Comparing learning outcomes. International assessment and education policy. London: Routledge Falmer, 81–107.
- Koulutuksen arvioinnin uusi suunta. Arviointiohjelma 2004–2007. Koulutuksen arviointineuvoston julkaisuja I. Jyväskylä.
- Leimu, K. 2001. The way to a strategic view on evaluation. Teoksessa K. Leimu, P. Linnakylä & J. Välijärvi (toim.) Merging national and international interests in educational system evaluation. University of Jyväskylä: Institute for Educational Research, 7–14.
- Leimu, K. (toim.) 2004. Kansainväliset IEA-tutkimukset Suomi-kuvaa luomassa. Jyväskylän yliopisto: Koulutuksen tutkimuslaitos.
- LeTendre, G. 2002. Advancements in conceptualizing and analyzing cultural effects in cross-national studies of educational achievement. Teoksessa A. Porter & A. Gamoran (toim.) Methodological advances in cross-national surveys of educational achievement. National Research Council Washington, DC: National Academy Press, 198–228.
- Lie, S., Linnakylä, P. & Roe, A. (toim.) 2003. Northern lights on PISA. Unity and diversity in the Nordic countries in PISA 2000. University of Oslo: Department of Teacher Education and School Development.
- Lindqvist, T. 1999. Evaluatation uskottavuus. Teoksessa R. Eräsaari, T. Lindqvist, M. Mäntysaari & M. Rajavaara (toim.) Arviointi ja asiantuntijuus. Helsinki: Gaudeamus, 106–118.
- Linnakylä, P. 2001. The national intertwined with the international. Teoksessa K. Leimu, P. Linnakylä & J. Välijärvi (toim.) Merging national and international interests in educational system evaluation. University of Jyväskylä: Institute for Educational Research, 63–76.
- Linnakylä, P. 2002. Kansainvälisten ja kansallisten oppimistulosten arviointien välisestä suhteesta. Teoksessa E. Olkinuora, R. Jaku-Sihvonen & E. Mattila (toim.) Koulutuksen arviointi: lähtökohtia, malleja ja tilannekatsauksia. Turun yliopiston Kasvatustieteiden tiedekunnan julkaisuja B, 31–52.
- Linnakylä, P., Malin, A. & Taube, K. 2004. Factors behind low reading literacy achievement. *Scandinavian Journal of Educational Research* 48 (3), 231–250.
- Lundgren, U. P. 2001. Strategic arenas of influence in pursuing quality of education: Some conceptual and general issues. Teoksessa K. Leimu, P. Linnakylä & J. Välijärvi (toim.) Merging national and international interests in educational system evaluation. University of Jyväskylä: Institute for Educational Research, 15–22.
- OECD 2000. Literacy in the information age. Final report of the International Adult Literacy Survey. Paris: OECD & Statistics Canada.
- OECD 2001. Knowledge and skills for life. First results from the OECD Programme for International Student Assessment (PISA) 2000. Paris: OECD.
- OECD & UNESCO Institute for Statistics 2003. Literacy skills for the world of tomorrow. Further results from PISA 2000. Paris.
- Owen, E. 2001. Educational indicators. Teoksessa K. Leimu, P. Linnakylä & J. Välijärvi (toim.) Merging national and international interests in educational system evaluation. University of Jyväskylä: Institute for Educational Research, 41–50.

- Owen, E., Stephens, M., Moskowitz, J. & Gil, G. 2004. Toward education improvement: the future of international assessment. Teoksessa J. Moskowitz & M. Stephens (toim.) *Comparing learning outcomes. International assessment and education policy*. London: Routledge Falmer, 3–23.
- Peschar, J. 2004. Cross-curricular competencies: developments in a new area of education outcome indicators. Teoksessa J. Moskowitz & M. Stephens (toim.) *Comparing learning outcomes. International assessment and education policy*. London: Routledge Falmer, 59–80.
- Plomp, T. 2001. The potential and challenges of international comparative studies of educational achievement. Teoksessa K. Leimu, P. Linnakylä & J. Välijärvi (toim.) *Merging national and international interests in educational system evaluation*. University of Jyväskylä: Institute for Educational Research, 23–40.
- Raudenbush, S. & Kim, J.-S. 2002. Statistical issues in analysis of international comparisons of educational achievement. Teoksessa A. Porter & A. Gamoran (toim.) *Methodological advances in cross-national surveys of educational achievement*. National Research Council Washington, DC: National Academy Press, 267–294.
- Stigler, J. & Hiebert, J. 1999. *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.
- Vuorenmaa, M. 2001. Ikkunoita arvioinnin tuolle puolen. Uusia avauksia suomalaiseen koulutusta koskevaan evaluaatiokeskusteluun. Jyväskylän yliopisto. *Jyväskylä studies in education, psychology and social research*.
- Välijärvi, J. 2001. Research in the context of national assessment. Teoksessa K. Leimu, P. Linnakylä & J. Välijärvi (toim.) *Merging national and international interests in educational system evaluation*. University of Jyväskylä: Institute for Educational Research, 113–122.
- Välijärvi, J. & Linnakylä, P. (toim.) 2002. *Tulevaisuuden osaajat. PISA 2000 Suomessa*. Jyväskylän yliopisto: Koulutuksen tutkimuslaitos.
- Välijärvi, J. & Malin, A. 2003. The two-level effect of socio-economic background. Teoksessa S. Lie, P. Linnakylä & A. Roe (toim.) 2003. *Northern lights on PISA. Unity and diversity in the Nordic countries in PISA 2000*. University of Oslo. Department of Teacher Education and School Development, 123–132.



Arviointitoiminta verkostoihin

Koulutuksen arviointineuvoston perustamisen myötä koulutuksen arviointi siirrettiin hallinnosta riippumattomien asiantuntijoiden verkostolle. Uuteen arviointijärjestelmään siirtymisellä haluttiin varmistaa riippumattomuuden ohella arvioinnin osuvuus, tehokkuus, taloudellisuus sekä vaikuttavuus. Arvioinnin verkostomalli mahdollistaa syvällisesti erikoistuneiden asiantuntijoiden käyttämisen kunkin arviointikohteen tarpeiden mukaisesti. Verkostomaisen toimintatavan avulla edistetään tieteen, opetushallinnon, käytännön opetustyön ja sidosryhmien sekä muiden tarpeellisten tahojen asiantuntijoiden yhteistyötä. Verkostomaisuus on kehittävän arvioinnin paradigmassa olennainen työkalu ja edellytys tärkeimmän päämäärän eli koulutuksen jatkuvan kehittämisen saavuttamiseksi.

Asiasanat: arviointi, koulutuksen arviointi, verkostot

Verkostotoiminnan lähtökohdat

Verkostona organisoitava koulutuksen arviointi perustuu lakiin. Koulutuksen arvioinnista on säädetty samansisältöisesti 1.1.1999 voimaan tulleissa perusopetuslain (628/1998) 21 §:ssä, lukiolain (629/1998) 16 §:ssä, ammatillisesta koulutuksesta annetun lain (630/1998) 24 §:ssä, ammatillisesta aikuiskoulutuksesta annetun lain (631/1998) 15 §:ssä sekä vapaasta sivistystyöstä annetun lain (632/1998) 7 §:ssä. Säännösten mukaan arvioinnin tarkoituksena on turvata koulutusta koskevan lainsäädännön tarkoituksen toteuttaminen sekä tukea koulutuksen kehittämistä ja parantaa oppimisen edellytyksiä. Lainsäädännössä arvioinnin tavasta säädetään seuraavasti: “Ulkopuolista arviointia varten opetusministeriön yhteydessä on erillinen koulutuksen arviointi-

neuvosto, joka organisoii arviointitoiminnan yliopistojen, opetushallituksen ja muiden arviointiasiantuntijoiden verkostona.”

Koulutuksen arvioinnista säädetyssä asetuksessa (SA 150/2003) arviointineuvoston yhtenä tehtävänä on “organisoida koulutuspolitiikkaan ja koulutuksen järjestäjien toimintaan liittyviä ulkopuolisia arviointeja arviointiasiantuntijoiden verkostona”. Koulutuksen arviointineuvoston arviointitoiminnan periaatteita ja vuosien 2004–2007 arviointiohjelmaa linjaava *Koulutuksen arvioinnin uusi suunta* toteaa: “Verkostomaisen toimintatavan avulla edistetään tieteen, opetushallinnon, käytännön opetustyön ja sidosryhmien sekä muiden tarpeellisten tahojen asiantuntijoiden yhteistyötä. Toiminta perustuu vastavuoroiseen kumppanuuteen, jossa osapuolet sitoutuvat toimimaan yhteisten periaatteiden mukaisesti. Arviointiverkostot muodostuvat sektorikohtaisesti parhaista kansallisista ja kansainvälisistä asiantuntijoista.”

Alunperin markkinataloudesta ja yritysmaailmasta tutun verkostoitumisen keskeiset vaikuttimet ovat tiedon tehokas käyttö, nopeus ja joustavuus (esim. Ollus & al. 1998). Samalla tavoin koulutuksen arvioinnin verkostomaisuuteen velvoitetaan siksi, että saadaan käyttöön

- kulloisenkin arviointikohteen ja tehtäväalueen ajantasaisin ja syvällisin erityisasiantuntemus,
- eri alojen ja kohteiden erityisasiantuntijoiden yhteen kokoamisen mukanaan tuomat synergiaedut,
- nopeasti laadukasta tietoa ja tuloksia,
- monipuolisesti kentän, sidosryhmien ja eri intressitahojen näkemykset tulkittoineen.

Arvioinnin verkostomalli mahdollistaa syvällisesti erikoistuneiden asiantuntijoiden käyttämisen kunkin arviointikohteen tarpeiden mukaisesti. Arvioinnin organisointitapojen jatkumolla ajatellen verkostomallin eräänlaisena vastinparina toisessa ääripäässä voidaan nähdä erillinen “arviointivirasto” kymmenine vakinaisine virkamiehinne suorittamassa arviointeja ja laatimassa niistä raportteja. (Vrt. Visakorpi 2002.) Tästä näkökulmasta tarkasteltuna arvioinnin uuden organisointitavan ero aiempaan on hyvin selkeä. Aiemmin Opetushallituksessa työskenteli yli 30 henkilöä arviointiasioissa ja arviointiraportit laadittiin virkatyönä samalla kun organisaatio oli keskeinen toimija koulutuksen kehittämistyössä. Koulutuksen arviointineuvostossa arviointitoimintaa organisoii pääsihteerin ja viiden pääsuunnittelijan muodostama arviointisihteeristö. Sen enempää arviointineuvosto kuin sen sihteeristöäkään ei tee arviointeja. Arviointit laaditaan ulkopuolisten riippumattomien asiantuntijoiden toimesta. Arviointineuvosto organisoii arviointitoiminnan ja luo mahdollisimman hyvät edellytykset, jotta arviointit tehdään laadukkaasti, tehokkaasti ja nopeasti (ks. Lyytinen & Hämäläinen 2004).

Verkoston koostumuksesta

Laissa arviointineuvoston toimintaa koskeva keskeinen sanoma kiteytyy “ulkopuolisen arvioinnin” ja “organisoinnin” ohella ilmaisussa “arviointiasiantuntijoiden verkosto”. Esimerkiksi paikallisesta arviointitoiminnasta ja sen kehittämisestä vastaavat ensisijaisesti koulutuksen järjestäjät, kuten kuntalaki edellyttää. “Organisoinnilla” laissa viitataan juuri siihen, että Koulutuksen arviointineuvosto itse ei arvioi, vaan organisoii arvioinnit toteutettavaksi verkoston avulla. Laissa verkoston toimijoista mainitaan erikseen yliopistot tieteellisen tutkimuksen tärkeimpänä instituutiona sekä Koulutuksen arviointineuvostoa edeltävän ajan keskeinen arvioinnin toteuttaja, Opetushallitus.

Koulutuksen arvioinnista säädetyssä asetuksessa (SA 150/2003) todetaan, että “Opetushallitus osallistuu osana koulutuksen arviointiasiantuntijoiden verkostoa opetussuunnitelmien ja tutkintojen perusteiden mukaisten oppimistulosten arviointiin”. Asetuksen mukaan Opetushallituksella on itseoikeutettu asema osallistua oppimistulosten arviointiin. Toisin sanoen Opetushallitus on mukana kaikissa oppimistulosten arviointihankkeissa yhtenä arvioijatahona verkoston muiden tahojen joukossa, mutta myös muut verkoston toimijat, kuten yliopistot, osallistuvat oppimistulosten arviointiin ydinosaamisensa pohjalta. Sen sijaan kaikissa muissa kuin oppimistulosten arvioinneissa kutakin arviointihanketta varten koottaviin arviointiryhmiin ei ole ennakoon säädetty varauksia, vaan arvioinnit toteutetaan siten, että verkostosta pyritään saamaan kohteen kannalta paras mahdollinen asiantuntemus. Arvioijina toimivat asiantuntijat voivat olla niin ulko- kuin kotimaisia.

Yksi yleisempi varaus koskee kuitenkin arviointitehtäviin valittavia tahoja: arviointitavan kohteen kannalta kehittämisestä vastaavat tahot eivät voi olla mukana toteuttamassa arviointia, saati arvioimassa. Arviointia koskevan lainsäädännön hengen mukaisesti riippumattomuus koskettaa Opetushallituksen mukana oloa arvioinnissa. Epäilyt uskottavuudesta ja riippumattomuudesta olivat nimittäin keskeinen syy koulutuksen arvioinnin siirtämiseen riippumattomalle neuvostolle pois hallinnon sisältä. Koulutuksen arviointineuvoston perustamiseen johtaneen lainsäädännön valmistelun yhteydessä eduskunnalle tehdyssä selvityksessä todettiin, että “arviointitoiminnan tulisi [—] olla itsenäistä kehittämistyötä, jotta samat ihmiset, jotka ovat valmistelleet tavoitteet ja sisällöt, eivät sitten myöhemmin arvioisi omaa tuotostaan. Tämän vuoksi eduskunnalle annettussa selvityksessä esitetäänkin erityisen opetus- ja koulutoimen arviointineuvoston perustamisesta. Neuvoston asettaisi opetusministeriö ja sen tehtävänä olisi edistää ja organisoida sen toimialaan kuuluvan koulutuksen arviointia ja arviointihankkeita, mutta varsinaisen arviointityön suorittaisivat erikseen asetettavat asiantuntijaryhmät.” (Selvitys... s. 30.)

Koulutuksen ja esimerkiksi opetussuunnitelmien ja tutkinnon perusteiden kehittäjänä Opetushallitus on periaatteessa arvioinnin kohde siinä kuin muutkin arviointikohteet koulutusjärjestelmässä ja sen kehittämisessä. Laissakin mainittujen yliopis-

tojen ohella arviointiasiantuntijoiden verkostoon kuuluvat luontevasti myös monet erilliset tutkimuslaitokset ja ammattikorkeakoulut sekä esimerkiksi lääninhallitukset, jotka huolehtivat kuntien peruspalveluiden arvioinnista alueellaan. Niin ikään monet koulutukseen kytkeytyvät sidosryhmät ovat luonteva elementti arvioinnin verkostoa. Esimerkiksi työelämän osapuolilla on keskuudessaan runsaasti sellaista asiantunte-
musta ja hiljaista tietoa sekä äänetöntä osaamista, jota ei esimerkiksi yliopistotutki-
muksen piirissä pääse juuri kehittymään. Sen saaminen arvioinnin käyttöön on ni-
menomaan osallistavan verkostotoiminnan yksi keskeinen funktio.

Uusi arviointijärjestelmä rakentuu siis jo olemassa olevan infrastruktuurin varaan. Koulutuksen arviointineuvoston arviointitoiminnan periaatteita ja vuosien 2004–2007 arviointiohjelmaa linjaava *Koulutuksen arvioinnin uusi suunta* paikantaa verkoston toiminnan kahtaalle: yhtäältä itse arviointihankkeisiin ja toisaalta arviointijärjestel-
män kehittämiseen.

Arviointihankkeiden lisäksi verkostoyhteistyössä kehitetään siis myös arviointijär-
jestelmää. Keskeisiä yhteistyöalueita ovat tietojärjestelmien, indikaattoreiden ja arvi-
ointimenetelmien ja -tutkimuksen kehittäminen, koulutuksen järjestäjien ja oppilai-
tosten henkilöstön kouluttaminen ja koulutuksen järjestäjille suunnattujen arviointi-
palveluiden kehittäminen. Arviointijärjestelmän kehittämisessä taas muun muassa
Opetushallitus on tärkeä yhteistyötaho esimerkiksi Tilastokeskuksen ja Stakesin
asiantuntijoiden ohella.

Toimijoiden identifiointista verkoston rakentamiseen

Kentän toimijat ovat verkoston solmukohtia. Toimijat voivat olla henkilöitä sekä
muodollisia tai epämuodollisia ryhmiä. Toimijoista arviointineuvoston tehtävänä on
aktiivisesti luoda ja ylläpitää verkostoja. Verkostojen luominen, ylläpitäminen ja säi-
lyminen perustuvat toimijoiden erikoistumiseen ja monenlaisen erikoisosaamisen
välttämättömyyteen kokonaisuuden kannalta. Verkosto on luonnehdittavissa saman-
aikaisesti sekä heterogeeniseksi että homogeeniseksi, minkä piirteen varaan sen vah-
vuus olennaisesti perustuu.

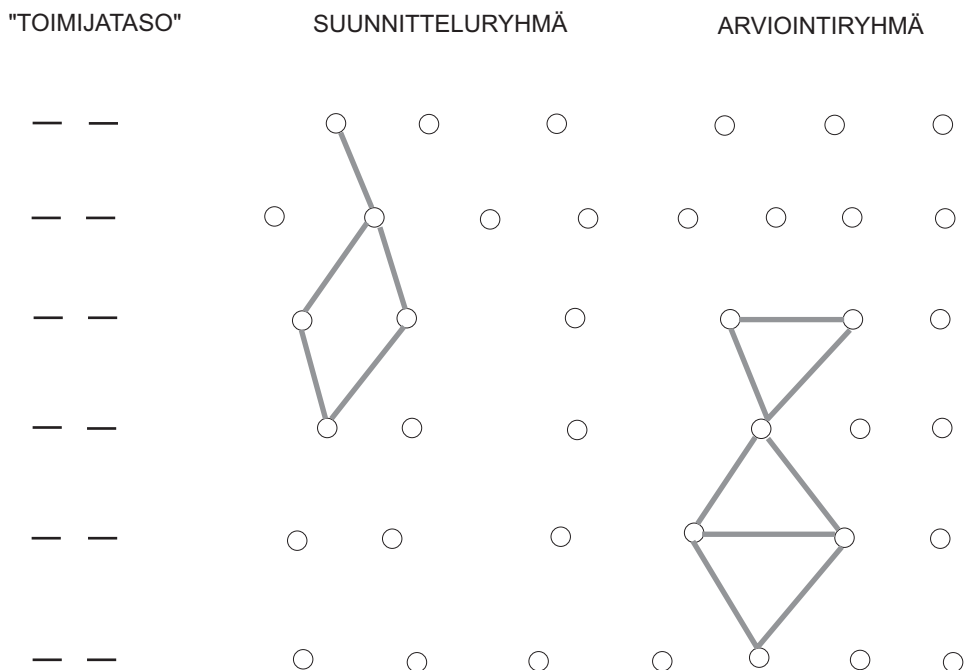
Verkoston rakentaminen alkaa arvioitavan kohteen kannalta relevanttien toimijoi-
den identifiomisella. Toimijat ovat organisaatioita ja ihmisiä, jotka voidaan konkreet-
tisimmillaan identifioida niminä. Taulukossa 1 havainnollistetaan kentän toimijoiden
identifiointia käyttämällä esimerkkinä vapaan sivistystyön kenttää.

Verkostomaisessa arviointimallissa on tärkeää, että kentän toimijat osallistuvat toi-
mintaan alusta lähtien. Arviointiryhmän ohella myös arviointisuunnitelman tekevä
suunnitteluryhmä rekrytoidaan verkostosta, joskin usein eri puolelta verkostoa kuin
arviointiryhmä, joka koostuu pääosin arviointi- ja metodiosaamisen sekä tieteellisen
tutkimuksen asiantuntijoista. Toki arviointiryhmässäkin voi olla “kentän”, esimerkiksi
työelämän osapuolten piiristä tulevia, asiantuntijoita.

Taulukko 1. Kentän toimijoiden identifiointi: esimerkkinä vapaan sivistystyön arviointi

<p>Opetusministeriö, lääninhallitukset</p> <p>Oppilaitosten ylläpitäjät, koulutuksen järjestäjät</p> <ul style="list-style-type: none">• kunnat, kuntayhtymät• järjestöt• muut koulutuksen järjestäjät. <p>Oppilaitokset</p> <ul style="list-style-type: none">• kansanopistot• kansalais- ja työväenopistot• opintokeskukset• kesäyliopistot• liikunnan koulutuskeskukset. <p>Koulutusorganisaatioiden ja niiden ylläpitäjien edunvalvontajärjestöt</p> <ul style="list-style-type: none">• KTOL, kansalais- ja työväenopistojen liitto• Suomen kansanopistoyhdistys• VSY, vapaan sivistystyön yhteisjärjestö• Suomen kesäyliopistot ry.• Kuntaliitto. <p>Opettajat</p> <ul style="list-style-type: none">• OAJ• Aikuiskouluttajien liitto• jne. <p>Koulutuksen käyttäjien edunvalvontajärjestöt</p> <ul style="list-style-type: none">• opiskelijajärjestöt• työnantajajärjestöt (EK, yrittäjien keskusjärjestöt jne.)• työntekijäjärjestöt (ammattiliitot, keskusjärjestöt jne.)• kansalaisjärjestöt. <p>Arviointia ja tutkimusta tekevät tahot</p> <ul style="list-style-type: none">• yliopistojen laitokset (HY, TY, TaY, JY, JoY, LY jne.)• koulutuksen arviointiin suuntautuneet organisaatiot (OPH, Efektia, Laatuokeskus jne.)• erilliset tutkimuslaitokset, joilla on asiantuntemusta koulutustutkimuksesta ja arvioinnista (VATT, PT, ETLA, Stakes jne.).
--

Kentän käytännön toimijoiden (ks. kuvio 1) käyttäminen jo arvioinnin suunnittelussa on omiaan osallistamaan koko kenttää, mikä puolestaan edesauttaa arviointitulos-
tosten hyödyntämistä ja vaikuttavuutta. Kun eri tahot tuovat yhteiseen pöytään oman
tulkintansa arvioitavasta kohteesta ja osallistuvat arvioinnin muotoiluun, tulosten hyö-
dynnettävyys ja sitä kautta käytön todennäköisyys lisääntyy (esim. Patton 1997; myös
Lyytinen ym. 2002). Toisaalta riippumatonta tutkimusta ja arviointiasiantuntemusta
edustava arvioijaryhmä on sillä tavoin koulutuksen kenttään nähden “ulkopuolises-



Kuvio 1. Arvioinnin suunnittelu ja toteutus verkoston avulla

sa” asemassa, että sen arviointiin, johtopäätöksiin ja mahdollisiin suosituksiin luoteetaan. Arviointiryhmä koostuu itsenäisistä asiantuntijoista, jotka omalla nimellään vastaavat arvioinnin tasosta arviointiraportissa.

Verkostomaisuus kytkeytyy elimellisesti *kehittävän arvioinnin* paradigmaan. Nykyisin jo vanhahtavaksikin luonnehdittua tarkastustoimintaa harjoittamaan saattaisi riittää hyvin esimerkiksi itse kentästä irrallinen virasto. Sen sijaan arvioinnin suunnittelu ja suorittaminen verkostona on kehittävän arvioinnin näkökulmasta mitä luontevin toimintamalli. Kehittävän arvioinnin paradigmassa verkostomaisuus on siis hyvin olennainen työkalu ja edellytys kaikkein *tärkeimmän päämäärän eli koulutuksen kehittämisen* saavuttamiseksi.

Arviointiverkoston toimintakyky

Verkoston rakentamisen ja toimintakyvyn arvioinnin keskeinen lähtökohta on kysymys: miksi jokin henkilö tai taho osallistuu verkostoon? Arviointihankkeen suunnitteluryhmän kohdalla osallistuminen ja toimintaan sitoutuminen eivät ensi alkuun näytä ongelmalliselta. Yleensä asianosaiset haluavat olla vaikuttamassa siihen, millainen kuva heidän omasta toiminnastaan luodaan ja miten he voivat olla itse kehittämässä

tuota kuvaa ja sen perustana olevaa todellisuutta. Ainakin he haluavat pitää huolen siitä, että arviointi on heidän oman toimintansa kannalta relevanttia. Niin kutsutussa kehittävässä arvioinnissa on tärkeää, että arvioinnin kohteet voivat käyttää arvioinnin tuloksia oman toimintansa kehittämiseen. Asianosaisten sitoutumista toimintaan edesauttaa se, että osallistumalla aktiivisesti he samalla huolehtivat siitä, että arvioinnissa otetaan huomioon heidän omasta näkökulmastaan relevantit tekijät.

Miten arviointiverkosto onnistutaan pitämään toimintakykyisenä? Tämä riippuu monista vaikeasti hallittavista tekijöistä, mutta suunta, jota kohti tulee pyrkiä on toki tiedossa. Toimintakykyisen verkoston keskeisiä ominaisuuksia on eritelty seuraavasti (Nikkanen & Lyytinen 1996, 59–60):

Täydentävyys: verkoston jäsenet voivat täydentää toisiltaan aineellisia ja henkisiä resurssiaan. Yksi toimija on vahva yhdellä, toinen toisella osaamisalueella.

Yhteensopivuus: Verkoston aluetoimijoiden, eri organisaatioyksiköiden jne. tulevaisuudenkuvat, päämäärät, toiminta-ajatukset, toimintakulttuuri ja toimintastrategiat edistävät yhdessä asetettujen tavoitteiden saavuttamista.

Yhteistyöhalukkuus: Hyvät olosuhteet ja monipuoliset valmiudet eivät riitä vaan tuloksellisuus edellyttää motivoitumista aitoon yhteistyöhön.

Sitoutuminen: Sitoutuminen ymmärretään ihmisen vapaaehtoiseksi kiinnostukseksi ja innostukseksi panna itsensä likoon jonkin asian puolesta ilman ulkopuolista houkutinta tai pakotetta. Taustalla voi olla henkilökohtainen intressi, inhimillinen uteliaisuus, vastuun- tai velvollisuudentunne, korkea suoritusmotivaatio, sosiaalinen päteminen, henkilökohtaisesti koetut tärkeät arvot jne. Verkosto-toimintaan sitoutuminen tapahtuu vaikuttamismahdollisuuksien sekä taustojen ja toiminnan vaihtoehtojen tuntemuksen kautta syntyvän ymmärryksen avulla.

Luottamus: Verkostoituminen edellyttää organisaatioilta sitä, että ne ovat toimintakulttuuriltaan avoimia eivätkä suljettuja. Luottamus ja avoimuus syntyvät päämäärien ja strategioiden yhteisen hyväksynnän kautta. Tiedonkulun avoimuus ja runsaus sekä informaatiokanavien avoimuus kaikille, arvoon ja asemaan katso-matta, vahvistaa luottamusta. Luottamus on verkostotoimintaa tehostavaa sosi-aalista pääomaa (esim. Ilmonen & Jokinen 2002; Ruuskanen 2003; Blomqvist 2004).

Miten tieteen ja tutkimuksen piiristä saadaan irrotettua asiantuntemusta arviointitoimintaan? Arviointihankkeiden keskeisiksi toimijoiksi oletettujen yliopistojen laitosten voi olla nopealla aikataululla hyvinkin vaikeata irrottaa henkilökuntaansa arviointihankkeisiin vähäisellä korvauksella.¹ Verkostomallin toimivuus edellyttää, että arviointihankkeet eivät jää asiantuntijaorganisaatioiden kannalta yksittäisiksi pyrhdyk-

¹ Tämä ongelma tuli usein esille arviointineuvoston ja -sihteeristön yliopistoihin ja tutkimuslaitoksiin vuonna 2004 tekemien verkostovierailujen yhteydessä.

siksi. Lakiin perustuva arviointitoiminta on kuitenkin jatkuvaa ja pitkäjänteistä toimintaa, jonka sujuva toteutuminen edellyttää kaiken aikaa aktivoitunutta verkostoa sekä tulevien hankkeiden ennakoitavuutta. Ennakoitavuutta parantaa arviointitoiminnan vakaa ja riittävä rahoitus, joka mahdollistaa arviointineuvoston useiksi vuosiksi eteenpäin laatiman, pitkäjänteiseen arviointiohjelman systemaattinen toteuttamisen.

Pidemmällä aikavälillä verkostomaisuuden toimivuuden kannalta on olennaista, muodostuuko Koulutuksen arviointineuvoston organisoimasta toiminnasta sellaista, johon osallistuminen koetaan muutenkin kuin taloudellisesti kannattavaksi. Arviointineuvoston johtaman toiminnan vetovoima ei riipu vain rahasta, vaan siinä on kysymys myös statuksesta: mikäli arviointitoiminta on korkealaatuista ja nauttii kentän sekä tiedeyhteisön laajaa luottamusta, esimerkiksi yliopiston laitoksen *pääsy* mukaan arviointihankkeisiin tuo niihin osallistuville henkilöille tai organisaatioille arvostusta ja mainetta alan asiantuntijana. Tunnustettu ja usein hyödynnetty arviointiosaaminen on meriitti tutkimuslaitokselle.

Verkostomallin mukaan organisoidun arviointijärjestelmän kannalta on strategisesti tärkeää, että sen piirissä tehtävät arvioinnit ovat paitsi verraten mittavia myös aiheiltaan koulutuspoliittisesti merkittäviä. Niukasti resurssoidut pienimuotoiset selvitykset eivät yksin riitä pitämään yllä korkeaa arvostusta nauttivaa arviointitoimintaa. Vaarana on, että verkoston toimintakyky rapautuu motivaation ja mielekkyyden puutteeseen. Alansa huippuasiantuntijoilla riittää kysyntää haasteellisiin ja koulutus- ja yhteiskuntapoliittisesti merkittäviinkin tehtäviin.

Uuden arviointijärjestelmän toimintaperiaatteet verkostonäkökulmasta

Lainvalmistelun yhteydessä eduskunnalle annetussa koulutuksen arviointia koskevassa selvityksessä arviointijärjestelmän toimivuutta ja kehittämistä koskevat näkökulmat olivat keskeisessä asemassa. Selvityksen mukaan arviointijärjestelmän toimivuuden kannalta kriittisiä kohteita olivat arvioinnin riippumattomuus, osuvuus, tehokkuus ja taloudellisuus sekä vaikuttavuus. Tuolloin kriittisinä pidetyt seikat ovat arvioinnin uuden organisoinnin keskeinen ohjenuora.

Koulutuksen arvioinnin uudeksi suunnaksi laadittujen toimintaperiaatteiden mukaan arviointi on (s. 16–18)

- riippumatonta,
- osallistavaa,
- kehittävää,
- tehokasta ja taloudellista sekä
- tähtää laadukkuuteen ja
- perustuu avoimuuteen.

Näiden toimintaperiaatteiden toteutuminen on kiinteässä yhteydessä nimenomaan arvioinnin verkostomaisuuteen.

Arviointi on ensisijaisesti koulutuksen ja opetuksen kehittämisen väline, jonka päämäärä on oppimisen edistäminen. Arvioinnin vaikuttavuus edellyttää arvioitsijan ja arviointien kohteiden välistä luottamusta, uskottavan arviointikulttuurin luomista. Luottamusta ja uskottavuutta on vaikea rakentaa ilman riippumattomuutta. Uudenlainen verkostomainen toimintatapa on keskeinen rakenteellinen edellytys sille, että koulutuksen arvioija on riippumaton opetushallinnosta (koulutuksen päätöksenteosta) ja koulutuksen kehittämiseen osallistuvista tahoista. Hallinnon ulkopuolisena tahona verkosto on myös riippumaton hallinnon sosiaalisesta järjestelmästä.

Toinen toimintaperiaate ”osallistavuus” nojaa siihen, että eri arvioijatahot ja toimijat suunnittelevat ja suorittavat arvioinnit asiantuntijaverkostona. Osallistava arviointi on kokemusten ja näkemysten vaihtoa ruokkiva toimintatapa. Osapuolten näkemysten ja tulkintojen esiin saaminen on välttämätön lähtökohta arviointiyhteistyön kehittämiseksi. Osallistuva verkosto myös pitää huolen siitä, että arviointi on osuvaa ja uudistuu kaiken aikaa. Päämääränä on, että koulutuksen järjestäjät ja oppilaitokset osallistuvat aktiivisesti arviointiin ja sitoutuvat avoimeen vuorovaikutukseen.

Verkostomainen ja osallistava toimintatapa ovat yhdessä hyvä lähtökohta arvioinnin kehittävyydelle. Osapuolten osallistuminen arviointihankkeisiin ja asiantuntijaverkoston rakentamiseen suunnitteluvaiheesta alkaen luo edellytyksiä sille, että arviointien tuloksia myös käytetään koulutuksen kehittämiseen. Näin arviointi antaa oppilaitoksille, koulutuksen järjestäjille, koulutuspoliittisille päätöksentekijöille sekä muille koulutuksen sidosryhmille relevanttia tietoa prosesseista, jotka edistävät yksilöllistä ja organisaatioiden oppimista. Ylipäätyäkään verkostomaisella arviointitoiminnalla edistetään yhteistä dialogia ja arvokeskustelua koulutusta ja kasvatusta järjestävissä työyhteisöissä ja koko yhteiskunnassa.

Arvioinnin organisoinnissa painotetaan ensisijaisesti arviointiprosessin keveyttä ja joustavuutta. Arviointiprosessien tehokkuutta ja taloudellisuutta edistetään kiinnittämällä huomiota arviointihankkeiden sisällölliseen ja ajalliseen rytmitykseen sekä niiden määrään, tiedonhankinta- ja analysointimenetelmiin, tulosten raportointiin, verkoston käyttöön, vaihtoehtoihin organisoitintapoihin ja olemassa olevan tiedon hyödyntämiseen.

Aktiivinen verkosto on periaatteessa tehokkain malli parhaiden asiantuntijoiden käyttöön. Arviointien käytännön toteutuksen laatu varmistetaan valitsemalla asiantuntijaverkostosta kulloisiinkin hankkeisiin osaavimmat toimijat.

Avoimuus on verkostossa kehittyvän luottamuksen keskeinen edellytys. Avoimuus edesauttaa verkostomuotoisen toimintatavan luomaa keskustelevaa ja vuorovaikutteista arviointikulttuuria. Sen lisäksi, että arviointien tulokset tuodaan julkiseen keskusteluun, verkostomaisuus edistää osapuolten aktiivista osallistumista koulutusta ja sen arviointia koskevaan yhteiskunnalliseen keskusteluun yleisemminkin.

Verkostomaisen toiminnan rajoituksista

Arviointihankkeiden suunnittelun monitahoisuudella tavoitellaan kentän sitoutumista ja arviointitulosten hyödynnettävyyttä, mutta tällaiseen suunnittelumalliin sisältyy myös rajoituksia ja riskejä. Mikäli monitahoisessa suunnittelussa mukana olevat näkevät oman tehtävänsä nimenomaan taustaryhmänsä edunvalvontana, mielekkään kokonaisuuden saavuttaminen vaikeutuu. Voi käydä niin, että kaivautuessaan omiin poteroihinsa eri intressiryhmiä edustavat henkilöt antavat oman asiantuntemuksensa suunnittelun käyttöön vain niiltä osin, joka hyödyttää heidän erityisintressejään. Näin he eivät toimi yhteisen päämäärän hyväksi vaan ainoastaan valvovat, ettei heidän etunäkökohtiaan jätetä hankkeessa huomiotta. Verkostoon toki aina osallistutaan omista etunäkökohdista lähtien, mutta oman edun tavoittelun pitäisi lopulta johtaa johonkin yhteiseen etuun. Jotta verkostoa voisi luonnehtia osaamispääoman kasaumaksi, sen pitäisi palvella päämäärää, joka on sovittu tavoitteeksi yhteisesti ja joka ylittää yksittäiset etunäkökohdat (Hailikari ym. 2000, 15). Muutoin verkosto jää vain kilpaken-täksi, jossa osapuolet pyrkivät maksimaalisesti hyötymään toisistaan antamatta kuitenkaan omaa osaamistaan yhteiseen käyttöön.

Paitsi ryhmäkohtaisia intressejä arviointi sivuaa läheisesti tietysti myös vallankäyt-töä. Vaikkakin arvioinnin päämääränä pidetään nykyisin toiminnan kehittämistä, arvioinnin kohteena oleva ei välttämättä halua ulkopuolisia arvioijia esittämään arviota toimintansa laadukkuudesta ja tasosta, prosessien tarkoituksenmukaisuudesta ja tu-loksellisuudesta tai tuloksettomuudesta. Kun pelätään, mitä arviointi ”paljastaa”, ei mielellään haluta edistää sen enempää verkostomaista kuin muutakaan ulkopuolista arviointia.

Jo pelkän uutuutensa johdosta verkostomaisella toimintatavalla on monia esteitä ylitettävänä. Kysymys on tottumattomuudesta uuteen toimintatapaan. Verkostokulttuuri on kehittymätöntä. Arviointitoiminnasta saatavaa yhteistä etua ja verkostoitumisen synergiaetuja voi olla vaikea nähdä. Synergiaetujen maksimoiminen edellyttää, että verkoston jäsenet tuovat toimintaan ydinosuamistaan. Ydinosuaminen on organisaation tarkoin valitsema keskittymisalue, johon se suuntaa olennaisen osan voimavaroistaan ja jota kilpailijoiden on yleensä vaikea kopioida. Jotkut voivat pelätä verkostoon osallistumisen kautta menettävänsä ”yksinoikeuden” kehittämäänsä ydinosuamiseen, johon heidän kykynsä tuottaa lisäarvoa asiakkaille paljolti perustuu (Viitala 2003, 175–182; vrt. Karppi 1994).

Luottamuksen puutteen lisäksi verkostotoiminnan kehittymistä voi estää myös pelko verkostoon osallistumisen aiheuttamasta lisätyöstä ja resursseja vievistä kustannuksista. Uuden arviointitoiminnan vähäisiksi tiedetyt resurssit eivät ole omiaan houkuttamaan yliopistojen, tutkimuslaitosten ja muiden arviointiosaajien osallistumista verkostomaiseen arviointiin.

Vaikka puolueettoman ja hallinnosta riippumattoman arvioinnin tarve on eduskunnan päätösten mukaan ilmeinen, menneinä vuosina arviointitoimintaan vakiintuneet

intressit eivät ole ylitettävissä ilman vastustusta. Kenenkään ei ole helppo luopua “saavutetuista eduista” ja vakiintuneista intresseistä varsinkin, mikäli se merkitsee vallan ja voimavarojen siirtymistä itseltä pois.

Verkostojen ylläpitäminen ja kehittäminen

Uuden toiminnan alkuvaiheessa keskeinen tehtävä on nimenomaan verkostojen luominen ja niiden toiminnan ja jatkuvuuden edellytysten rakentaminen. Neuvosto ja sihteeristö on luonut perustaa verkostomaiselle toimintatavalle ja kartoittanut suomalaisen arviointiosaamisen kenttää tekemällä “verkostovierailuja” yliopistoissa ja tutkimuslaitoksissa sekä muissa koulutustutkimuksen ja arvioinnin asiantuntijaorganisaatioissa. Näissä verkostovierailuissa arviointisihteeristö tapasi ja kuuli noin kahtasataa asiantuntijaa vuosina 2004 ja 2005.

Koulutuksen eri sektoreilla (perusopetus, lukiokoulutus, ammatillinen koulutus, aikuiskoulutus, ruotsinkielinen koulutus) arvioinnin suhteen relevantit verkostot poikkeavat toisistaan. Sektoreilla on tiettyjä ominaispiirteitä, millä voi olla vaikutusta myös siihen, millaisia ovat verkoston elinvoimaisuuden kannalta keskeiset toiminnot:

- verkoston luominen,
- toimijoiden sitouttaminen,
- verkoston johtaminen,
- verkoston ylläpitäminen ja
- verkoston kehittäminen.

Verkoston johtamisessa tulisi kyetä aktivoimaan verkon valmiutta rakentaa itseään ja täydentää eri osiensa kapasiteettia asetettujen tavoitteiden saavuttamiseksi. Kun useinkaan ei ole mahdollista luoda kiinteää yhteisöllisyyttä, tarvitaan toiminnan tuloksellisuuden takaajaksi yhteistä tavoitteenasettelua ja yhteisiä strategisia linjauksia ja vahvaa sitoutumista niihin. Koska verkosto on sosiaalisen teknologian tuote, jonka ylläpitäminen, muokkaaminen ja muuttaminen on jatkuva prosessi, johtaminen on keskeisesti nimenomaan yhteistyön edistämistä. Yhteistyö on omien kykyjen ja voimavarojen täydennystä. Samalla kun toimijat saavat lisää voimavaroja, he saavat yhteistyösuhteen keskinäisessä työnjaossa roolin, joka vapauttaa suuntaamaan omat voimavarat parhaiten osattuihin asioihin, ydinosaamiseen. Yhteistyösuhteen onnistumiseen vaikuttaa keskeisesti strateginen yhteensopivuus, organisaatiokulttuurin piirteet ja henkilösuhteet sekä operatiivinen yhteensopivuus, mihin verkoston rakentamisessa ja ylläpitämisessä tulee erityisesti kiinnittää huomiota. (Nikkanen & Lyytinen 1996.)

Miten koko verkostoa (solmukohtineen) pidetään koko ajan sillä tavoin aktiivisena, että sieltä voidaan rekrytoida asiantuntijaresursseja mahdollisimman nopealla aikataululla? Tämä on nopeaan reagointiin kykenevän arviointijärjestelmän yksi avainky-

symys. Verkoston toimijoiden valmiutta suhteessa Koulutuksen arviointineuvostoon on tarpeen pitää yllä koko ajan. Tässä tarkoituksessa muun muassa järjestetään seminaareja, tapaamisia ja keskustelutilaisuuksia sekä osallistutaan aktiivisesti muiden järjestämiin tilaisuuksiin. On myös ensisijaisen tärkeää, että verkostoon kuuluvat saavat kaiken relevantin tiedotus- ym. materiaalin, jota Koulutuksen arviointineuvosto tuottaa. Verkoston toimijoiden on luonnollisesti saatava tuntee, että heitä pidetään ajan tasalla arvioinnin nykytilasta, tulevaisuudesta ja kehittämisestä. Erityisen tärkeää tämä on nykyisissä oloissa, kun koko arvioinnin kenttä on monien mielestä vaikeasti hahmotettavassa käymistilassa eri suuntiin vetävine intresseineen.

Lähteet

- Blomqvist, K. 2004. Partnering in the dynamic environment: The role of trust in asymmetric technology partnership formation. Lappeenranta University of Technology. Acta Universitatis Lappeenrantaensis 122.
- Hailikari, M., Immonen, S., Kokko, N., Herrala, M., Salminen, A. & Ahola, M. 2000. Osaamiskeskusmallin kehittäminen – kuinka osaaminen saadaan leviämään. Helsinki: Työministeriö.
- Ilmonen, K. & Jokinen, K. 2002. Luottamus modernissa maailmassa. Jyväskylän yliopisto. SoPhi 60.
- Karppi, J. I. 1994. Onko yhteistoiminta mahdollista? Yleinen strategia julkisen intressin ilmentymänä. Teoksessa I. Raatikainen & J. Ahopelto. Verkostoajattelusta verkostotoimintaan. Kuopion yliopiston julkaisuja E, Yhteiskuntatieteet 21.
- Lyytinen, H.K. & Hämäläinen, K. 2004. Koulutuksen kansallisen arvioinnin kehittäminen Suomessa. Saatavilla www-muodossa <<http://www.koulutuksenarviointineuvosto.fi/linja.htm>>
- Lyytinen, H.K., Vaso, J., Kangasoja, M. & Rönkä, A. 2002. Arvioi ja opi, opi ja arvioi: arvioimalla yhteisölliseen oppimiseen. Vantaa: Printel.
- Nikkanen, P. & Lyytinen, H.K. 1996. Oppiva koulu ja itsearviointi. Jyväskylän yliopisto: Koulutuksen tutkimuslaitos.
- Ollus, M., Ranta, J. & Ylä-Anttila, P. (toim.) 1998. Yritysverkostot: kilpailua tiedolla, nopeudella ja joustavuudella. Helsinki: Sitra & Taloustieto.
- Patton, M.Q. 1997. Utilization-focused evaluation. Thousand Oaks: Sage.
- Ruuskanen, P. 2003. Verkostotalous ja luottamus. Jyväskylän yliopisto. SoPhi 78.
- Viitala, R. 2003. Henkilöstöjohtaminen. Helsinki: Edita.
- Visakorpi, J. 2002. Koulutuksen arviointineuvoston ja sen sihteeristön vaihtoehtoiset sijoitusratkaisut. Selvitysmiehen raportti. Helsinki: Opetusministeriö.
- Vuorenmaa, M. 2001. Ikkunoita arvioinnin tuolle puolen: uusia avauksia suomalaisen koulutusta koskevaan evaluaatiokeskusteluun. Jyväskylän yliopisto. Jyväskylä Studies in Education, Psychology and Social Research 176.



Arviointi järjestelmälliseksi

Artikkelissa avarretaan ja visualisoidaan lyhyesti Koulutuksen arviointineuvoston aikaisemmassa julkaisussa esitettyä arvioinnin kokonaisuuden rakentumisista. Etsitään vastausta kysymykseen miksi, jotta voitaisiin edetä kysymyksiin mitä ja miten. Artikkelissa esitettävän tikkataulumallin mukaisesti olennaisinta olisi arvioinninkin keinoin pitää huolta siitä, että opettajien ja oppilaiden jokapäiväiset älylliset ja emotionaaliset kontaktit pysyvät tuottavina. Kyseessä on kompleksinen prosessi. Sen osuva ja oikeudenmukainen arviointi edellyttää kokonaisuuden, syy–vaikutus-suhteiden, hallintaa. Arvioinnilla on oltava ympäristö ja historia. Tarvitaan sekä ajallisia, sisällöllisiä että koulutussektoreittaisia kytkeitä. Niistä ei selvitä yhden toimijan satunnaisten selvitysten ja testien voimin. Tarvitaan monen asiantuntijatahon kiinteää yhteistyötä.

Avainsanat: koulutuksen arviointi, lukioarviointi

Koulutuksen arvioinnin yksityiskohtia ja kokonaisuutta voidaan lähestyä eri tarkoituksissa, eri tavoin, eri termejä käyttäen, sekä niitä eri tavoin painottaen ja visualisoiden. Kaikki tämä ilmentää vastausta kysymykseen, miksi arvioidaan, ja myös vaikuttaa siihen, mitä ja miten arvioidaan. Valittu tarkastelukulma ja terminologia synnyttävät mielikuvat, tai päinvastoin. Ne johtavat tiettyihin painotuksiin käytännön toimissa ja luovat toimintakulttuurin.

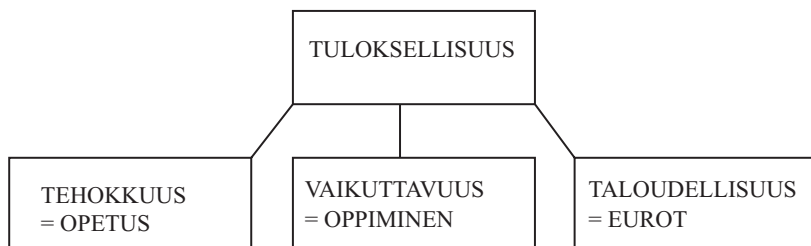
Arviointitiedon hankinnalle ja käytölle on yleisesti ottaen kolme erilaista tehtävää: 1) tilivelvollisuutta palveleva, päätösten seuraamuksiin liittyvä hallinnollinen tai oikeusturvatehtävä; 2) päätöksentekoa palveleva, suunnitteluun, ohjelmointiin ja toiminnan ohjaukseen liittyvä tehtävä sekä 3) johtopäätöksiin tähtäävä, opetus- ja oppimistapahtuman lainalaisuuksien tms. yleistettävien ilmiöiden tutkiva selvittely, joka

luo edellytyksiä opetustoimen säätelyprosessien soveltavaan kehittämiseen. (Leimu 2005, 9.) Nämä eivät ole toisensa poissulkevia kategorioita.

Valittu toimintalinja on tähän saakka painottunut kahteen ensimmäiseen. Arvioinnissa tämä on ilmennyt yksittäisinä erillisinä selvityksinä ja tiuhaan toistuvina oppimistulosten testauksina, joiden painoarvo suhteessa koulutuksen kokonaisuuteen ei näy eikä tunnu. Tieto ei ole kumuloitunut siinä määrin kuin olisi mahdollista ja suorastaan tarpeen. Arviointi on aavaa ja rannatonta. Syy ei välttämättä ole ollut arviointisijoissa. Arviointikohteiden valinta/valikoituminen ei ole noudattanut pitkäjänteistä suunnitelmallisuutta, ohjelmaa, jossa olisi nähty sekä metsä että puut.

Koulutuksen tuloksellisuus arvioinnin ensisijaisena määrittäjänä

Opetushallituksessa valittiin aikoinaan arvioinnin keskeiseksi käsitteeksi taloustieteen puolelta tuttu tuloksellisuus, jota on käytetty valtionhallinnossa ja myös eri maiden koulutusjärjestelmissä yleisemminkin. Tosin tuottavuus näyttää nykyisin olevan syrjäyttämässä sen. Tuloksellisuus koostuu kolmesta ulottuvuudesta: tehokkuus (= opetus), vaikuttavuus (= oppiminen) ja taloudellisuus (= eurot) (ks. Peltonen ym. 1992). Nuo sanat kertoivat koulutuksen kentälle, miksi arvioidaan. Ne myös osaltaan määrittivät mitä ja miten arvioidaan.



Kuvio 1. Opetushallituksen arviointimalli

Koulutuksen eri tasoilla määriteltiin eri ulottuvuuksia kuvaamaan arviointikohteita ja saatiin arvioinnin kehikko, joka muuttuu sitä mukaa kun esille tulee uusia tarpeelliseksi koettuja arviointikohteita. Tältä perustalta ja lyhyesti sanottuna vaikuttavuus tarkoittaa oppimista eri tavoin mitattuna sekä koulutustarpeen ja tarjonnan vastaavuutta. Tehokkuus tarkoittaa sekä yksittäisen oppilaitoksen että koko koulutusjärjestelmän toimivuutta ja vaikuttavuutta, ja sitä ehkäisevien ja edistävien tekijöiden arviointia. Taloudellisuus tarkoittaa rahoitusjärjestelmien arviointia ja mm. koulutuksen vaihto-

ehtoisten tuotantotapojen pohdintaa. Käytännössä on rajoitettu poikkileikkaustyyppi-
sesti oppimistuloksiin ja joihinkin erillisiin teema-arviointeihin.

OECD (Education at a Glance 2003) on pyrkinyt määrittämään määrällisiä indikaat-
toreita, jotka parhaiten kuvaavat koulutuksen kokonaisuutta, ja on käyttänyt siinä
samantapaista lähtökohtaa kuin edellä. Määrittelyssä on kuitenkin päädytty koulutuk-
sen arvioinnin kokonaisuuden hahmottamisen kannalta ilmaisuvoimaisempaan kehik-
koon (kuvio 2).

	Tuloksellisuus- tieto	Tuloksellisuuteen vaikuttavat tekijät (keinot ja puitteet)	Toimintaa mää- rittävät edellytykset tai -rajoitukset
Koulutukseen osallistujat	1A	2A	3A
Oppimisympäristöt	1B	2B	3B
Koulutuksen järjestäjät	1C	2C	3C
Kansallinen koulutus- ja ohjausjärjestelmä	1D	2D	3D

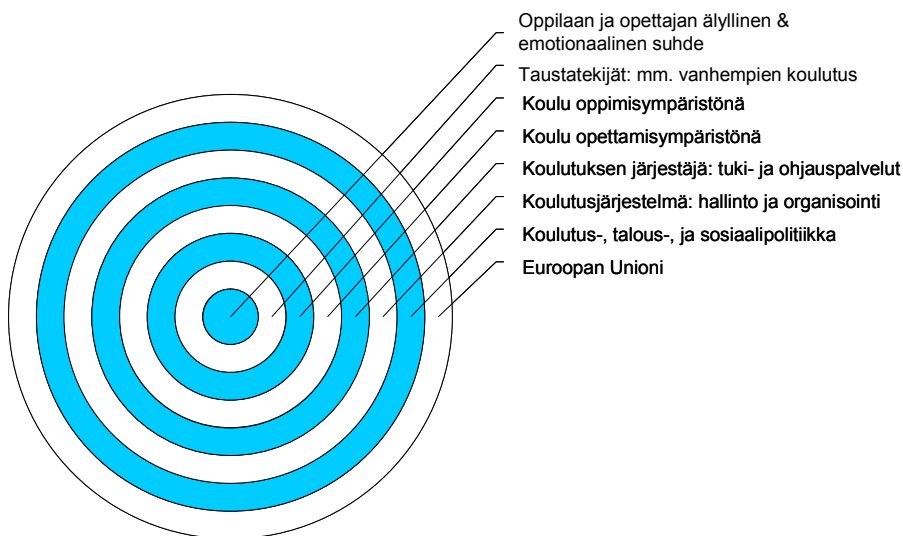
Kuvio 2. Koulutuksen arvioinnin kokonaisuus, OECD

OECD:n hahmotelmassa tarkastellaan tuloksellisuutta, tuloksellisuuteen vaikuttami-
sen mahdollisuuksia sekä noudatettavan toimintapolitiikan kontekstuaalisia edellytyk-
siä ja rajoituksia suhteessa koulutusjärjestelmän eri tasoihin. Olennainen ero verrat-
tuna Opetushallituksen perusmalliin on koulutusjärjestelmän purkamisen tasoihin.
Kuhunkin laatikkoon 1A–3D etsitään mahdollisimman kuvaavat indikaattorit; esimer-
kiksi 1A: koulusaavutukset, asenteet, oppimisen strategiat; 2A: sukupuolierot, kyky
ja halu elinikäiseen oppimiseen.

Tämänkaltaisen kehikon rakentaminen johtaa ainakin periaatteessa siihen, että kou-
lutuksen arviointia ajatellaan laajasti suhteessa tuloksellisuuteen ja ettei joitakin kes-
keisiä osatekijöitä jätetä huomiotta. Oppiminen ja sen eritasoiset tuotokset (vaikutuk-
set) sekä edellytykset nousevat vahvasti esiin. Vastauksena miksi-kysymykseen voi-
tisiin yksinkertaistaen sanoa, että kyse on koulutusjärjestelmän tuloksellisuudesta
huolehtimisesta. Tarkoitus on tunnistaa tuloksellisuuteen sekä negatiivisessa että po-
siitiivisessa suhteessa olevat tekijät sekä arvioida niiden vallitsevuutta ja vaikuttavuut-
ta.

Suuntana selityskykyiset kehikot

Jos lähdetään siitä, mikä koulutusjärjestelmän tuloksellisuuden (tuottavuuden) kannalta on olennaisinta, siis itse opetus- ja kasvatustehtävä, voidaan päätyä esimerkiksi kuvion 3 kaltaiseen kuvaukseen. Kuvioista voidaan käyttää vaikka nimitystä samankeskisten kehien vaikutuskenttä (ks. Leimu 2005), arvioinnin sisäkkäiset kehät tai tikkataulu. OECD:n kehikon ajatusta mukailleen se sisältää sekä eri tasot että koulutusjärjestelmän tulokseen vaikuttavia tekijöitä ja edellytyksiä.



Kuvio 3. Koulutuksen arvioinnin sisäkkäiset kehät

Sisäkkäisten kehien perusidea edellisiin kuvioihin verrattuna on, että kun suunnitellaan mille kehälle tahansa sijoittuvaa tai enemmän painottuvaa arviointia, katse – ja toivottavasti myös ajatus – hakeutuu sisimpään kehään, koko koulutusjärjestelmän ytimeen pohtimaan, miten aiottu arviointi edistää sen hyvinvointia. Samoin on katsottava arvioinnin suhdetta muihin kehiin, kokonaisuuteen. Toivottavasti nämä ajatukset osaltaan auttaisivat vähentämään vähemmän tarpeellisen arvioinnin määrää ja lisäämään arvioinnin osuvuutta. Kehät myös havainnollistavat, kuinka monentasoiset tekijät vaikuttavat koulutusjärjestelmän ja yksittäisen oppilaitoksen tuotokseen. Tässä luvussa käytetyt esimerkit ja terminologia viittaavat perusopetukseen, mutta kehäidea on sovellettavissa muihinkin koulutussektoreihin.

Kuvion 3 napakymppissä – *oppilaan ja opettajan älyllisessä ja emotionaalisisessa (tai koulumaisemmin ilmaistuna pedagogisessa) suhteessa* – luodaan suomalaisen koulutusjärjestelmän monimuotoinen tulos, jokaisena koulupäivänä. Tästä myllystä tule-

vat ne jauhot, joista kansakunnan leipä varsin pitkälti leivotaan. Tavoitteet vaihtelevat ajan myötä mutta perusprosessori on ja pysyy, ainakin toistaiseksi. Tästä lähtökohdasta *koulutuksen arvioinnin peruskäsitteet* ovat samat kuin koulutuksenkin: *oppiminen ja opettaminen*. Näiden tilaa ja tilaan vaikuttavien tekijöiden merkitystä arvioidaan. Huolimatta perusopetukseen jatkuvasti liittyvästä negatiivisesta uutisoinnista ja mielipidekirjoittelusta mylly tuottaa tasa- ja korkealaatuista tulosta niillä mittareilla mitattuna, joita toistaiseksi on käytetty. Nykyistä perusteellisemmän tutkimisen arvoista olisi, miksi tämä oppilaan ja opettajan suhde suomalaisessa koulujärjestelmässä on niin hyvä ja tuottava kuin on. Kun tiedetään, miksi se on hyvä, välttään ehkä hyvän muuttamiselta huonommaksi.

Koulujärjestelmältäkin odotetaan yhä vain parempaa tuottavuutta. Käytännössä kysymys on, kuinka suuri teho tästä myllystä, opettajan ja oppilaan välisestä suhteesta, voidaan puristaa särkemättä sitä. Arvioinnin yksi keskeinen tehtävä olisi silloin kertoa paitsi tuloksen suhteesta tuotantokustannuksiin myös siitä, kuinka lähellä murtumispiste on ja millä keinoin se voitaisiin välttää. Tämä on yksi, hieman julmaltakin vaikuttava, vastaus kysymykseen miksi. Inhimillisemmin ilmaistuna arviointia olisi tehtävä siksi ja vain sen verran, että tiedetään millä tasolla mylly toimii. Jos se ei toimi kunnolla, voidaan arvioida, mitä on tehtävissä ja miltä kehiltä ongelmat ja ratkaisut löytyvät. Viimeisin PISA-tutkimus (Kupari ym. 2004) kertoo taas kerran tuloksen olevan erinomainen, mutta muut viestit koulutuksen kentältä kertovat opettajien väsymyksestä ja oppilaiden viihtymättömyydestä. Onko tehoraja tulossa vastaan vai onko jälkimmäisissä viesteissä kysymys jostakin muusta?

Oppilaan ja opettajan älyllisen ja emotionaalisen suhteen tuotoksiin – tiedollisiin, taidollisiin, asenteellisiin jne. oppimistuloksiin – liittyvissä arvioinneissa minimiväahtimus on, että kontrolloidaan niitä *taustatekijöitä*, joiden jo olemassa olevan tutkimuksen perusteella tiedetään keskeisesti vaikuttavan tuotoksiin. Kun on etsitty sekä yksilöllisten että koulun ja jopa aluetason erojen syytä, vahvimmat selittävät tekijät ovat liittyneet oppilaiden ominaisuuksiin (kognitiiviset taidot, sukupuoli, motivaatio jne.) sekä vanhempien sosioekonomiseen taustaan. Tähän liittyy myös yleensä koulutustaso. (Blom 1994; Kirjavainen & Loikkanen 1994; Kuusela 2003; Kuusela & Hautamäki 2004.) Nämä eivät ole mullistavia löydöksiä mutta tuntuvat unohtuvan etenkin silloin, kun lukioita julkisesti laitetaan tuloksellisuusjärjestykseen. Jos ja kun yleensäkin puhutaan koulun tuloksesta tai tuloksellisuudesta, yleisenä lähtökohtana voidaan pitää seuraavaa: koulun tulos on vähintään hyvä silloin, kun *oppilaat edistyvät pidemmälle kuin mitä heidän lähtötilanteensa antaisi olettaa* (Mortimore 1991). Kyseessä on silloin koulun tuottama lisäarvo. Saman voidaan ajatella pätevän myös alueellisiin tai kansainvälisiin vertailuihin.

Oppimistulosten arviointi kertoo, mitä oppilaiden ja opettajien jokapäiväiset kontaktit, taustatekijät kontrolloiden, tuottavat. Arvioinnin osuvuutta ja selitysvoimaa täydentää koulun ottaminen huomioon *oppimis- ja opettamisympäristönä*. Näillä kehillä on arvioinneissa erityinen merkitys, koska tästä kokonaisuudesta on löydetä-

vissä olennaisia tekijöitä, jotka kertovat oppilaiden ja opettajien suhteen laadusta ja siihen vaikuttavista tekijöistä koulun tasolla. Tärkeää olisi löytää ja määrittää muutama keskeinen “hälytysindikaattori”. Olisiko opettajien väsymys yksi hälytysindikaattori, ja onko se ehkä seurausta joistakin muista opettamisympäristöön liittyvistä tekijöistä? Nykytietämyksen mukaan silloin kun puhutaan koulun tehokkuudesta oppimistuloksilla mitattuna, on koulun tasolla kyse usean tekijän yhdysvaikutuksesta. Vaikutukset ulottuvat johtamisen jämakkydestä ja johdonmukaisuudesta aina luokkaprosessien perusvarmuuteen ja turvallisuuteen oppilaan kannalta katsottuna. Kaikessa toiminnassa pitäisi korostua oppimisen ja opettamisen merkitys koulun perustehtävänä (Mortimore 1991).

Oppimis- ja opettamisympäristöön kietoutuvat oppimiselle ja opettamiselle asetetut taloudelliset ja henkiset reunaehdot, jotka ovat riippuvaisia *koulutuksen järjestäjien* mahdollisuuksista ja toimista. Näitä ovat mm. aineellinen ja henkinen työympäristö, henkilökunnan pätevyys ja pysyvyys, tasavertaiset oppimismahdollisuudet sekä erityistä tukea tarvitsevien oppilaiden opetuksen järjestäminen. Tässä suhteessa myös koulutuksen järjestäjät ovat keskenään eriarvoisessa asemassa riippuen muun muassa koulutuksen arvostuksesta kuntatalouden kokonaisuudessa. Olisiko opettajien väsymys resurssien puutetta tai väärin kohdentumista?

PISA-tulosten arvioinneissa on tuloksen korkeaan laatuun keskeisesti vaikuttavina tekijöinä tuotu esiin yhtenäiskoulujärjestelmä ja opettajien korkea koulutustaso. Nämä ovat esimerkkejä *koulutusjärjestelmän* kehästä. Muita olennaisia samalle kehälle kuuluvia tekijöitä ovat mm. hallinto ja ohjaus, rahoitus, opetussuunnitelmajärjestelmä ja kehittämistoimet. Viimeistään tässä vaiheessa pitäisi herätä kysymys, onko opetus-hallinto itse oikea toimija arvioimaan tekemistensä tai tekemättä jättämistensä vaikutuksia, sijoittuvat ne mille kehälle tahansa. Yksittäisissä sisällöllisesti rajatuissa arvioinneissa koulutusjärjestelmään liittyvien tekijöiden osuuteen voidaan yleensä viitata vain periaatteella “saattaa olla yhteydessä”. Vai olisiko sittenkin niin, että opettajat opettavat hyvin ja oppilaat oppivat hyvin koko kehästä riippumatta? Vai olisiko niin, että opettajien väsymys on seurausta yhtenäiskoulujärjestelmän opetuksen tasolle tuomasta heterogeenisuudesta? Arvioinneissa tiukempi linkitys olisi tarpeen jos senkin vuoksi, etteivät luokkatodellisuus ja hallintotodellisuus erkane toisistaan tarpeettoman kauas.

Valtakunnallisella tasolla arvioinnin osuvuuden viimeinen ja ajan oloon ehkä ratkaiseva kehä löytyy *koulutus-, talous- ja sosiaalipolitiikasta*. Paitsi toimintaolosuhteet ja resurssointi myös tulosta kuvaavat tekijät määrittyvät pitkälti uloimmalta kehältä; mitä asioita kulloinkin pidetään isänmaan tulevaisuuden kannalta tärkeimpinä. Ne ovat sopimuksenvaraisia. Koulutusjärjestelmä ei ole suljettu, vaan taloudellisten, poliittisten, sosiaalisten ja kulttuuristen ehtojen säätelemä instituutio, osa yhteiskunnan kokonaisjärjestelmää, jonka osien välillä vallitsevat monimutkaiset rajapintasuhteet (Raivola 2000). Koulutus on yhteiskunnan osa ja voi hyvin tai huonosti sen mukana, ellei sille jostain syystä myönnetä erityisasemaa. Kuriositeettina mainittakoon,

että viimeisimmän arvion mukaan huomattava joukko suomalaislapsista ja -nuorista elää oloissa, jotka vaarantavat terveen kasvun ja kehityksen. Lastensuojelun piirissä olevien perheiden lapsia oli viime vuonna yli 60 000, lähes yksi kouluun vuosittain astuva ikäluokka (Söderholm ym. 2004). Olisiko niin, että tällä asialla ja opetukseen käytettävissä olevien varojen vähenemisellä on jokin yhteys, ehkä ratkaiseva, opettajien väsymiseen? Näkyykö jälki ehkä seuraavassa PISA-tutkimuksessa?

Uloimpana kehänä on Euroopan Unioni ja kansainvälistyminen yleensäkin. EU vaiuttaa jo paitsi kehittämishankkeiden sisältöön ja rahoitukseen myös ennen pitkää ja vaivihkaa koulutuksen rakenteisiin ja opetussuunnitelmien perusteisiin. EU:sta tulee myös arvioinnin malleja suuremmassa ja pienemmässä mittakaavassa. Suomi yhtenäistyy Eurooppaan, vaikka kouluasioissa ehkä näyttäisi olevan aiheellisempaa toimia päinvastoin. Väsähtääkö koko koulutusjärjestelmä eurooppalaisiin normeihin?

Miksi, mitä, miten?

Asetuksella (n:o 150/2003) on koulutuksen arvioinnin tavoitteiksi asetettu

- hankkia ja analysoida tietoa valtakunnallisen koulutuspoliittisen päätöksenteon ja koulutuksen kehittämisen pohjaksi;
- hankkia ja analysoida tietoa koulutuksen paikallisen kehittämistyön ja päätöksenteon pohjaksi sekä
- tukea opiskelijoiden oppimista, opetustoimen henkilöstön työtä ja oppilaitosten kehittämistä.

Nämä ovat viralliset vastaukset miksi-kysymyksiin ja kattavat Leimun (2005) esittämät arvioinnin perustarkoitukset. Epävirallisemmin ajateltuna *olennaisinta olisi tulostiedon tuottamisen ohella myös arvioinnin avulla pitää huolta siitä, että opettajien ja oppilaiden jokapäiväiset kontaktit pysyvät tuottavina sekä omalta osaltaan taa-ta opettajien ja oppilaiden hyvinvointi*. Tällöin on katsottava paitsi yksityiskohtia myös koulutusjärjestelmää kokonaisuudessaan. Tässä työssä keskeinen teonsana ei ole tuomitseminen vaan tukeminen. Idealistinen ajatus, myönnettäköön.

Tikkataulu arvioinnin lähtökohtakehikkona kertoo, mikä arvioinnissa on tärkeintä. Entä miten? Arviointineuvoston arviointinäkemys yksi peruslähtökohta on ollut, että ollakseen pätevää arvioinnilla on oltava ympäristö ja historia. Tietoa on suhteutettava johonkin jollakin perusteella. Tarvitaan sekä ajallisia, sisällöllisiä että koulutus-sektoreittaisia kytkentöjä (ks. Korkeakosken artikkeli tässä julkaisussa). Lukioarviointissa tämä tarkoittaisi mm. sitä, että suhteessa lukion monipuolisiin tavoitteisiin ylioppilaskirjoitusten tulos on koulukohtaisestikin tarkasteltuna yksipuolinen mittari. Sekä koulujen omiin että valtakunnan tarpeisiin olisi vähimmäistavoitteena saata-va arviointitietoa, jossa ylioppilaskirjoitusten tulokseen yhdistyvät muiden tavoitteiden saavuttamiseen liittyvä tieto, aikaisempaan kouluhistoriaan (perusopetus) liittyvä tieto, kouluun oppimis- ja opettamisympäristönä liittyvä tieto ja viimein jatko-opin-

toihin tai työelämään sijoittumista koskeva tieto. Tähän perusrunkoon kytetään kulloinkin vain aivan välttämättömiksi katsotut tila- ja teema-arvioinnit. Tämän tyyppistä laajaa ja pitkittäistä, noin neljän-viiden vuoden välein toteutettavaa, tiedon keruuta ja arviointiperustaa on esitetty lukioarvioinnin rungoksi itse asiassa jo kymmenkunta vuotta (ks. Blom 1994 ja Välijärvi 1995).

Vastaukset kaikkein perimmäisiin miksi-kysymyksiin ja myös syyt kulloinkin valitseviin koulutuksen toimintaolosuhteisiin ovat vaikeasti tavoitettavissa ja eri tutkimuksiin ja arviointeihin hajautetusti kätkeytyneitä. On selvää, että arvioinnin kokonaisuutta ei hallita yhden toimijan voimin. On hyödynnettävä kaikkea sitä asiantuntemusta mitä valtakunnasta löytyy. On myös selvää, että paketoimalla arvioinnit ja tutkimukset toisistaan erillisiin pieniin nyytteihin ja kilpailuttamalla niiden tekijät pyssy arviointi edelleenkin aavana ja rannattomana.

Lähteet

- A 150/2003 koulutuksen arvioinnista n:o 150/2003.
- Blom, H. 1994. Ylioppilastutkinto erojen ilmentäjänä. Teoksessa R. Jakku-Sihvonen & H. Blom (toim.) Lukion tila 1994. Opetushallitus. Arviointi ja seuranta 5. Helsinki, 201–224.
- Education at a Glance 2003. OECD.
- Jakku-Sihvonen, R. & Blom, H. (toim.) 1994. Lukion tila 1994. Opetushallitus. Arviointi ja seuranta 5. Helsinki.
- Kirjavainen, T. & Loikkanen, H. A. 1994. Lukioiden menestystä selittävät tekijät. Teoksessa R. Jakku-Sihvonen & H. Blom (toim.) Lukion tila 1994. Opetushallitus. Arviointi ja seuranta 5. Helsinki, 247–266.
- Koulutuksen tuloksellisuuden arviointimalli. 1998. Arviointi 7. Opetushallitus. Helsinki.
- Kupari, P., Välijärvi, J., Linnakylä, P., Reinikainen, P., Brunell, V., Leino, K., Sulkunen, S., Törnroos, J., Malin, A. & Puhakka, E. 2004. Nuoret osaajat. PISA 2003 -tutkimuksen ensituloksia. Jyväskylän yliopisto: Koulutuksen tutkimuslaitos.
- Kuusela, J. 2003. Lukioiden tuloksiin vaikuttavista tekijöistä. Opetushallitus. Moniste 13. Helsinki.
- Kuusela, J. & Hautamäki, J. 2005. Mitataan oppilaita, mutta päätellään koulusta. Artikkelitässä julkaisussa.
- Leimu, K. 2005. Kansainvälisyys koulutuksen arviointitutkimuksessa. Teoksessa K. Leimu (toim.) Kansainväliset IEA-tutkimukset Suomi-kuvaa luomassa. Jyväskylän yliopisto: Koulutuksen tutkimuslaitos, 7–26.
- Mortimore, P. 1991. The nature and findings of research on school effectiveness in the primary sector. Teoksessa S. Riddell & S. Brown (toim.) School effectiveness research: Its message for school improvement. Edinburgh: HMSO.
- Peltonen, M., Laitinen, J. & Juuti, P. 1992. Koulutuksen tuloksellisuus. Johtamistaidon opisto. Aavartta-sarja 34. Tampere.
- Raivola, R. (toim.) 2000. Vaikuttavuutta koulutukseen. Suomen Akatemian koulutuksen vaikuttavuusohjelman tutkimuksia. Suomen Akatemian julkaisuja 1/2000. Helsinki: Edita.

- Raivola, R., Valtonen, P. & Vuorensyrjä, M. 2000. Käsitteet, mallit ja indikaattorit koulutuksen tehokkuutta ja vaikuttavuutta arvioitaessa. Teoksessa R. Raivola (toim.) Vaikuttavuutta koulutukseen. Suomen Akatemian koulutuksen vaikuttavuusohjelman tutkimuksia. Suomen Akatemian julkaisuja 1/2000. Helsinki: Edita, 11–28.
- Söderholm, A., Halila, R., Kivitie-Kallio, S., Mertsola, J. & Niemi, S. (toim.) 2004. Lapsen kaltoinkohdeltu. Helsinki: Duodecim.
- Väljjarvi, J. 1994. Lukiolaisten korkeakouluvalmiudet. Teoksessa R. Jakku-Sihvonen & H. Blom (toim.) Lukion tila 1994. Opetushallitus, Arviointi ja seuranta 5. Helsinki, 283–288.
- Väljjarvi, J. 1995. Yo-tutkinto ja arviointi. Ylioppilastutkinnon tietokannan hyödyntäminen koulutuksen arvioinnissa. Muistio Opetushallitukselle. Jyväskylän yliopisto. Koulutuksen tutkimuslaitos.



Med siktet på samverkan i utvärdering

I denna artikel presenterar och diskuterar vi det nya utvärderingssystemet ur finlandssvensk synvinkel. Vad innebär de riktlinjer som lagstiftningen ger och de verksamhetsprinciper som det nya rådet för utvärdering slagit fast för utvärderingen av den svenska utbildningen? Rådets uppgift är att få till stånd ett nationellt fungerande utvärderingsnätverk, med uppgift att bidra till att utveckla utbildningen. För Svenskfinlands del skall rådets verksamhet innefatta specifika svenska frågor. Därmed medverkar rådet till att upprätthålla och utveckla svenskspråkig utbildning och ett levande tvåspråkigt samhälle.

Nyckelord: svenskspråkig utbildning, utvärdering, utvärderingssystem, deltagande utvärdering, utvecklande utvärdering

Vad inbegrips i fenomenet utvärdering?

Svaret på frågan vad utvärdering innebär är omfattande och komplext. Vi anknyter vår definition i första hand till den tolkning som rådet gör. På en mer principiell nivå kan emellertid två processer särskiljas. Exempelvis Huotari (2000, 216), som ingående har analyserat begreppet utvärdering, innefattar i det både en beskrivning av och ett värdeomdöme om det objekt som är föremål för utvärderingen. Rådet gör ungefär samma tolkning (jfr *Utbildningsutvärderingens nya inriktning*, 2004). Utvärdering betyder att bestämma värdet på ett objekt eller en verksamhet som granskas genom en värderande tolkningsanalys. Objektet för utvärderingen varierar och kan gälla individer och system. Det kan således inbegripa inlärningsresultat, men också verksamheten i vidare mening, t.ex. dess förutsättningar. Objektet kan innefatta ett antal konkreta fall som exempelvis enskilda skolor eller också ett visst utbildningsstadium.

Utvärdering kan ses som en specifik form av vetenskaplig forskning. Forskningen strävar till vetenskapligt sann kunskap, medan utvärderingen ger tillförlitlig kunskap för något praktiskt syfte (Karjalainen & Huusko 2004). Därför är utvärderingen som idé alltid knuten till åtgärder och förändring. Utvärderingen ställer diagnos och kan ge mer eller mindre konkret underlag för praktiska åtgärder som rättar till missförhållanden eller som utvecklar verksamheten. Detta betyder att utvärdering är nära sammanbunden med kommunikation mellan den som utvärderar och den som blir utvärderad. I dag betonas inom forskningen deltagande utvärdering (participatory evaluation) som understryker kommunikationens betydelse. Fokus ligger på att involvera inblandade parter i diskussioner och överväganden (Cousins & Earl 1995).

Generellt sett har utvärderingen i utbildningsdiskussionen fått en alltmer framskjuten roll både nationellt och internationellt. Perioden från 90-talets början till i dag kan i Europa med fog benämnas utvärderingens era. Under denna tid har tilltron till utvärderingens möjligheter att tjäna som ett rationellt, effektivt och allmänt accepterat instrument för slutsatser och kontroll av utbildningen vuxit sig stark (Broadfoot & Black 2003). Detta hänger uppenbarligen samman med det allmänna politiska klimatet i Europa och de ökade ansträngningarna att förenhetliga den europeiska utbildningen.

I Finland fick begreppet utvärdering en ny innebörd med införandet av lokala läroplaner (Svedlin 2003). På 1980-talet talade vi om evaluering och innefattade i begreppet i huvudsak elevbedömning, men under följande decennium gav vi det en bredare innebörd och övergick till att tala om utvärdering som rymmer både individ och system. Kommunerna blev skyldiga att utvärdera den verksamhet de ansvarar för och skolorna fick i uppgift att utvärdera sig själva (lag 628/1998, 21§).

Olika uppfattningar om vad utvärdering är eller borde vara återspeglas i den metodiska uppläggning som används (Nikkanen & Lyytinen 1996). Utvärderingen kan såsom vi framhållit betraktas som en form av *tillämpad forskning*, men också som ett led i *styrningen av utbildningen som system*, som *ställningstaganden av specialister* eller som *politik* (Smith & Glass, 1987). Bakom dessa uppfattningar ligger också olika sätt att definiera kunskapens natur.

Om utvärdering definierats som en form av *tillämpad forskning*, har den ofta inneburit att traditionella kvantitativa forskningsmetoder utnyttjats, dvs. kvantitativt mätbara fakta har samlats in och statistiskt tillförlitliga slutsatser har dragits. Utvärdering som *styrning av utbildningen som system* innebär att insatser, aktiviteter och resultat beskrivs och relateras till varandra och informationen används för att vägleda valet av åtgärder så att utbildningen når de mål som uppställts. Om utvärdering karakteriserats som *ställningstagande av specialister*, är utgångspunkten att samlad sakkunskap om ett utvärderat objekt ger färdigheter för kvalitativa bedömningar av objektet. Då utvärdering uppfattas som *politik*, har tron på objektiv, värdeneutral kunskap minst inflytande.

Under de senaste åren har en rad forskningsresultat minskat tilltron till den förmen- ta objektiviteten och effektiviteten i konventionella summativa utvärderingstekniker,

som på ett enkelt och oproblematiskt sätt ansetts inverka motiverande på verksamheten. (Broadfoot & Black 2003). På individuell nivå har det i flera undersökningar kunnat konstateras en negativ inverkan av kunskapsmätning på lågpresterande individer (Harlen & Crick 2003), vilket är fatalt då det yttersta syftet med individuell utvärdering måste vara att främja lärande.

Sättet att genomföra utvärdering har emellertid visat sig vara avgörande för hurdana konsekvenser den får. De negativa effekterna verkar kunna minskas med hjälp av en rad åtgärder såsom just parternas medverkan (participatory evaluation) i planeringen av utvärderingen (jfr. Stobart 2003; Robinson & Cousin 2004). Utvärderingen kan utformas så att den stöder utvecklingen av deltagarnas självvärderingsförmåga och fokuserar på själva utvecklingen. Dylika omständigheter behöver beaktas då utvärderingar planeras och även då system för nationell utvärdering byggs upp. Detta har också varit en uttalad ambition, när det nya utvärderingssystemet i vårt land utarbetats och dess verksamhet konkretiserats.

Ett unikt utvärderingssystem

Det nya utvärderingssystemet bygger på en lag från 2003 (Lag om ändring av 21 § lagen om grundläggande utbildning, 24.1.2003/32). Ett särskilt råd för utvärdering av utbildning tillsattes samma år med uppgift att bistå undervisningsministeriet och organisera nationella utvärderingar. Rådet består av representanter för olika intressentgrupper inom utbildningen. Genom det nya systemet anses Finland för närvarande gå i spetsen för utvecklingen av utvärderingen i Europa. Systemet ger åtminstone enligt lagen större möjligheter än någon annanstans till en oavhängig utvärdering i förhållande till utbildningsförvaltningen. Inom utbildningsförvaltningen finns det dock ett visst motstånd mot det nya systemet, vilket lett till att ibruktagandet förhalats. Enligt lagstiftningen (Statsrådets förordning om utvärdering av utbildning 20.2.2003/150) svarar Rådet för utbildningsutvärdering tillsammans med undervisningsministeriet för den externa utvärderings-verksamheten och utvecklingen av den. Till rådets uppgifter hör att organisera externa utvärderingar med hjälp av nätverk som består av experter inom olika områden.

Svenskspråkig utbildning – jämbördig men ibland annorlunda

Den bärande principen för utvärderingen av den svenskspråkiga utbildningen är att den är jämbördig med i de nationella och internationella utvärderingarna. Genom att den svenska utbildningen är integrerat med i alla utvärderingar, vet vi var vi står nationellt och internationellt sett. Det kräver förstås att samplet i de svenska skolorna är

så stort att tillförlitliga jämförande analyser kan göras. I praktiken innebär detta en procentuellt sett större andel svenska än finska skolor. När utvärderingarna är av deltagande natur (participatory evaluation), garanterar medverkan att också den svenska utbildningen får verktyg och impulser till utveckling av de objekt som utvärderas.

Ibland kan resultaten av internationella och nationella utvärderingar lyfta fram frågor som verkar vara problematiska för enbart den svenskspråkiga utbildningen och som därför behöver utvärderas närmare. I dessa fall är det berättigat att genomföra egna utvärderingsprojekt för att belysa specifikt finlandssvenska utbildningsförhållanden.

De åländska skolorna utgör en viktig del av den svenskspråkiga utbildningen i Finland. De är emellertid inte underställda undervisningsministeriet. Det är Ålands landskapsregering som har lednings- tillsyns- och utvecklingsansvaret för utbildningen på Åland (3§ GrL 18/1995 och 3§ LL 52/1997 ä. 82/2002). De utvärderingar som är relevanta också ur den åländska utbildningens synvinkel är det möjligt och önskvärt att genomföra också på Åland. Det sker i samarbete med Ålands landskapsregering.

Konsekvenser av de valda utvärderingsprinciperna

Rådet för utvärdering har formulerat en rad principer som skall styra utvärderingen. Principerna är ideologiskt genomsyrade av bestämda värden som rådet enats om och som uttrycks med hjälp av ett antal begrepp. Till de styrande principerna hör oavhängighet, hög klass, utvecklande utvärdering, deltagande och effekt. Principerna får något olika konsekvenser beroende på om vi talar om interna finlandssvenska utvärderingsprojekt eller om vi tänker på projekt där den svenskspråkiga utbildningen utvärderas integrerat med den finskspråkiga.

Oavhängig utvärdering innebär att den myndighet som ansvarar för utveckling och styrning av utbildningen inte själv utvärderar. I stället utnyttjas grupper av sakkunniga vid universitet och andra oberoende organisationer. Mångfasetterad utvärdering, dvs. utvärdering som utförs av olika intressentgrupper med olika synvinklar på fenomenet, är ett annat sätt att främja oavhängighet. Då exempelvis svensk och finsk utbildning utvärderas tillsammans, blir utvärderingen mångsidigare och mera oavhängig, ifall experter på finlandssvensk utbildning anlitas.

När specifikt finlandssvenska projekt planeras blir problematiken en annan. Beroende på utvärderingsobjekt kan det vara svårt att samla ihop tillräckligt med oberoende och sakkunniga utvärderare. För att arbetet inte skall få alltför starka drag av självutvärdering är det därför skäl att kalla in utvärderare med hemvist utanför Svenskfinland. Exempel på ett dylikt förfarande kan vi hitta i andra länders praxis. Utländska utvärderare utnyttjas till exempel i de danska utvärderingsprojekten (Evalueringsgrupp 2004).

Om *utvecklande* utvärdering (developmental evaluation) talas det när det yttersta syftet med utvärderingen uppfattas vara att främja utveckling och när utvärderings-

processen läggs upp med detta i sikte. Detta i sin tur innebär att utvärderingen inte ensidigt kan begränsas till resultaten. Kvalitativt orienterade utvärderingar som inriktas på såväl utbildningsförlopp, kontext, som resultat är lika viktiga. Då blir det möjligt att formulera förslag till åtgärder på nationell nivå. Utvecklande utvärdering medför också behov av att engagera alla de grupper som berörs av utvärderingen, dvs. av *deltagande* utvärdering (participatory evaluation). Beroende på utvärderingsobjekt engageras elever, lärare, övrig skolpersonal, utbildningsanordnare och vid systemutvärdering också beslutsfattare på riksnivå. Utvärderaren uppfattas ur den utvecklande utvärderingens synvinkel som en medlem i ett utvecklingsteam och med uppgift att ställa utvärderande frågor och stimulera teamet till databaserat beslutsfattande (Patton 1997). Oberoende av om utvärderingen av den svenskspråkiga utbildningen utförs integrerat i den finska eller om projektet är specifikt svenskt krävs det tillräckligt med utvärderare, dvs. resurser för att utvärderingen skall vara utvecklande.

Hög klass eller kvalitet kan uppfattas och definieras på många sätt. Vi kan se det som kvalitet när verksamheten är exceptionell, är felfri eller jämn, är ändamålsenlig, ger valuta för pengarna eller får till stånd förändring (Harvey & Green 1993). Men frågan om kvalitet är mer komplex än så. Ordet kvalitet är högfrekvent i dagens utbildningsdiskussion och uttrycker egenskaper som betraktas som centrala i all utbildning. Den egentliga innebörden i begreppet fortsätter dock att vara svårfångad. Så mycket kan emellertid sägas att kvalitet är knutet till bestämda värden som uttrycker utbildningens ideal. I lexikalisk mening är en kvalitet en utmärkande egenskap eller beskaffenhet hos något. Inom skola och utbildning hänför sig också olika slag av sociala innebörder till ordet kvalitet, som fördelning, särskiljning, klassificering, gränsdragning, rangordning osv. (Hamilton 2004). Därför knyts betydelsen lätt till värdemässiga omdömen i betydelsen bättre och sämre, lyckat och misslyckat osv. I utvärderingsrådets verksamhet är avsikten att i diskussionen om kvalitet inte fastna i enkla dikotomier av den typ som antytts. Principen om deltagande utvärdering (participatory evaluation) skall förhoppningsvis bidra till att ge en nyanserad och innebördsrik bild av de komplexa fenomen som utgör objekt för utvärdering.

Vad högklassig utvärdering beträffar ser utvärderingsrådet träffsäkerhet och effektivitet som centrala element. Högklassig utvärdering förutsätter att de aspekter som fokuseras i utvärderingsarbetet är relevanta. Ur denna synvinkel är det därför motiverat att representanter för den svenskspråkiga utbildningen är med också vid planeringen av utvärderingarna och att det vid behov kan genomföras projekt som inriktas på specifikt finlandssvenska frågor. Dessutom betonar rådet sakkunskapens roll. Högklassig utvärdering förutsätter att landets främsta experter utnyttjas för utvärderingsarbetet. För de integrerade projekten medför detta att det i utvärderingsgruppen måste finnas personer med kännedom om den svenskspråkiga utbildningskontexten. Däremot kan det beroende på tema vara svårt att hitta oberoende finlandssvenska experter på vissa objekt som utvärderas. För de specifikt svenska projekten kan kvaliteten på utvärderingen höjas genom att utomstående experter kompletterar den interna sakkunskapen.

Effektfull utvärdering förutsätter att alla de grupper inom utbildningen som behöver utvärderingsresultaten för styrning av utvecklingsarbetet är med och planerar utvärderingar. I planeringsgrupperna bör så långt som möjligt ingå representanter för alla intressenter i utbildningen. Eftersom planeringsgrupperna inte kan göras hur stora som helst, är det ett pusselarbete att få alla berörda parter företrädta. I integrerade utvärderingsprojekt kommer den svenska representationen i planeringsgruppen att vara liten – oftast endast en representant. För att främja effektfull utvärdering bör denna representant ha en så bred förankring som möjligt i den svenska utbildningen. Dessutom behövs det såväl i planeringsskedet som i det skede då de nationella resultaten kan nyttiggöras en fungerande intern finlandssvensk debatt.

Tillämpningen av utvärderingsresultaten främjas också av att uppläggningsen av utvärderingen utarbetas i interaktiv riktning. Genom deltagande utvärdering i form av samtal stimuleras utbytet av tankar och utformningen av en gemensam referensram för arbetet (Robinson & Cousins 2004). Utvärderingen har sålunda effekt redan under arbetets gång.

Rapporteringen av utvärderingsresultaten är en annan kritisk faktor med tanke på nyttiggörandet. Då en utvärdering berör olika grupper inom utbildningen är det till fördel om resultaten kan formuleras skilt för de olika grupperna så att varje grupps utvecklingsansvar beaktas. Detta innebär skilda rapporteringar med delvis varierande innehåll. Också formen kan variera: skriftliga rapporter, seminarier, diskussioner på nätet. Detta i sin tur innebär en omfördelning av utvärderingsresurserna så att betydligt mer tid och ekonomiska resurser än tidigare reserveras för rapportering och spridning av resultaten. I integrerade utvärderingsprojekt är det viktigt att resultaten sprids effektivt också till stöd för utvecklingen av den svenskspråkiga skolan på olika plan.

En ny utvärderingskultur med nätverk som bas

En av de viktigaste principerna för rådets arbete är nätverksfunktionen. Till rådets uppgifter hör att organisera externa utvärderingar med hjälp av nätverket av utvärderingsaktörer och att främja samarbete rörande utvärdering inom olika områden. Idén är att förhindra att knappa resurser slösas bort på överlappande funktioner.

För att utvärderingsnätverken skall fungera krävs en ny utvärderingskultur, präglad av öppenhet och samarbete mellan de inblandade. Med öppenhet avses ett fungerande informationsutbyte, så att alla aktörer känner till varandras verksamhet och planer. Likaså behövs beredskap att diskutera även kontroversiella frågor. Det gäller att skapa förutsättningar för och vilja till samarbete såväl enheter och organisationer som sakkunniga emellan.

För Svenskfinlands del innebär detta ett utökat samarbete i utvärderingsfrågor både internt och över språkgränsen. Samarbetet kommer att beröra specifika projekt, men också utbildning i och utveckling av själva utvärderingsmetodiken, dvs. av förmågan

att planera och genomföra utvärderingsprojekt. Därmed kommer också anknytningen av utvärderingsverksamheten till universiteten att förstärkas.

Till den nya utvärderingskulturen hör också att bygga broar mellan nationell och lokal utvärdering för att effektivera utvärderingen och den konkreta tillämpningen av resultaten. Detta kräver en kontinuerlig dialog mellan olika parter på olika nivåer inom utbildningen. Modeller utvecklade på finskt håll kan utnyttjas, men för att modellerna skall fungera krävs det att de anpassas till de egna utbildningsförhållandena.

Utvärderingsprocesser

Utvärderingsrådet har som uppgift att utveckla modeller för utvärdering. I rådets första utvärderingar kan processen som helhet beskrivas med hjälp av tre faser.

Den första fasen utgörs av *valet av utvärderingsobjekt*. Detta är en utbildningspolitisk fråga av stor principiell betydelse. Vilket eller vilka objekt skall väljas? Enligt lagstiftningen skall rådet göra upp ett utvärderingsprogram för sin mandatperiod. Detta har rådet nu gjort och i huvudsak baserat sig på undervisningsministeriets utvecklingsplan för utbildning och forskning samt på de riktlinjer undervisningsministeriet gett rådet och slutligen på en egen analys av verksamhetsmiljön. Centrala organisationer i nätverket fick i uppläggningsskedet ge sitt utlåtande om utvärderingsprogrammet. Nätverkets roll accentueras även av det faktum att rådet i sig består av representanter för olika intressentgrupper inom utbildningen som på sätt och vis själva är nyckelpersoner i utvärderingsnätverket.

Valet av utvärderingsobjekt för rådets fyraåriga mandatperiod är alltså baserat på en konsensusinriktad förhandlingsprocess med många parter. Utvärderingspolitiken har synliggjorts och öppnats för påverkan. Då valet av utvärderingsobjekt fastställs i avtalsförhandlingar mellan undervisningsministeriet och enskilda utvärderande instanser är det svårare att få insyn i valet av utvärderingsobjekt.

Den andra fasen utgörs av *planering*. Planeringskedet innebär att rama in och beskriva utvärderingsobjekt. Hur väl en utvärdering lyckas beror mycket på hur träffsäkert valda objekt kan beskrivas och hur relevanta de valda sidorna av ett objekt är (Huotari 2000). I denna fas mobiliseras nätverksarbetet. Olika modeller kommer att prövas, men i de första utvärderingsprojekten har rådet valt att tillämpa den modell som Rådet för utvärdering av högskolorna ofta utnyttjar. För de enskilda projekten har rådet tillsatt planeringsgrupper som så brett som möjligt representerar de grupper inom utbildningen som berörs av utvärderingens tema. Planeringsgruppen gör upp en detaljerad plan, där begrepp definieras, relevanta frågor utkristalliseras, indikatorer utarbetas och metoderna slås fast. Det är dock rådet som i sista hand godkänner planen.

Den tredje fasen utgörs av själva genomförandet av *utvärderingen*. Uppgiften är att från individuella värdeomdömen komma fram till en gemensam värdekonklusion (Huotari 2000, 217). En förutsättning för detta är en ingående diskussion och argu-

mentation kring valda värden (House 1995, 31). Utvärderingen sköts av en utvärderingsgrupp bestående av sakkunniga med substanskännedom men som är oavhängiga, dvs. icke-ansvariga för utvecklingen av utvärderingsobjektet. Planeringsgruppen föreslår medlemmarna i utvärderingsgruppen, men rådet utnämner gruppen. Utvärderingsgruppen ansvarar för genomförandet av utvärderingen och för slutsatserna och rapporteringen av den. Eftersom utvärdering är en starkt värdebunden verksamhet och knuten till valda perspektiv, räcker det inte med att enbart förlita sig på utomstående experters observationer och omdömen. I den nationella utvärderingsprocessen går man även in för att beakta den självutvärdering som görs inom de organisationer som utvärderingen berör.

Specifikt svenska behov

Den svenskspråkiga utbildningen utvecklas såsom framhållits parallellt med den finskspråkiga. Undervisningsministeriets plan för utveckling av utbildning och forskning styr i lika utsträckning den svenskspråkiga utbildningen som den finskspråkiga (med undantag för den åländska). Merparten av utvärderingsobjekten är därför de samma. Det finns dock frågeställningar som inte berör den finskspråkiga utbildningen på samma sätt som den svenskspråkiga. I dessa fall är det alltså befogat med utvärdering av enbart den svenskspråkiga utbildningen.

För den svenskspråkiga utbildningen har språket och därmed förknippade fenomen relativt sett en större betydelse. En språklig minoritet har i allmänhet en starkt utvecklad medvetenhet om språkets betydelse inom utbildning och kultur. Den svenskspråkiga utbildningen strävar ytterst efter att garantera finlandssvenskarnas och finlandssvenskans ställning och att förverkliga en samlad finlandssvensk utbildningsstrategi. Traditionellt sett har språkfrågan i utvärderingssammanhang närmast fokuserats på inlärningsresultaten, dvs. på hur inläringen av modersmålet framskrider och hur tvåspråkigheten utvecklas. Språket inverkar dock på många andra plan och andra fenomen inom utbildningen. I rådets program för perioden 2004-2007 ingår därför planer på en mer övergripande utvärdering av den svenskspråkiga utbildningens förutsättningar och villkor med hänsyn till språksituationen – en utvärdering som beaktar den interna språkliga heterogeniteten.

Denna övergripande utvärdering, inriktad på alla stadier från den grundläggande utbildningen till vuxenutbildningen, samlar – om den kan förverkligas – nyckelpersonerna inom svenskspråkig utbildning till samarbete kring planering av projektet eller kring själva utvärderingen. Som sidoeffekt kan då en fördjupad kunskap och förståelse för det nya utvärderingssystemet skapas. Samtidigt samlar den svenskspråkiga utbildningens sakkunniga nyttig erfarenhet som kan brukas både då de deltar i integrerade nationella projekt och i lokala utvärderingsprojekt.

Ett exempel på frågor som utgör ett betydligt större problem på svenskt håll än på finskt är lärarnas formella kompetens. Bristande kompetens bland lärarna har konstaterats i flera undersökningar (t.ex. Lithen & Salo 2000) och för närvarande görs satsningar för att åtgärda problemet. Efter något år är det skäl att genom en uppföljande utvärdering granska hur man lyckats i dessa utbildningspolitiska ansträngningar.

Utvärdering i utvecklingens tjänst

Utvärdering betraktas i dag som en nödvändig förutsättning för utvecklingsarbete samtidigt som dess begränsningar också måste beaktas. Verkligheten är komplex och den fulla sanningen om den framstår som en hägring. Sanningen måste det oaktat eftersträvas. Mångfasetterad och genomskinlig analys utgör viktiga redskap i sanningsökandet.

Utvärdering får inte i något skede bli ett självändamål. Den viktiga fråga som behöver ställas lyder (jfr Hamilton 2004): Håller vi rätt kurs? Vad rätt står för i detta sammanhang bestäms av de mål som formulerats och de värden målen är förankrade i. Det krävs mångsidigt deltagande och engagemang i planeringen och förverkligandet av utvärderingarna, för att dessa skall tjäna utvecklingsarbetet på alla plan. Nätverksutvärdering utgör ett sätt att engagera hela fältet.

Källor

- Broadfoot, P. & Black, P. 2003. Redefining assessment? The first ten years of Assessment in Education. *Assessment in Education: Principles, Policy and Practice II (I)*, 17–23.
- Cousins, J. B. & Earl, L.M. 1995. The case for participatory evaluation: Theory, research, practice. I J.B. Cousins & L.M. Earl, *Participatory evaluation in education*. London: Falmer.
- Evalueringsgruppe. 2004. <http://www.eva.dk/swwwwing/app/cm/Browse/.jsp?PAGE=1141>
- Hamilton, D. 2004. Att återinföra begreppet kvalitet i undervisningen. I K. Holmlund (red.) *Vad har kvalitet med skolan att göra?* Lund: Studentlitteratur, 127–145.
- Harlen, W. & Crick, R.D. 2003. Testing and motivation for learning. *Assessment of Education: Principles, Policy and Practice 10 (2)*, 169–207.
- Huotari, V. 2000. Asiantuntijuus arvioinnissa. *Aikuiskasvatus* 3, 215–219.
- House, E. 1995. Evaluoinnin futuurinen perfektii. I S. Takala (red.) *Arviointi ja koulutuksen laadun kehittäminen*. Jyväskylän yliopisto: Kasvatustieteiden tutkimuslaitos, 23–32.
- Karjalainen, A. & Huusko, M. 2004. *Arviointiajattelun pikaopas*. Oulun yliopisto: Opetuksen kehittämisyksikkö.
- Lithen, M. & Salo, P. 2000. *Lärarsituationen i de finlandssvenska grundskolorna och gymnasierna samt yrkesläroanstalterna*. Helsingfors: Utbildningsstyrelsen.
- Nikkanen, P. & Lyytinen, H.K. 1996. *Oppiva koulu ja itsearviointi*. Jyväskylän yliopisto: Koulutuksen tutkimuslaitos.
- Patton, M.Q. 1997. *Utilization-Focused Evaluation: The New Century Text*. London: SAGE Publications.



- Robinson, T.T. & Cousins, J.B. 2004. Internal participatory evaluation as an organizational learning system: A longitudinal case study. *Studies in Educational Evaluation* 30 (2), 1–22.
- Smith, M. & Glass, G. 1987. *Research and evaluation in education and the social sciences*. NJ: Englewood Cliffs, Prentice Hall.
- Stobart, G. 2003. Editorial. The impact of assessment: intended and unintended consequences. *Assessment of Education: Principles, Policy and Practice* 10 (2), 139–140.
- Svedlin, R. 2003. *Utvärdering i skolan – mellan skolutveckling och styrpolitik*. Helsingfors universitet. Pedagogiska institutionen. Forskningsrapport 185.
- Utbildningsutvärderingens nya inriktning 2004. *Utvärderingsprogram för perioden 2004–2007*. Publikationer från Rådet för utbildningsutvärdering 2. Jyväskylä.



Arviointitiedot hyötykäyttöön

Artikkelin tarkoituksena on avata näkökulmia koulutuksen arvioinnin hyödyntämiseen. Alkuosassa tarkastelen koulutuksen arviointijärjestelmää ja säädösten asettamia vaatimuksia, jotka koskevat arvioinnin hyödyntämistä. Arviointitutkimus laajentaa arvioinnin hyödyntämisen teemaa. Kysymys ei ole enää arviointitulosten ja -tiedon mekaanisesta käytöstä, vaan koko arviointiprosessin vaikutuksesta osallisten ajatteluun, vuorovaikutukseen ja toimintaan. Arvioinnin toteuttaminen ja menetelmälliset valinnat vaikuttavat siihen, miten arviointia käytetään ja hyödynnetään. Parhaimmillaan arviointi on luovasti ja joustavasti toteutettava prosessi, joka ottaa huomioon erilaiset käyttäjien tarpeet, toimintaympäristöt ja menetelmälliset vaihtoehdot. Hyötyvaikutukset näkyvät sekä päätöksenteossa että kannustuksena ja tukena kaikilla tasoilla tapahtuvassa kehittämisessä. Hyödyntämistä voidaan edistää muun muassa osallistamalla käyttäjät ja sidosryhmät arvioinnin eri vaiheisiin, linkittämällä arviointi osaksi johtamista ja strategiatyötä, analysoimalla ja tiedostamalla arvioinnin hyötyvaikutukset sekä soveltamalla arvioinnin tuloksia aktiivisesti käytäntöön. Merkittävä rooli on myös arvioinnin jatkuvalla seurannalla ja kehittämisellä, johon paneudun kirjoituksen loppuosassa.

Asiasanat: arviointitiedon hyödyntäminen, levittäminen, soveltaminen, hyväksikäyttö, vaikuttavuus, arviointiprosessi, kehittävä arviointi, realistinen arviointi, hyödyntämisen seuranta, meta-arviointi

Arviointijärjestelmä ja hyödyntämisen lähtökohdat

Arviointi tuli kaikkien koulumuotojen lakisääteiseksi tehtäväksi vuoden 1999 alusta. Tämä laajensi yhä enemmän paikallisen tason vastuuta oppimisen ja opetuksen laadusta. Säädösten valmistelun yhteydessä kohdistui paljon odotuksia arvioinnin kehittämistä tukeviin vaikutuksiin. Eduskunnan sivistysvaliokunta (3/1998) muun muassa painotti, että “kaikkein oleellisinta on arvioinnin tulosten määrätietoinen hyväksikäyttö koulutuksen kehittämisessä” ja että “arviointiin kytkeytyy läheisesti kokemusten ja erityisesti hyväksi todettujen käytäntöjen levittäminen kaikkien hyödynnettäväksi.” Muutamaa vuotta myöhemmin osoittautui, että odotuksista huolimatta tehdyllä arviointityöllä oli ollut hyvin vähän vaikutusta. Myös sen seuranta, mitä toimenpiteitä arviointitulosten perusteella oli toteutettu ja miten ne vaikuttivat koulutuksen järjestäjän ja koulun tasolla, puuttui. (Valtioneuvoston selonteko 2002, 42–44; Eduskunnan sivistysvaliokunta 11/2002.)

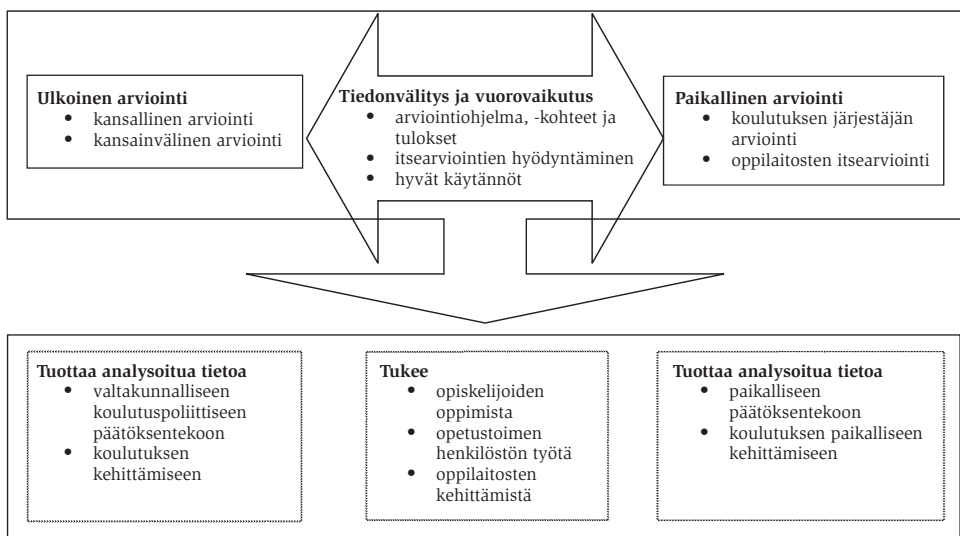
Lähtökohtaisesti arviointitoiminnan tulosten tulee olla hyödynnettävissä käytännön opetustyössä ja koulutuspoliittisessa päätöksenteossa. Arvioinnin avulla pyritään varmistamaan koulutuksen tasa-arvoinen ja yhdenmukainen kehittyminen ja opetuksen tavoitteiden saavuttaminen. Asetus koulutuksen arvioinnista (150/2003) korostaa arvioinnin asemaa koulutuksen kehittämisen välineenä. Sen mukaan arvioinnin tavoitteena on:

- hankkia ja analysoida tietoa valtakunnallisen koulutuspoliittisen päätöksenteon ja koulutuksen kehittämisen pohjaksi;
- hankkia ja analysoida tietoa koulutuksen paikallisen kehittämistyön ja päätöksenteon pohjaksi sekä
- tukea opiskelijoiden oppimista, opetustoimen henkilöstön työtä ja oppilaitosten kehittämistä.

Arviointiin perustuvan kehittämistoiminnan on oltava monentasoista ja -suuntaista. Koulutuksessa arvioinnin kokonaisuus muodostuu kansallisen tason arvioinnista sekä koulutuksen järjestäjien ja oppilaitosten suorittamasta paikallisen tason arvioinnista, johon kuuluu mukaan itsearviointi ja osallistuminen ulkopuoliseen arviointiin. Lisäksi kunnat tekevät kuntalain mukaista arviointia, lääninhallitusten toimesta tehdään alueellista peruspalvelujen arviointia ja eri tutkimuslaitosten toimesta opetushallinnon ulkopuolista koulutukseen kohdistuvaa arviointia ja tutkimusta. Eri tasoilla arviointitietoa voidaan tarvita esimerkiksi tilivelvollisuuden ja läpinäkyvyyden vaatimusten toteuttamiseen, johtamisen avuksi, toiminnan parantamiseen, taloudellisuuden seurantaan ja pedagogiseen kehittämiseen. Aiemmin koulutuksen säädös-, normi-, rahoitus- ja informaatio-ohjaus on ollut pääosin ylhäältä alas suuntautuvaa. Sen sijaan arvioinnissa pyrkimys on vuoropuheluun, joka tukee paikallista itsehallintoa ja -ohjautuvuutta. Paikallistasolla seurattavat tavoitteet ja kriteerit ovat koulutuksen järjestäjän päätettävissä. Arvioinnin yleisen viitekehyksen muodostaa lainsäädäntö, valta-

kunnalliset kehittämissuunnitelmat, opetussuunnitelmien ja tutkintojen perusteet, arviointineuvoston arviointiohjelma sekä paikalliset kehittämis- ja opetussuunnitelmat.

Koulutusjärjestelmän ja päätöksenteon monitasoisuus asettaa arvioinnin hyödyntämiselle suuria vaatimuksia. Hyödyntämistä edesauttaa arvioinnin, kehittämistoimien ja päätöksenteon koordinoitu suunnittelu ja tavoitteiden määrittely. Suunnitelmallisen ja määrätietoisen toiminnan vaatimus sisältyy myös arviointia ohjaavaan lainsäädäntöön. Haasteena on saada järjestelmä toimimaan kokonaisuutena siten, että arviointitiedon ja -tulosten välittäminen ja hyödyntäminen useaan eri suuntaan on mahdollista. (Ks. kuvio 1.)



Kuvio 1. Arviointijärjestelmä ja hyödyntämisen lähtökohdat (ks. Asetus koulutuksen arvioinnista 150/2003.)

Kansallisen arviointitiedon hyödyntäminen paikallisen päätöksenteon tarpeisiin edellyttää tietoa valtakunnallisesta arviointiohjelmasta, -kohteista ja tulosten käytöstä. Vastavuoroisesti järjestelmän tulee mahdollistaa paikallisen arviointitiedon kerääminen ja laajamittainen hyödyntäminen kansallisessa arvioinnissa. Arviointikokonaisuuden hahmottaminen ja kehittämistarpeiden jäsentäminen edellyttää asiantuntemusta paikallistasolla. Vastuuhenkilöiden olisi kyettävä johtamaan sekä arviointitiedon tuottamista että kansallisesti tuotetun tiedon tulkintaa (Eduskunnan sivistysvaliokunta 6/2002). Myös pedagogisen näkemyksen tulisi välittyä päätöksentekoon ja suunnitteluun. Jos tähän ei kyetä, on vaarana, että arviointi päinvastaisesta pyrkimyksestä huolimatta lisää oppilaiden, oppilaitosten, koulutuksen järjestäjien ja alueiden eriarvoistumista. Koulutuksen arviointia koskevan asetuksen valmistelun yhteydessä edelly-

tettiin toimenpiteitä paikallisen arvioinnin tukemiseksi ja kehittämiseksi (Eduskunnan sivistysvaliokunta 11/2002). Lakimuutoksia voidaan kuitenkin pitää varovaisina konkreettisten toimenpiteiden osalta.

Julkisten palvelujen arvioinnin arvopohjan muodostavat kansalaisten perusoikeudet, yhdenvertaisuus ja tasa-arvo sekä demokratia (Kohti julkisten palvelujen yhteistä arviointia 2001, 10). Koulutuksen avulla pyritään nostamaan kansalaisten sivistystasoa ja osaamista sekä lisäämään hyvinvointia. Arviointi on väline koulutuksen laadun kehittämiseksi ja tärkeä osa toimivaa koulutusjärjestelmää. Koulutuksen laatua on puolestaan määritelty monin eri tavoin, muun muassa kustannustehokkuuden, tarkoituksenmukaisuuden, tuloksellisuuden ja muutoskykyisyyden näkökulmista (Korkeakoulutuksen laadunvarmistus 2004, 49; Oksanen 2003, 75–87). Tässä yhteydessä koulutuksen laadulla tarkoitetaan lähinnä kykyä vastata koulutukselle asetettuihin tavoitteisiin ja asiakkaan tarpeisiin (Räisänen & Korpi 2002, 73). Toiminnan lähtökohtia ovat muun muassa asiakastyytyväisyys, palvelujen saatavuus ja palvelukyky. Hallinto- ja virkamiespainotteisuuden sijaan pyrkimyksenä on kansanvaltaisuuden lisääminen. Tämä edellyttää erilaisten arviointikäytäntöjen yhteistä kehittämistä eri käyttötarkoituksiin siten, että koulutuspalvelujen käyttäjien, koulutuksen järjestäjien sekä eri sidosryhmienkin näkökulmat ja käyttötarpeet otetaan huomioon.

Arvioinnin hyödyntämisen tavat

Arviointitiedon hyödyntämisen viitekehys pohjautuu suurelta osin tutkimustiedon hyödyntämisen tutkimukseen. Tällöin käsite hyödyntäminen (utilization, evaluation use) ymmärretään päämäärätietoisena tieteellisen tutkimuksen käyttönä toimintapolitiikkojen muotoilussa ja yhteiskunnallisissa päätöksenteossa. Hyödyntämisen lähi-käsitteitä ovat levittäminen (disseminaatio), leviäminen (diffuusio), soveltaminen (application) sekä hyväksikäyttö (exploitation) (Lampinen 1992, 5, 29). Englanninkielisessä arviointikirjallisuudessa käytetään myös käsitettä evaluation influence, kun korostetaan arvioinnin moninaisia ja usein vaikeasti ennakoitavia vaikutuksia (Alkin 2005, 436). Yleisesti arvioinnilla tarkoitetaan julkilausuttujen kriteerien pohjalta tapahtuvaa ilmiöiden ja toiminnan arvon määrittämistä, suhteuttamista arviointiperusteeseen (esim. House 1993, 4; House & Howe 1999). Arvottaminen erottaa arvioinnin ja arviointitutkimuksen muusta tutkimuksesta kytkien sen tiiviimmin tietoperustaiseen päätöksentekoon. Niin ikään tulosten yleistettävyyden on arvioinnin alueella rajatumpaa kuin tieteellisessä tutkimuksessa. Kuten Raivola (2000a, 65–66) toteaa, “arvioinnin tehtävänä on tuottaa systemaattista informaatiota, johon liitetään arvo- ja hyötypäätelmiä niin, että tietoa voidaan käyttää ohjaamaan jotakin sosiaalista toimintaa pääsemään entistä paremmin ja tehokkaammin sille asetettuihin tavoitteisiin.” Arviointitiedon hyödyntäminen on osa koulutuksen arvioinnin vaikuttavuutta. Koulutuksessa vaikuttavuudella (effectiveness) tarkoitetaan useimmiten toiminnalla tavoitelta-

vaa, myönteistä koulutuksen onnistumista ja sen tavoitteiden ja tehtävien täyttymistä (Raivola 2000b, 12).

Arviointitutkimuksessa vakiintuneen jaottelun mukaisesti arviointitiedon hyödyntäminen voi olla suoraviivaista (instrumental utilization), käsitteellistä ja näkökulmia avaavaa (conceptual utilization) tai aiempia päätöksiä epäsuorasti tukevaa ja legitimoivaa (persuasive utilization) (Rossi ym. 2004, 411; Lampinen 1992, 30–37; Leviton & Hughess 1981, 528–548). Chelimskyn (1997, 10–18) mukaan arvioinnilla voidaan pyrkiä kolmeen tarkoitukseen: tilivelvollisuuden täyttämiseen, uuden tiedon tuottamiseen tai kehittämistyön edistämiseen. Koulutuksen arvioinnilla on kaikki edellä mainitut tehtävät. Koulutusta arvioitaessa kyse on siitä, miten hyvin koulutusjärjestelmälle ja siihen liittyvälle toiminnalle asetetut tavoitteet saavutetaan. Arviointitiedon odotetaan tukevan johtopäätösten tekoa, kehittämistoimenpiteiden suunnittelua ja tavoitteiden asettamista luokkahuonetasolta aina kansalliselle tasolle saakka. Lähtökohdat ja tavoitteet arviointityöhön ovat erilaiset ja määräytyvät sen mukaan, millaista tietoa arvioinnilla tavoitellaan ja mitkä ovat arvioinnin motiivit. Tulosten hyödynnettävyyden kannalta on oleellista, että arviointi kykenee vastaamaan kehittämis- ja tiedontarpeisiin sekä kohdistuu ajankohtaisiin kysymyksiin. Tiedon on oltava käytettävissä silloin ja siinä muodossa kuin sitä kulloinkin tarvitaan (Rossi ym. 2004, 416). Arvioinnin toteuttajilta tämä edellyttää tietämystä nykytilanteesta ja näkemystä tulevaisuudesta.

Arviointia ja sen vaikutuksia voidaan tarkastella eri aikaulottuvuuksilla, sen mukaan keskittyykö se toiminnan aikaiseen (ex nunc), etukäteiseen (ex ante) vai jälkikäteiseen (ex post) arviointiin (Evaluating EU Expenditure Programmes 1997). Arviointi voi lisätä ymmärrystä nykyhetkestä ja edesauttaa muutoksen havaitsemista sekä siihen sopeutumista. Tämä on mahdollista, mikäli yhteys johtopäätöksistä päätöksiin ja konkreettisiin toimenpiteisiin on suora ja tapahtuu lyhyellä aikavälillä. Suoraviivainen arviointitiedon hyödyntäminen pohjautuu usein muuhun reaaliaikaiseen tietoperustaan, kuten tieto- ja tilastointijärjestelmiin. Laukkanen (2002, 267–268) varoittaa liian nopeasta arviointirytmistä, joka ei jätä riittävästi aikaa kehittämistoimien suunnitteluun. Arviointi saattaa saada selvityksen luonteen, mikä ei johda sisällölliseen keskusteluun, vaan pelkästään asioiden hallinnolliseen tarkasteluun. Toisaalta arviointi voi olla jatkuvaa toiminnan oikeansuuntaisuuden arviointia tai tuottaa pidemmän aikavälin kuluessa tietoa, joka toimii keskustelun herättäjänä ja sovittaa yhteen ihmisten ajattelutapoja sekä asenteita. Tällöin arviointi on pikemminkin eri sidosryhmiä osallistava neuvottelumuoto kuin totuuden etsimistä.

Arvioinnin avulla voidaan myös pyrkiä luomaan ja hallitsemaan tulevaisuutta. Tällöin sen toivotaan avaavan uusia näköaloja, herättävän yhteiskunnan muutoksiin ja ennakoiviin toimenpiteisiin. Vaikka tutkimuksen piirissä etukäteisarvioinnin (ex ante) merkitystä kritisoidaan, tulevaisuuden näkökulma ja ennakointi on päätöksenteon kannalta keskeistä. Etukäteisarvioinnilla voidaan luoda edellytyksiä toiminnan suunnitteluun ja päätöksentekoon, kun se kytketään osaksi strategiatyöskentelyä, tu-

losohjausta ja -seuranta sekä budjettiprosessia. Tässä mielessä evaluointia esimerkiksi strategisen suunnittelun tukena ja hallinnon läpinäkyvyyden edistäjänä ei ole vielä riittävästi hyödynnetty. Etukäteisarviointia tai siihen verrattavaa toimintaa on kuitenkin tehty jossain määrin, muun muassa koulutusta koskevien hanke-arviointien, erilaisten koulukokeilujen ja pilottihankkeitten yhteydessä sekä lakien valmisteluun liittyen. (Etukäteisarviointi Suomessa ja EU-maissa 2000, 21, 29, 44.)

Useimmiten arviointitulosten hyödyntäminen suuntautuu kuitenkin ajallisesti taaksepäin (ex post). Menneisyyteen keskittyvällä arvioinnilla voidaan pyrkiä osoittamaan, kuinka hyvin järjestelmälle asetetut tavoitteet on pidemmällä aikavälillä saavutettu. Lisäksi voidaan pyrkiä kiinnittämään huomiota laajemmin toiminnan vaikutuksiin sekä syy- ja seuraussuhteisiin. Kyse voi olla esimerkiksi aiemmin tehtyjen suunnitelmien tai toteutettujen ohjelmien arvioinneista. Tämänkaltaisen retrospektiivisen arvioinnin motiivit liittyvät usein tulosohjaukseen ja tehokkuusajatteluun sekä poliittisiin päämääriin. Arviointi, joka toteutetaan täyttämään tilivelvollisuuden vaatimukset tai toiminnan oikeuttamiseksi (legitimative utilization), palvelee pääasiassa arvioinnin teettäjän tarpeita, eikä tarvetta arvioinnin laajaan hyödyntämiseen ole. Chelmsky (1997, 15–16) toteaaakin, että arviointeja hyödynnetään usein poliittishallinnollisessa päätöksenteossa ja kontrollin välineenä, piittaamatta juurikaan arvioinnin ulkopuolisista tiedon hyödyntäjistä. Tällöin arviointi saattaa jäädä vain kaavamaiseksi tapahtumien kirjaamiseksi, saada hallinnollisen valvonnan luonteen tai jäädä symboliseksi (symbolic use), osoitukseksi tehdyistä toimenpiteistä (Alkin 2005, 435).

Arviointitieto päätöksenteossa

Arviointi- ja tutkimustiedon hyödyntämistä poliittis-hallinnollisessa päätöksenteossa on tutkittu runsaasti erityisesti Yhdysvalloissa. 1970-luvun empiiriset tutkimukset osoittivat suunnitelmallisen ja välineellisen hyödyntämisen olevan harvinaista päätösten perustana (Cook 1997, 40–41). Huolimatta siitä, että tavoitteena on toimintojen järjeistäminen arvioinnin avulla, päätöksentekoa ohjaavat yhteisen harkinnan ja analyttisen prosessin lisäksi paljolti politiikka ja valtasuhteet (Khakee 1998, 364). Vaihtoehtoisia tapoja toimia ovat esimerkiksi pitäytyminen hallinnollisissa rutiineissa ja tavoissa, mallien kopiointi, sopeutuminen ympärillä tehtäviin päätöksiin ja improvisointi (Weiss 1982, 26). Arvioinnista saatetaan myös poimia irrallisia argumentteja vahvistamaan haluttuja poliittisia tai henkilökohtaisia pyrkimyksiä (Vedung 1997, 275–276). Ongelmana on myös, että vaikka arviointitietoa kysytään yhä enemmän, sen kysyjät ja hyödyntäjät eivät välttämättä luota tieteellisiin vastauksiin (Rajavaara 1999, 39).

Koulutuksen kansallisilla arvioinneilla pyritään saamaan vastauksia poliittisiin, taloudellisiin ja sosiokulttuurisiin ongelmiin sekä määrittelemään, miten yhteiskunta on pystynyt vastaamaan julkisille palveluille asetettuihin tavoitteisiin ja vaatimuksiin

(Oksanen 2003, 22). Kansallinen arviointi on koostunut järjestelmäarvioinneista, tila- ja tema-arvioinneista, oppimistulosten arvioinneista ja koulutusindikaattoreista (Koulutuksen tuloksellisuuden arviointimalli 1998, 19–21). Tietoa kansallisten arviointien hyödyntämisestä ja niiden vaikutuksista on varsin niukasti. Kuten Temmes ym. (2002, 91) toteavat, “suuri ongelma maamme koulutusjärjestelmän hallinnon arvioinnin kannalta on epätietoisuus siitä, mitä tuloksia arvioinneista on ollut, tulisi olla ja voisi olla, mikä on arvioinnin suhde muuhun informaatio-ohjaukseen, strategiseen tavoitteiden asettamiseen ja policyanalyysiin ja kuinka arviointia itseään tulisi arvioida ja kehittää.” Tuoreen Tiede- ja yhteiskuntatyöryhmän muistion (2004, 28) mukaan myös tutkimustieto välittyy Suomessa riittämättömästi päätöksentekoon, tai käytettävissä ei ole riittävästi riippumatonta tietoa päätöksenteon tueksi. Lumijärvi ja Kaarlejärvi (2000, 33) ovat puolestaan todenneet, että opetusministeriön hankkimalla arviointitiedolla on heikko yhteys hallinnonalan tulosohtaukseen, poliittiseen päätöksentekoon ja päivittäisjohtamiseen. Arvioinnin näkökulma on painottunut tilivelvollisyyteen ja tiedontuotantoon, jolloin kehittäminen ja hyödyntäminen eivät ole olleet keskeisinä lähtökohtina. Myös sen seuranta, miten tehdyt arvioinnit vaikuttavat arviointien toimintaan ja koulutuksen kehittämiseen, on puuttunut (Eduskunnan sivistysvaliokunta 5/2001; 11/2002.) Jakku-Sihvosen ja Heinosen (2001, 145) mukaan kansallisia ja kansainvälisiä arviointeja on hyödynnetty poliittisessa päätöksenteossa, päätetäessä lisätutkimus- tai selvitysehdoituksista ja toimenpiteistä sekä käytännön opetustyössä. Niemen (2004, 194) mukaan perusopetuksen oppimistulosten arviointien osalta hyödyntäminen on ollut kaksijakoista. Arviointitietoa hyödynnetään keskushallinnossa ja opettajankoulutuksessa mutta hyöty käytännön opetustyölle nähdään kuitenkin vähäiseksi. Rinteen ym. (2004, 43–48) mukaan OECD:n tuottamat maa- ja teematutkinnat ovat vaikuttaneet koulutuspoliittisiin toimintalinjauksiin ja valintoihin, koululainsäädäntöön ja uudistuksiin sekä voimavarojen kohdentamiseen. Ne ovat myös kiinnittäneet huomiota koulutukseen arviointiin ja laatuun.

Paikallinen arviointi perustuu koulutuksen järjestäjän pitkäaikaissuunnitelmiin ja valtakunnallisiin linjauksiin. Sen tehtävänä on tuottaa tietoa, miten hyvin koulutuksen järjestäjä on onnistunut tavoitteiden saavuttamisessa ja tehtäviensä hoitamisessa. Yhdessä oppilaitosarvioinnin kanssa se voi merkittävästi tukea oppilaitosten kehittämistä ja luoda parempia edellytyksiä oppimiselle. Arviointitietoa käsitellään yleensä toimialasta vastaavassa lautakunnassa, kuntayhtymän hallituksessa, kunnanhallituksessa ja -valtuustossa. Tietoa hyödynnetään esimerkiksi sivistys- ja koulutuspoliittisten ohjelmien laadinnassa, talous- ja toimintasuunnitelmien käsittelyn yhteydessä sekä paikallisia opetussuunnitelmia tarkennettaessa. Paikallisen arvioinnin tilasta ei ole käytettävissä ajankohtaista tutkimustietoa. Yleisesti voidaan kuitenkin todeta, että paikallistason arviointistrategiat ovat vielä keskeneräisiä, ja sen myötä arvioinnin asema päätöksenteossa vaihtelee runsaasti koulutuksenjärjestäjien kesken (Eduskunnan sivistysvaliokunta 2001, 2004; Rajanen 2000, 88–89). Järjestäjien ja oppilaitosten arviointeja ei kansallisissa arvioinneissa ole juurikaan kyetty hyödyntämään. Ar-

vioinnin toteuttamisesta ovat niin ikään puuttuneet verkostomaisen toiminnan mallit ja vuorovaikutuskanavat, jotka mahdollistaisivat arviointien ja arviointiosaamisen hyödyntämisen ja kehittämisen sekä kansallisella että paikallisella tasolla (Eduskunnan sivistysvaliokunta 2004).

Koulutuksen järjestäjä luo puitteet oppilaitosten toiminnan arvioinnille ja sen kehittämiselle. Oppilaitostasolla omien toimintatapojen ja -prosessien tunteminen on suunnitelmallisen kehittämisen lähtökohta ja kytkettävissä muun muassa oppivan organisaation viitekehykseen. Arviointi voidaan toteuttaa ulkoisena arviointina tai ulkoisen arvioinnin ja itsearvioinnin sisältävänä yhdistelmäarviointina. Itsearviointi ja ulkoinen arviointi ovat toisiaan täydentäviä arvioinnin näkökulmia. Itsearvioinnissa syntynyttä tietoa voidaan hyödyntää ulkopuolisessa arvioinnissa, vastaavasti ulkoista arviointia voidaan perustella itsearvioinnin tukena (Lyytinen 2001, 7). EU:n (Euroopan parlamentin ja neuvoston suositus 2000, 7) toimesta korostetaan, että ulkoinen arviointi ei saa rajoittua pelkkään hallinnolliseen valvontaan, vaan sen on kannustettava oppilaitoksia jatkuvaan kehitysprosessiin. Ulkoisella arvioinnilla voidaan edistää muutosprosessien käynnistämistä lisäämällä tietoisuutta omista ja toisten toimintatavoista. Ulkopuolinen arviointi voi myös ohjata koulutuksen järjestäjiä suuntaamaan itsearviointia niille koulutuksen osa-alueille, jotka kulloinkin ovat valtakunnallisen arvioinnin kohteena. Näin mahdollisesti löytyvät selittävät tekijät täydentävät arvioinnin antamaa kokonaiskuvaa. Itsearviointi puolestaan voi edesauttaa tunnistamaan realiteetit ja madaltaa kynnyksiä mahdollisiin muutoksiin ja toimenpiteisiin. Korkeakoulujen arvioinneista saatujen kokemusten mukaan itsearviointi yhdistettynä ulkoiseen arviointiin on usein merkittävä impulssi kehittämistyölle (Hämäläinen 2002, 7–9).

Korkeakosken (2000, 143) mukaan arviointijärjestelmän kehittäminen oppilaitostasolla on toiminut tehokkaana sisäisen koulutautumisen välineenä, tehostanut yhteistä oppimista ja tiimityöskentelyä sekä luonut edellytyksiä arvioinnin kehitystyölle. Oppilaitostasolla arviointia tulee edelleen kehittää siten, että se parhaalla mahdollisella tavalla tukee koulun johtamista, opetusta ja oppimista. Keskeistä on, että arviointityö on käytännönläheistä, ja siitä on toiminnalle osoitettavissa olevaa konkreettista hyötyä. Vaikutukset voivat näkyä esimerkiksi opetuksen muuttuneissa käytännöissä ja järjestelyissä. Olisikin tärkeää saada jaettua kokemuksia ja tietoa, miten oppilaitoksissa hyödyntäminen ja kehitysprosessit etenevät, missä asioissa onnistutaan, mitkä ovat kehittämistyön organisoimisen heikkoudet ja vahvuudet ja miten kehittämistyössä edelleen kehitytään. Tällaisina oppilaitosympäristöön soveltuvia tiedonkeruun menetelminä voisivat toimia esimerkiksi oppimispäiväkirja tai yhteisöportfolio. Portfolion käytöstä on kokemuksia opettajankoulutuksessa muun muassa oppimisen, itsearvioinnin ja ammatillisen identiteettityön välineenä (Heikkinen 2002, 117–118).

Kehittävä ja käyttäjälähtöinen arviointi

Aiemmin positivismia ihannoiva arviointi pohjautui kokeellisiin ja kvantitatiivisiin tutkimusasetelmiin, ja tiedon oletettiin olevan suoraviivaisesti hyödynnettävissä päätöksenteossa. Arviointia käytettiin usein ylhäältä alas suuntautuvan johtamisen välineenä, auktorisoimaan johdon asemaa ja päätöksiä. Kehitys on kuitenkin kulkenut kohti pragmaattista ja konstruktivistista arviointinäkemystä. Arviointi nähdään tällöin yhtenä vuorovaikutuksellisenä oppimisprosessina, jossa hyödyntäjät eivät ole vain passiivisena osapuolena (Cook 1997, 40–41). Arviointitiedolle, havainnoille ja johtopäätöksille annetaan yksilö- ja yhteisötasolla omaan kokemuspiiriin suhteutettuja selityksiä ja tulkintoja. Khakeen (1998, 361) mukaan ajatus hyvästä arvioinnista on enemmänkin keskustelemaa kuin pyrkimystä objektiivisuuteen. Arvioinnissa korostuvat myös tavoitteellisuus, käytännöllisyys ja hyödyllisyys. Kun arviointi on siirtynyt yhä enemmän toimijoitten itsensä tehtäväksi, he myös kantavat vastuun tiedon hyödyntämisestä (Rajavaara 1999, 52).

Patton (1997, 20–22, 104–105) käyttää käsitteitä kehittävä arviointi (developmental evaluation) ja käyttäjälähtöinen arviointi (utilization-focused evaluation). Näkemyksissä korostuu arvioinnin prosessimaisuus, itsearviointi ja eri osapuolten osallistuminen arviointiin. Arviointitiedon hyödyntäminen ja käyttäjälähtöisyys kulkevat käsi kädessä. Pattonin mukaan arviointi on oppimisen väline silloin, kun se vastaa yksilön tai organisaation tiedon tarpeisiin ja mahdollistaa oman toiminnan reflektoinnin. Reflektiivisyyteen kuuluu omien arvojen olettamusten ja uskomusten tutkimista ja niiden suhteuttamista toisten käsityksiin. Oppiminen ja poisoppiminen tapahtuvat näin lisääntyneen tietämyksen ja toiminnan muutoksen kautta. Sitoutuminen ja osallisuus arviointiprosessin suunnitteluun ja toteutukseen, mahdollisuuksien mukaan myös päätöksentekoon, motivoi hyödyntämään tuloksia ja toteuttamaan kehittämistoimenpiteitä. Sitoutumista edistäviin, sisäisiin motiiveihin, kuuluu tarve ja halu kehittää omaa työtä ja auttaa organisaation työyhteisöä oppimaan.

Arviointi on itsereflektion ja kollegiaalisen reflektion väline. Käyttäjälähtöisyys arvioinnissa edellyttää tämän merkityksen oivaltamista sekä kykyä uudistua, keskustella ja ratkoa asioita yhdessä. Arvioinnin onnistumisen ja hyödyntämisen kannalta on tärkeää, että kaikki osalliset hyväksyvät arviointitehtävän ja ovat aktiivisesti sitä toteuttamassa. Huomio kiinnittyy tällöin arviointiprosessin ja arviointitiedon hyödyllisyyteen yksittäisen käyttäjän näkökulmasta. Lähtökohtana voidaan pitää kysymystä “mitä tekisit, jos sinulla olisi vastaukset kysymyksiin, joita esität?” (Patton 2002, 174). Analyttistä tietoa voidaan arvioida yksilötasolla sen uskottavuuden ja luotettavuuden mukaan, mutta myös sen mukaan, miten se vastaa senhetkisiin tiedontarpeisiin. Hyödyllisyyden arviointi pakottaa yksilöä reflektoimaan myös omaa työtään, toiminnan edellytyksiä, menettelytapoja, tuloksia ja suuntaa. (Rossi ym. 2004, 415; Weiss & Bucuvalas 1980.)

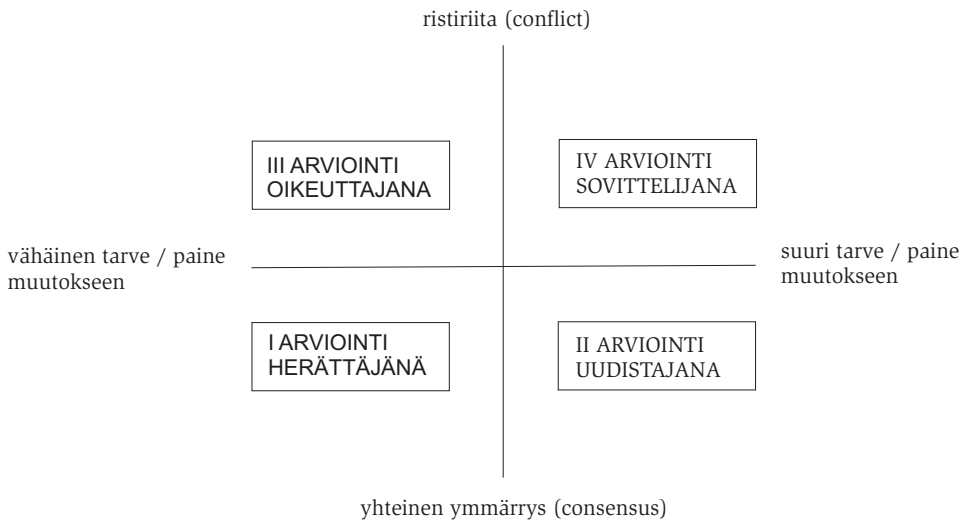
Koulutuksen alueella toiminnan arviointi ja laadunvalvonta on jäänyt pääasiassa hallintovirkamiesten ja oppilaitosten rehtoreiden tehtäväksi. Arvioinnin kautta tapahtuva oppimiskulttuurin kehittäminen edellyttääkin opettajilta ja koulun sidosryhmiltä uudenlaista ajattelua ja vastuunottoa. Se edellyttää myös kykyä oman työn ja toiminnan arviointiin, jolloin oppiminen ei tapahdu vastausten vaan kysymysten kautta. Arvioinnin tehtävä on herättää kysymyksiä ja kyseenalaistaa asioita. Tässä tullaan lähelle tutkivan oppimisen mallia, joka korostaa yksilön ja yhteisön tiedon rakentamista sekä tutkimuksen ja hyödyntämisen tasapainoa. Hakkaraisen (2004, 298, 352) mukaan tutkimukselle on tyypillistä etsiä, luoda, kehittää ja keksiä uusia ajatuksia ja innovaatioita. Hyödyntäminen vastaavasti liittyy siihen, kuinka ihminen täsmentää, soveltaa ja käyttää tietoaan ja muodostaa sen varassa rutiineja. Kokemuksen muuttuminen rutiineiksi vapauttaa voimavaroja uusien haasteiden kohtaamiseen ja tutkimiseen. Oksanen (2003, 243) korostaa, että jos arviointi on osa prosessia, jossa sitä käytetään välineenä koulun ongelmien kartoittamiseksi ja prosessia jatketaan siten, että arviointia seuraa kehittämissuunnitelma ja sen toimenpiteiden toteuttaminen, niin arvioinnin merkitys aukeaa kouluille. Vasta tätä kautta luodaan tartuntapintaa arviointitiedon hyödyntämiselle ja kehittämistoimenpiteiden jalostamiselle siten, että arvioinnin vaikutukset ulottuvat pedagogiikkaan ja oppimiseen asti.

Realistinen arviointi ja kontekstitekijät

Realistinen arviointi (realistic evaluation) edustaa uudempaa arviointitutkimuksen näkökulmaa. Se korostaa vaikutusyhteyksien selvittämistä teorian ja käytännön vuoropuhelun kautta. Syy- ja seuraussuhteet ovat sidoksissa olosuhteisiin ja toimintaympäristöön. Arvioijan on pyrittävä ymmärtämään konteksti ja vaikutusprosessit (mekanismit), jotka johtavat toivottuihin tai ei-toivottuihin tuloksiin. Peruskysymyksiä ovat: mikä toimii, kenen kohdalla ja millaisissa olosuhteissa? (Pawson & Tilley 2005, 362–367; 2000, 215–217; Tilley 2000, 4.) Arvioinnin hyödyntäminen on sidoksissa yksilön ja organisaation kykyyn ja tarpeeseen uudistua.

Valovirran (2000, 11–13) esittämässä mallissa (ks. Pettersson & Wallin 1995; Virtanen & Uusikylä, 2000) on tarkasteltu arvioinnin roolia ja kontekstitekijöitä kahdella jatkumolla, joiden ääripäissä ovat konflikti ja konsensus sekä vähäinen ja suuri muutostarve. (Ks. kuvio 2.) Yhteisössä, jossa vallitsee laaja yhteisymmärrys ja konsensus (I ja II) yhteisistä arvoista sekä tavoitteista, arvioinnilla on kehittävä rooli. Jos organisaation tarve muutokseen on pieni (I), arvioinnin kautta tarkistetaan toimintaa ja herätellään uudistuksiin, mikäli puutteita ilmenee. Kun tarve ja paine muutokseen on suuri (II), arviointi toimii tehokkaana liikkeelle panevana voimana ja uudistavan toiminnan käynnistäjänä. Ristiriitaisessa ympäristössä (III ja IV) arviointia käytetään usein kiistelyn ja valtataistelun välineenä vallitsevan toiminnan oikeuttamiseksi, ilman asioiden perusteellisempaa tarkastelua. Voimakkailla argumenteilla voidaan täl-

löin tukahduttaa kehittämään pyrkivä vuoropuhelu, eikä saatuja kokemuksia käytetä hyväksi toiminnan kehittämisessä. Konfliktisessa ympäristössä, jossa tarve muutokseen on suuri (IV), arvioinnille voi muodostua sovitteleva ja rauhaa rakentava rooli. Ongelmana on tällöin osapuolten sitouttaminen arvioinnin kautta tapahtuvaan kehittämiseen. (Valovirta 2000, 11–13.)



Kuvio 2. Arvioinnin rooli ja konteksti (Valovirta 2000, 11–13; ks. myös Pettersson & Wallin 1995; Virtanen & Uusikylä 2000)

Edellä esitettyssä mallissa toimintaympäristöä tarkasteltiin organisaatiotasolla. Näkökulmaa voi kuitenkin laajentaa koskemaan koko arviointijärjestelmää. Vaikka vastuu oppimisprosessin kautta tapahtuvasta jatkuvasta kehittämisestä on yhteinen, johdolla ja osallistavalla ihmisten johtamisella on tärkeä merkitys yhteisen tahtotilan, konsensuksen ja kehittämistyön edellytysten luojana. Osallistumista voidaan edistää esimerkiksi avoimuudella sekä rohkaisemisella vastuunottoon. Turvallinen ja kannustava ilmapiiri tukee myös uusien ideoiden esittämistä, vuorovaikutusta ja yhdistää osamista. Käytännössä kehittämishalukkuutta säätelevät monet emotionaaliset, normatiiviset ja yhteisölliset tekijät. Erityisesti toimintatapaan tai kulttuuriin liittyvien uudistusten läpivieminen on usein vaikeampaa kuin konkreettisten ideoiden toteuttaminen. Eteen tulevat ongelmat voivat liittyä esimerkiksi muutosvastarintaan, puutteellisiin resursseihin tai tukeen.

Arviointityön johtaminen ja suunnitelmallisuus

Patton korostaa, että evaluatiivisen ajattelun kehittymisellä voi olla kestävämpi vaikutus kuin yksittäisen arvioinnin tuloksilla (Patton 1997, 88). Tavoitteen selkiyttäminen ja paneutuminen siihen, miten tuloksia tullaan käyttämään jo suunnitteluvaiheessa, motivoi arviointiprosessiin ja madaltaa kynnystä mahdolliseen muutokseen. Sen kautta voidaan varmistua, että kaikki arvioinnin osapuolet ymmärtävät suunnittelun, toteutuksen, hyödyntämisen ja seurannan käsitteet sekä prosessin ja vastuut samalla tavalla. Samassa yhteydessä on myös määriteltävä keitä arviointitiedon hyödyntäjät ovat. Arvioinnin rajaaminen ja suhteuttaminen kustannuksiin, työmäärään, käytettävissä olevaan aikaan ja muihin voimavaroihin on myös tärkeää. Arvioinnin muodostuessa liian raskaaksi tai pitkäksi hyödyntäminen voi jäädä puolitiehen ja johtaa turhautumiseen arviointia ja kehittämistä kohtaan.

Patton (1997, 58) tuo esille seitsemän arvioinnin suunnitteluvaiheen “syntiä”. Hänen mukaansa arviointitietoa ei hyödynnetä muun muassa seuraavista syistä:

1. Arvioinnin tekijät, eivät niinkään tiedon hyödyntäjät, nähdään keskeisinä päätöksentekijöinä.
2. Hyödyntäjiä ei tunnisteta tai kyetä tarkemmin määrittelemään.
3. Hyödyntäjiksi tulkitaan organisaatiot, ei ihmiset.
4. Hyödyntäjät pyritään tunnistamaan vasta arviointiprosessin loppuvaiheessa, kun tieto on jo kerätty.
5. Ensisijaiseksi hyödyntäjäksi tulkitaan automaattisesti arvioinnin teettäjä tai rahoittaja.
6. Toiminnassa keskitytään päätöksiin, päätöksentekijät eivät ole selvillä prosessista ja sen eri vaiheista.
7. Arviointitiedon tuottajat eivät kohtaa hyödyntäjiä tasavertaisina, vaan asettavat itsensä näiden yläpuolelle.

Arviointitutkimus nostaa esille myös useita muita tiedon hyödyntämiseen vaikuttavia tekijöitä, jotka liittyvät arviointiprosessiin, tiedon käyttäjiin tai ympäröivään organisaatioon. Käytännön esteenä ovat oppilaitostasolla muun muassa kiire ja ajanpuute, tulosten ymmärrettävyys ja soveltamisen vaikeus sekä työyhteisön valtahierarkiaan liittyvät tekijät. Patton korostaa (1997, 48) käyttäjien arviointia kohtaan osoittaman kiinnostuksen, asennoitumisen sekä positiivisten kokemusten merkitystä. Tähän voidaan pyrkiä vaikuttamaan lisäämällä arviointiosaamista esimerkiksi täydennyskoulutuksen kautta. Samalla voidaan myös levittää positiivista arviointi- ja kehittämiskulttuuria ja luoda mahdollisuudet hyvien käytäntöjen välittämiseen.

Arviointitulosten ja johtopäätösten julkistaminen avoimesti ja selkeästi, sovittujen periaatteiden mukaisesti edesauttaa arvioinnin hyödyntämistä. Perusvaatimuksena on tulosten ymmärrettävyys ja uskottavuus. Käyttäjystävällinen ja kohdennettu arviointitieto, -ohjelmat, raportit ja lähdeaineistot lisäävät arviointien käyttömahdollisuuksia. Arvioitsijan tulisi myös olla selvillä tavoista, joilla päätöksentekijät muodostavat

käsityksensä (Rossi ym. 2004, 416). Arviointitutkimuksen piirissä on korostettu, että arviointitulokset eivät ole yksiselitteisiä, vaan edellyttävät aina yhteistä dialogia ja tulkintaa (Lindqvist 1999, 113). Kun vuoropuhelu kykenee yhdistämään sekä teoreettisen että käytännön asiantuntemuksen, kehittämistyön edellytykset paranevat. Yhdessä tapahtuva kehittäminen edellyttää muun muassa kykyä yhteistyöhön, laaja-alaiseen asioiden tarkasteluun ja rakentavaan kriittisyyteen. Toimijoitten, tutkimuksen, hallinnon, opettajankoulutuksen ja muiden sidosryhmien erilaiset näkökulmat monipuolistavat tulosten hyödyntämismahdollisuuksia. O'Sullivan (2004, 26) käyttää käsitettä yhteistoiminnallinen arviointi (collaborative evaluation). Erityisesti paikallistason arviointiin tämä viitekehys voi tuoda uusia toteutusvaihtoehtoja. Se pyrkii kokoamaan sidosryhmät yhteen arvioinnin eri vaiheissa sekä jakamaan vastuuta, koskien ei vain arviointia, vaan myös päätöksentekoa. Eri osapuolten tasavertainen kuuleminen mahdollistaa yhteisen kehittymisen. Myös erilaiset osallistavat vertaisarviointimenetelmät voivat tarjota vertailutietoa ja toimia keskustelun käynnistäjänä. Menetelmät antavat mahdollisuuden peilata omaa tekemistä toisten toimintaan. Samalla ne tukevat yhteistä oppimista ja levittävät tietoa hyvistä käytännöistä.

Koulutus voi kehittyä lisääntyneen tietämyksen, yhteisesti sovittujen tavoitteiden ja toiminnan muutoksen kautta. Pelkkä lisääntynyt tietämys kehittämistarpeista ja -mahdollisuuksista ei kuitenkaan takaa toimintaa. Jatkuvan kehittämisen kannalta ratkaisevaa on, mitä arvioinnin esiin nostamille asioille tapahtuu ja miten tulokset liitetään johtamisjärjestelmään, strategiseen suunnitteluun ja arkityöhön. Arvioinnin pohjalta tehtävät johtopäätökset ja kehittämisideat voivat olla nopeasti toteutettavia tai vastaavasti vaatia runsaasti resursseja ja tukea. Kaikkea ei kuitenkaan voi toteuttaa kerralla, vaan pidemmän aikavälin kuluessa ja suunnitelmallisena prosessina. Kehittämisen ohella on myös pyrittävä säilyttämään ja jalostamaan sitä, mikä on hyvää ja vastaavasti luopumaan siitä, mikä on tarpeetonta. Arvioinnin käyttöarvon kannalta on keskeistä, että sen esiintuomiin kehitys- ja tuentarpeisiin kyetään myös vastaamaan ja toimenpiteet saatetaan näkyviksi. Korkeakoski korostaa (2000, 150), että arviointitoiminnan organisoinnissa on olennaista selkeä pitkän aikavälin suunnitelmallisuus, joka alkaa tavoitteista ja niiden analyysistä ja johtaa suunnitelmallisen toiminnan kautta arviointiin ja sen tulosten analyysiin kehittämis ehdotuksineen sekä lopulta tarvittaessa kehittämisohjelmaan. Keskeisessä asemassa ovat päättäjät, kouluviranomaiset ja rehtorit sekä muut tahot, joilla on mahdollisuus muutosten eteenpäin viemiseksi. Omalla esimerkillä ja sitoutumisella he voivat edistää kehittämistoimenpiteiden etenemistä.

Hyödyntämisen seuranta ja meta-arviointi

Toiminnassa tapahtuneen muutoksen ja kehityksen säännönmukainen varmistaminen ja seuranta (follow up) edesauttaa arvioinnin hyödyntämistä. Seurannassa pyritään

analysoimaan, mitä toimenpiteitä arviointitulosten perusteella toteutetaan ja miten ne vaikuttavat. Ollakseen kattavaa, seurannan tulisi ulottua päätöksenteosta oppimisprosessiin asti. Seurannan tavoitteena on todentaa tapahtuneet muutokset ja tukea kehittämissuunnitelmien toteutusta. Kehittämistoimenpiteistä kantavat päävastuun siitä vastaavat viranomaistahot ja välillisesti opettajankoulutuslaitokset. Arvioinnin vaikuttavuuden lisäämiseksi on esitetty alueellisen tuki- ja seurantajärjestelmän luomista, jonka kautta kyettäisiin seuraamaan ja analysoimaan arviointitulosten perusteella toteuttavia toimenpiteitä koulutuksenjärjestäjän ja koulun tasolla (Eduskunnan sivistysvaliokunta 11/2002). Myös kansallisella tasolla on edelleen kehitettävä arviointien hyödyntämistä, vaikuttavuuden seurantaa ja analysointia sekä arvioinnin laadun kehittämistä. Jotta arviointi olisi jatkuvasti kehittyvää dynaamista toimintaa ja toimisi muutoksen välineenä, myös sitä tulee jatkuvasti arvioida ja kehittää. Laadukas ja uskottava, menetelmiltään hyväksyttävä arviointi tuottaa hyödynnettäviä ja luotettavia tuloksia. Yhdistelemällä eri menetelmiä, aineistoja tai lähestymistapoja voidaan lisätä arvioinnin luotettavuutta ja laajentaa hyödyntämismahdollisuuksia. Esimerkiksi ulkoinen arviointi voidaan kohdentaa itsearviointitulosten hyödyntämiseen, kehittämissuunnitelman toteutumiseen tai koko arviointijärjestelmän arvioimiseen. Laadun arvioinnin organisointimuotoina voidaan käyttää esimerkiksi vertaisryhmää (peer group) tai asiantuntijaryhmää (auditointi / expert group). Arviointitulosten hyödyntämisen edistämiseksi tarvitaan myös menetelmiä yhteisötason reflektointiin sekä toiminnan perusteiden, laadun ja tulosten pohtimiseen koko yhteisön kehittämistaroituksessa.

Käsitteellä metaevaluaatio tarkoitetaan arvioinnin arviointia (Scriven 2005, 249; Patton 2002, 211). Arvioinnin arviointi lisää suunnitelmallisuutta ja antaa tietoa arviointien hyödyntämisestä ja toteutuksesta. Arvioinnin arviointi voi keskittyä arviointien tulosten syntetisoivaan tarkasteluun, vertailla keskenään arviointeja, jotka on toteutettu erilaisin menetelmin, kriteerein ja erilaisista näkökulmista. Vertailulla voidaan pyrkiä kehittämään arviointikäytäntöjä, laajentaa näkökulmaa ja edistää tiedon kumuloitumista sekä pyrkimystä “konsensukseen”. Metaevaluaatio voi keskittyä myös arviointiprosessin laatuun ja onnistuneisuuteen, tulosten ja sisällön arviointiin toiminnalle asetettujen kriteerien pohjalta. Arviointisuunnitelmassa asetettujen tavoitteiden saavuttamista voidaan arvioida muun muassa käytettävyys-, toteutettavuus-, asianmukaisuus- ja täsmällisyysstandardien mukaisesti (Patton 1997, 17; Joint Committee 1994). Arviointi voidaan kohdistaa myös vaikuttavuuteen ja pyrkiä selvittämään, millaisia vaikutuksia arvioinneilla on ollut päätöksentekoon ja jatkuvaan kehitysprosessiin. Vaikutukset voivat olla esimerkiksi ennakoituja tai ennakoimattomia, hyödyllisiä tai haitallisia sekä merkitykseltään olennaisia tai epäoleellisia (Kuitunen & Hyytinen 2004, 22). Raivolan (2000a, 93) mukaan meta-arvioinnin yhteydessä pohditaan usein neljää arvioinnin pääaluetta 1) hyödyllisyyttä, 2) käyttökelpoisuutta ja soveltuvuutta, 3) oikeutta ja kohtuutta sekä 4) tarkkuusstandardeja. Arviointitiedon hyödyntämistä voidaan peilata myös suhteessa tulevaisuuden tarpeisiin tai siihen,

mikä on tuloksista saatu hyöty suhteessa kuluneiden resurssien määrään. Mark ym. (2000, 47) mukaan hyödyntämisen sijaan tarkastelun tulisi ensisijaisesti keskittyä seuraaviin näkökulmiin:

- Kuinka hyvin arvioitsija on kyennyt määrittelemään, mikä tieto tukee parhaiten yhteiskunnallista hyvinvointia ja kehitystä?
- Kuinka laadukasta hankittu tieto on?
- Kuinka onnistuneesti tulokset on kyetty levittämään asiaankuuluville tahoille, organisaatioille, arviointiin osallisille ja julkisuuteen?

Arviointiprosessin kriittinen arviointi, analysointi ja sen pohjalta tapahtuva arviointijärjestelmän kehittäminen on haasteellinen tehtävä. Arvioinnin erottaminen muista tekijöistä on hankalaa, ja kokemukset vaikuttavuuden seurantaan käytettävissä olevista menetelmistä ovat vielä vähäisiä. Syy- ja seuraussuhteiden arviointi on myös ongelmallista jo pelkästään siksi, että arviointiin perustuvat uudistukset eivät näy heti, vaan askel askeleelta etenevissä valinnoissa ja päätöksissä. Vaikutukset saattavat ilmetä vasta vuosien päästä esimerkiksi elämänlaadun paranemisena, tasa-arvon ja oikeudenmukaisuuden lisääntymisenä tai kilpailukyvyyn kasvuna. Tarkastelussa on myös otettava huomioon, että tiedon käyttö perustuu kielelliseen vuorovaikutukseen, jossa käsitteillä, symboleilla ja retoriikalla on keskeinen sijansa (Valovirta 2001, 11). Lisäksi arvioinnin arviointi voi nostaa esille myös arvioinnin seurauksena syntyneitä negatiivisia ja toimintaa heikentäviä seikkoja. Epäluottamus, mahdolliset väärät käsitykset, arvoriistiriidat tai aiemmat kokemukset arviointitiedon väärinkäytöstä kyseenalaistavat arvioinnin mielekkyyden. Eettiseltä perustaltaan kyseenalainen arviointi voi kääntyä arviointia ja kehittämistä vastaan ja vahingoittaa monin tavoin koko koulutusjärjestelmäämme.

Yhteenveto

Usein jo se, että asiat otetaan tarkastelun kohteeksi, herättää ihmiset pohtimaan, kyseenalaistamaan ja perustelevaan niitä. Tämä puolestaan edistää kehittämistä ja muutokseen sopeutumista. Pelkkään arviointiprosessin hyödyllisyyteen tyytyminen on kuitenkin voimavarojen hukkaamista. Lisäksi tulee pyrkiä siihen, että arviointitulokset ovat hyödynnettäviä, ne analysoidaan avoimesti ja perusteellisesti ja niiden perusteella tehdään johtopäätöksiä ja kehittämissuosituksia. Koulutuksen arviointineuvosto on asettanut tavoitteekseen tietoperustaisen päätöksenteon mallin, jossa arviointitieto saa entistä keskeisemmän aseman koulutuspoliittisessa ohjauksessa ja paikallisen koulutuksen kehittämisessä. Tämä tarkoittaa myös arvioinnin antamaa tukea reformien toimeenpanon ja kehittämisprosessien kulun ohjauksessa sekä arvioinnin osuvuutta, oikea-aikaisuutta ja monipuolisuutta. (Koulutuksen arvioinnin uusi suunta 2004, 29–30.)

Vaikka odotukset tietoperustaiselle päätöksenteolle ovat korkeat, toimintaa näyttäsivät ohjaavan usein muut seikat, muun muassa eri intressit ja valtapyrkimykset sekä olemassa olevat rakenteet ja rutiinit. Mikäli päätöksenteko perustuu kuitenkin kokonaan muille asioille, arviointi menettää merkityksensä. Arvioinnin hyödynnettävyyttä olisikin nostettava uudella tavalla suunnittelun ja toteutuksen lähtökohdaksi. Kuten Patton (2004, 277) toteaa, “oleellista on se miten todelliset ihmiset, todellisessa maailmassa, soveltavat arviointituloksia ja kokemuksia.” Arvioinnin kehittäminen yhä tasapuolisemmaksi, oikeudenmukaisemmaksi ja luotettavammaksi lisää uskoa arviointiin ja edesauttaa hyödyntämistä. Arvioinnin hyödyntäminen puolestaan edellyttää aikaa, malttia ja harkintaa sekä yhteistä näkemystä, mihin toiminnalla pyritään. Laadukas arviointi on kokonaisuus, joka syntyy eri osapuolten ja tahojen osaavan ja hyvän yhteistyön tuloksena. Siltä on lupa odottaa uudenlaisia työvälineitä, näkökulmia ja ideoita arvioinnin toteuttamiseen, tulosten hyödyntämiseen ja toimenpiteiden seurantaan. Tätä kautta arviointi voi muodostua todelliseksi kannustimeksi kaikilla tasoilla tapahtuvalle koulutuksen kehittämislle.

Lähteet

- AI50/2003 koulutuksen arvioinnista. Suomen sädöskokoelma.
- Alkin, M. 2005. Utilization of evaluation. Teoksessa S. Mathison (toim.) *Encyclopedia of Evaluation*. Thousand Oaks, London, New York, New Delhi: Sage Publications, 434–436.
- Chelimsky, E. 1997. The coming transformations in evaluation. Teoksessa E. Chelimsky & W. R. Shadish (toim.) *Evaluation for the 21st century. A Handbook*. Thousand Oaks, London, New York: Sage Publications, 1–26.
- Cook, T. 1997. Lessons learned in evaluation over the past 25 years. Teoksessa E. Chelimsky & W. R. Shadish (toim.) *Evaluation for the 21st Century. A Handbook*. Thousand Oaks, London, New York, New Delhi: Sage Publications, 30–52.
- Eduskunnan sivistysvaliokunnan lausunto 5/2001.
- Eduskunnan sivistysvaliokunnan lausunto 6/2004.
- Eduskunnan sivistysvaliokunnan mietintö 3/1998.
- Eduskunnan sivistysvaliokunnan mietintö 6/2002.
- Eduskunnan sivistysvaliokunnan mietintö 11/2002.
- Etukäteisarviointi Suomessa ja EU-maissa. 2000. Valtiovarainministeriö. Tutkimukset ja selvitykset 7. Helsinki: Edita.
- Euroopan parlamentin ja neuvoston suositus eurooppalaisesta yhteistyöstä kouluopetuksen laadunarvioinnissa. 2000. Euroopan yhteisöjen komissio. Yhteinen kanta C5-0565/2000, 00/0022 (COD).
- Evaluating EU Expenditure Programmes: A Guide. 1997. Ex post and Intermediate Evaluation Directorate-General XIX – Budgets, European Commission. European Commission. First edition.
- Hakkarainen, K., Lonka, K. & Lipponen, L. 2004. Tutkiva oppiminen: järki, tunteet ja kulttuuri oppimisen sytyttäjänä. Helsinki: WSOY.
- Heikkinen, H. 2002. Elämää kansiossa. Portfolio narratiivisena identiteettinä opettajankoulutuksessa. Teoksessa H. Heikkinen & L. Syrjälä (toim.) *Minussa elää monta tarinaa*. Kirjoituksia opettajuudesta. Kansanvalistusseura. Vantaa: Dark, 116–126.
- House, E. 1993. *Professional evaluation*. Newbury Park, California: Sage Publications.

- House, E. & Howe, K. 1999. Values in evaluation and social research. Thousand Oaks, California: Sage Publications.
- Hämäläinen, K. 2002. Arvioinnin onnistumisen edellytyksiä. Teoksessa K. Hämäläinen & M. Kaarti-
nen-Koutaniemi (toim.) Benchmarking korkeakoulujen kehittämisvälineenä. Korkeakoulujen arviointineuvoston julkaisuja 13:2002. Helsinki, 7–9.
- Jakku-Sihvonen, R. & Heinonen, S. 2001. Johdatus koulutuksen uudistuvaan arviointikulttuuriin. Opetushallitus. Arviointi 2. Helsinki.
- Joint Committee on Standards for Educational Evaluation. 1994. The program evaluation standards. Thousand Oaks, California: Sage Publications.
- Khakee, A. 1998. Evaluation and planning: inseparable concepts. *Town Planning Review* 69, 359–373.
- Kohti julkisten palvelujen yhteistä arviointia. 2001. Suomen Kuntaliitto, Sisäasiainministeriö, Valtio-
varainministeriö. Kansallinen tuottavuusohjelma. Tuottavuudella tulevaisuuteen. Helsinki.
- Korkeakoski, E. 2000. Kuntien ja koulujen itsearviointikäytäntöjä. Opetushallitus. Arviointi 7. Helsin-
ki.
- Korkeakoulutuksen laadunvarmistus. Opetusministeriön työryhmämuistioita ja selvityksiä 2004:6. Helsinki.
- Koulutuksen arvioinnin uusi suunta. 2004. Arviointiohjelma 2004–2007. Koulutuksen arviointineu-
voston julkaisuja 1. Jyväskylä.
- Koulutuksen tuloksellisuuden arviointimalli. 1998. Helsinki: Opetushallitus.
- Kuitunen, S. & Hytinen, K. 2004. Julkisten tutkimuslaitosten vaikutusten arviointi. VTT tiedotteita. Espoo.
- Lampinen, O. 1992. The utilization of social science research on public policy. Suomen Akatemian julkaisuja 4/92. Helsinki: VAPK-kustannus.
- Leviton, L. & Hughes, E. 1981. Research on the utilization of evaluations. *Evaluation Review* 4 (5), 525–548.
- Laukkanen, R. 2002. Evaluaatio hallinnan välineenä. Teoksessa R. Honkonen (toim.) Koulutuksen lumo. Retoriikka, politiikka ja arviointi. OKKA – Opetus-, kasvat- ja koulutusalojen säätiö.
- Lindqvist, T. 1999. Evaluaation uskottavuus. Teoksessa R. Eräsaari, T. Lindqvist, M. Mäntysaari & M. Rajavaara Arviointi ja asiantuntijuus. Tampere: Gaudeamus.
- Lumijärvi, I. & Kaarlejärvi, J. 2000. Ministeriöt arviointitutkimuksen hyödyntäjinä. Tampereen yliopis-
to. Hallintotieteen laitos. Hallintotiede 2000 A 23.
- Lyytinen, H. K. 2000. Ulkoinen arviointi ja oppilaitoksen kehittämissykli. Teoksessa M. Kangasoja & R. Kinnunen Arvioinnin arkea ja teoriaa. Länsi-suomen lääninhallituksen julkaisusarja 20/2000, 7.
- Mark, M., Henry, G. & Julnes, G. 2000. Evaluation. An integrated framework for understanding, guiding, and improving policies and programs. San Francisco: Jossey-Bass.
- Niemi, E. 2004. Perusopetuksen oppimistulosten kansallinen arviointi ja tulosten hyödyntäminen koulutuspoliittisessa kontekstissa. Perusopetuksen matematiikan oppimistulosten kansallinen arviointi 6. vuosiluokalla vuonna 2000. Turun yliopiston julkaisusarja 216/C.
- Oksanen, R. 2003. Laadun määrittely perusopetusta koskevassa kunnallisessa päätöksenteossa. Jo-
ensuun yliopisto. Kasvatustieteellisiä julkaisuja.
- O'Sullivan, R. 2004. Practicing evaluation. A collaborative approach. Thousand Oaks, California: Sage Publications.
- Patton, M. Q. 1997. Utilization-focused evaluation: The new century text. Third revised edition. London: Sage Publication.
- Patton, M. Q. 2002. Qualitative research & evaluation methods. London: Sage Publications.
- Patton, M. Q. 2004. The roots of utilization-focused evaluation. Teoksessa M. Alkin Tracing theo-
rists' views and influences. Thousand Oaks, California: Sage Publications.

- Pawson, R. & Tilley, N. 2000. *Realistic Evaluation*. London, Thousand Oaks and New Delhi: Sage Publications.
- Pawson, R. & Tilley, N. 2005. *Realistic Evaluation*. Teoksessa Mathison, S. (toim.), *Encyclopedia of Evaluation*. Thousand Oaks, London, New York, New Delhi: Sage Publications, 434–436.
- Petterson, S. & Wallin, E. 1995. *Utvärderingsmakt*. Teoksessa Mombach & Sahli-Andersson (toim.) *Från sanningssökande till styrmedel. Moderna utvärderingar i offentlig sektor*. Göteborg: Nerenius & Santerus Förlag.
- Raivola, R. 2000a. *Teho vai laatua koulutukseen?* Helsinki: WSOY.
- Raivola, R. 2000b. *Vaikuttavuutta koulutukseen. Suomen akatemian koulutuksen vaikuttavuusohjelman tutkimuksia*. Helsinki: Edita.
- Rajanen, J. 2000. *Selvitys koulutuksen paikallisen arvioinnin tilasta. Opetushallitus II/2000*. Helsinki.
- Rajavaara, M. 1999. *Arviointitutkimuksen hyödynnettävyys*. Teoksessa R. Eräsaari, T. Lindqvist, M. Mäntysaari & M. Rajavaara 1999. *Arviointi ja asiantuntijuus*. Tampere: Gaudeamus.
- Rinne, R., Kallo, J. & Hokka, S. 2004. *Liian innokas mukautumaan? OECD:n koulutuspolitiikka ja Suomen vastauksia*. *Kasvatus* 35 (1), 34–54.
- Rossi, P., Lipsey, M. & Freeman, H. 2004. *Evaluation: A systematic approach*. Seventh Edition. Thousand Oaks, California: Sage Publications.
- Räisänen, A. & Korpi, A. 2002. *Euroopan laatupalkintomallin sovellus ammatilliseen koulutukseen*. Opetushallitus. Arviointi 3. Helsinki.
- Scriven, M. 2005. Teoksessa S. Mathison *Encyclopedia of evaluation*. Thousand Oaks, California: Sage Publications.
- Temmes, M., Ahonen, P. & Ojala, T. 2002. *Suomen koulutusjärjestelmän hallinnon arviointi*. Helsinki: Opetusministeriö.
- Tiede ja yhteiskunta – työryhmän muistio. 2004. *Opetusministeriön työryhmämuistioita ja selvityksiä 2004:28*. Helsinki.
- Tilley, N. 2000. *Realistic evaluation: An overview*. Presented at the Founding Conference of the Danish Evaluation society, September 2000. Saatavana [www-muodossa: <http://www.danskevalueringsselskab.dk/Materiale_fra_DES.asp>](http://www.muodossa: <http://www.danskevalueringsselskab.dk/Materiale_fra_DES.asp>)
- Valovirta, V. 2000. *Evaluation utilization as argumentation*. A paper to be presented in Fourth European evaluation Society conference in Lausanne, Oct. 12–14 th 2000.
- Valovirta, V. 2001. *Metaevaluaatio arvioinnin kehittämisen välineenä*. *Finsoc-lehti* 1, 10–11.
- Valtioneuvoston selonteko eduskunnalle uusien koululakien vaikutuksista ja laeissa asetettujen tavoitteiden toteutumisesta 13.3.2002. Saatavana [www-muodossa: <http://www.minedu.fi/opm/koulutus/koulutusselonteko2002.pdf>](http://www.muodossa: <http://www.minedu.fi/opm/koulutus/koulutusselonteko2002.pdf>)
- Vedung, E. 1997. *Public policy and program evaluation*. New Brunswick, New Jersey Transaction publishers.
- Virtanen, P. & Uusikylä, P. 2000. *Agency level evaluations in Finland*. Paper presented in the CES conference in Montreal, 17.5.2000.
- Weiss, C. & Bucuvalas, M. 1980. *Truth tests and utility tests: Decision-makers' frame of reference for social science research*. *American Sociological Review* 45 (2), 302–313.
- Weiss, C. 1982. *Policy research in the context of diffuse decision Making*. *Policy Studies Review annual* 6, 19–36.



Kunta koulutuksen arvioijana

Arvioinnin tehtävä kunnassa on palvelujen ja toiminnan kehittäminen ja niitä koskevan päätöksenteon tukeminen. Koulutuksen paikallisen arvioinnin tärkein osa on kunnanvaltuuston asettamien tavoitteiden arviointi. Kunnan järjestämän koulutuksen arviointi on kokonaisuus, johon sisältyy myös erityislainsäädännön mukainen koulutuksen järjestäjän arviointi sekä koulujen ja oppilaitosten itsearviointi. Paikallisarvioinnin menettelytavat ja organisointi ovat kunnan vastuulla. Koulutuksen kansallisen ja paikallisen arvioinnin tulee täydentää toisiaan. Koulutuksen kehittämiseksi tarvitaan monipuolista arviointitietoa, jonka tuottaminen edellyttää tasavertaisten toimijoiden vuoropuhelua, avoimuutta ja vastavuoroisuutta.

Asiasanat: koulutuksen arviointi, paikallinen arviointi, valtuuston asettamien tavoitteiden arviointi, erityislainsäädännön mukainen arviointi, koulutuksen järjestäjän arviointi, EFQM-arviointi

Koulutuksen kuntajärjestäjän paikallinen arviointi samastetaan usein koulujen ja oppilaitosten itsearviointiin tai joskus jopa tiedonkeruuseen, jota kunnissa tehdään kansallisia arviointeja ja arviointitutkimusta varten. Se on kuitenkin laaja tehtäväkenttä ja kokonaisuus, joka läpäisee koulutuksen toimialan kaikki kunnalliset toimijat: valtuuston ja lautakunnan, viranhaltijajohdon, koulut ja oppilaitokset. Paikallisen arvioinnin tärkein osa on kunnanvaltuuston koulutukselle asettamien tavoitteiden arviointi. Näiden tavoitteiden lähtökohtina ovat kunnan sivistystoimen lakisäätteiset perustehtävät sekä kunnan strategiat.

Kunnan arviointi pohjautuu kuntalain mukaiseen arviointijärjestelmään, jota “rikastutetaan eri johtamis- ja arviointijärjestelmistä saaduilla vaikutteilla” ja täydennetään

erityyppisillä erillisselvityksillä (Laitila 2002, 93). Siihen sisältyy myös koulutuksen erityislainsäädännön mukaisen arviointivelvoitteen mukainen arviointi, joka velvoittaa koulutuksen järjestäjää arvioimaan antamaansa koulutusta ja sen vaikuttavuutta.

Kuntalaki korostaa nimenomaan valtuuston asemaa kunnan keskeisten toiminta- ja taloustavoitteiden ja kauaskantoisten päämäärien asettajana. Arviointi liittyy kiinteästi kunnan käytössä olevien voimavarojen riittävyteen ja kohdentamiseen. Se on keskeinen osa tulohajautuksen ja -johtamisen käytäntöä, johon kunnissa on siirrytty aiemmasta resurssihajautuksesta (Meklin 2001). Tuloksellisuutta tarkastellaan laajasti toiminnan onnistumisena: miten hyvin kunta täyttää tehtävänsä asukkaidensa palveluiden tuottajana?

Kunta on koulutuspoliittinen toimija

Yhteiskunnan nopean muuttumisen vuoksi koulutus saa jatkuvasti uusia, entistä vaativampia tehtäviä. On katsottu, että näihin haasteisiin voidaan vastata parhaiten väljentämällä koulutusta koskevaa lainsäädäntöä, hajauttamalla hallintoa ja uudistamalla koulutuksen rakenteita. Näin kuntien itsenäisyys koulutuksen järjestämisessä on lisääntynyt, ja niillä on vastuu koulutuksen kehittämisestä paikalliset ja alueelliset olosuhteet huomioon ottaen.

Paikallista koulutuspolitiikkaa tarvitaan, koska kunnat ja alueet ovat keskenään erilaisia. Kuntalaisten koulutustarpeet, maantieteelliset etäisyydet sekä elinkeino- ja koulutusrakenteet vaihtelevat. On välttämätöntä, että kunnilla ja alueilla on mahdollisuus luoda koulutuspolitiikkaa, joka perustuu niiden omiin lähtökohtiin ja paikalliseen koulutustarpeeseen. Välttämätöntä on myös riittävä liikkumavara koulutuspalveluiden järjestämisessä. Kunta ei enää ole ainoastaan koulutuspalveluiden järjestäjä, vaan myös koulutuspoliittinen toimija. Myös kunnallinen itsehallinto edellyttää, että kunnalla on oma sivistyspoliittinen linjansa. (Suomen Kuntaliiton yleiskirje 13/80/2001.) Kuntatason merkitys korostuu tutkimushavaintojen perusteella myös oppilaitosten johtamisessa, jota kuntien toimintalinjaukset määrittävät yhä suuremmassa määrin (Ojala 2004, 203).

Kunnanvaltuustokaan ei ole enää hyvinvointivaltiollisen kauden hallinto-orgaani, jonka tuli vain toteuttaa mahdollisimman täsmällisesti lainsäädännön ja valtion keskushallinnon määrittämät tavoitteet. Tällä hetkellä kyseessä on selvästi paikallinen poliittinen tahdonmuodostus, ei enää pelkästään täytäntöönpano. (Ryynänen 2001, 48).

Valtuustolla keskeinen rooli paikallisessa arvioinnissa

Nykyisessä kuntalaissa korostetaan valtuuston vahvaa asemaa sekä kokonaisvastuuta kunnan toiminnasta ja taloudesta. Kunnanvaltuusto vastaa strategisesta päätöksenteosta ja koko kuntakonsernin tavoitteiden asettamisesta, toiminnan ja talouden tasapainosta, toiminnan arvioinnista ja seurannan järjestämisestä. Lisäksi valtuuston on pidettävä huolta kuntalaisten osallistumis- ja vaikuttamismahdollisuuksista. (Majoinen 2001, 9.) Käyttäjädemokratian myötä myös kunnallisten koulutuspalvelujen asiakkaat osallistuvat palvelujen suunnitteluun.

Kuntayhtymien perustehtävä on samoin palvelu omistajakuntiaan koulutuksen järjestämisessä alueellisen kehittämisen näkökulma huomioon ottaen. Vertaamalla asettamia tavoitteita ja toiminnasta syntyviä tuloksia toisiinsa peruskunnanvaltuusto arvioi palvelujen laatua ja tuloksellisuutta (Markkanen 2004, 26). Myös silloin, kun kunta järjestää koulutuspalveluja esimerkiksi ostopalveluna, sillä on kokonaisvastuu niiden saatavuudesta, laajuudesta ja laadusta.

Myös kunnan koulutuspoliittiset strategiat ja linjaukset sekä paikallisten opetussuunnitelmien tavoitteet hyväksytään valtuustossa. Opetussuunnitelmien varsinaisista sisällöistä päätetään yleensä joko asianomaisessa lautakunnassa tai johtokunnissa. Kouluissa laaditaan työsuunnitelmat. Näiden asiakirjojen ja niiden arvioinnin on muodostettava johdonmukainen ja toimiva kokonaisuus.

Paikallisten koulutuspalvelujen arvioinnin lähtökohtana ovat lakisääteisten perustehtävien pohjalta nousevien kansallisten tavoitteiden lisäksi kunnan omat, valtuuston asettamat paikalliset tai alueelliset tavoitteet. Arvioinnin tärkein perustelu kunnassa on palvelujen ja toiminnan kehittäminen ja sitä koskevan paikallisen päätöksenteon tukeminen. Myös koulutuksen toimialalla jokaisella kunnalla ja alueella on omat kehittämishaasteensa.

Aluetason, toisin sanoen kunnan rajat ylittävän yhteistyön, tarkastelu ja arviointi on kunnissa yhä välttämättömämpää. Esimerkiksi lukiokoulutuksen ja ammatillisen koulutuksen järjestämiseen ja mitoittamiseen joudutaan kunnissa ottamaan kantaa silloinkin, kun kunnassa ei ole toisen asteen koulutusta. Myös kansalaisopistotoiminnassa tukeudutaan yhä useammin alueellisiin ratkaisuihin.

Asukkaiden hyvinvointi on kunnan palvelujen kokonaisvaltainen, sektorirajat ylittävä tavoite, jota ei mikään kunnan hallinnonala voi toimillaan yksin saavuttaa. Lap-suuden historiaa tutkineen Mirja Satkan mukaan lapset elävät aiemman ydinperhelapsuuden sijasta nykyisin ”verkostolapsuutta”, jolloin koulun vaikutusmahdollisuudet ovat vähäisemmät kuin aiemmin. Kunnan eri toimialoilla on yhteinen vastuu esimerkiksi koulunsa keskeyttäneistä, työttömistä nuorista. Siksi tarvitaan yhä kattavampaa arviointitietoa perinteisten hallintokuntarajojen yli menevän ohjauksen ja johtamisen tarpeisiin.

Koulutuspalveluja tuotetaan yhteistyössä muiden tahojen kanssa, mikä merkitsee myös sidosryhmä- ja kumppanuuslähtöisen ja poikkihallinnollisen, kokonaan kunnan

omista kysymyksenasetteluista lähtevän arviointitarpeen lisääntymistä. Tästä esimerkkinä on aamu- ja iltapäivätoiminta.

Kunnille kokonaisvaltaiset arvioinnit ovat välttämättömiä myös siksi, että niiden on samoilla kokonaisvoimavaroilla vastattava esimerkiksi usean ministeriön toimialaan kuuluvista asioista ja sovitettava yhteen eri tehtävät ja voimavarat.

Koulutuksen paikallinen arviointi sisältää ensinnäkin poliittishallinnollisesta näkökulmasta suoritettua arvioinnin, jossa on keskeistä koulutuksen lainsäädännön toteutumisen seuranta sekä talouden tasapainottaminen, priorisoinneista vastaaminen ja toiminnan voimavarojen turvaaminen. Päävastuussa on luottamusmies- ja viranhaltijajohto. Lisäksi koulutuksen paikalliseen arviointiin kuuluu koulutuksen professionaalista intresseistä käsin tehtävä arviointi, joka painottuu laadukkaaseen opettamiseen ja oppimiseen sekä oppilaitosten johtamisen parantamiseen. Tässä arvioinnissa kouluilla, rehtoreilla, opettajilla ja oppilailta sekä koulujen sidosryhmillä on keskeinen rooli. Paikallisen arvioinnin tulee vastata niin politiikan, hallinnon, opetusprofession kuin kuntalaistenkin käytännöllisiin tietotarpeisiin.

Paikallisen arvioinnin toimintamalli

Kunnanvaltuusto määrittää kunnan sivistys- tai koulutusvision, toiminta-ajatuksen ja arvovalinnat ja asettaa tavoitteet näiden pohjalta. Valtuusto päättää sekä pitkän aikavälin keskeisistä tavoitteista että vuosittain talousarvioon ja taloussuunnitelmaan sisältyvistä tavoitteista.

Kunnan sivistystoimen yleiset tavoitteet ja linjaukset, jotka lähtevät kunnan kokonaisstrategian kulloisistakin haasteista ja painotuksista, kootaan useimmissa kunnissa sivistys- tai koulutuspoliittiseen ohjelmaan. Vaihtoehtoisesti ne esitellään osana yleistä kuntastrategiaa. Jotta kunnan koulutuspoliittisella ohjelmalla on riittävä painoarvo, sen tulee olla valtuuston hyväksymä. Jotta myös sen toteuttaminen onnistuisi, sillä tulee olla laaja hyväksyntä. Siksi on tärkeää, että sekä luottamusmiehiä että viranhaltijoita osallistuu monipuolisesti ohjelmatyöhön.

Opetukseen liittyviä tavoitteita voidaan asettaa esimerkiksi koulutuksen vaikuttavuudelle, opetuksen voimavaroille ja kunnan opetussuunnitelmalle. Opetuksen osalueiden tavoitteet voivat koskea esimerkiksi kouluverkkoa, erityisopetusta, koulujen välistä yhteistyötä, esiopetusta, oppilashuoltoa, koulukuljetuksia, kieliohjelmaa, kaksikielisyyttä, täydennyskoulutusta, työyhteisön kehittämistä, johtamista, tietotekniikkaa ja niin edelleen.

Valtuuston asettamien toiminnallisten ja taloudellisten tavoitteiden toteutumisen arviointi on kuntalaissa säädetty tarkastuslautakunnan tehtäväksi. Tarkastuslautakunta on valtuuston apuelin ja sillä on riippumaton asema suhteessa muuhun hallintoon. Tarkastuslautakunnan arviointi pohjautuu kunnanhallituksen toimintakertomukseen ja muuhun keskeiseen tietoaaineistoon. Arviointitulokset esitetään valtuustolle tilinpää-

töksen yhteydessä arviointikertomuksessa. Siinä tarkastuslautakunta tuo esiin käsitöksensä, miten kunnanhallitus ja kunnan muut viranomaiset ovat onnistuneet toteuttamaan valtuuston tahdon (Suomen Kuntaliiton yleiskirje 13/80/2001).

Koulutuksen arviointi on myös erityislainsäädännössä asetettu velvoite, joka koskee suurta osaa kunnan sivistystoimintaa (perusopetusta, toisen asteen koulutusta, ammatillista aikuiskoulutusta, taiteen perusopetusta, vapaata sivistystyötä sekä kirjasto- ja tietopalveluja). Lainsäädäntö velvoittaa koulutuksen järjestäjiä arvioimaan antamansa koulutusta sekä osallistumaan opetusministeriön järjestämään ulkoiseen arviointiin, jota toteuttavat niin lääninhallitukset kuin Koulutuksen arviointineuvostokin. Asetuksessa, jolla arviointineuvostosta säädetään, todetaan että paikallisesta arviointitoiminnasta ja sen kehittamisestä vastaavat ensisijaisesti koulutuksen järjestäjät (VN:n asetus 150/2003). Lainsäädännössä koulutuksen järjestäjällä tarkoitetaan kuntaa, kuntayhtymää, yksityistä yhteisöä tai säätiötä.

Kunnan hallinnossa valtuustolla, hallituksella, lautakunnalla sekä viranhaltijaorganisaatiolla on koulutuksen järjestämisessä kullakin omat tehtävänsä ja vastuunsa. Ne määrittävät kuntalaissa sekä kunnan hallinto- tai johtosäännössä ja muissa määräyksissä. Paikallisesti ratkaistaan, miten valtuuston tavoitteet ja muut arvioinnit limitetään toisiinsa käytännön arviointitoiminnassa.

Koulutusta koskevissa erityislaeissa ei ole säädetty arvioinnin suorittamistapaa. Paikallisella tasolla arvioinnin menettelytavat, organisaatio ja tulosten julkistaminen ovat kunnan tai kuntayhtymän vastuulla. Arviointimenetelmät räätälöidään arvioinnin kulloisiinkin tavoitteisiin, kohteisiin ja osallistujille sopiviksi.

Kunnan suunnittelu- ja arviointijärjestelmän kaikkien osien tulee ohjata toiminta keskeisissä asioissa samaan suuntaan. Kunnan strategian mukaisten valtuustotason tavoitteiden, talousarvion tavoitteiden, koulutuspoliittisten suunnitelmien, opetussuunnitelmien, koulujen työsuunnitelmien, tietostrategioiden ja suunnitelmien tulee siis tukea ja täydentää toisiaan. Opetustoiminnassa näiden pitää lisäksi olla sopusoinnussa laissa määriteltyjen tavoitteiden ja valtakunnallisen arvioinnin kanssa. (Nikander 2002.)

Kunnan opetuksen arviointijärjestelmän tulee ennen kaikkea kyetä muuttamaan ja kehittämään toimintaa nopeasti. Sen tulee olla koko toiminnan kattava ohjausjärjestelmä, joka toimii yhteen sovitettuna joka tasolla. Hyvällä arviointitoiminnan organisoinnilla voidaan välttää kuntalain ja erityislainsäädännön mukaisen arviointityön päällekkäisyys tai irrallisuus sekä kahden erillisen ohjausjärjestelmän syntyminen.

Oppilaitosten laatuarvioinnin kehittäminen

Koulutuksen paikallisarvioinnin tarpeisiin on kehitetty monia menetelmiä, ja jokainen kunta valitsee menetelmänsä itse. Käytännössä kunnassa on usein yhtä aikaa meneillään monia erilaisia teema- ja hankearviointeja sekä asiakas- ja henkilöstöky-

selyjä, työilmapiirikartoituksia ja niin edelleen.

Laatupolitiikka otettiin Suomessa virallisesti hallintopolitiikan osaksi jo 1990-luvun lopussa. Valtioneuvosto antoi tuolloin hallinnon kehittämisen periaatepäätöksen, jossa esitetään julkisten palvelujen laadun ja saatavuuden parantamista. Valtio ja kunnat laativat yhdessä samana vuonna Julkisten palvelujen laatustrategian. (Ks. esim. Lumijärvi 199, 192.) Siinä suositeltiin EFQM-laatupalkintokriteeristöä julkisten palvelujen laatuarvioinnin työkaluksi. EFQM-kriteeristöä käytetäänkin nykyisin kunnissa lähes ainoana laadun arviointimenetelmänä. Laatupalkintomalli on hyväksi koettu laadun elementtien arviointikriteeristö, jonka avulla voi itsearviointissa verrata esimerkiksi oman oppilaitoksen sijaintia laadun portailla. Teknisellä sektorilla ja terveydenhuollossa on otettu jossain määrin käyttöön myös ISO-standardiin perustuva toiminnan laadun sertifiointi.

Strategiseen suunnitteluun tarkoitettu “tasapainotettu mittaristo” (Balanced Scorecard, BSC) on myös yleistynyt kunnissa nopeasti. Strategisen suunnittelun ja johtamisen menetelmänä BSC sopii hyvin EFQM-mallin rinnalle. BSC:n näkökulmat sisältävät organisaation tavoitteet ja keinot niiden saavuttamiseksi. EFQM-arvioinnin työkaluna kertoo puolestaan, miten tavoitteiden saavuttamisessa on onnistuttu.

Koulutuksen toimintaympäristö ja itse koulutus kansainvälistyy päivä päivältä. Kansainvälisessä koulutuskentässä laadunosoittamistarpeet ovat viime aikoina lisääntyneet. Erityisesti EU:n laatupolitiikka vaikuttaa kansallisen laadunarvioinnin kehittämiseen myös Suomessa.

Oppilaitoksen itsearviointi – esimerkkinä EFQM-menetelmä

Itsearviointin tulee ensinnäkin antaa kokonaiskuva oppilaitoksen tilanteesta ja osoittaa parantamiskohteet. Toiseksi sen tulee tuottaa tietoa sekä itsearviointin että ulkoisen arvioinnin tarpeisiin. Itsearviointin on tarjottava mahdollisuus myös oppilaitosten ja muiden organisaatioiden väliseen vertailuun esimerkiksi kunnassa ja sovittava kaikenkokoisille oppilaitoksille. Sen tulee olla osa normaalia työtä, se ei toisin sanoen saa vaatia kohtuutonta työpanosta. Itsearviointin täytyy olla myös kustannuksiltaan sopiva.

Nämä vaatimukset osoittavat, että pelkkä opetus- ja oppimisprosessin arviointi ei riitä oppilaitoksen itsearviointin tarpeisiin. Arviointimenetelmän avulla täytyy pystyä hankkimaan tietoa kaikista oppilaitoksen toiminnan osa-alueista. Päällekkäisen arviointityön välttämiseksi on suotavaa, että sama menetelmä tuottaa tietoa, joka tyydyttää mahdollisimman monenlaiset arviointitarpeet. Arviointitiedon tulee sopia sekä oppilaitosten että niiden taustayhteisön, kunnan, tarpeisiin. Kunnissa halutaan nykyään vertailla hyvin erilaisia toimintoja keskenään. Suotavaa on, että samalla menetelmällä saadaan tietoa myös ulkoista arviointia varten.

Menetelmän pitää olla niin joustava, että sillä voidaan arvioida sekä suuria oppilaitoksia, joissa on paljon päätoimista henkilökuntaa että pieniä oppilaitoksia, joissa ei välttämättä ole juuri lainkaan päätoimisia työntekijöitä. Tällaisia ovat esimerkiksi pienet kansalaisopistot. Kustannuksiltaan sopiva menetelmä on silloin, kun sen käytöstä aiheutuvat kustannukset eivät ylitä toiminnan parantumisesta saatuja tuloja. Esimerkiksi varsin kallis ISO-standardi saattaa olla kustannuksiltaan sopiva arviointimenetelmä sellaiselle oppilaitokselle, joka sen avulla onnistuu myymään riittävän määrän koulutuspalvelujaan jollekin ISO-standardia vaativalle yritykselle. Tavalliselle oppilaitokselle se on varmasti liian kallis, puhumattakaan sen vaatimasta työmäärästä.

Näillä perusteilla varsin käyttökelpoisia itsearviointimenetelmiä ovat erilaisiin laatupalkintoihin perustuvat menetelmät, joiden käyttö on jatkuvasti yleistynyt myös julkishallinnossa.

Kun työskennellään Euroopan laatupalkintokriteeristön (EFQM) pohjalta joudutaan huomio kiinnittämään seuraaviin arviointialueisiin: johtajuus, toimintaperiaatteet ja strategia, henkilöstö, kumppanuudet ja resurssit, prosessit, asiakastulokset, henkilötulokset, yhteiskunnalliset tulokset sekä keskeiset suorituskykytulokset. Itse arviointitoiminta etenee tässä mallissa seuraavien vaiheiden kautta:

- Osana strategista suunnitteluprosessia määritellään tulokset, jotka halutaan saavuttaa. Nämä tavoiteltavat tulokset kattavat oppilaitoksen koko suorituskyvyn, mukaan lukien varsinaiset oppimistulokset sekä taloudelliset ja toiminnalliset näkökohdat. Myös sidosryhmien näkemykset otetaan huomioon.
- Seuraavaksi vastataan kysymykseen, miten nämä tulokset saavutetaan ja tunnustetaan syyt siihen, miksi toimitaan juuri näin. Erinomaisesti toimivassa oppilaitoksessa toimintatavoilla on järkevät perusteet, niihin liittyvät prosessit ovat hyvin määriteltyjä ja ne ottavat huomioon sidosryhmien tarpeet. Toimintatavat ovat myös yhtenäisiä ja ne tukevat toimintaperiaatteita ja strategiaa.
- Käytännön soveltamisessa keskeistä on valittujen toimintatapojen järjestelmällinen soveltaminen kaikilla toiminnan osa-alueilla, koska vasta se takaa arvioinnille kunnollisen perustan.
- Arvioinnissa ja parantamisessa vastataan kysymykseen, miten oppilaitos arvioi ja parantaa toimintatapojaan ja niiden käytännön soveltamista. Tämän edellytyksenä on, että toiminnan tuloksia mitataan säännöllisesti ja että niitä käytetään parantamistoimenpiteiden tunnistamiseen, priorisointiin, suunnitteluun ja toteuttamiseen.

Näin ymmärrettynä oppilaitoksen itsearviointi tukee oppilaitoksen johtamista. Mutta kuten edellä todettiin, oppilaitoksia arvioidaan myös muista syistä. Olennaisen tärkeää silloin on, miten itsearviointiprosessissa saatua tietoa voidaan hyödyntää esimerkiksi oppilaitokseen kohdistuvassa ulkoisessa arvioinnissa, jotta voidaan välttää päällekkäinen työ. Tämä edellyttää muun muassa, että sovitaan keskeisistä toimintaa ku-

vaavista indikaattoreista, jotka palvelevat sekä oppilaitoksen itsearviointia että ulkoista arviointia.

Kansallisen ja paikallisen arvioinnin yhteistyö

Arvioinnin tehtävä ja asema koulutuksen kansallisessa ohjausjärjestelmässä on ollut melko epäselvä (Vuorenmaa 2001, 202). Lähtökohta on, että koulutuksen kansallisella arvioinnilla tuotetaan tietoa, miten kansalliset koulutuspoliittiset tavoitteet on saavutettu. Lisäksi tuotetaan tietoa kuntiin kohdistuvan informaatio-ohjauksen ja muun ohjauksen tarpeisiin. Arvioinnilla pyritään siis sekä kontrolloiviin että kehittäviin tavoitteisiin. Systemaattinen ja analysoitu koulutuspalvelujen tilaa koskeva arviointitieto auttaa tunnistamaan ja paikallistamaan kehitystarpeet.

Informaatio-ohjauksessa olennaista on, miten arviointien avulla tuotettua tietoa osataan rakentavalla tavalla syöttää takaisin paikallisiin prosesseihin. Tiedon vastaanottajan kannalta merkityksellinen tieto johtaa tehokkaimmin toivottuihin toimenpiteisiin. Tuloksellisessa informaatio-ohjauksessa korostuvat tiedon kaksisuuntaisuus ja vuorovaikutuksellisuus. Kansallisen arviointijärjestelmän tulee siksi olla paikallista koulutusjärjestelmää tukeva, avoin ja monitasoinen.

Arviointi on aina arvo- ja näkökulmasidonnaista. Se on ajan, paikan ja tilanteen mukaan vaihtelevaa, kontekstuaalista toimintaa. Siksi kansallisen ja paikallisen arvioinnin tavoitteet ja kysymyksenasettelut ovat väistämättä osittain erilaiset, ja siitä johtuen käytetään myös erilaisia tiedonkeruu- ja analyysimenetelmiä. Kansallisten arviointikohteiden valinnan ja niiden varsinaisen arvioinnin tulee perustua ensisijaisesti valtakunnallisten tavoitteiden toteutumisen seurantaan siihen sopivin menetelmin. Paikallisen arvioinnin tulee vastaavasti perustua paikallisten ja alueellisten tavoitteiden toteutumisen arviointiin.

On hyvin tärkeää, että koulutuksen arviointitietoa tuotetaan erilaisista lähtökohdista, erilaisilla metodeilla ja tutkimustyön universaaleja periaatteita kunnioittaen. Sen tulee hyödyttää erilaisia tiedonintressejä ja erilaista arviointiosaamista. Arvioinneissa tulee myös noudattaa ”hyvää arviointitapaa”, johon kuuluu muun muassa eettisten periaatteiden mukaan toimiminen ja kuntiin kohdistuvissa arvioinneissa niiden itsehallinnon kunnioittaminen. Arviointien yhteydessä on myös aina kysyttävä, onko ”tilitelvollinen” tosiasia voinut vaikuttaa asiaan vai ei. Uudistusten toimeenpanolle on myös jätettävä riittävästi aikaa ennen kuin niitä voidaan luotettavasti arvioida.

Valtakunnallisenkin arvioinnin eetos voi olla ymmärtävä ja esimerkiksi kuntien erilaisia koulutusolosuhteita, järjestämistapoja ja koulutustarpeita tasapuolisesti esiin tuova. Käytössä olevat mittarit ja indikaattorit kuvaavat usein vain arvioitavan kohteen yksittäisiä piirteitä, eivät kokonaisuutta. Tarvitaan olosuhteet ja toiminnan erityispiirteet tuntevaa laadullista asiantuntemusta, jotta tulkinnot ja vertailut eivät jää liian kapealle pohjalle. Arviointihavaintojen tulkitsemisessa saatetaan joissain tapa-

uksissa tarvita osapuolten yhteistä neuvotteluprosessia.

Pelkkä asioiden tilan todentaminen kansallisella arvioinnilla ei anna parasta tietopohjaa ja välineitä kun pohditaan, millaisia koulutuspoliittisia ratkaisuja kansallisella tasolla olisi tehtävä. Siksi tarvitaan paikallisarvioinnin tuottamaa tietoa koulutuspoliittisten uudistusten onnistumisesta, mutta myös mahdollisista ennakoimattomista tai kielteisistä vaikutuksista. Monipuolisen paikallisen arvioinnin välittämän tiedon avulla on mahdollista tarkastella ongelmia uusista näkökulmista ja ratkaista niitä uudella tavalla.

Koulutuksen kehittäminen on kansallisen arvioinnin perimmäinen tavoite. Vastuu kansallisten arviointitulosten pohjalta tehtävistä muutoksista palautuu suurelta osalta paikalliselle ja alueelliselle tasolle. Muutosten tekemisessä tarvitaan informaatio-ohjauksen kautta saadun tiedon lisäksi kuitenkin myös toimijoiden omaa tahtoa ja sitoutumista. Asian ytimessä ei ole yksinomaan tietäminen, vaan myös tahtominen (Pylkkänen 2003). Tämä korostuu samalla kun kuntien koulutuspoliittinen rooli edelleen vahvistuu. Teknisluonteiset ulkoiset arvioinnit eivät ole omiaan käynnistämään muutoksia, vaan siihen tarvitaan tasavertaisten toimijoiden vuoropuhelua, avoimuutta ja vastavuoroisuutta. Arviointitulokset tulee aina julkistaa, mutta sen lisäksi niistä täytyy pystyä keskustelemaan avoimesti.

Koulutuksen kehittäminen ei voi perustua yksinomaan harvoin virallisesti asetettuihin tavoitteisiin ja niiden saavuttamisen mittaamiseen. Siksi keskeistä on paikallisen arvioinnin itseohjautuvuus sekä aito alueellisuus ja paikallisuus. Paikallinen kehittämispotentiaali menetetään myös, jos arviointi yksipuolistuu. Jos paikallisarviointi tuottaisi vain siltä tilattua tietoa, vaarana olisi niin sanottu ”tunneliefekti”.

Kansallisen poliittisen päättäjän käytössä tulee olla tietoa myös suuren yleisön koulutukseen kohdistamista odotuksista ja tyytyväisyydestä. Tämä liittyy muun muassa koulutuksen tavoitteiden laatuun ja oikeellisuuteen sekä siihen, miten hyvin koulutus onnistuu palvelemaan yksilöiden omaa, henkilökohtaista kehitystä. Tämä tieto välittyy parhaiten kansallisen arvioinnin käyttöön kuntatasolta, koska kunta on kansalaisia lähinnä oleva koulutuksen hallinnon taso. Alue- ja kuntatasolta saadaan luotettavinta tietoa myös esimerkiksi työelämän vaatimista koulutustarpeiden muutoksista.

Koulutuksen arviointineuvosto on ensimmäisessä arviointiohjelmassaan (Koulutuksen arvioinnin uusi suunta 2004, 28) kiinnittänyt huomiota siihen, että arviointia koskevan lainsäädännön ja sen perustelujen mukaan koulutuksen järjestäjien ja oppilaitosten arviointien tuloksia pitää hyödyntää kansallisissa arvioinneissa. Vastaavasti on oikeus saada kansalliset arviointitulokset paikalliseen käyttöön. Lisäksi arviointien tulee täydentää toisiaan. Tavoitteena ei arviointineuvoston mukaan kuitenkaan ole arviointien samanlaisuus vaan yhdensuuntaisuus, jota edistetään avoimesti yhdessä tapahtuvan kehittämisen ja jatkuvan toisilta oppimisen avulla.

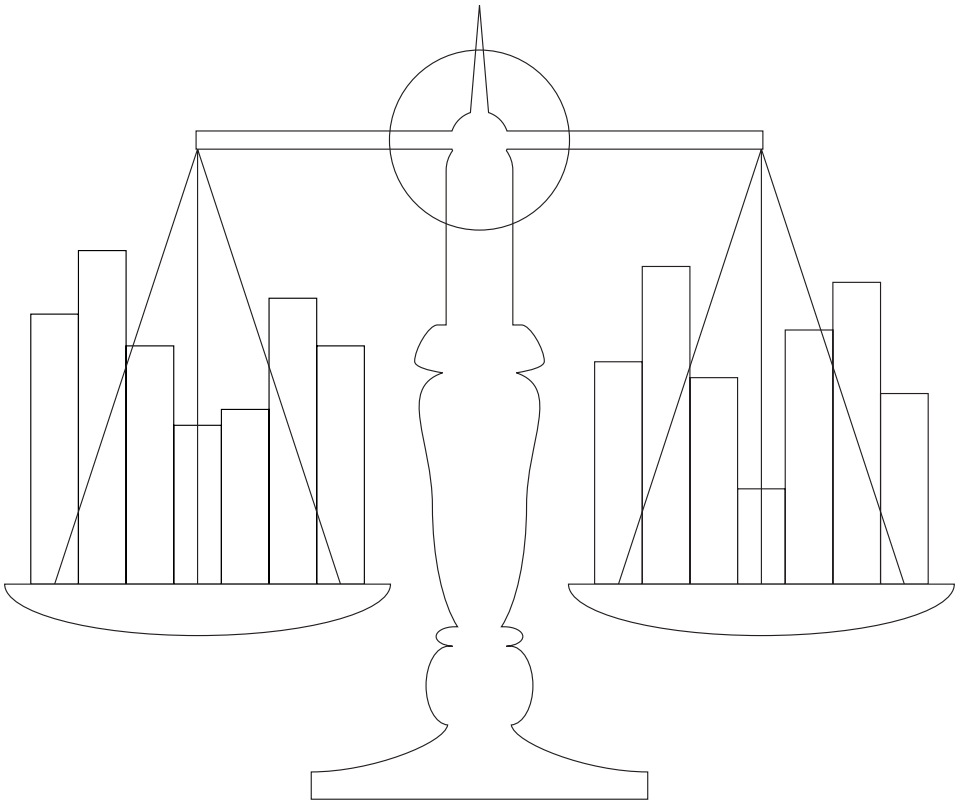
Aika on ajanut viranomais- ja hallintokeskeisen, ylhäältä-alas-strategiaan perustuvan, autoritaarisen ja teknisluonteisen arvioinnin ohi. Monipuoliseen arviointijärjes-

telmään ja kansallisen ja paikallisen arvioinnin hedelmälliseen yhteistyöhön voidaan päästä, kun noudatetaan kaikessa arvioinnissa symmetriseen viestintään ja vuoropuheluun perustuvan arvioinnin toimintaperiaatteita.

Lähteet

- A 150/2003 Asetus koulutuksen arvioinnista.
- Koulutuksen arvioinnin uusi suunta. Arviointiohjelma 2004–2007. Koulutuksen arviointineuvoston julkaisuja I. Jyväskylä.
- Laitila, T. 2002. Koulutuksen ohjauksen kipupisteitä ja kehittämishaasteita. Teoksessa T. Holopainen, E. Lukkarinen & P. Rautama (toim.) Paluu kansalliseen koulutuspolitiikkaan. Kirjoituksia suomalaisesta koulutuksesta ja koulutuspolitiikasta. Vaasa: Valtion alueellisen sivistyshallinnon virkamiehet VSV, 67–105.
- Lumijärvi, I. 1999. Laatujohtamisen soveltuvuus julkiselle sektorille. Hallinnon tutkimus 3/1999, 180–193.
- Majoinen, K. 2001. Mitä virkaa valtuustolla? Kuntalailla säädetyn valtuuston perustehtävän monitahoarviointi. Suomen kuntaliitto. Acta-väitöskirjasarja 2.
- Markkanen, M. 2004. Koulutuskuntayhtymän arviointimalli toimii. Kuntalehti 13/2004, 26.
- Meklin, P. 2001. Tavoitteiden saavuttamisen arviointi kuntataloudessa. Teoksessa A. Myllymäki & J. Vakkuri Tulos, normi, tilivelvollisuus. Näkökulmia tilintarkastukseen ja arviointiin. Tampere University Press, 91–116.
- Nikander, H. 2002. Vuorovaikutteinen arviointi kunnassa. Teoksessa V. Granö-Suomalainen & M. Lovio (toim.) Mihin me pyrimme? Miksi arvioida kunnan koulutus- ja kirjastopalveluja? Helsinki: Suomen Kuntaliitto, 56–59.
- Ojala, I. 2004. Oppilaitosjohtamisen haasteet toimintaympäristön muutoksessa. Kunnallistieteellinen aikakauskirja 3/2004, 200–208.
- Pylkkänen, P. 2003. Pakkopullasta rakentavaan työkaluun. Maaseudun uusi aika 2:2003, 63–68.
- Ryytänen, A. 2001. Tarkastuslautakunta valtuuston apuelimenä. Valtuuston aseman vahvistamista vai kuntien valtionvalvonnan kunnallistamista. Teoksessa A. Myllymäki & J. Vakkuri Tulos, normi, tilivelvollisuus. Näkökulmia tilintarkastukseen ja arviointiin. Tampere: University Press, 41–61.
- Stenvall, J. 2001. Hyvä arviointi ei ole elvistelyä. Hallinnon tutkimus 3:2001.
- Suomen Kuntaliiton yleiskirje 13/80/2001: Sivistystoimen arviointi on osa valtuustotason arviointia.
- Suomen Kuntaliiton yleiskirje 14/80/2003: Sivistystoimen arvioinnin indikaattorit – keskustelualoite.
- Vuorenmaa, Maritta 2001. Ikkunoita arvioinnin tuolle puolen. Uusia avauksia suomalaisen koulutusta koskevaan evaluaatiokeskusteluun. Jyväskylä: Jyväskylän yliopisto.

ARVIOINTI KEHITTÄVÄKSI



Koulutuksen arvioinnin keskeisenä tehtävänä on kehittää koulutusta ja sen osia. Arviointi voi siten suunnata tulevaisuutta niin sen prosessin kuin tulosten avulla. Arviointiprosessi on itsessään tehokas oppimisen väline ja arviotavasta kohteesta saadun tiedon avulla voidaan ennustaa tulevaisuuteen ja asettaa uusia tavoitteita.

Arviointi on kehittävää silloin, kun se sitoo toiminnan arvioinnin sen aktiiviseen kehittämiseen. Tämän kaltainen arviointi voi kohdistua koulutuksen eri osiin tai kokonaisuuteen. Vaikka erilaisten arviointien tehtävänä on johtaa kehittämiseen, erityisen hyvin tämän kaltainen arviointi soveltuu prosessuaalisiin toimintoihin.

Kehittävän arvioinnin ensisijaisena tehtävänä on tukea päätöksentekoa, johtamista ja toiminnan jatkuvaa parantamista, kehittämistä ja hallintaa. Haasteena on, että kehittämissuunnat, tavoitteet ja eritasoiset päätökset perustuisivat entistä enemmän myös arvioinnin ja tutkimuksen avulla osoitettuihin näyttöihin. Kehittävässä arvioinnissa ei ole kyse yksittäisestä menetelmästä, vaan laajasta menetelmien kirjosta, joita voidaan käyttää ja soveltaa tarkoituksenmukaisella tavalla ottaen huomioon myös arviointiin kohdistuvat tarpeet ja odotukset ja kontekstin vaatimukset. Kehittävässä arvioinnissa korostetaan arviointikohteenä olevien aktiivista osallistumista arvioinnin suunnitteluun ja arviointiin sekä arviointitulosten perusteella tapahtuvaan jatkuvaan parantamiseen (subjektius). Kyse on tarpeisiin räätälöidystä arviointitavasta.

Kehittävä arviointi on luonteeltaan ohjaavaa ja varsin praktista - sekä arviointitiedon että -prosessin tulee hyödyttää kehittämistä. Tässä merkityksessä kehittävä arviointi painottaa yhteisöllisyyttä, konsensushakuisuutta ja sekä tulosten että prosessissa syntyvän tietämyksen ja ymmärryksen hyödynnettävyyttä. Kehittävän arvioinnin avulla ei haeta "mitattavaa" totuutta eikä sen avulla saatu näyttö (evidence based) täytä kaikilta osin tieteellisen tutkimuksen vaatimuksia, vaikkakin se on systemaattisesti ja luotettavasti kerättyä ja osoitettu.

Kehittävä arviointi on osallistavaa ja usein myös monitahoista. Monitahoarvioinnille on luonteenomaista eri menetelmien hyväksikäyttö, arvioitavan yksilön tai yhteisön itsearviointi ja sidosryhmien osallistuminen arviointiin. Tämän kaltaisen arvioinnin perusta on sen laaja-alaisuudessa, neuvottelevuudessa, sosiaalisessa oikeudenmukaisuudessa ja evoluutionaarisuudessa. Myös konsultatiivinen arviointi, vertaisarviointit (esim. peer evaluation, benchmarking), yksilö- ja yhteisöportfoliot, moderation samoin kuin kokonaisvaltaiseen laadunhallintaan perustuvat arviointimallit (esim. EFQM-arviointi) sisäiset ja ulkoiset auditoinnit mukaan lukien ovat luonteeltaan kehittävää arviointia.

Koulutuksen arviointineuvoston määrittämän arvioinnin uuden suunnan mukaan kansallisissa koulutuksen arvioinneissa painotetaan kehittävää arviointia. Kehittävä arviointi on tulkittu tässä yhteydessä käyttäjälähtöiseksi ja osallistavaksi prosessiksi, jossa arviointimenetelmä räätälöidään arvioinnin tavoitteiden ja arvioitavan teeman mukaan. Arvioinnit perustuvat itsearviointiin ja arviointia tehdään monitahoisuuteen perustuvien arviointiryhmien toimesta. Arvioinnit johtavat myös kehittämissuosituksiin.

Anu Räisäsen ensimmäisessä artikkelissa **Näytöt kehittämisen perustaksi** tarkastellaan arvioinnin, tutkimuksen ja muilla keinoin osoitettujen näyttöjen merkitystä kehittämisen lähtökohtina. Artikkelissa korostetaan todennetun tiedon ja todistusaineiston merkitys-

tä kehittämisen lähtökohtana uskomusten, tottumusten ja intressien sijasta. Artikkelissa tarkastellaan näyttöön perustuvaa toimintaa ja sen merkitystä koulujärjestelmä ohjauksessa, päätöksenteossa ja kehittämisessä sekä opettajuuden kehitymisessä. Artikkelissa esitetään myös toimenpide-ehdotuksia näyttöperusteisen toimenpidekulttuurin luomiseksi.

Anu Räisänen toisessa artikkelissa **Kehittävään arviointiin** tarkastellaan kehittävän arvioinnin luonnetta ja menetelmiä. Kehittävä arviointi luonnehditaan osallistavaksi, avoimeksi, läpinäkyväksi, käytännönläheiseksi ja oppimista edistäväksi mutta kuitenkin arvioinnin vaatimukset täyttäväksi. Arvioinnin keskeisenä lähtökohtana on hyöty ja vaikuttavuus. Artikkelissa korostetaan itsearvioinnin merkitystä arvioinnin vaikuttavuuden lisäämisessä.

Heikki K. Lyytinen ja **Pentti Nikkanen** tarkastelevat artikkelissaan **Kehittämisoimaa arvioinnista** arviointia yksilöiden ja yhteisöjen voimaantumista edistävänä tekijänä. Kirjoittajat kiinnittävät voimaannuttavan arvioinnin oppivan organisaation ajattelutapaan. Voimaannuttava arviointi ymmärretään toiminnan kehittämisprosessina - osana itseohjautuvaa toimintaa, päätöksentekoa, muutoksen hallintaa ja hyvinvointia. Arvioinnissa kiinnitetään erityistä huomiota vahvuuksiin ja arvioinnin tietoiseen hyödyntämiseen prosessin eri vaiheissa. Kirjoittajat kiinnittävät suurimmat haasteet arviointitulosten tulkitsemiseen sekä kehittämiseen, itsemääräytymiseen ja tuloksellisuuden optimoimiseen.

Päivi Atjonen tarkastelee artikkelissaan **Arviointi opetuksen ja oppimisen tueksi** arvioinnin tehtäviä, välineitä ja ilmenemismuotoja oppimis-opetusprosessien ja niiden ohjaamisen kannalta. Artikkelissa esitetään kriittinen kanta kvantitatiiviseen arviointitietoon ja indikaattoreihin ja korostetaan kehittävän arvioinnin merkitystä. Artikkelissa nostetaan keskiöön arviointi, joka kartuttaa opettajan pedagogista ajattelua ja sisältötietoa. Artikkelin kirjoittaja suosittelee konstruktivistisen oppimiskäsityksen, optimistisen ihmis- ja kasvatuskäsityksen ja kommunikatiivisen arvioinnin huomioon ottamista arviointikäytäntöjen ja -välineiden kehittämisessä.

Renata Svedlin tarkastelee artikkelissaan **Utvärdering till stöd för lokalt beslutfattande** arviointien aidon vaikuttavuuden merkitystä paikallisella tasolla. Kirjoittaja korostaa kansallisen ja paikallisen tason yhteistyön tärkeyttä jo arviointiprojektien suunnitteluvaiheessa. Artikkelissa on myös ehdotuksia siitä, miten arviointituloksia tulisi paikallisesti hyödyntää.

Olli Hietanen ja **Jari Kaivo-oja** nostavat artikkelissaan **Ennakoivaan arviointiin** esiin arvioinnin ja ennakkoinnin keskinäisen suhteen uutena näkökulmana. Kirjoittajat katsovat, että ennakkoinnilla voidaan parantaa arvioinnin vaikuttavuutta. Kirjoittajien mielestä artikkelissa esitetyt ennakkointimenetelmät sopivat kehittävän arvioinnin ryhmätyömenetelmiksi.



Näytöt kehittämisen perustaksi

Artikkelissa tarkastellaan näyttöön perustuvaa toimintaa ja sen merkitystä koulutusjärjestelmän ohjauksessa, päätöksenteossa ja jatkuvassa kehittämisessä sekä viime kädessä koulutyössä toimivan johtajan ja opettajan toiminnassa. Artikkelin valottaa näyttöön perustuvan toiminnan luonnetta ja muun muassa sitä, millaisia näyttöön perustuvat päätöksenteko-, kehittämis- ja toimintakäytänteet ovat. Näyttöön perustuvan toiminnan hyödyt, edellytykset ja rajoitukset käsitellään niin ikään. Artikkelissa painotetaan näyttöjen hyödyntämistä opettajien työssä ja oppilaitoksen johtamisessa. Artikkelin lopussa tehdään myös ehdotuksia näyttöorientoituneen toiminnan kehittämiseksi. Ydinkysymys artikkelissa on, miten näyttöön perustuva toiminta liittyy koulutuksen arviointiin.

Asiasanat: evidence based -toiminta, näyttöön perustuva toiminta.

Mistä näyttöperustaisessa toiminnassa on kyse?

Näyttöön perustuva käytäntö (*evidence based practice*) on syntynyt lääketieteen piirissä 1990-luvulla. Lääketieteen piiristä se on levinnyt laajemmin muun muassa terveydenhuoltoon ja muille yhteiskunnan alueille, kuten kasvatustieteeseen (Hammersley 2004; Davies 2000; Eitel & Steiner 1999). Näyttö antaa perusteet uskoa ja luottaa, että asiantila on kuvattuunlainen, joskin sillä voi olla erilaisia piirteitä ja arvostuksia sen mukaan, millaisessa yhteydessä ja millä tieteenalalla käsitettä käytetään. Juridikkassa, luonnontieteissä, lääketieteessä tai humanistisissa tieteissä näytön käsitteessä painottuvat erilaiset seikat. Yhteistä kuitenkin on monipuolisen ja luotettavan tiedon sekä perustelujen etsiminen (mm. Thomas 2004).

Näyttö tarkoittaa tietoa, joka on ilmeistä, varmaa ja totta. Vahvimmillaan näyttöperustaisuus sisältää oletuksen, että kulloinkin valittu toimintatapa, tehty päätös tai linjaus perustuu todennettuun tietoon ja todistusaineistoon. Näyttöperustaisuuteen sisältyy myös oletus, että toiminta ja sen perusteet ovat läpinäkyviä ja siten uskottavasti todistettavissa ja arvioitavissa. Toiminnan ja sen kehittämisen näyttöperustaisuus merkitsee vaatimusta valita kulloinkin paras ja vaikuttavin toimintatapa ja etenemisen tie.

Näyttöön perustuva toiminta voidaan määrittää koulutusympäristössä huolellisesti valitun ja parhaan ajan tasalla olevan tiedon käytöksi päätöksenteossa, joka koskee yksittäisen opiskelijan, opiskelijaryhmän tai väestön elinikäistä oppimista ja koulutusta ja koko koulutusjärjestelmää ja sen eri tasoja. Tietoa käytetään päätöksenteon lisäksi myös itse toiminnassa. (Vrt. Saclett ym. 1996; Thomas 2004.) Näytön arvioinnin kriteereitä voivat Thomasin mukaan (emt.) olla **relevanssi riittävyys ja totuudellisuus**. Lisäksi näytön voimakkuuden arvioinnissa tulee ottaa huomioon sosiaalinen ja tulkinnallinen konteksti, jossa näyttöä kulloinkin arvioidaan. Näyttö voi vaihdella inspiraatiosta, tunteesta ja aavistuksesta rationaaliseen uskoon ja tieteellisesti todennettuun tietoon.

Näytön lajit

Näyttö voidaan jakaa kolmeen pääluokkaan (Leino-Kilpi & Lauri 2003; Clarke 1999; Sebba 2004; Eraut 2004). *Näytön perustaksi tarvitaan joko tutkimustietoa (tieteellinen näyttö) ja/tai muuta sekä kokemukseen, monipuoliseen tietoon että asiantuntijoiden käsityksiin perustuvaa tietoa (hyväksi havaittu toimintanäyttö) ja/tai asiakkaan omaan tietoon ja kokemuksiin liittyvää tietoa (kokemukseen perustuva näyttö).*

Tieteellisesti havaittu tutkimusnäyttö tarkoittaa tieteellisiin kriteereihin perustuvan tutkimuksen avulla saavutettua näyttöä toiminnasta tai sen vaikuttavuudesta. Tätä tietoa tuotetaan muun muassa *koulutuksen arviointitutkimuksen avulla*.

Tieteellisesti havaittu tutkimusnäyttö tuotetaan eri menetelmien avulla sen mukaan, millaiseen tiedon yleistettävyyteen pyritään. Näytön tuottaminen voi perustua erilaisiin lähestymistapoihin, ja tietoa voidaan tuottaa sekä kvantitatiivisin että kvalitatiivisin menetelmin. Perinteisesti kvantitatiivisen tutkimuksen tuottamaa yleistettävää tietoa ja laajojen kirjallisuuskatsausten tuottamaa tietoa on pidetty vahvan näytön merkinä. Sen sijaan kvalitatiivisen tutkimuksen pieniin otoksiin perustuvan tiedon näyttöastetta on pidetty edellistä heikompana.

Näyttöperustaisen koulutuksen kehittämisen ja päätöksenteon perustaksi tarvitaan enenevästi erilaiseen tutkimustraditioon perustuvaa tietoa. Tutkimustiedon lisäämisen tarve on perusteltavissa ensiksi siksi, että päätöksenteon ja kehittämisen perustaksi ei useinkaan ole saatavissa tutkimustietoa. Toinen peruste on, että koulutus alana ja

järjestelmänä on niin dynaaminen, moniulotteinen ja kompleksinen, ettei se tule ymmärretyksi yksipuolisten tutkimusstrategioiden ja lähestymistapojen avulla. Muun muassa koulutus- ja kasvatustoiminnan moniulotteisuudesta, kompleksisuudesta ja dynaamisuudesta johtuen perinteiseen mittaamiseen ja usein vahvaan näyttöön johtavien koe- ja kontrolliasetelmien luominen on melko vaikeaa. Koulutuksen alueella ei myöskään riitä, että tutkimus kohdistetaan opettamiseen ja oppimiseen, vaan vankkaa näyttöä tarvitaan myös koko koulu yhteisön toiminnasta, yhteistyösuhteista ja alueellisista kytkennöistä sekä koko ohjaus- ja toimintajärjestelmän toimivuudesta.

Hyväksi havaitun toiminnan näyttö voidaan osoittaa systemaattisesti kootun tiedon sekä muun muassa laadunarvioinnin avulla. Hyväksi havaitun toiminnan näytön osoittamiseen ei vaadita tutkimustiedon edellyttämiä tieteellisiä kriteereitä, vaikka siltäkin vaaditaan systemaattisuutta ja luotettavuutta. Esimerkkejä tämänkaltaisista näytöistä ovat erilaiset arvioinnit, määrälliset tunnusluvut ja muutoin eri yhteyksistä toistuvasti kerätyt tiedot. Eri yhteyksissä tehdyt konsensusarvioinnit ja vertaisarviointit kuuluvat tämän näytön alueeseen. Sebban (2004) mukaan myös tarkastustoiminnan tuottamat tiedot kuuluvat tämän näyttökategorian piiriin. Konsensuslausumien näytön vahvuutta voidaan lisätä perustamalla kyseiset arvioinnit tutkimuksiin, systemaattisesti kerättyihin tietoihin, monipuolisiin kokemuksiin ja ulkopuolisten asiantuntijoiden lausuntoihin (vrt. myös Eraut 2004). Niiden luonteeseen kuuluu pyrkimys löytää systematiikkaa kaikkien edellä mainittujen pohjalta ja esittää se konsensusmuodossa.

Kokemukseen perustuva näyttö nojaa käytännön kokemuksiin ja niiden kautta syntyneisiin käsityksiin ja mielikuviin. Toisaalta tässä ajattelutavassa näytöksi riittävät omat mielipiteet ja käsitykset käytännössä hyväksi havaituista toiminnoista sekä tarpeista ja odotuksista, joita oppilaat ja opiskelijat sekä heidän vanhempansa kohdistavat toiminnan organisointiin. Omat käsitykset voivat perustua muun muassa pitkäaikaiseen työkokemukseen ja sen myötä kehittyneeseen asiantuntijuuteen (eksperittisyys), työyhteisön toimintakulttuuriin sosiaalistumiseen ja ulkopuolisten tahojen työhön kohdistamiin odotuksiin. Koulutuspalveluiden käyttäjän – perusopetuksessa olevan lapsen ja hänen vanhempiansa, ammatillista koulutusta käyvän nuoren, eri elämänvaiheissa opiskelevan aikuisen ja koulutuspalveluita käyttävien muiden tahojen kuten työelämän organisaatioiden – käsitys “ideaalisesta ja todellisesta näytöstä” perustuu myös usein omakohtaisiin kokemuksiin tai mielikuviin koulutus- ja opetuskäytännöistä.

Kokemukseen perustuva näyttö voi ilmentää toimintakulttuurissa vallalla olevaa vakiintunutta tapaa sekä tiedostettuja ja tiedostamattomia odotuksia, käsityksiä ja mielikuvia. Perusteena voivat olla mielikuvat, miten tulee toimia ja millainen toiminta tuottaa parhaimmat tulokset. Työyhteisöjen kulttuuriperimä saattaa siirtyä sukupolvelta toiselle ilman, että toiminnan tarkoituksenmukaisuutta tai sen vaikuttavuutta kyseenalaistetaan muuttuneissa olosuhteissa. Tämänkaltaisen kokemukseen perustuva näyttö ei myöskään yleensä perustu mihinkään tiedossa olevaan kriteeriin.

Mitä näyttöperustainen toiminta voisi merkitä koulutuksen alueella?

Näyttöperustaisuuden vaatimus on voimistunut koulutusjärjestelmää koskevassa yhteiskunnallisessa keskustelussa viimeisen kymmenen vuoden aikana. Keskustelun ytiminä ovat muun muassa olleet koulutuksen voimavarat ja saavutetut tulokset sekä näiden väliset yhteydet. Lisäksi keskustelu koulutusjärjestelmän sekä siihen liittyvän päätöksenteon ja toiminnan läpinäkyvyydestä on lisääntynyt. Edelleen myös tuloksellisuuden ja tuottavuuden vaatimukset sekä väestön tasa-arvoon ja alueelliseen eriarvoistumiseen liittyvät kysymykset ovat olleet keskusteluiden asialistalla.

Koulutuksen alueella ei kuitenkaan terveydenhuollon tapaan ole vielä kovin näkyvästi nostettu esiin näyttöperustaisuuden vaatimuksia ja mahdollisuuksia. Tämä näkökulma voisi tuoda uusia avauksia keskustelun lisäksi myös päätöksentekoon, kehittämiseen, arviointiin ja tutkimukseen, eri toimijoiden rooleihin ja tehtäviin sekä ylipäätään koulujen, oppilaitosten ja siellä toimivien professioiden toimintaan. Näyttöön perustuva toiminta vaatii tarkastelemaan kriittisesti myös ohjausmekanismeja.

Onko edellytyksiä enemmän kuin rajoituksia ja esteitä?

Edellytyksiä

Näyttöön perustuvan toiminnan edellytys on, että tutkittua, arvioitua ja analysoitua tietoa on kattavasti, ja se on helposti saatavilla. Lisäksi edellytys on, että tätä tietoa käytetään tietoisesti hyväksi päätöksenteossa, johtamisessa, kehittämisessä ja ylipäätään omaan työhön liittyvässä päätöksenteossa. Muut edellytykset liittyvät toimintakulttuuriin, osaamiseen, tutkimus- ja arviointimyönteisiin asenteisiin, esimiehien ja kollegoiden kannustukseen, aineellisiin voimavaroihin ja koko infrastruktuuriin. (Esim. Oranta, Routasalo & Hupli 2002.)

Esteitä

Monet tutkijat (esim. Sebba 2004; Gallagher 2004; Kuuppelomäki & Tuomi 2003) pitivät suurimpana näyttöön perustuvan toiminnan esteenä sitä, ettei aiemmin tutkittua ja arvioitua tietoa joko ole saatavissa tai sitä on liikaa tai että tieto on vaikeasti hyödynnettävissä. Hyödyntämisen ongelmat johtuvat usein myös vaikeudesta ymmärtää tutkimus- ja arviointitietoa. Tutkimus- ja arviointitulosten hyödyntämistä voivat haitata myös tutkimusten ja arviointien heikko taso ja niiden kohdentuminen käytännön työn kannalta epäolennaisiin seikkoihin. Epäselvä näytön kuvaaminen estää näytön hyödyntämisen kehittämisessä ja päätöksenteossa. Yhtenä syynä pidetään myös sitä, etteivät tutkijat ja arvioijat panosta raportoinnin ja tulosten levittämisen yhteydessä

tulosten käytäntöön soveltamiseen (esim. Sebba 2004).

Näyttöperustaisen toimintakulttuurin kehittämisen esteitä löytyy myös organisaatiokulttuurista, vallitsevista arvoista, asenteista, uskomuksista, motiiveista sekä tutkimukseen, arviointiin ja tulosten hyödyntämiseen liittyvien taitojen puutteista. Ongelmana voi olla johdon sitoutumattomuus, johdon ja kollegoiden tuen puute tai henkilöstökoulutuksen heikkoudet. Lisäksi riittämättömät voimavarat oman toiminnan arviointiin ja tutkimiseen, muualla tehtyyn tutkimus- ja arviointitietoon perehtymiseen ja hyödyntämiseen ja innovatiiviseen kehittämiseen aiheuttavat myös ongelmia. Lisäksi ongelmia voi aiheuttaa näyttöjen dokumentoinnin, seurannan ja arvioinnin vähäisyys, tietojärjestelmien hataruus ja moniammatillisen yhteistyön ohuus.

Työyhteisön kulttuurinen este on myös mieltymys pitäytyä totutuissa käytänteissä. Arviointi- ja tutkimustiedon merkitystä voidaan jopa vähätellä kehittämisen lähtökohdana ja välineenä. Esteitä voivat luoda myös epäselvyydet, jotka koskevat koko yhteisön ydintehtävää ja ammattikuntien rooleja ja vastuita. Tällöin voi myös hämärtyä, millä alueella oma näyttö pitää ensisijaisesti saada esiin. Lisäksi este voi olla työntekijän puutteellisissa mahdollisuuksissa muuttaa käytäntöjä. Näyttöön perustuvan johtamisen esteenä on Peiposen (2001) mukaan hallinnossa toimivien henkilöiden kokemattomuus etsiä ja käyttää tutkittua tietoa päätöksenteon pohjana, ei niinkään tutkimustiedon puute. Vakiintuneita käytänteitä suositusten antamiseen näyttöjen perusteella ei myöskään ole.

Rajoituksia

Vahvoin näyttöihin perustuvassa toiminnassa korostuu tutkitun/arvioidun tiedon ylivoimaisuus. Ajattelumalli voi synnyttää mielikuvan toiminnan ja siihen liittyvän päätöksenteon liiallisesta rationaalisuudesta, joka perustuu kirjoitettuun tietoon. Näin ei kuitenkaan ole. Tutkimus- ja arviointitiedon tuottaminen ja näiden tulosten hyödyntäminen ei voi estää asiakaslähtöisten toimintatapojen ja luovien ongelmanratkaisutoiminta- ja ajattelumallien kehittämistä. Se ei myöskään ole este uusien käytänteiden kokeilulle tai syvällisen käytännön asiantuntemuksen (eksperttiys) ja sen myötä syntyvän hiljaisen tiedon ja siihen sisältyvän intuition ja luovuuden hyödyntämiselle niin päätöksenteossa, kehittämisessä kuin myös päivittäisessä toiminnassa. Päätöksenteossa ja toiminnassa voi muodostua jopa ongelmalliseksi se, että luotetaan liian yksipuolisesti määrällisen tutkimuksen tuottamaan tietoon ja väheksytään muunlaista, joko omaa tai ulkopuolisten asiantuntijoiden ja toimintaan jollakin tavalla kytkeytyvien tahojen kokemuseräistä tietoa (mm. Marks-Maran 1999; Hammersley 2004; Thomas 2004). Olennaista on muistaa, että kunkin näytön vahvuuteen vaikuttaa aina myös näytön oma konteksti (Pring 2004).

Millaista näyttöön perustuva toiminta on?

Näyttöön perustuva päätöksenteko

Koulutusjärjestelmän eri tasoilla tehdään jatkuvasti päätöksiä. Sen lisäksi, että näyttöperustaisuuden osoittamiseksi tarvitaan tietoa päätöksentekijöiden toiminnasta, tarvitaan tietoa myös siitä, mihin heidän päätöksensä perustuvat. On melko vaikea tehdä johtopäätöksiä arviointi- ja tutkimustiedon vaikutuksista päätöksentekoon muun muassa vaikutusten arviointimekanismien kehittymättömyyden vuoksi. Silti monien tutkijoiden mukaan (mm. Lumijärvi & Kaarlejärvi 2000; Hillage et al. 1998; Sebba 2004) on selvää näyttöä siitä, että poliittisessa päätöksenteossa hyödynnetään hyvin vähän tutkimusten ja arvioinnin avulla tuotettua tietoa.

Perustaakseen päätöksensä näyttöön poliittinen päätöksentekijä tarvitsee Sebban (emt.) mukaan *ensiksi* seurannan mahdollistavia tilastotietoja ja erilaisten tietojen yhdistämiseen perustuvia arviointi- ja ennakointianalysejä (survey). Nämä tiedot ovat pääosin kvantitatiivisia. *Toiseksi* poliittinen päätöksentekijä tarvitsee sekä arviointiin että mahdollisesti tarkastustoimintaan liittyvää tietoa, joka kohdistuu päätösten toimeenpanoon ja päätöksentekoon sinänsä. *Kolmanneksi* poliittinen päätöksentekijä tarvitsee riippumattomien tutkijoiden ja arvioijien tuottamaa arviointi- ja tutkimustietoa koko järjestelmän ja siellä olevien yksiköiden toimivuudesta.

Tutkimus- ja arviointitiedon hyödyntämiseksi tarvitaan monien tahojen yhteistyötä sekä välittäviä mekanismeja ja foorumeita, jotka edistävät hyödyntämistä. Sebba (2004) erottaa tulosten levittämisen ja vaikutukset toisistaan. Levittämisessä on kyse prosessista, jossa tulokset saatetaan ulkopuolisten tietoon. Levittämisen avulla voidaan vaikuttaa tietoon, asenteisiin ja uskomuksiin. Levittäminen tehdään tavallisesti raporteina ja konferensseina tutkimuksen jälkeen.

Tutkimusten ja arvioinnin vaikuttavuutta voidaan lisätä strategisesti oikeilla valinnoilla. Ongelmallista voi olla, että levittäminen on tehokasta, mutta vaikuttavuus heikkoa. Tästä syystä huomio on kohdistettava uuden tiedon tuottamisen, levittämisen ja käytön väliseen yhteyteen (esim. Sebba 2004). Vaikuttavuutta ja muutosten toteutumista voidaan edistää kehittämällä erilaisia eri tahoja osallistavia prosesseja ja turvaamalla muutosten edellyttämät voimavarat. Arviointi- ja tutkimustiedon saatavuus, oikea-aikaisuus ja hyödynnettävyys käytännössä on myös päätöksenteossa hyödyntämisen välttämätön ehto.

Keskeistä on myös kiinnittää huomiota, miten arviointi- ja tutkimustieto vaikuttaa epäsuorasti poliittisiin päätöksentekijöihin. On myös muistettava, että arviointi- ja tutkimustiedon lisäksi päätöksentekijällä on – ja tulee olla – myös muuta tietoa käytettävissään päätöksentekonsa perustana.

Näyttöön perustuva kehittäminen

Vahvaan näyttöön perustuvan toiminnan kehittäminen nojaa samoihin lähtökohtiin kuin edellä kuvattiin päätöksenteon osalta. Tämän lisäksi arviointi- ja tutkimustieto voi muodostaa kivijalan tuloshakuiselle jatkuvalla parantamiselle ja uusien tavoitteiden määrittämiselle. Näytön osoittamisessa voidaan käyttää erilaisia arviointimenetelmiä (esimerkiksi Euroopan laatupalkintokriteeristö).

Tuloshakuisuus voidaan osoittaa muun muassa siten, että tehdään näkyväksi, miten tavoitteet, jotka organisaatio itse tai ulkopuolinen taho (esim. eduskunta, valtioneuvosto, opetusministeriö) on toiminnalle asettanut, on saavutettu. Tuloshakuinen organisaatio käyttää ennakointi-, tutkimus- ja arviointitietoa organisaation toimintaperiaatteiden, strategioiden, toimintasuunnitelmien ja niihin sisältyvien tavoitteiden asettamisen perustana. Lisäksi tietoa käytetään organisaation toiminnan laadun pitkäjänteiseen parantamiseen.

Kumulatiivisen näytön merkitys

Arviointien ja tutkimusten vaikuttavuuden lisäämistarpeen myötä nousee myös esiin kumulatiivisen näyttöperustan tarve (vrt. Sebba 2004; Hillage et al. 1998). Kumulatiivisen tiedon tarpeen puolesta esitetyt argumentit liittyvät *ensiksi* tutkimus- ja arviointitiedon tuottamisen ennakoitavuuteen (mitä on tutkittu/arvioitu aiemmin ja mitä tullaan tutkimaan/arvioimaan). *Toiseksi* heikkoon vaikuttavuuteen johtavien yksittäisten tutkimusten sijaan tulisi painopisteen olla laajoissa kirjallisuus- ja tutkimuskatsauksissa, jotka tuottavat synteesejä käytännön toimijoille ja päätöksentekijöille. *Kolmanneksi* kumulatiivinen näyttöperustaisuus mahdollistaa yksittäisten tutkimusten/arviointien kriittisen arvioinnin ja niiden suhteuttamisen muiden tutkimusten/arviointien tuottamiin evidensseihin. *Neljänneksi* on tarvetta kehittää menettelyitä, joiden kautta jokin taho voisi luoda tutkimustulosten perusteella ennusteita ja näkökulmia tulevaisuuteen. Tulevaisuuden ennakointi edellyttää painopisteen muutosta *reaktiivisesta toiminnasta proaktiiviseen toimintaan*. Tämä on myös haaste luoda yhteyksiä ja synergiaa ennakointi- ja arviointitoiminnan välille.

Näyttöön perustuva toiminta paikallisella tasolla

Näyttöön perustuvan toiminnan vahvistaminen paikallisella tasolla edellyttää edellä kuvatun kaltaista kehitystä myös paikallisella tasolla. Näyttöön perustuva toiminta edellyttää paikallisilta toimijoilta kykyä ja halua tutkimus- ja arviointitiedon hankintaan, seurantaan ja analysointiin. Lisäksi se edellyttää kykyä ja halua käyttää tätä tietoa henkilöstön koulutuksessa, omien tuotteiden, palveluiden ja työyhteisön kehittämisessä, kehittämis- ja kokeiluhankkeissa sekä toimintamallien ja menetelmien uudistamisessa ja vakiinnuttamisessa. Johdon tehtävänä on mahdollistaa sekä henki-

löstön että käytänteiden kehittyminen. Toimintakulttuurin kehittäminen edellyttää jatkuvaa arviointia.

Näyttöön perustuvassa opettajuudessa on kyse siitä, että opettaja ja opettajakollegiivi hyödyntää aktiivisesti tutkimuksia ja arviointeja sekä asiantuntijoiden näkemyksiä omassa työssään ja siihen liittyvässä päätöksenteossa (mm. Eraut 2004).

Näyttöön perustuva opettajuus

Opettaja tai opettajayhteisö, joka on orientoitunut näyttöön perustuvan ajattelutavan mukaan, joutuu etsimään sekä pedagogiikkaan että opetuksen sisältöön liittyviä näyttöjä. Näyttöä voidaan etsiä tutkimusten ja arviointien lisäksi muusta kirjallisuudesta, ammattilehdistä, tietokannoista, tilastoista, mahdollisista suosituksista ja alan asiantuntijoilta. Näyttöön perustuvassa opettajuudessa korostuu kyky käyttää muiden tekemiä tutkimuksia ja arviointeja sekä kyky arvioida omaa toimintaa (antaa näyttö). Osaamisalueista muita painavampia ovat tietojen hankinta-, analysointi- ja hyödyntämistäidot sekä arviointiosaaminen (Cordingley 2004; Eitel & Katz & Tesche 2001; Elomaa 2001; Kalpio 2004; Leino-Kilpi 2001). Olennaista on myös, miten opettaja kykenee yhdistämään tutkitun ja kokemusperäisen tiedon, ja miten hän osaa hyödyntää uutta tietoa työssään (Kalpio 2004; Leino-Kilpi 2001).

Näyttöön perustuva opettajuus haastaa opettajan sekä käyttämään tutkimuksen ja arvioinnin avulla tuotettua tietoa että tuottamaan tätä tietoa omasta ja koko yhteisön työstä. Näyttöperustaisuus haastaa myös kehittämään opettajan mahdollisuuksia toimia oman työnsä tutkijana ja arvioijana. Tulevaisuuden haaste on myös se, että päättäjät, tutkijat ja opettajat kehittävät yhteistyöhön ja verkostoitumiseen perustuvia toimintamalleja lisätäkseen arvioinnin ja tutkimuksen monitahoisuutta ja vaikuttavuutta (Torrance 2004). Tutkivan opettajan erityispiirteitä ovat aktiivisuus, jatkuva oppiminen, opiskelu ja itsensä kehittäminen, yhteistyö ja yhteisöllisyys ja kansainvälisyys sekä oman toiminnan jatkuva reflektointi. Omaa työtään tutkiva ja arvioiva opettaja kokeilee ja kehittää omaa työtään ja on kiinnostunut olemaan mukana tutkivien opettajien yhteistyöverkostossa (Korpinen 2003).

Opettajuuden arviointi

Näyttöön perustuva opettajuus näyttää nousevan yhdeksi 2000-luvun keskeisimmistä haasteista. Nolanin ja Hooverin (2004) mukaan opettajuuteen liittyvän näytön arviointijärjestelmä sisältää kolme osiota: *opettajan toiminnan laatuvaatimukset, arviointiperiaatteet ja arviointimenettelyt* (myös Kalpio 2004). *Opettajan toiminnan laatuvaatimukset* on kuvattu suosituksenomaisesti laatustandardina:

- TASO 1 Selkeä näyttö siitä, että opettajan toiminta täyttää laatuvaatimukset.
- TASO 2 Epäselvä näyttö, jolloin arvioijan tulee yhdessä opettajan kanssa yrittää etsiä parempaa, sopivampaa näyttöä, jota voidaan käyttää arvioinnin pohjana.
- TASO 3 Selkeä näyttö siitä, että opettajan toiminta ei ole laatuvaatimusten mukaista.

Nolanin ja Hooverin (emt.) esittämät opettajan toiminnan arvioinnin periaatteet on laadittu sovellettavaksi lasten opetuksen kontekstissa. Heidän mukaansa opettajan toiminnan arvioinnissa voisivat pääpiirteissään painottua seuraavat näkökohdat:

- Opettajan näyttöjen arviointi on laaja-alaista ja se kattaa kaikki opettajan vastualueet.
- Luotettavan kuvan saamiseksi tarvitaan monipuolista tietoa opettajan näytöistä.
- Pätevät, koulutetut esimiehet ovat sopivia arvioimaan opettajien näyttöjä.
- Kaikille opettajille taataan arviointiin pohjautuva jatkuva ammatillinen kehittyminen.
- Opettajan toiminnan arviointi on osallistavaa ja avointa eri sidosryhmille.
- Opettajan toiminnan arviointi on syvälistä ja se kohdistuu sekä työn sisältöön että opetukseen.
- Opettajien oikeuksia tulee arvioinnissa suojella.
- Kokoneiden opettajien ja noviisiopettajien arviointiprosessit ovat erilaisia.

Arviointimenettelyesimerkkinä Nolan ja Hoover (emt.) käyttävät lomaketta (liite 1). Myös Darling-Hammondin (1989) mukaan opettajan toiminnan arvioinnin tulee olla monitahoista, ja siinä voidaan käyttää erilaisia menetelmiä (esim. *opettajien pätevyys ja vaikuttavuus, vertaisarviointit, koulun itsearviointi, opettajien näytöt, opettajien haastattelu, opiskelijoiden arviot ja opiskelijoiden saavutukset sekä epäsuorat muut mittarit*). Muita opettajan toiminnan arviointimenettelyitä voivat olla sidosryhmien käyttö arvioinnissa, opettajan opetusmateriaaliin perustuva arviointi, opetustilanteiden videointi arvioinnin pohjaksi, opiskelijoiden ja vanhempien kokemukset, opettajan ammatillisen kasvun dokumentit (esim. portfoliot) sekä esimiehen ja opettajan yhdessä tekemä kokoava arviointi kirjallisine raportteineen (Nolan ja Hoover 2004).

Edellä kuvatun arviointimenettelyn soveltaminen ammatillisessa koulutuksessa tai opettajakoulutuksessa edellyttää laatuvaatimusten ja kriteereiden muokkaamista näiden kontekstien vaatimusten mukaisesti. Kontekstit eroavat muun muassa oppimisympäristöjen (esim. työelämä), sisällön sekä opiskelu- ja opetusjärjestelyiden suhteen.

Näyttöön perustuva johtaminen

Näyttöön perustuvassa johtamisessa ja päätöksenteossa on kyse systemaattisesti hankitun ja luotettavana pidetyn tiedon käytäntöön soveltamisesta. Johtajien tehtävänä on etsiä tietoa tutkimustuloksista ja muista varmoista lähteistä johtopäätöksensä perustaksi (Simoila 2001). Johtajien päätösten läpinäkyvyys myös lisääntyy, kun he pystyvät osoittamaan, millaista tietoa he ovat käyttäneet päätöksiä tehdessään. Lisäksi johtajien on luotava arvioinnille, tutkimukselle, jatkuvalle parantamisella ja yhteisön oppimiselle otolliset olosuhteet ja edellytykset. Johtajien on myös luotava edellytyksiä sille, että henkilöstö voi toimia omassa työssään näyttöön perustuvasti. Tämä puolestaan edellyttää henkilöstöltä kykyä hyödyntää tutkittua ja arvioitua tietoa, mahdollisuuksia päästä tiedon lähteille, aikaa tutkitun ja arvioinnin avulla tuotetun tiedon analysointiin, mahdollisuuksia osallistua tutkimus- ja arviointihankkeisiin sekä työyhteisön yhteisiä mahdollisuuksia oman työn reflektointiin, kyseenalaistamiseen sekä toisilta oppimiseen (empowerment-kulttuuri).

Näyttöön perustuvan johtamisen suurimmat haasteet liittyvät koko toiminta- ja johtamiskulttuurin muuttamiseen tutkimus- ja arviointiorientoituneeksi. Tämä edellyttää muun muassa päämäärä- ja tavoitetietoisuutta, tietojärjestelmien tarkoituksenmukaisuutta ja ajantasaisuutta, tulosten mittaamista (esim. määrälliset ja laadulliset indikaattorit ja kyselyt), kehityksen seuranta ja kriittiseen pohdintaan ja oppimiseen kannustavaa ilmapiiriä sekä laadun jatkuvaa parantamista. Olennaista johtamisessa on myös, että tulosten mittaaminen on suhteutettu yksikön ja siellä toimivien omiin tarpeisiin.

Näyttöön perustuvan johtamisen vaatimuksena on johtajien oma kyky ja halu hyödyntää tutkimus- ja arviointitietoa johtamisessaan. Näyttöön perustuvan johtamisen vaatimuksena on, niin ikään, että johtajat etsivät näyttöä niin ydintoiminnan kehittämisen kuin johtamista koskevan päätöksentekonsa pohjaksi (Peiponen 2001). Näyttöön perustuva johtajuus vaatii onnistuakseen kulttuurimuutosta ja uutta johtajasukupolvea, joka luottaa omiin tiedonhankinta- ja analysointitaitoihinsa. Organisaation ilmapiiriin tulee olla avoin, hyväksyvä, kannustava ja kyselemään rohkaiseva. (Vrt. Oranta, Routasalo & Hupli 2002.)

Eri toimijoiden välisen avoimen dialogin välttämättömyys

Avoin dialogi tutkijoiden, arvioijien, poliittisten päätöksentekijöiden, kehittäjien ja koulutuksen järjestäjien välillä voi edistää positiivisia kokemuksia arvioinnista/tutkimuksesta ja tulosten hyödyntämisestä. Lisäksi se voi edistää ymmärrystä näyttöön perustuvan toiminnan merkityksestä. Samalla dialogi voi lisätä arvioijien ja tutkijoiden syvällistä ymmärrystä poliittisista prosesseista, kehittämisdynamiikasta ja muusta toiminnasta. On luotava yhteisiä foorumeita ja uusia yhteistyöhön perustuvia toimintamalleja. Tarvitaan interventioita, jotka auttavat vastavuoroista oppimista, syväl-

listä ymmärrystä ja ydinprosessien kehittämistä ja niiden sovittamista yhteen toiminnalliseksi kokonaisuudeksi. Tarvitaan uudenlaisia verkostomaisia toimintatapoja. Vastavuoroisuuteen ja kumppanuuteen perustuva verkostomalli voi olla enemmän kuin pelkkä dialogi poliittisten päätöksentekijöiden, käytännön toimijoiden ja tutkijoiden ja arvioijien välillä. Se edellyttää muun muassa päätöksentekijöiden, kehittäji- en ja muiden toimijoiden aktiivista osallistumista ja mahdollisuutta vaikuttaa arviointi- ja tutkimusprosessiin (kysymykset, design, tietojen koonti, tulkinta ja levittäminen) sekä arvioijien ja tutkijoiden osallistumista kehittämisprosesseihin. Opettajien ja rehtoreiden osallistuminen subjekteina tähän dialogiin ja yhteistyöhön poliittisten päätöksentekijöiden ja muiden toimijoiden kanssa korostuu myös.

Näyttöön perustuvat laatusuosituks

Näyttöjen vaikuttavuutta voidaan lisätä laatimalla erilaisia toimintoja koskevia laatusuosituksia (mm. Torrance 2004). Muun muassa terveydenhuoltoala on koulutusala enemmän laatinut arviointi- ja tutkimustietoon perustuvia laatusuosituksia (esim. käypähoitosuosituks

Laatusuositusten perustana on tutkimuksen ja arvioinnin avulla tuotettu validi tieto. Näyttö suositusten perustaksi haetaan ensisijaisesti tutkimuksista ja hyväiksi havaituista toimintanäytöistä (arviointi).

Näyttöperustaisen toimintakulttuurin kehittämis

Näyttöperustaista päätöksenteko- ja toimintakulttuuria voidaan vahvistaa muun muassa siten, että tuotetaan näytöt näkyviksi osoittavaa arviointi- ja tutkimustietoa, parannetaan näyttöjen hyödyntämistä päätöksenteossa, kehittämisessä ja muussa toiminnassa, kehitetään hyvien käytäntöjen levittämistä ja hyödyntämistä ja laaditaan näyttöihin perustuvia laatusuosituksia. Näyttöperustaisen toimintakulttuurin edistäminen voisi tarkoittaa muun muassa seuraavaa:

Päätöksenteko ja kehittäminen

- Syvennetään tutkimus- ja arviointitiedon hyödyntämistä koulutuspoliittisessa päätöksenteossa ja kehittämisessä.

- Parannetaan päätöksentekijöiden ja muiden toimijoiden kykyä hyödyntää tutkimuksia ja arviointeja omassa toiminnassaan.

Arviointi- ja tutkimustiedon tuottaminen

- Varmistetaan, että päätöksenteon ja kehittämisen perustana on käytettävissä arviointi- ja tutkimustietoa. Varmistetaan arviointien ja tutkimusten kattavuus ja oikea-aikaisuus.
- Varmistetaan pitkällä aikavälillä päätöksenteon, kehittämisen ja arvioinnin ja tutkimuksen integroituminen toisiinsa ja parannetaan niiden ajallista vaiheistusta.
- Tehostetaan päätöksenteon, kehittämis-, tutkimus- ja arviointitoiminnan vaikuttavuutta koskevan tiedon tuottamista.
- Luodaan seurantatutkimusten ja -arviointien (follow up) käytänne jatkuvaksi toimintatavaksi.
- Parannetaan arvioijien ja tutkijoiden valmiuksia osoittaa, mitä voidaan tehdä ja mitä pitäisi tehdä. Ei riitä, että kiinnitetään huomiota siihen, mitä pitäisi *tehdä enemmän*, vaan on osattava ja rohjettava *luopua ja tehdä uusia asioita* (muutos) ja *tehdä asioita paremmin kuin aiemmin* (vrt. Dyson ja Desforges 2002).

Yhteiset foorumit ja prosessit

- Luodaan päätöksentekijöiden, kehittäjien, arvioijien ja koulutuksen järjestäjien ja muiden toimijoiden yhteisiä pysyviä foorumeita, jotka tukevat näyttöperustaisen toimintakulttuurin voimistamista.
- Luodaan eri organisaatioiden rajat ylittäviä verkostoituvia toimintamalleja arviointi- ja tutkimustiedon analysoimiseksi, levittämiseksi ja hyödyntämiseksi kehittämisessä.
- Mahdollistetaan koulutuksen järjestäjien ja stakeholdereitten osallistuminen koulutusta ja sen kehittämistä koskevaan päätöksentekoon ja arviointiin.

Julkaisut

- Julkaistaan tietoa, joka koskee näyttöön perustuvaan toimintaan liittyviä peruskäsitteitä ja arviointien ja tutkimusten tuloksia.
- Nimetään vastuuhenkilöt organisoimaan tutkimus- ja arviointitulosten seuranta ja levittämistä. Perustetaan verkosto hakemaan kansallista ja kansainvälistä tietoa ja laatimaan tutkimuksia ja arviointeja koskevia tiivistäviä katsauksia. Kehitetään välineitä tiedon hankintaan.

- Tuetaan toimintakulttuurien kehittymistä kriittisiksi ja reflektiivisiksi ja edistetään myönteisen asenneilmaston kehittymistä tutkimus- ja arviointitiedon hyödyntämistä kohtaan.

Arviointiosaaminen

- Mahdollistetaan tutkijoille ja arvioijille oman arviointi- ja tutkimusosaamisen uudistaminen luomalla edellytyksiä kursseihin osallistumiseen, valmennukseen, mentorointiin, tutkija- ja arvioitsijavaihtoihin, study visit –ohjelmiin sekä erilaisiin kehittämistoimintoihin ja kansainvälisiin arviointihankkeisiin osallistumiseen.
- Parannetaan tutkimus- ja arviointitoiminnan laatua, tehostetaan tutkija- ja arvioijaverkoston toimintaa ja informaatioteknologian hyödyntämistä ja edistetään toiminnan taloudellisuutta.
- Kehitetään arviointi- ja tutkimustiedon tuottamiseen ja levittämiseen tarkoitettuja web-ympäristöjä, laaditaan arviointi- ja tutkimustuloksista synteesejä ja luodaan foorumeita, joissa poliittisilla päätöksentekijöillä on mahdollisuuksia kommunikoida kehittäjien, arvioijien, tutkijoiden ja muiden toimijoiden kanssa.
- Edistetään kehittäjien, koulutuksen järjestäjien ja opettajien arviointi- ja tutkimusosaamisen kehittymistä koulutuksen, arviointeihin osallistumisen, valmennuksen ja erilaisen ohjauksen ja mentoroinnin avulla. Osaamista voidaan kehittää muun muassa luomalla oppimista edistäviä verkko-ympäristöjä, mallintamalla verkosto-osaamista ja sen edellyttämiä toimintamalleja sekä luomalla erilaisen dialogin mahdollistavia foorumeita.

Voimavarat

- Varmistetaan voimavarat arviointi- ja tutkimustiedon tuottamiseen ja hyödyntämiseen (taloudelliset, älylliset, inhimilliset ja materiaaliset voimavarat).
- Kartoitetaan päällekkäisyyksiin liittyvät kriittiset kohdat ja vältetään niitä.
- Allokoidaan koulutuksen nykyisiä voimavaroja siten, että sekä kehittäminen että siihen kiinteästi liittyvä arviointi ja tutkimus ovat mahdollisia. Edellytysten tasapaino edellyttää voimavarojen ohjaamista kehittämisestä arviointiin ja tutkimukseen.
- Voimavarojen uudelleenallokointi edellyttää myös arvioinnin ja tutkimuksen pitämistä tulevaisuudessa nykyistä keskeisempänä ja tehokkaampana kehittämisen mekanismina. Tämä puolestaan edellyttää tutkimus- arviointi- ja kehittämistoiminnan monitahoisuutta ja instituutio- ja toimintarajat ylittäviä uusia toimintamalleja. Lisäksi tarvitaan uusia näkökulmia kvantitatiivisten ja kvali-

tatiivisten lähestymistapojen ja metodien käytön integrointiin – samoin kuin arvioinnin ja arviointitutkimuksen integrointiin.

- Parannetaan kaikkien edellä kuvattujen toimintojen tehokkuutta ja taloudellisuutta, jotta mahdollistetaan voimavarojen uudelleenallokointi.

Laatusuositukset

- Laaditaan arviointi- ja tutkimusnäyttöön perustuvia laatusuosituksia laajoissa monitahoisuutta edustavissa yhteistyöverkostoissa.

Johtaminen

- Kehitetään näyttöön perustuvaa johtamista ja johtamiskulttuuria.

Opettajien ja muiden professioiden toiminta

- Varmistetaan opetushenkilöstön ja muun henkilöstön tutkimus- ja arviointitiedon hyödyntämiseen sekä näyttöön perustuvan toiminnan periaatteiden ymmärtämiseen ja toteuttamiseen liittyvä osaaminen. Osaamisvaatimukset liittyvät muun muassa tutkimusten luvun ja kriittisen arvioinnin hallintaan, tutkimus- ja arviointimenetelmien hallintaan, tietojen hallintaan, muutosten hallintaan sekä kielitaitoon.
- Varmistetaan, että sekä opetuksessa käytetyt menetelmät että opetuksen sisältö perustuvat uusimpaan tutkimus- ja arviointitietoon.
- Kehitetään ydintoiminnan vaikuttavuuden arviointimenetelmiä ja käytänteitä.

Lähteet

- Clarke, J. B. 1999. Evidence-based practice: a retrograde step? The importance of pluralism in evidence generation for the practice of health care. *Journal of Clinical Nursing* 8, 89–94.
- Cordingley, P. 2004. Teachers using evidence: using what we know about teaching and learning to reconceptualize evidence based practice. Teoksessa G. Thomas & R. Pring (toim.) 2004. Evidence-based practice in education. London: Open University Press, 77–90.
- Darling-Hammond, L., Wise, A. E. & Pease. S. R. 1989. Teacher evaluation in the organisational context: A review of the literature. Teoksessa E. R. House (toim.) New directions in educational evaluation. London: Farmer Press, 203–253.
- Davies, P. 2000. Approaches to evidence-based teaching. *Medical Teacher* 22 (1), 14–21.
- Davies, P. 2004. Systematic reviews and the Campbell Collaboration. Teoksessa G. Thomas & R. Pring Evidence-based practice in education. London: Open University Press, 21–33.
- Dyson, A. & Desforjes, C. 2002. Building research capacity: some possible lines of action. Paper commissioned by the National Education Research Forum at www.nerf-uk.org.

- Eitel, F., Kanz, K. & Tesche, A. 2000. Training and certification of teachers and trainers: the professionalization of medical education. *Medical Teacher* 22 (5), 517–526.
- Eitel, F. & Steiner, S. 1999. Evidence-based learning. *Medical Teacher* 21 (5), 506–512.
- Elomaa, L. 2001. Näyttöön perustuva hoitotyö koulutuksessa: tutkimustiedon käyttö ja opetus. Teoksessa P. Voutilainen, H. Leino-Kilpi, T. Mikkola & A. Peiponen 2001. *Hoitotyön vuosikirja 2001. Hygienia*. Helsinki: Tammi, 127–133.
- Eraut, M. 2004. Practice-based evidence. Teoksessa G. Thomas & R. Pring (toim.) *Evidence-based practice*. London: Open University Press 91–101.
- Gallagher, D. 2004. Educational research, philosophical orthodoxy and unfulfilled promises: the quandary of traditional research in US special education. Teoksessa G. Thomas & R. Pring (toim.) 2004. *Evidence-based practice in education*. London: Open University Press, 119–132.
- Hammersley, M. 2004. Some questions about evidence-based practice in education. Teoksessa G. Thomas & R. Pring (toim.) *Evidence-based practice in education*. London: Open University Press, 133–149.
- Hillage, J., Pearson, R., Anderson, A. & Tamkin, P. 1998. *Excellence in research in schools*. London: Department for Education and Employment, Institute for Employment Studies.
- Kalpio, R. 2004. Näyttöön perustuva (evidence based) opettajuus. Teoksessa A. Räisänen *Arvioinnin teoriaa ja käytäntöjä*. Opiskelijoiden kurssitehtävät. Jyväskylän yliopiston kasvatustieteen laitos, 29–36.
- Korpinen, E. 2003. Mikä ihmeen tutkiva opettaja? Teoksessa E. Korpinen & J. Hyvärinen *Tutkiva opettaja*. *Journal of teacher researcher* 3/2003, 4–6.
- Kuuppelomäki, M. & Tuomi, J. 2003. Miksi tutkitun tiedon käyttö on hoitotyössä vähäistä? *Sairaanhoitaja* 12/2003, 11–12.
- Leino-Kilpi, H. 2001. Näyttöön perustuva hoitotyö: onko sitä? Teoksessa P. Voutilainen, H. Leino-Kilpi, T. Mikkola & A. Peiponen *Hoitotyön vuosikirja 2001. Hygienia*. Helsinki: Tammi, 9–20.
- Leino-Kilpi, H. & Lauri, S. 2003. Näyttöön perustuvan hoitotyön lähtökohdat. Teoksessa S. Lauri (toim.) *Näyttöön perustuva hoitotyö*. Helsinki: WSOY, 7–20.
- Lumijärvi, I. & Kaarlejärvi, J. 2000. Ministeriöt arviointitutkimuksen hyödyntäjinä. *Tampereen yliopisto. Hallintotieteen laitos. Hallintotiede* 23.
- Marks-Maran, D. 1999. Reconstructing nursing: evidence, artistry and curriculum. *Nurse Education Today* 19, 3–11.
- Nolan, J. & Hoover, L. 2004. *Teacher supervision and evaluation*. Wiley: John Wiley & Sons.
- Oranta, O., Routasalo, P. & Hupli, M. 2002. Sairaanhoitaja tutkimustiedon hyödyntäjänä – estävät ja edistävät tekijät. *Hoitotiede* 14 (1), 26–37.
- Peiponen, A. 2001. Näyttöön perustuva johtaminen. Teoksessa P. Voutilainen, H. Leino-Kilpi, T. Mikkola & A. Peiponen, A. 2001. *Hoitotyön vuosikirja 2001. Hygienia*. Helsinki: Tammi, 120–126.
- Pring, R. 2004. Conclusion: evidence based policy and practice. Teoksessa G. Thomas & R. Pring *Evidence-based practice in education*. London: Open University Press, 201–212.
- Saclett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B. & Richardson, W. S. 1996. Evidence-based medicine: what it is and what isn't it. *British Medical Journal* 312, 71–72.
- Sebba, J. 2004. Developing evidence-informed policy and practice in education. Teoksessa G. Thomas & R. Pring *Evidence-based practice in education*. London: Open University Press, 34–43.
- Simola, R. 2001. Hoitotyön johtaja näyttöön perustuvan hoitotyön mahdollistajana. *Ylihoitajalehti* 29 (2), 4–14.
- Terveyttä ja hyvinvointia näyttöön perustuvalla hoitotyöllä. Kansallinen tavoite- ja toimintaohjelma 2004–2007. Sosiaali- ja terveysministeriön julkaisuja 18. Helsinki.
- Thomas, G. 2004. Introduction: evidence and practice. Teoksessa G. Thomas & R. Pring (toim.) 2004. *Evidence-based practice in education*. London: Open University Press, 1–20.

- Thomas, G. & Pring, R. 2004. Evidence based practice in education. Teoksessa G. Thomas & R. Pring (toim.) Evidence-based practice in education. London: Open University Press, 21–33.
- Torrance, H. 2004. Using action research to generate knowledge about educational practice. Teoksessa G. Thomas & R. Pring (toim.) Evidence-based practice in education. London: Open University Press, 187–200.

Liite 1. Nolanin ja Hooverin opettajan toiminnan arvioinnin laatustandardit ja esimerkkejä indikaattoreista (Nolan & Hoover 2004; myös Kalpio 2004)

1. Opettajalla on syvä ymmärrys opettamansa aiheen sisällöstä

- A. Ymmärtää tieteenalansa keskeiset käsitteet ja tutkimusmenetelmät tarkasti
- B. Esittää sisällön selkeästi opiskelijoille
- C. Liittää sisällön opiskelijoiden kokemuksiin
- D. Ajantasaistaa tietonsa alansa kehityksen mukaan

2. Opettaja kehittää pikän aikavälin ja päivittäisen suunnitelman, joka on opiskelijoille sopiva ja pohjautuu sekä alueelliseen että valtakunnalliseen opetussuunnitelmaan

- A. Laatii selkeät ja saavutettavat oppimistavoitteet
- B. Jakaa oppitunnit sopiviin osioihin
- C. Perustaa tuntisuunnitelmansa opiskelijoiden aiempaan tietoon ja taitoon
- D. Käyttää toimintatapoja, jotka sopivat sekä kokonais- että osatavoitteisiin
- E. Käyttää erilaisia resursseja, kuten teknisiä apuvälineitä opetuksensa suunnittelussa

3. Opettaja käyttää monipuolisia opetusmenetelmiä, mm. tehokasta kyselytekniikkaa, osallistaakseen opiskelijat oppimiseen ja edistääkseen sisällön syvällistä ymmärtämistä

- A. Käyttää monipuolisia opiskelijoiden kehitystasoon sopivia strategioita
- B. Käyttää toimintoja, jotka tähtäävät syvälliseen, ei vain pinnalliseen oppimiseen
- C. Käyttää erilaisia tapoja sisällön esittämisessä, mm. esimerkkejä, analogioita, demonstraatioita sekä taidollisia että tiedollisia aktiviteetteja
- D. Oivaltaa “opetukselliset tuokiot”, jotka maksimoivat opiskelijan sitoutumisen
- E. Auttaa opiskelijoita tunnistamaan opiskeltavan aihepiirin tärkeyden
- F. Kommunikoi opiskelijoiden kanssa selkeästi ja tehokkaasti
- G. Käyttää tehokasta kyselytekniikkaa edistääkseen ja arvioidakseen opiskelijoiden sisällön ymmärtämistä
- H. Sisällyttää teknologian sopivalla tavalla opetukseensa

4. Opettaja arvioi opiskelijan oppimista säännöllisesti, antaa palautetta sekä käyttää palautetta opetuksensa suunnittelun pohjana

- A. Käyttää monipuolisesti sekä muodollisia että epämuodollisia arviointitekniikoita, kuten kokeita, pistokokeita, projekteja, tarkistuslistoja ja luetteloita
- B. Arvioi opiskelijan ymmärtämistä jatkuvasti opetuksen aikana
- C. Käyttää opetussuunnitelman mukaisia ja opiskelijoiden kehitystasoon sopivia arviointimenetelmiä
- D. Antaa laadukasta palautetta sopivalla tavalla

- E. Antaa tukiovetusta, jos siihen on arviointitiedon perusteella tarvetta
- F. Ylläpitää ajan tasalla olevaa arviointirekisteriä ja käsittelee rekisteriä luottamuksellisesti

5. Opettaja luo myönteisen oppimisympäristön, joka on hyvin järjestetty ja oppimista edistävä ja jolla on selvä vaikutus opiskelijan käyttäytymiseen

- A. Luo luokkahuoneeseen ilmapiirin, joka on opiskelijaa kunnioittava ja kasvattava
- B. Luo rutiineja ja toimintatapoja, jotka tukevat oppimista
- C. Keskittyy keinoihin, joilla saa opiskelijat sitoutumaan tuloksekkaaseen oppimiseen
- D. Käyttää sopivasti siirtymävaiheita maksimoidakseen opetuksellisen ajan
- E. Kohdatessaan epäasiallista käyttäytymistä puuttuu asioihin siten, että se auttaa opiskelijaa kontrolloimaan omaa käyttäytymistään.

6. Opettaja ymmärtää ja ottaa huomioon opiskelijoiden kulttuurisen erilaisuuden sekä erityistarpeet

- A. Eriyttää opetusta kohdatakseen kaikkien opiskelijoiden tarpeet ja vahvuudet
- B. Tarjoaa opiskelijoille mahdollisuuden tarkastella opittavia asioita monesta näkökulmasta
- C. Toimii yhteistyössä avustavan henkilökunnan kanssa mukauttaakseen opetuksensa erityisopiskelijoiden tarpeisiin
- D. Kannustaa ja arvostaa kaikkia opiskelijoitaan

7. Opettaja on elinikäinen oppija, joka etsii mahdollisuuksia reflektoida toimintaansa sekä kasvaa ammatillisesti

- A. Reflektoi kysyessään toiminnastaan ja parantaa toimintaansa sen pohjalta
- B. Käyttää opiskelija- ja kollegapalautetta parantaakseen omaa opetustoimintaansa
- C. Etsii ja hyödyntää mahdollisuudet ammatilliseen kasvuun
- D. Käyttää ammatillisia organisaatioita, ammattilehtiä jne. kasvun ja kehittymisen lähteenä
- E. Toimii kollegoiden voimavarana, kun se on mahdollista

8. Opettaja toimii ammatillisesti vanhempien ja kollegoiden kanssa sekä edistää kouluorganisaation toimintaa

- A. Toimii hienotunteisesti ja eettisesti kollegoiden ja vanhempien kanssa
- B. Keskustelee säännöllisesti vanhempien kanssa heidän lapsensa edistymisestä
- C. Ymmärtää ja kunnioittaa perheiden erilaisia uskomuksia, perinteitä ja arvoja
- D. Hoitaa ei-opetukselliset velvollisuutensa tehokkaasti
- E. Työskentelee laaja-alaisesti koulu yhteisön hyväksi



Kehittävään arviointiin

Tässä artikkelissa kuvataan arvioinnin kehittävää tehtävää sekä kehittämiseen sidostuvan arvioinnin luonnetta ja metodiikkaa. Otsikko viittaa arvioinnin kehittävään näkökulmaan, ei yksittäiseen metodiin. Artikkelisiidotaan kehittävän arvioinnin käsitteisiin ja periaatteisiin. Artikkelisi sisältää lyhyitä kuvauksia joistakin kehittämissä arvioinneissa sovellettavista arviointimenetelmistä. Arviointimethodoja käsittelevä kuvaus johdattelee monitahoarvioinnin, osallistavan arvioinnin, vertaisarvioinnin, vertailuun perustuvien arviointien, portfolioarvioinnin, moderaatiomenetelmän (yhteensovittava), konsultatiivisen arvioinnin ja arviointi-istuntojen hyödyntämiseen arvioinnissa.

Artikkelissa korostetaan kehittäville arvioinnille luonteenomaisia piirteitä: avoimuutta, läpinäkyvyyttä, osallistavuutta ja eri tahojen välistä vuoropuhelua. Osallistavuuden vaatimukset kohdistuvat niin arvioinnin kohteena oleviin tahoihin kuin sidosryhmiin ja yhteistyökumppaneihin. Arviointitulosten ohella kehittämissä arvioinneissa painotetaan prosessia ja sen aikana syntyvää oppimista. Arvioinneissa ei korosteta vertailuun perustuvaa järjestykseen asettamista (ranking-listat) eikä vertailuihin perustuvia sanktioita.

Artikkelissa käsitellään myös kehittävän arvioinnin hyödyntämistä kansallisissa arvioinneissa. Tällöin arvioinnit räätälöidään tavoitteiden ja arvioitavan teeman mukaan, arvioitavat osallistuvat aktiivisesti arviointiprosessiin ja arviointia suunnittelevat ja toteuttavat monitahoisuuteen ja verkostomaisuuteen perustuvat suunnittelu- ja arviointiryhmät. Arvioinnit johtavat myös parantamisalueiden esille nostamiseen tai suosituksiin ja kehittämisehdotuksiin.

Asiasanat: kehittävä arviointi, itsearviointi, ulkoinen arviointi, arviointimenetelmät

Kehittävän arvioinnin luonne

Kehittävän arvioinnin käsitteen määrittäminen on vaikeaa sen moni-ilmeisyyden vuoksi. Käsittemäärittelyä vaikeuttaa toisaalta se, että määrittelyitä on kirjallisuudessa vähän ja toisaalta se, että asiantuntijoiden ymmärrys kehittävän arvioinnin käsitteestä vaihtelee. Laajimmillaan kehittävän arvioinnin käsite viittaa kehittämiseen sidostuvaan ja osallistavuutta korostavaan ja menetelmiltään monimuotoiseen arviointiin. Kapeimmillaan käsite viittaa toiminnan kehittämisprosessiin sisältyvään näkökulmaan arvioinnista.

Kehittävä arviointi ymmärretään yleisesti syvälliseen oppimiseen tähtääväksi toiminnaksi. Hakkaraisen (1983) mukaan kehittävä arviointi on toiminnan osa, ja se itsessään tähtää muutokseen yksilön ja yhteisön ajattelussa ja ajattelun suhteessa todellisuuteen. Myös Patton korostaa em. piirteitä. Hänen mukaansa myös ulkopuoliset arvioijat ovat osa arvioitavasta kohteesta vastuussa olevaa tiimiä ja työskentelevät yhdessä pitkän aikaa tarkoituksenaan kehittää organisaation toimintaa tai muuta kohdetoimintaa (Patton 1997). Tässä merkityksessä kyse on pitkäaikaisesta prosessista, jossa arvioijat ovat kehittämisprosessin käynnistäjiä, ylläpitäjiä, opastajia, valmentajia, selventäjiä ja kyseenalaistajia. Tällöin he tuovat kehittämiseen arvioinnin näkökulman. Arvioinnin tulos ei ole arviointiraportti vaan ensisijaisia tuloksia ovat oppiminen, kokemus ja oppimisen avulla syntyvä uusi ajattelu. Arvioinnissa painottuu reflektiivinen vuorovaikutus, prosessin aikana kehittynyt ymmärrys ja muutos, jota arviointi tukee.

Kehittävä arviointi on luonteeltaan formatiivista (vrt. summatiivinen). Pattonin mukaan (1997 ja 1982; Hämäläinen & Kauppi 2000; Jokiranta 1993) arvioinnissa painotetaan osallistavuutta ja jatkuvaa vuorovaikutusta käytännön toiminnan ja arvioinnin välillä. Pattonin (emt.) mukaan arviointiin osallistuvat tahot ja yksilöt määrittävät yhdessä arvioinnin tavoitteet, päämäärät ja menetelmät. Tämänkaltainen arviointi ei perustu standardeihin ja ennalta määritettyihin kriteereihin eikä siinä korosteta arviointitulosten yhdenmukaisuutta tai kausaalisia yhteyksiä. Siinä ei myöskään painoteta yksinomaan lopputulosta. Arvioinnissa painotetaan eri tahojen välisen vuorovaikutuksen ohella monitahoisuutta ja yhteisymmärrystä, konsensusta. Tämänkaltaisessa arvioinnissa todellisuus kuvataan ja pyritään ymmärtämään sellaisena kuin se on. Keskeistä on nähdä jatkuvan kehityksen ja muutoksen luonne ja syvällisyys. Arvioinnin odotetaan ennen kaikkea tukevan kehitystä ja muutosta. Formatiivisen luonteensa mukaisesti kehittävässä arvioinnissa pyritään osoittamaan kehityksen luonne ja laatu. Kehittävä arviointi paljastaa, mitä on muuttunut ja miten.

Kehittävä arviointi soveltuu luonteensa mukaisesti parhaiten toiminnan muutosten arviointiin ja pitkäkestoisiin hankkeisiin. Koska kehittävässä arvioinnissa painottuu arvioinnin prosessimaisuus, se soveltuu koulu yhteisöissä erityisesti ydintoimintojen, kuten koulutusohjelmien ja opetussuunnitelmien, opetuksen ja ohjauksen ja muiden opiskelijapalveluiden arviointiin sekä toimintakulttuurin arviointiin. Muita sopivia

arviointikohteita ovat koulutuksen järjestäjien koko opetustoimi tai sen osa, strategiat, laadunhallinta, johtamisjärjestelmät, henkilöstövoimavarat, vanhempien ja koulun yhteistyö ja yhteistyösuhteet. Lisäksi tämänkaltaista arvioinnin lähestymistapaa voidaan käyttää erittäin hyvin järjestäjien tukemiseen, kun kehitetään paikallista arviointia.

Kehittävä arviointi on luonteeltaan varsin käytännöllistä – sekä arviointitiedon että arviointiprosessin tulee hyödyttää kehittämistä. Tällöin sekä tulosten että prosessissa syntyvän tietämyksen ja ymmärryksen oletetaan olevan hyödyksi. Kehittävän arvioinnin avulla ei haeta absoluuttista totuutta eikä sen avulla saatu näyttö (evidence based) aina täytä tieteellisen tutkimuksen vaatimuksia, vaikka se onkin systemaattisesti ja luotettavasti kerättyä ja osoitettua (esim. Sebba 2004). Olennaista kuitenkin on, että arviointi täyttää arvioinnille asetetut vaatimukset – arvioinnissa otetaan kantaa ja tehdään arvioivia johtopäätöksiä. Kehittävässä arvioinnissa voidaan nostaa esiin, monistaa muista lähestymistavoista poiketen, parantamisalueita ja esittää kehittämissuosituksia.

Lähtökohtana arviointiin osallistuvien tarpeet ja odotukset

Kehittävässä arvioinnissa otetaan aina huomioon arviointiin liittyvät tarpeet ja odotukset sekä kontekstin vaatimukset. Kehittävässä arvioinnissa korostetaan arviointikohteena *olevien aktiivista osallistumista arvioinnin suunnitteluun ja arviointiin sekä arviointitulosten perusteella tapahtuvaan jatkuvaan parantamis- ja kehittämistyöhön*. Arvioinnissa ei korosteta arviointiin osallistuvien olemista arvioinnin kohteena, vaan aktiivista osallistumista arvioinnin *subjekteina* arviointiin ja sen suunnitteluun. Kyse on tarpeisiin räätälöidystä arviointitavasta (esim. Kauppi ym. 2003; Patton 1997; Rossi, Lipsey & Freeman 2004; Sullivan 2004; Stake 2004; Torrance 1995a ja Torrance 1995b).

Tehtävänä kehittäminen

Kehittäväällä arvioinnilla on useita *tehtäviä*, joskin muita tärkeämpää on auttaa arviointiin osallistuvia kehittämään työkäytäntöjä. Arvioinnin tehtävänä on siten tukea työyhteisön ja siellä toimivien yksilöiden voimaantumista ja motivoitumista. Lisäksi arvioinnin tehtävänä on tukea päätöksentekoa, johtamista ja toiminnan jatkuvaa parantamista ja kehittämistä.

Kehittävän arvioinnin katsotaan kuuluvan niiden arviointien ryhmään, joissa yksilöiden ja yhteisöjen voimaantuminen mahdollistuu korostamalla vahvaa sitoutumista, reflektioon perustuvaa oppimista, arviointiin osallistuvien oikeutta osallistua arviointia koskevaan päätöksentekoon ja arviointiin ja mahdollisuutta ulkopuolisten asian-

tuntijoiden interventioihin ja eri tahojen yhteistyötä (Nolan & Hoover 2004; York-Barr ym. 2001). Pattonin (1997) mukaan kehittävää arviointia voidaan käyttää:

- *Syventämään yhteisön tai yksilöiden omaa ymmärrystä ja jakamaan tämä ymmärrys yhteisöllisesti.* Tällöin huomio voi kiinnittyä yhteisön tai yksilön oman arvioinnin osuvuuteen ja erilaisten perspektiivien esille tuloon ja heikoimman äänen kuuluvuuteen.
- *Tukemaan, suuntaamaan ja muuttamaan kohdetoiminnan interventioita.* Tällöin voidaan painottaa arvioinnin prosessiluonteen ymmärtämistä ja osallistujien seurantaa ja tarkkailua sekä arviointitulosten analysointia yhdessä arviointiin osallistuvien kanssa.
- *Lisäämään sitoutumista, omaa vastuuta ja omistajuutta.* Tällöin huomio voidaan kohdentaa arvioinnin osallistavuuteen ja yhteistyöperustaisuuteen, voimaantumiseen, yhteisön reflektiivisyyteen ja itsearviointiin.
- *Kohdetoiminnan ja organisaation toiminnan kehittämiseen.* Tällöin apuna voidaan käyttää kehittävän arvioinnin tapaan myös toimintatutkimusta, misio-orientoitunutta strategista arviointia, arvioinnin arviointia ja kehittämistoiminnan mallintamista.

Lähtökohtana hyöty ja vaikuttavuus

Arvioinnin vaikuttavuus on sitä parempi mitä enemmän arvioinnista on hyötyä päätöksenteossa ja johtamisessa, työyhteisön toiminnassa sekä viime kädessä myös yksittäisen työntekijän, tiimin ja työyhteisön työssä. Arvioinnin on siten autettava kehittämään arviointikohteena olevaa toimintaa ja toimijoita. Painottamalla monitahoista yhteistä suunnittelua ja arviointia annetaan myös arvioinnin kohteena oleville toimijoille mahdollisuus löytää arviointitoiminnalle henkilökohtainen ja yhteisöllinen merkitys. Samalla arvioinnista muodostuu koko koulutuksen kehittämisen keskeinen väline.

Kansallisella tasolla arviointitoiminnan vaikuttavuudesta voidaan etsiä näyttöä kiinnittämällä huomiota esimerkiksi arviointitulosten ohjaavuuteen sekä päätöksenteon että kehittämisen kannalta. Myös kehittävä arviointi voi olla tehokas hallinnon ja johtamisen väline.

Hyötysuhdetta voidaan niin ikään tarkastella arviointitulosten relevanssina kehittämistarpeeseen nähden. Hyötysuhde lienee jo silloin kohtuullinen, jos arvioinneista voidaan todeta olleen enemmän hyötyä kuin haittaa. Korkeakoulujen arviointineuvosto on omissa arvioinneissaan pyrkinyt lisäämään kouluille koituvaa hyötyä mm. mahdollistamalla koulujen aloitteenteon ja osallistumisen arvioinnin suunnitteluun sekä järjestämällä ennen arvioinnin aloittamista arviointiseminaareja (Moitus & Saari 2004).

Avoimuutta ja läpinäkyvyyttä julkista tilivelvollisuutta enemmän

Kehittävässä arvioinnissa ei korosteta ulkoista kontrolliin perustuvaa tilivelvollisuutta monien muiden arvioinnin lähestymistapojen tapaan. Arvioinnissa painottuvan subjektiivisuuden korostamisesta seuraa myös, että ”tilivelvollisuus” syntyy erityisesti itselle ja arviointiin osallistuville tahoille. Tästä näkökulmasta tarkasteltuna arvioinnissa korostuu ahdasta ulkoista tilivelvollisuutta enemmän toiminnan läpinäkyvyyden ja avoimuuden vaatimus suhteessa sekä arviointiin osallistuviin tahoihin että ulkoiseen todellisuuteen. Avoimuuteen perustuvassa kehittävässä arvioinnissa toiminta tehdään läpinäkyväksi itsen lisäksi myös muille: opiskelijoille, sidosryhmille, kump-paneille ja päätöksentekijöille.

Kehittävässä arvioinnissa joudutaan ottamaan kantaa myös arviointien läpinäkyvyyteen ja julkisuuteen. Svan (2001) pitää läpinäkyvyyttä ja julkisuutta eri asioina. Svanin mukaan kehittävässä arvioinnissa painotetaan läpinäkyvyyttä julkisuuden sijaan. Svan korostaa muita enemmän tarkoituksen läpinäkyvyyttä (miksi arviointi toteutetaan, mikä on oletettu lopputulos, miten tulokset vaikuttavat päätöksentekoon jne.). Vaikka arvioinnin läpinäkyvyys mahdollistaa myös muiden kuin arviointia tekevien tiedon saannin arvioinnista ja sen tuloksista, arviointitieto ei itsessään palvele median luomia julkisuusintressejä, kuten arviointikohteena olevien organisaatioiden välistä vertailuja.

Monitahoisuus monipuolisuuden varmistajana

Kehittävä arviointi on luonteeltaan yhteisöllistä ja monitahoista ja muiden arviointien tapaan aina myös arvosidonnaista. Yhteisöllisyys painottuu, kun korostetaan kohde-toiminnan ja sen arvioinnin arvosidonnaisuutta ja kontekstuaalisuutta sekä arvioinnin ja kehittämisen integroitumista. Monitahoisuutta puolestaan kuvaa (mm. Vartiainen 1994; Patton 1997; Nolan & Hoover 2004) eri menetelmien hyväksikäyttö, arviointitavan yksilön tai yhteisön oma arviointi ja sidosryhmien (stakeholder) osallistuminen arviointiin. Tämänkaltaisen arvioinnin perusta on sen *laaja-alaisuudessa, neuvottelevuudessa, sosiaalisessa oikeudenmukaisuudessa ja evolutionaarisuudessa*. Kehittävässä arvioinnissa painotetaan täydellisen riippumattomuuden tavoittelun sijaan eri tahojen yhtäaikaista osallistumista ja tässä kohtaamisessa syntyvää syvällistä yhteisymmärrystä (konsensus).

Monitahoisessa arvioinnissa toimijoiden ja erilaisten sidosryhmien arvomaailmat kohtaavat toisensa. Arvioinnin monitahoisuus mahdollistaa erilaisten näkökulmien ja käsitysten esilletulon ja huomioon ottamisen. Eri tahojen osallistuminen arviointiin lisää myös arviointitulosten tulkinnan monipuolisuutta ja luotettavuutta, arviointitulosten hyödynnettävyyttä ja arvioinnin vaikuttavuutta. Keskeistä on, että monitahoi-

suus estää arvioijien omien, joko tiedotettujen tai tiedostamattomien intressien, vaikutuksen ja arvioinnin näkökulmien yksipuolisuuden.

Itsearviointi edellytyksenä

Kehittävässä arvioinnissa painotetaan myös kohteena olevien tahojen ja organisaatioiden ja yksilöiden omaa arviointia. Itsearviointi on välttämätöntä arvioinnin luotettavuuden kannalta, koska subjektilla itsellään on paras käsitys arvioitavasta toiminnasta ja siihen vaikuttavista tekijöistä. Itsearvioinnin avulla voidaan myös läpivalaista toiminnan taustalla olevaa hiljaista tietoa ja nostaa myös se arvioinnin keskiöön. Yksilön ja yhteisön tekemän reflektiivisen analyysin avulla voidaan samaan aikaan tuottaa uutta tietoa ja kehittää uusia toimintamalleja. Arvioinnille ei ole määriteltävissä selkeää alku- ja päätepistettä, vaan se on jatkuvaa. Sen lisäksi, että yksilön ja yhteisön itsearviointi on omaehtoista ja itsenäistä, se voi olla myös ohjattua (vrt. konsultatiivinen arviointi). Konsultatiivisessa arvioinnissa ulkopuolinen asiantuntija voi auttaa arviointia tekeviä näkemään asioita, jotka ovat sekoittuneet arjen rutiineihin ja rituaaleihin ja hiljaiseen tietoon.

Stakeholderit mukana – perusteluita ja vastaväitteitä

Monet asiantuntijat pitävät keskeisten stakeholdereiden (sidosryhmät/intressitahot) osallistumista arviointiin välttämättömänä (Cuba & Lincoln 1989; Nolan & Hoover 2004; Patton 1997; Smith 1980; Weiss 1989). Näiden intressitahojen osallistuminen arviointiin lisää erityisesti arvioijien mahdollisuuksia selvittää ja syventää omia tulkintojaan. Samalla kun arvioinnissa tulee esiin erilaisia näkökulmia, arviointi monipuolistuu. Näiden tahojen osallistuminen lisää myös empiirisiä evidenssejä. Sen lisäksi, että sidosryhmien osallistumisen kautta arviointiin tulee lisää näkökulmia, hyötyjä saavutetaan jo pelkästä osallistumisesta ja vastuun jakamisesta. Välittävänä linkkinä toimivat sidosryhmät lisäävät myös mahdollisuuksia vaikuttaa suoraan päätöksentekoon, valtaapitäviin eliitteihin ja kehityksen suuntaan.

Sidosryhmien käyttöä on kirjallisuudessa myös vahvasti kritisoitu (Patton 1997, 1982). On nähty, että sidosryhmien käytöllä arvioinneissa on myös rajoituksia. Smithin (1980) mukaan mahdollisuus selvittää arvioijien omia tulkintoja on varsin vähäinen ja sidosryhmistä saatava hyöty näennäinen. Hänen mielestään sidosryhmät eivät yleensä paranna arvioinnin metodista laatua eikä niiden sitoutuminen arviointiprosessiin ole taattua ja itsestään selvää. Edelleen Smithin (emt.) mielestä lisäevidenssien saanti on mahdollista vain, jos arviointiin on onnistuttu valitsemaan korkeatasoiset ja osaavat sidosryhmien asiantuntijat. Arviointivastuun jakamisessa on Smithin mukaan myös kaksi puolta: toisaalta sidosryhmät jättävät arvioinnin arvioijille ja

toisaalta arvioijat eivät saisi delegoida arviointivastuutaan sidosryhmille, koska vastuu kuuluu heille. Myös sidosryhmien vaikuttaminen päätöksentekoon on Smithin mukaan mahdollista vain erityistapauksissa.

Edellä mainituista rajoitteista ja kriittistä huolimatta Patton (1982) pitää sidosryhmien osallistumista arviointiin välttämättömänä, koska ne selventävät, kyseenalaistavat ja täydentävät arvioijien näkökulmia ja tuovat arviointiryhmään arvioijilta usein puuttuvan kokemuksen ja syvällisen ymmärryksen. Sidosryhmien keskeinen merkitys korostuu Pattonin mukaan (emt.) myös oppimisen edistämässä sekä toimimisessa sanansaattajina ja puolestapuhujina arvioinnin jälkeen (vrt. asianajajat).

Kehittävä arviointi ulkoisen arvioinnin osana

Kehittävää arviointia ei ole kovin helppo yhdistää ulkoisista intresseistä ohjautuvaan arviointiin, joskin korkeakoulujen arvioinneissa tässä lienee onnistuttu hyvin. Kehittävän arvioinnin viitekehyksessä toteutetun ulkoisen arvioinnin kannalta on keskeistä eri toimijoiden päämäärien, arviointitehtävien ja arviointiin kohdistuvien odotusten ja tarpeiden sovittaminen yhteen. Kehittävä arviointi ei perustu ulkoiseen kontrolliin eikä sen perusteella seuraa sanktioita. Kehittävän arvioinnin soveltamisen ensisijaisena lähtökohtana on tuottaa sellaista tietoa, jota koulutuksen järjestäjät ja oppilaitokset ja opettajat voivat käyttää hyväksi kehittäessään omaa toimintaansa. Kehittävän arvioinnin tehtävänä on myös välittää arvioinnin avulla esiin tulleita hyviä käytänteitä (best practises) eri toimijoiden tarpeisiin. Tämän ensisijaisen tavoitteen lisäksi kehittävä arviointi tuottaa tietoa myös koulutuspoliittisen päätöksenteon ja kansallisen kehittämisen perustaksi.

Koulutuksen arviointineuvoston organisoimissa kansallisissa arvioinneissa kehittävä arviointi voidaan tulkita monitahoisuutta ja osallistavuutta painottavaksi toiminnaksi, jossa arviointimenetelmät räätälöidään arvioinnin tavoitteiden ja arvioitavan teeman mukaan. Arvioinnit perustuvat *itsearviointiin*, ja arviointia suunnittelevat ja toteuttavat monitahoisuuteen ja *verkostomaisuuteen perustuvat suunnittelu- ja arviointiryhmät (ulkopuolinen asiantuntija-arviointi)*. Arvioinnit sisältävät aina myös *arviointivierailuja* (vuorovaikutteisuus). *Arviointiraportit ovat julkisia*. Arvioinnit johtavat myös parantamisalueiden esille nostamiseen tai suosituksiin ja kehittämisehdotuksiin. Vertaisarviointien sisällyttäminen arviointeihin on myös mahdollista. Eduskunta on niin ikään edellyttänyt, että arviointien osaksi kehitetään myös *arviointien vaikuttavuuteen kohdistuvaa systemaattista seuranta* (follow up).

Kehittävässä arvioinnissa painotetaan avoimuutta. Yhtenä avoimuuden piirteenä on, että koulutuksen järjestäjillä ja muilla arvioinnin kohteena olevilla tahoilla on mahdollisuus vaikuttaa arvioinnin kohteisiin, painotuksiin, menetelmiin ja kriteereihin. Kehittävässä arvioinnissa painotetaan niin ikään eri tahojen sitoutumista vuoropuheeseen.

Suunnittelu- ja arviointiryhmät

Arvioinnin suunnittelu ja arviointi erotetaan toisistaan. **Suunnitteluryhmä** edustaa laajasti sekä arvioinnin että arviointikohteen tuntemusta ja eri intressitahoja. Suunnitteluryhmä

- laatii projektisuunnitelman ja hyväksyttää sen Koulutuksen arviointineuvostolla ennen toimeenpanoa,
- huolehtii koulutuksen järjestäjien ja oppilaitosten sekä opettajien kuulemisesta,
- kokoaa tausta-aineistot ja suunnittelee arviointimenetelmät, mittarit, indikaattorit ja kriteerit sekä arviointiprojektin kaikki vaiheet,
- tekee esityksen arviointiryhmästä arviointineuvostolle ja
- vastaa järjestäjien ja oppilaitosten sekä arvioitsijoiden koulutuksesta ja valmentamisesta arviointeihin.

Arviointiryhmä vastaa arvioinnista ja arviointiraporttien laadinnasta. Arviointiryhmä nimetään monitahoisuutta korostavan periaatteen mukaan. Arviointiryhmissä otetaan tarpeen mukaan huomioon myös monitieteisyyden vaatimukset. Arviointiryhmän nimeämisessä hyödynnetään arviointiverkoston. Arviointiryhmä valitaan seuraavien **kriteereiden** perusteella: *arviointikohteen ja kontekstin syvälinen tuntemus, syvälinen arviointiosaaminen, riippumattomuus/subjektius sekä koko ryhmän monitahoisuus. Arviointiryhmässä on aina myös ruotsinkielisen koulutuksen sidosryhmiä edustava asiantuntija (stakeholder)*. Arviointikohteena olevien tahojen näkemys voidaan parhaiten turvata nimittämällä myös järjestäjien, oppilaitosten ja/tai opettajien edustajia arviointiryhmään (subjekti). Tällöin voidaan myös soveltaa vertaisuuteen perustuvan arvioinnin menettelyitä (peer evaluation, peer review).

Arviointiasiantuntijoista ja sidosryhmistä koostuvat arviointiryhmät tekevät arvioinnin itsenäisesti ja laativat arviointiraportin. Suunnitteluryhmä, arviointineuvosto ja -sihteeristö eivät osallistu arviointiin.

Itsearviointi

Kehittävässä arvioinnissa arvioinnin kohteena olevat yksiköt tekevät itsearvioinnin. Arviointia varten laaditaan itsearviointia ohjaavat ja jäsentävät kysymykset, mutta arvioinnin organisointia ei ohjeisteta kovin tarkasti. Arviointikohteiden määrittämisellä ja arviointikysymyksillä pyritään helpottamaan niin itsearviointia kuin ulkoisten arvioitsijoiden työtä. Itsearvioinnin ohella järjestäjiltä ja oppilaitoksilta kootaan tarvittaessa määrällistä tietoa taustoittamaan arviointia. Tämä aineisto voidaan koota tietokannoista ja/tai kyselylomakkeilla.

Itsearviointi tehdään ulkoisen arvioinnin tapaan monitahoisuutta edustavien ryhmien toimesta. Arviointi perustuu toimintakuvauksiin, jotka jokainen organisaatio laatii

arviointimetodin mukaisesti. Arvioinnissa tehdään konsensukseen perustuvat arvioivat johtopäätökset ja määritetään omaa toimintaa koskevat vahvuudet ja kehittämis-kohteet.

Ulkoinen arviointi

Ulkoisissa arvioinneissa hyödynnetään itsearvioinnin ja tausta-aineistoksi tarvittaessa kootun määrällisen tiedon lisäksi myös aiemmin koottua ja analysoitua arviointi-, tutkimus- ja seurantatietoa (dokumenttianalyysi).

Ulkoinen arviointi voidaan organisoida hyvin eri tavalla ja siinä voidaan soveltaa erilaisia menettelyitä. Keskeistä on, että arviointiryhmä muodostaa yhteisen näkemyksen arvioinnin kokonaisuudesta ja arvioinnin kohdealueista. Arviointiryhmä täydentää itsearvioinneista ja muusta tausta-aineistoista luomaansa kuvaa arviointivierailujen ja arviointiseminaarien avulla. Ulkoisessa arvioinnissa voidaan niin ikään käyttää apuna erilaisia arviointiraateja, esimerkiksi alueellisia arviointiraateja. Arviointivierailut antavat arviointiryhmälle mahdollisuuden varmistaa omien tulkintojen luotettavuus ja täydentää saamaansa käsitystä. Arviointiryhmän tulee niin ikään antaa näissä tilanteissa palautetta ja tehdä kehittämissuhteita arviointivierailun kohteena olevalla organisaatiolla.

Arviointivierailut kestävät vähintään yhden päivän. Arviointikäynti valmistellaan etukäteen ja se sisältää eri tahojen haastattelut (esim. johto, opettajat, opiskelijat, muu henkilöstö, työelämän edustajat ja sidosryhmät) ja tutustumisen opetus- ja oppimisympäristöön.

Arvioinneissa voidaan käyttää apuna erilaisia arviointiraateja, parivertailuja ja vertaisarvioiteja (myös vertailuarvioinnit kuten benchmarking) sekä arviointiseminaareja. Lisäksi arviointitietoa voidaan tarvittaessa täydentää delfoi-metodin avulla. Näiden avulla voidaan tuottaa täydentävää tietoa ja mahdollistaa järjestäjien, oppilaitosten ja opettajien osallistuminen arviointiin. Arviointiseminaareja ja -raateja tms. järjestetään tarpeen mukaisesti.

Arviointiraportit ovat julkisia

Arviointiraportit laaditaan itsearviointiaineistojen, arviointivierailujen ja muiden arviointiin liittyvien arviointitietojen ja aktiviteettien perusteella. Arviointiraportit sisältävät arviointiryhmän tekemät tulkinnot sekä arvioivat johtopäätökset. Keskeistä on, että arviointiraportissa esitetään myös arviointiryhmän käsitys arviointikohteeseen liittyvistä vahvuuksista ja parantamisalueista (myös SiVM 5/2001). Lisäksi arviointiraportit sisältävät kuvauksia parhaista käytänteistä, jotka voidaan valita vertaisarviointien, erillisten arviointiraatien tai arviointiryhmän toimesta. Parhaiden käytänteiden arvioinneissa voidaan soveltaa portfolio-metodia.

Arvioinnin vaikuttavuutta voidaan edistää sillä, että arvioinnin kohteena olevat tahot voivat kommentoida arviointiraporttia ennen sen lopullista hyväksymistä arviointiryhmässä. Arviointi on kehittävää vain, jos siihen osallistuvat tahot ja yksilöt oppivat prosessin aikana ja jos arvioinnista on hyötyä myös arvioinnin jälkeen. Arvioinnin vaikuttavuutta pyritään lisäämään korostamalla arviointiin osallistuneiden oikeutta saada heitä koskevat tulokset ja niihin liittyvät vertailutulokset, kiinnittämällä huomiota arviointiraporttien selkeyteen ja panostamalla arviointitulosten levittämiseen.

Arviointiraportit ovat julkisia. Arviointiraporteissa ei kuitenkaan tehdä erityisesti julkisuuden näkökulmasta usein painottuneita vertailuja ja niihin perustuvia ranking-listoja. Lainsäätäjän tahto on (A150/2003), että koulutuksen järjestäjille ja oppilaitoksille tulee toimittaa niitä koskevat arviointitulokset oman kehittämistyön perustaksi. Tästä jouduttaneen aikakin ensimmäisissä arviointihankkeissa tinkimään määrärahojen vähäisyyden vuoksi.

Arvioinnin seuranta

Eduskunta on edellyttänyt, että arviointien vaikuttavuuden arviointiin kehitetään systemaattinen seuranta-arviointi (follow up). Vastaava menettely on olemassa korkeakoulujen arvioinneissa. Seuranta-arviointi tehdään yleensä muutaman vuoden kuluttua varsinaisesta arvioinnista. Seuranta kohdistetaan yleensä arvioinnissa esitettyjen suositusten toteuttamiseen (vrt. Moitus & Saari 2004).

Kehittävän arvioinnin menetelmiä

Eduskunta on edellyttänyt, että koulutuksen arvioinnissa lisätään kehittävää arviointia korkeakoulujen arviointien tapaan (SiVL 5/2001; Koulutuksen arvioinnin uusi suunta 2004). Korkeakoulujen arviointineuvoston tekemissä kansallisissa arvioinneissa on painottunut kehittävän, erityisesti kommunikatiivisen arvioinnin periaatteet ja menetelmät (Moitus & Mustonen 2004; Moitus & Saari 2004). Arvioinneissa on painotettu menetelmällistä riippumattomuutta (Moitus & Saari emt.), joskaan menetelmällistä riippumattomuuden ihannetta ei ole pidetty arvona sinänsä, vaan on korostettu menetelmien tarkoituksenmukaisuutta tavoitteiden kannalta.

Kehittävässä arvioinnissa ei ole kyse yksittäisestä menetelmästä, vaan laajasta menetelmien kirjosta. Eri menetelmiä voidaan käyttää ja soveltaa tarkoituksenmukaisella ja kehittämistä tukevalla tavalla. Kehittävä arviointi voidaan tehdä erityisesti toimintatutkimuksen (action research) avulla, mutta kehittävät elementit ovat mukana kaikissa osallistavuutta ja kommunikatiivisuutta painottavissa arviointimenetelmissä. Tämänkaltaisia menetelmiä ovat muun muassa vertaisarviointit (esim. peer evaluation, benchmarking), parivertailut, erilaiset arviointikeskustelut, yksilö- ja yhteisöportfoliot ja arviointiseminaarit ja moderatio-mallin sovellukset kuten myös koko-

naisvaltaiseen laadunhallintaan perustuvat arviointimallit (EFQM). Kehittävän arvioinnin keskeisiä tunnuspiirteitä ovat:

- arvioinnin räätälöinti tarpeiden ja odotusten mukaisesti
- kohteena olevien tahojen aktiivinen osallistuminen suunnitteluun ja arviointiin sekä niiden perusteella tapahtuvaan jatkuvaan kehittämiseen (subjektius, vuorovaikutus, avoimuus)
- jatkuvaa laadun parantamista ja kehittämistä tukeva arviointiprosessi ja tulos (hyödynnettävyys, käytännönläheisyys ja palautemenettely)
- monitahoinen reflektointi ja konsensus “absoluuttisen totuuden” mittaamisen sijaan
- asiantuntijavastuun jakaminen ja osallisuus (toimijat, sidosryhmät, suunnittelijat, arvioijat).

Menetelmäkuvauksia

Tässä luvussa esitetään pääpiirteitä joistakin kehittävää arviointia edustavista lähestymistavoista ja menetelmistä sekä niiden käyttömahdollisuuksista itsearvioinneissa ja ulkoisissa arvioinneissa. Kuvaus on luonteeltaan yleinen ja orientoiva.

1. Monitahoarviointi

Monitahoarviointia voidaan soveltaa sekä itsearvioinneissa että ulkoisissa arvioinneissa. Monitahoarvioinnissa otetaan Vartiaisen (1994; myös Patton 1997) mukaan laaja-alaisesti huomioon erilaisia näkemyksiä, kannanottoja ja arviolauseita. Arvioinnissa painotetaan avainryhmien välistä vuorovaikutusta ja neuvottelevuutta sekä sosiaalista oikeudenmukaisuutta ja kaikkien osapuolten äänen kuulumista (vähäisinkin hyödynsaaja). Lisäksi arvioinnissa tarkastellaan arviointikohteen muutosta suhteessa ympäristön muutokseen (evolutionaarisuus).

Monitahoarviointi edellyttää eri menetelmien samanaikaista soveltamista. Lisäksi edellytetään, että avainryhmät, kuten eri toimijatahot ja sidosryhmät, osallistuvat arvioinnin suunnitteluun ja arviointiin (esim. rahoittajat, päätöksentekijät, kehittäjät, toimijat, intressiryhmät ja vertaiset). Arvioinnin monitahoisuutta voivat toteuttaa myös monitahoisuutta edustavat arviointiraadit. Arviointi nähdään sitä parempana mitä paremmin siinä otetaan huomioon eri tahojen näkemykset. Arvioijat tekevät arviointinsa omien kokemustensa ja näkemystensä perusteella (henkilökohtaisuus), ja monitahoisuus syntyy näiden yksilöarvioinnin yhteen sovittamisesta. Näkemykset arvioinnin perusteista ja johtopäätösten oikeutuksesta voivat vaihdella suuresti, mutta arviointiryhmän tehtävänä on löytää neuvottelevuuden avulla sellainen yhteinen ymmärrys (konsensus), johon se ryhmänä voi sitoutua.

2. Osallistava arviointi

Kehittävä arviointi nojaa osallistavuuteen ja vuorovaikutukseen. **Osallistavassa arvioinnissa** on yhteisiä piirteitä monitahoarvioinnin kanssa. Monitahoarvioinnit edellyttävät yleensä osallistavuutta, mutta kaikki osallistavat arvioinnit eivät välttämättä ole monitahoisia. Myös yksilöarviointi voi olla osallistavaa. Osallistavaa arviointia voidaan tehdä sekä yksilöiden että yhteisöjen itsearviointeina. Yhteisötasolla tehdyssä arvioinnissa korostuu eri tahojen välinen vuoropuhelu ja pyrkimys yhteisymmärrykseen sekä työyhteisön ja sen jäsenten oppimiseen ja voimaantumiseen (empowerment).

Osallistava arviointi voidaan myös organisoida tapaustutkimuksena (action research). Näkökulmaa voidaan hyödyntää myös ulkoisissa arvioinneissa. Osallistavuus nousee keskiöön myös konsultatiivisissa arvioinneissa ja silloin kun tuetaan järjestäjiä, oppilaitoksia ja opettajia arviointiin liittyvissä asioissa. Tällöin ulkopuolisen asiantuntijan tehtävänä on opastaa ja motivoida arviointiin sekä edistää luottamusta ja arviointiprosessin hallintaa. Tällaisissa tilanteissa asiantuntija ei toimi niinkään arvioijana vaan neuvottelijana, opastajana ja kouluttajana. Konsultatiivisessa arvioinnissa korostetaan prosessia, jossa arvioija tukee arvioinnin keinoin yhteisöä löytämään itse ratkaisun ja kyvyn arvioida omaa toimintaansa.

3. Vertaisarviointi

Vertaisarviointien erilaisia sovelluksia voidaan soveltaa niin itsearvioinneissa kuin ulkoisissa arvioinneissakin. Vertaisarvioinnin juuret ovat akateemisten julkaisujen arvioinneissa, joissa riippumattomat asiantuntijat arvioivat artikkeleita, tutkimuksia tms. (peer review) (mm. Armstrong 1997; Brown 2004). Vertaisarviointi on kuitenkin laajentunut muille arvioinnin alueille samalla kun sen menetelmät ovat monipuolistuneet. Yleisesti *vertaisarvioinnilla (peer evaluation, peer assessment) tarkoitetaan samaan ammattiryhmään kuuluvien tai saman ongelman parissa työskentelevien kesken tapahtuvaa työn kehittämistä, jossa vertaisten muodostama ryhmä arvioi kriittisesti ja järjestelmällisesti toisiaan* (Vertaisarviointi... 2003). Arvioijana voi toimia saman koulutuksen saanut tai samalla organisaation tasolla toimiva henkilö (kollega) tai samaa työtä tekevä ryhmä (vertaisryhmä).

Vertaisarvioinnin menetelmissä korostetaan yleisesti arvioinnin kehittävyttä ja arvioinnin kohteena olevien aktiivista osallistumista arviointiin (mm. Nolan & Hoover 2004; MacBeath & McGlynn 2002; Jakku-Sihvonen & Heinonen 2001).

Vertaisarvioinnissa sovitaan vertaisten kanssa yhdessä arvioitavista kohteista, kriteereistä ja arvioinnin vastavuoroisuudesta sekä prosessin etenemisestä.

Vertaisarviointi voi perustua parivertailuun tai ryhmien ja työyhteisöjen sekä yksiköiden väliseen vastavuoroisuuteen. Vertaisia voidaan käyttää arvioinneissa myös ilman vastavuoroisuuden vaatimusta esimerkiksi ulkoisissa arvioinneissa. Työyhteisöissä vertaisarviointi soveltuu erityisen hyvin tilanteisiin, joissa halutaan parantaa ilmapiiriä, kehittää laatua tai edistyä ammatillisessa kasvussa. Lisäksi opetus- ja oppimismenetelmien kehittäminen voi perustua vahvaan vertaisuuteen ja vertaisarvioinnin hyödyntämiseen. Opetuksen kehittämisessä voidaan niin ikään käyttää vertaisohjauksen (vrt. valmentaminen) menetelmiä (peer coaching). Vertaisohjauksessa on kyse asiantuntijoiden toistensa toiminnan havainnoista, palautteen annosta ja tuesta (Nolan & Hoover 2004).

Vertaisuutta voidaan hyödyntää arviointipaneeleissa. Samoin benchmarking-käynnit voivat perustua vertaisuuden hyödyntämiseen. Ulkopuoliset arvioinnit ja arviointikäynnit voidaan järjestää niin ikään vertaisten toimesta (esim. Moitus & Saari 2004). Myös konsultatiivinen arviointi voidaan järjestää vertaisten toimesta.

4. Vertailuarviointi

Vertailuarvioinnissa on kyse kahden tai useamman kohteen vertaamisesta toisiinsa ja tähän vertaamiseen perustuvasta tulkinnallisesta analyysistä. Vertailuarviointi soveltuu erityisesti itsearviointiin ja jatkuvaan toiminnan laadun parantamiseen. Sitä voidaan myös käyttää täydentävänä metodina kansallisissa arvioinneissa erityisesti hyvien käytänteiden kuvaamisessa ja niistä oppimiseen liittyvissä arvioinneissa ja arviointiesimerkkien esille nostamisessa.

Vertailuarvioinnin metodeista on eniten sovellettu benchmarking-metodia. Se on menetelmä, jossa systemaattisesti opitaan hyviltä esikuvilta toimialasta riippumatta. Benchmarking-prosessin tavoite on oman toiminnan parantaminen ja jatkuva kehittäminen (Hotanen, Laine & Pietiläinen 2001). Benchmarking-prosessi perustuu vuorovaikutteiseen oppimiseen, jossa on vahvasti mukana arvioivia elementtejä. Benchmarking-prosessissa analysoidaan omaa toimintaa, tehdään vertailuja parhaaseen käytäntöön (omaa parempaan) ja kehitetään tämän vertailun perusteella edelleen omaa toimintaa. Myös vertailun kohteena olevalle organisaatiolle annetaan arvioiva palaute.

Benchmarking-vertailuja voidaan tehdä tunnuslukujen/indikaattoreiden avulla. Arviointimenetelmä sopii erityisen hyvin prosessien sekä strategioiden kehittä-

miseen. Organisaatioiden välisessä oppimisessa voi olla kyse myös vastavuoroisuudesta (kahdenvälinen benchmarking). Kumppanuussuhteissa voidaan niin ikään oppia toinen toisiltaan (kumppanuusbenchmarking). Ryhmäbenchmarking-menettelyt soveltuvat niin ikään eri organisaatioiden ja asiantuntijaryhmien yhteiseen oppimiseen. (Hotanen, Laine & Pietiläinen 2001; Lecklin 2002.) Tämän kaltainen arviointi voidaan toteuttaa verkostona.

5. Parivertailut

Parivertailut soveltuvat sekä saman ammattikunnan, opiskelijoiden ja opettajien, opettajien ja työelämän asiantuntijoiden, opettajien ja vanhempien sekä muiden vastaavien ryhmien tai yhteisöjen keskinäiseen arviointiin. Parivertailut edellyttävät osapuolten omaa arviointia ja siihen perustuvaa vertailua. Parivertailu voidaan tehdä myös eri tahojen kuten opiskelijoiden ja opettajien ryhminä laatimien itsearviointien perusteella (Moitus & Saari 2004). Parivertailu perustuu tällöin itsearviointiraportteihin, ja arvioinneissa voidaan tuottaa arvioivat lausunnot toinen toisensa arvioinneista.

Parivertailuja voidaan täydentää ulkoisella arvioinnilla. Tällöin ulkoinen arviointi perustuu sekä itsearviointeihin että parivertailuihin. Lisäksi voidaan järjestää näitä arviointia täydentäviä, eri tahoja osallistavia arviointifoorumeita.

6. Portfolio

Portfolio on autenttisen arvioinnin muoto, jonka avulla voidaan edistää arviointikohteen oppimista omasta toiminnastaan (Bernhard et al. 2000). Portfoliossa ei ole kyse yhdestä mallista, vaan tarkoitus, tekijä ja tilanne ohjaavat portfolion muotoa, sisältöä, toteutusta ja arviointia. Portfolio voi olla dokumentoitu työkansio tai -salkku, johon kootaan arvioidut työnäytteet, suunnitelmat, työpapereit ja keskeneräisetkin työt. Portfolion sisältämät pitkän ajanjakson aikana syntyneet työnäytteet antavat mahdollisuuden myös kohteen kehityksen arviointiin (prosessiluonne). Portfolio voi sisältää yksilösuorituksia, mutta myös yhteistoimintaan perustuvia töitä (Pollari ym. 1996).

Portfolioarviointi soveltuu opiskelijoiden ja opettajien oman arvioinnin lisäksi myös koko kouluyhteisön itsearviointiin. Opettajaportfoliossa voi olla näytteitä opetuksen suunnittelusta, työjärjestyksestä, opetustunneista, opiskelija-arvioinnista, oppimis- ja opetusympäristön sosiaalisesta rakenteesta, sosiaalisten tapahtumien ohjelmista, yhteistyöstä oppilaitoksen sisällä ja muiden koulujen kanssa, opiskelijoiden ja sidosryhmien palautteista, artikkeleista ja erilaisista innovaatioista (esim. Stronge & Tucker 2003; Brown & Irby 2001). Yhteisöä kos-

kevassa portfolioissa voi Bernhardin (1999, 2003) ja Bernhardin (et al., 2000) mukaan olla kuvauksia missiosta, visiosta, arvoista ja jatkuvasta parantamisesta, tavoitteista ja strategioista. Edelleen portfolio voi sisältää näytteitä yhteistyö- ja sidosryhmäsuhteista, voimavarojen hallinnasta ja ydintoimintaan liittyvistä innovaatioista sekä kehittämishankkeista. Portfolio voi sisältää myös keskeisten tulosten kuvauksia ja kuvauksia tulevaisuudesta.

Portfolioarvioinnin vahvuus on, että sitä voidaan soveltaa kehityksen eri vaiheissa olevien yksilöiden ja yhteisöjen osaamisen, kasvun, oppimisen sekä kokemusten näkyväksi tekemiseen, itsearviointiin ja yhteisölliseen pohdintaan (Kankaanranta 2001). Portfolion laadinnan avulla voidaan synnyttää pohdintaa hyvän opetuksen kriteereistä, arvioinnista, arvioijista ja päämääristä. Portfolio on erittäin hyvä keino yhteisölliseen reflektointiin (Tenhula ym. 1996).

Portfolioarviointia voidaan soveltaa erityisesti eritasoisissa itsearvioinneissa mutta myös ulkoisessa arvioinneissa. Itsearviointi ulkoisen arvioinnin perustaksi voidaan tuottaa portfolioarvioinnin metodin mukaan. Moituksen ja Saaren (2004) mukaan arviointiprosessi etenee arviointiaineiston koonnista, sen valinnasta ja arvioinnista. Keskeistä on osoittaa kehityskaari ja tulevaisuuden kehityksen suunta.

7. Erilaiset arviointikeskustelut

Kehittävissä arvioinneissa voidaan painottaa myös erilaisia keskusteluita. Näistä Moitus ja Saari (2004) tuovat esiin teemakeskustelut ja fokusryhmäkeskustelut. Teemakeskustelut voivat heidän mukaansa korvata ja täydentää ulkoisen arviointiryhmän arviointivierailuja. Teemakeskusteluihin voivat osallistua eri tahot sidosryhmistä arviointikohdetta edustaviin asiantuntijoihin ja opiskelijoihin. Teemakeskusteluita järjestetään useita, jolloin niiden sisällöt vaihtelevat. Teemakeskustelu perustuu yleensä itsearviointien perusteella tehtyihin johtopäätöksiin ja jatkokysymyksiin, jotka arviointiryhmä toimittaa teemakeskusteluun osallistuville etukäteen.

Fokusryhmäkeskusteluilla Moitus ja Saari 2004 tarkoittavat ryhmän ohjaajan johdolla tapahtuvaa keskustelua ja haastattelua. Fokusryhmät ovat homogeenisiä. Myös näissä ryhmien istunnoissa voidaan käsitellä teemoja, joita on vaikea saada muutoin esille. Fokusryhmissä voidaan niin ikään keskustella hyvistä käytänteistä ja kehittämiskohteista. Fokusryhmäkeskustelut perustuvat monien muiden arviointimetodien tapaan työyhteisöjen itsearviointeihin.

8. Moderation – sovitteleva arviointi

Moderaatiossa on kyse kokonaisvaltaisesta menetelmästä, joka perustuu eri tahojen yhtäaikaiseen osallistumiseen ja konsensuksen synnyttämiseen (mm. Beywl & Potter 1998; Radnon & Shaw 1996). Menetelmä on syntynyt reaktiona perinteisten kokouskäytäntöjen varassa tapahtuvalle kehittämiselle. Kyse on kommunikatiivisesta ja demokraattisesta toimintatavasta, joka itsessään takaa toiminnan ja siihen liittyvän päätöksenteon läpinäkyvyyden ja kehittämisen analyttisyyden (Beywl & Potter emt.). Menetelmässä painotetaan myös osallistujien monitahoisuutta, erityisesti sidosryhmien osallistumista.

Moderaatioon perustuvissa arvioinneissa on keskeistä saada aikaan eri tahojen yhteisymmärrys tarkastelun kohteena olevasta asiasta. Tämä yhteisymmärrys syntyy kommunikaation ja kohtuullistamisen avulla. Olennaista on myös tuoda esiin eri tahojen erilaiset arvot ja pyrkimykset, myös julkilausumattomat intressit. Lisäksi tarkoituksena on saada kuva osallistujien näkemyksistä ja mielenkiinnon kohteista arviointiteemaan liittyen. Mallissa painotetaan arvioinnin positiivisuutta ja arvioinnin etenemistä ryhmän kypsyyden mukaan. Malli tunnustaa myös ulkopuolisen arvioinnin tai arvioitsijoiden merkityksen ja erilaisten ulkopuolisten tahojen tekemän laadun varmistuksen. Ulkoisen asiantuntijan ensisijaisin tehtävä on edistää arviointiprosessia, ei osallistua sisällön käsittelyyn. Ulkoinen asiantuntija on ensisijaisesti metodiikan asiantuntija ja havaintojen tekijä.

Moderaatioon perustuva arviointiprosessi on pitkä ja se sisältää useita istuntoja. Olennaista on, että arviointiprosessiin osallistujat eivät vaihdu istuntojen välillä. Moderaatioprosessiin osallistuvat henkilöt toimivat myös tiedonlevittäjinä omissa organisaatioissaan. Moderaatio-prosessissa voi yksinkertaisimmillaan olla kyse siitä, että organisaatiokohtaiset ryhmät analysoivat ja arvioivat omaa toimintaansa ensin, minkä jälkeen alueelliset vertaisryhmät kokoontuvat arvioimaan kohdetoimintaa yhdessä. Konsensukseen tähtäävien istuntojen lisäksi voidaan järjestää myös arviointiseminaareja. Arviointiprosessin aikana voidaan käyttää erilaisia menetelmiä, kuten aivoriihimetodia. Lisäksi voidaan määrittää vahvuuksia ja parantamisalueita, tehdä arvottavia yhteenvetoja ja kriteereihin perustuvia tulkintoja jne. (Beywl & Potter 1998; Kuhlmann, S. 1998). Moderaatio-menetelmän avulla voidaan arvioida muun muassa opetussuunnitelmia, koulutusohjelmia, henkilöstövoimavarojen hallintaa, opiskelijavirtoja, erilaisia yhteistyösuhteita ja opetusprosesseja.

Moderaatiometodi soveltuu erityisesti itsearviointeihin, mutta sitä voidaan käyttää apuna myös ulkoisissa arvioinneissa. Itsearviointien lisäksi alueellisten moderaatioistuntojen raportit voivat olla ulkoisen arvioinnin lähteinä. Lisäksi ulkoinen arviointi voi perustua moderaatioistuntojen havainnointiin ja ryhmien haastatteluun.

9. Konsultatiivinen arviointi

Konsultatiivisessa arvioinnissa hyödynnetään arviointiasiantuntijoiden osaamista. Asiantuntijan tehtävänä on keskustella arviointia suunnittelevan ja kehittävä henkilöstön kanssa tehtävästä ja siihen liittyvistä odotuksista ja päämääristä. Näiden keskusteluiden tulokset ohjaavat arviointituen tai arvioinnin räätälöinnin kyseisten tarpeiden ja odotusten mukaiseksi. Konsultaatio perustuu sekä avuntarvitsijoiden että konsultatiivisessa roolissa toimivan arvioijan henkilökohtaisiin käsityksiin ja niiden perusteella syntyvään yhteiseen käsitykseen päämääristä ja konsultaation tarpeesta. Konsultatiivisessa roolissa toimivien arvioijien on perehdyttävä hyvin organisaatioon ennen kuin he alkavat käsitteellistää keskustelua tai arvioinnin kehittämisiongelmiä. Ongelmien identifiointin jälkeen arvioijat jäsentävät arvioinnin ja tuen muodon ja mallin. Konsultatiivisen tuen lähestymistapa ja malli räätälöidään tarpeiden pohjalta.

Konsultatiivinen arviointitehtävä voi liittyä järjestäjien tukemiseen arviointiin liittyvissä asioissa, arviointitiedon tuottamiseen tai kehittämisehdotusten tekemiseen. Arviointitiedon tuottamiseen tähtäävässä prosessissa arvioija kerää tarvitsemansa tiedon ja räätälöi arviointitulosten analysoinnin tavan ja prosessin. Arvioija ei raportoi kaikkia ongelmia, vaan keskittyy konsultatiivisen tehtävän kannalta keskeisiin. Päätös raportoitavista kohteista tehdään yhdessä kohteen tai muun tilaajan kanssa. Konsultatiivisessa arvioinnissa arvioijat kuulevat arvioinnin teettäjän ja tilaajan näkemyksiä, mutta tekevät aina itse arviointia koskevat päätökset.

Lähteet

- AI50/2003 koulutuksen arvioinnista.
- Amstrong, J. S. 1997. Peer review for journals: Evidence on quality control, fairness, and innovation. *Science and Engineering Ethics* 3, 63–84.
- Bernhard, V., von Blanckensee, L. L., Lauck, M. S., Rebello, F. F., Bonilla, G. L. & Tribbey, M. M. 2000. The example school portfolio. A companion to the school portfolio: A comprehensive framework for school improvement. Larchmont: Eye on Education.
- Bernhard, V., V. L. 1999. The school portfolio. A comprehensive framework for school improvement. Larchmont: Eye on Education.
- Bernhard, V. 2002 The school portfolio TOOLKIT. A planning, implementation, and evaluation guide for continuous school improvement. Larchmont: Eye on Education.
- Beywl, W. & Potter, P. 1998. RENOMO – A design tool for evaluation. *Evaluation* 4 (1), 53–71.
- Brawn, G. & Irby, B. J. 2001. The principal portfolio. California: Corwin Press.
- Brown, T. 2004. PEER REVIEW and the acceptance of new scientific ideals. Discussion paper from a Working Party on equipping the public with and understanding of peer review.
- Cuba, E. G. & Lincoln, Y. S. 1989. Forth generation evaluation. London: Sage Publications.

- Hakkarainen, P. 1983. Opetuksen ja oppimisen arviointi. Teoksessa *Aikuisopetuksen perusteita. Vaapaan sivistystyön XXVII vuosikirja*. Helsinki: Aikuiskasvatuksen Tutkimusseura ja Kansanvalistusseura.
- Hotanen, J., Laine, R. O. & Pietiläinen, S. 2001. Benchmarking-opas. Helsinki: Laatukeskus.
- Hämäläinen, K. & Kauppi, A. 2000. Evaluation in today's education – from control to empowerment. *Lifelong Learning in Europe V (1)*, 68–75.
- Jaku-Sihvonen, R. & Heinonen, S. 2001. Johdatus koulutuksen uudistuvaan arviointikulttuuriin. Opetushallitus. Arviointi 2/2001. Helsinki.
- Jokiranta, H. 1996. Arviointi ja sosiaalityö. Teoksessa R. Granfeld, H. Jokiranta, S. Karvinen, A.-L. Matthies & A. Pohjola *Monisärmäinen sosiaalityö*. Helsinki: Sosiaaliturvan keskusliitto, 285–314.
- Kauppi, A., Hanski, K., Nousiainen, R. & Lahtiranta, K. 2004. Arviointi pedagogisen kehittämisen välineenä. Teoksessa H. Kotila (toim.) *Ammattikorkeakoulupedagogiikka. Ajankohtaisia puheenvuoroja*. Helsinki: Edita, 255–269.
- Koulutuksen arvioinnin uusi suunta. 2004. Arviointiohjelma 2004–2007. Koulutuksen arviointineuvoston julkaisuja I. Koulutuksen arviointineuvosto. Jyväskylä.
- Kuhlmann, S. 1998. Moderation of policy-making? Science and technology policy evaluation beyond impact measurement – The case of Germany. *Evaluation. The International Journal of Theory, Research and Practice* 4 (2), 130–148.
- Lecklin, O. 2002. Laatu yrityksen menestystekijänä. Jyväskylä: Gummerus.
- MacBeath, J. & McGlynn, A. 2002. *Self Evaluation. What's in for schools*. London: Routledge Falmer.
- Moitus, S. & Saari, S. 2004. Menetelmistä kehittämiseen. Korkeakoulujen arviointineuvoston arviointimenetelmät vuosina 1996–2003. Helsinki: Korkeakoulutuksen arviointineuvosto.
- Nolan, J. Jr & Hoover, L. 2004. *Teacher supervision and evaluation. Theory into practice*. United States of America: John Wiley & Sons.
- Patton, M. Q. 1982. *Practical evaluation*. Beverly Hills: Sage.
- Patton, M. Q. 1997. *Utilization Focused Evaluation*. 3rd ed. The New Century Text. Thousand Oaks, CA: Sage.
- Radnor, H. & Shaw, K. 1995. Developing a collaborative approach to moderation. Teoksessa H. Torrance (toim.) *Evaluating authentic assessment*. Buckingham: Open University Press, 124–165.
- Rossi, P. H., Lipsley, M. W. & Freeman, H. E. 2004. *Evaluation. A systematic approach*. London: SAGE Publications.
- Ryan, C. W. & Kick, F. 1995. *Portfolios across the curriculum and beyond*. California: Corwin Press, A. Sage Publications.
- Sebba, J. 2004. Developing evidence-informed policy and practice in education. Teoksessa G. Thomas & R. Pring *Evidence-based practice in education*. London: Open University Press, 34–43.
- Sivistysvaliokunnan lausunto 5/2001 opetusministeriön selvitykseen (MINS 4/2001) koulutuksen arviointitoiminnasta.
- Smith, N. L. 1980. Studying evaluation assumptions. *Evaluation Network Newsletter*, 39–40.
- Sullivan, R. G. *Practising evaluation. A collaborative approach*. London: Sage Publications.
- Stake, R. E. 2004. *Standard. Based & Responsivise Evaluation*. London: Sage Publications.
- Stronge, J. H. & Tucker, P. D. 2003. *Handbook on teacher evaluation. Assessing and improving performance*. Larchmont: Eye on Education.
- Swan, J. 2001. Transparency – an essential characteristic of the evaluation of higher education? Paper presented in INQAAHE-2001 Conference of Quality. Standards and Recognition.
- Vartiainen, P. 1994. *Palveluorganisaatioiden tuloksellisuusarviointi, teoriat ja käytäntö. Päivähoitoon, erityisesti päiväkoteihin kohdistuva monitahoarviointi*. Vaasan yliopisto. Acta Wasaensia 40.
- Vartiainen, P. 2003. *Intressiryhmät evaluatioprosessissa. Metodologista pohdintaa monitahoarvioinnin avulla*. Hallinnon tutkimus I, 29–39.

- Vertaisarviointi hyvinvoinnin työkaluna. Stakesin ideakortti I/03. Stakes. Saatavana www-muodossa:
<<http://www.stakes.fi/hyvinvointi/khs/ideakortit/vertaisarviointi.pdf>>
- Weiss, C. H. 1989. The stakeholder approach to evaluation: Origins and promise. Teoksessa E.R. House (toim.) New directions in educational evaluation. London: The Falmer Press, 145–157.
- York-Barr, J., Sommers, W. A., Ghere, G. S. & Montie, J. 2001. Reflective practice to improve schools. California: Corwin Press.



Kehittämisvoimaa arvioinnista

Artikkelissa tarkastellaan suomalaisessa arviointikeskustelussa varsin vähälle huomiolle jäänyttä voimaannuttavan arvioinnin lähestymistapaa. Voimaannuttavan arvioinnin prosessin nähdään tukevan yksilöiden ja yhteisöjen itseohjautumista, itsenäistä päätöksentekoa, muutoksen hallintaa, hyvinvointia sekä arviointiosaamista. Voimaantumisen ja voimaannuttava arviointi kiinnitetään artikkelissa oppivan organisaation ajattelutapaan. Itsearviointi ja ulkoinen arviointi palvelevat yhteisöllistä oppimista ja kehittämistä. Itsearviointi ja ulkoinen arviointi nähdään kiinteästi toisiinsa liittyvinä ja toisiaan täydentävinä.

Voimaannuttavan arvioinnin vaiheet kuvataan artikkelissa toiminnan kehittämisprosessina. Siinä kiinnitetään huomiota erityisesti toiminnan vahvuuksiin, ja arviointia hyödynnetään tietoisesti prosessin eri vaiheissa. Voimaannuttavan arvioinnin keskeisenä haasteena artikkelissa pidetään sitä, miten arviointitulokset tulkitaan kehittämisprosessien ja toiminnan tuloksellisuuden kannalta ja miten toisaalta kehittäminen, itsemääräytyminen ja tuloksellisuus pystytään optimoimaan. Niin ikään pidetään tärkeänä yhteisön jäsenten ja ulkoisen arviointisijan intressien välistä tasapainoa. Artikkelin tarkoitus on tukemaan paikallista ja oppilaitoskohtaista itsearviointia lähinnä lähestymis- ja ajattelutapana ja tarjoamaan uutta näkökulmaa myös kansallisen ulkoisen arvioinnin ja paikallisen arvioinnin suhteeseen.

Avainsanat: voimaantuminen, voimaannuttava arviointi, oppiva organisaatio, kehittäminen



Arvioimalla yhteisölliseen oppimiseen

Oppivan organisaation merkitys ja oikeutus perustuvat muun muassa siihen, että oppimisyhteisöjen jäsenillä on aiempaa paremmat mahdollisuudet oppia enemmän ja tehokkaammin, koska oppimisen esteitä pyritään tietoisesti poistamaan ja yhteisöllisen oppimisen edellytyksiä pyritään parantamaan. Oppivalle organisaatiolle ominainen kannustava ilmapiiri rohkaisee jäseniä yhtäältä lisäämään työenergiaansa ja kokemaan työn iloa. Toisaalta se kannustaa jäseniä innovatiivisuuteen ja luovuuteen sekä olemaan kykenevämpiä toimimaan ja ottamaan vastaan uusia haasteita, tyydyttämään kasvutarpeita sekä parantamaan elämänsä laatua. Oppivaa organisaatiota pidetään lupauksena asioista, joita toivotaan ja odotetaan. Se nähdään uutena asioiden tarkastelutapana, tavoitteiden saavuttamista edistävänä tilana sekä yksilöitä yhdistävänä mahdollisuutena ja työhön sitoutumisen vahvistajana. Oppivan organisaation nähdään antavan toivoa jatkuvuudesta ja yleensä paremmasta tulevaisuudesta. (Ks. Nikkanen & Lyytinen 1996.)

Oppivan organisaation kannustavassa ilmapiirissä puhuminen vaihtuu helpommin tekemiseksi. On entistä tärkeämpää, että ihmiset sitoutuvat työhönsä ja organisaatioonsa. Path-goal-johtamisteorian mukaan johtajan tehtävänä organisaatiossa on kirkaistaa tavoitteita työyhteisölle, "avata ovia", rohkaista ja kannustaa henkilöstön jäseniä, jotta jokainen löytäisi polkunsä omiin ja organisaation päämääriin ja tavoitteisiin. Johtaminen on taitoa ja tahtoa tehdä asioita yhdessä. Voimaantuminen-käsitteen synonyymina jotkut tutkijat ovatkin käyttäneet *jaettu johtajuus* -käsitettä (Appelbaum 1999; Beach 1996; Koivunen 2003). Voimaantuminen tapahtuu sosiaalisessa vuorovaikutustilanteessa. Ihmissuhteisiin liittyvässä jaetussa johtajuudessa on kyse vaikutusvallan antamisesta. Työyhteisön jäsenille annetaan työtehtävien hoidon edellyttämää valtaa. Voimaantumista vahvistavan organisaatiokulttuurin edistämässä ja ylläpitämässä hyvän itsetunnon omaavalla johtajalla on tärkeä merkitys (Bennis & Nanus 1985). Oman sisäisen voimansa löytäminen (empowerment), sitoutuminen sekä organisaation toiminnan tehostaminen ovat nyt ajankohtaisia ja tärkeitä, kun vastataan työyhteisön hyvinvoinnin ja tuloksellisuuden vaatimuksiin.

Oppivaa organisaatiota luotaessa arvioinnilla on keskeinen merkitys. Arviointi on sinällään jo kehittämisprosessi, joka yhdessä systemaattisesti toteutettuna johtaa yhteisölliseen oppimisprosessiin. Samoin nykytila-arviointien ja kehittämistoimien myötä siirrytään uuteen ja tuntemattomaan. Siksi opetustoimen arvioinnissa on korostettu paikallisen ja oppilaitoskohtaisen arvioinnin merkitystä. Paikallisella tasolla on oleellisin toimintaan vaikuttava tieto sekä mahdollisuudet tiedon hyödyntämiseen ottaen huomioon tapahtumiin vaikuttavat asiat. Myös päätöksenteko kehittämistoimien käytettävistä voimavaroista tapahtuu paikallisella tasolla. Näin arvioinnin tulokset voivat parhaimmillaan olla nopeasti vaikuttavia ja luoda siten tahtoa ja uskoa omiin vaikuttamismahdollisuuksiin. Oppivien organisaatioiden verkostoituessa toimimaan yhdessä esim. paikallisesti, prosessia on alettu luonnehtia oppivan alueen

metaforalla. Tavoitteena on tällöin tehostaa oppimista paikallisesti tai alueellisesti. Tässä yhteydessä esimerkiksi oppiva kunta tarkoittaa organisaatioiden välistä oppimista, joka tapahtuu tavoitteellisesti kunnan eri yhteisöjen vuorovaikutusverkostossa. Tästä näkökulmasta tarkasteltuna arviointi palvelee laajassa merkityksessä oppivan alueen kehittymistä. Tämä tarjoaa uuden ja hedelmällisen näkökulman myös paikalliseen arviointiin.

Koulutuksen arviointineuvoston uudessa arviointiohjelmassa lähtökohtana on, että sisäinen ja ulkoinen arviointi täydentävät toisiaan ja niitä voidaan hyödyntää yhdessä. Tavoitteena on näin edistää arviointien vaikuttavuutta. Sitoutuminen arviointiin tapahtuu itsearviointivaiheessa osallistumisen kautta. Ulkopuolinen arviointi taas toimii palautemekanismina. Toiminnan vahvuuksia korostavana se voi vahvistaa voimaantumisedellytyksiä. Näin ulkoista arviointia ei tulkita ulkokohtaiseksi väliintuloksi, vaan kehittämistä tukevaksi. Tässä suhteessa voimaannuttavan arvioinnin lähestymistavalla voi olla merkitystä myös, kun edistetään kansallisen arvioinnin tavoitteita. Ulkoinen arviointi ei tuota vain “kosmeettisia kullissimuutoksia”, vaan jotta parhaimmillaan syvälliseen kehittämiseen.

Käsillä olevassa artikkelissa palaamme aikaisemmin (1996) kirjoittamamme kirjan “Oppiva koulu ja itsearviointi” aihepiiriin ja tutkimme, miten sisäistä ja ulkoista arviointia voitaisiin toteuttaa niin, että ne auttaisivat yksilöitä ja työyhteisöjä löytämään oman sisäisen voiman tunteensa ja siten ymmärtämään ja tukemaan arvioinnin kehittämistehtävää.

Arviointi vahvistamaan sisäistä voimantunnetta ja itsemääräytymistä

Ihmiselle ominaista merkitykselliseksi kokemisen tarvetta on lähestytty voimaantuminen-käsitteen kautta. Kun yksilö ei pysty enää kokemaan työnsä mielekkyyttä, merkityksellisyyttä tai arvoa joko yhteisöstä, rakenteista, tilanteista tai yksilöstä itsestään johtuvista syistä, yksilön voimaantumisen/voimaantumisprosessista tulee yksilölle ja yhteisölle tärkeä. Siitonen ja Robinson (1998) painottavat, että koska arvostuksen ja merkityksellisyyden tunteen tarve on ihmisessä niin syvä inhimillinen välttämättömyys/pakko, tämän tunteen vähäiselläkin voimistumisella on niin yksilöä kuin myös suurta ryhmää kehittävä vaikutus. Se voi synnyttää runsaasti energiaa oppimiseen ja kasvuun.

Voimaantuminen-käsitteen käyttö on noin kahden vuosikymmenen aikana lisääntynyt kirjallisuudessa. Se on nopeasti vakiintunut useille tieteen aloille. Kirjallisuudessa englanninkielinen käsite *empowerment* ja suomenkieliset käsitteet *voimaantuminen*, *valtautuminen*, *valtaistuminen*, *voimaannuttaminen* ja *valtauttaminen* esiintyvät synonyymeina. Voimaantuminen on yksi tärkeä näkökulma, kun halutaan edistää yksilön ja työyhteisön hyvinvointia. Alan tutkimuskirjallisuudessa on kuitenkin

varsin niukasti vakavia pyrkimyksiä voimaantumista koskevaan käsiteanalyysiin.

Voimaantumisen lähtökohdasta, sen alulle panevista tekijöistä, seikoista tai ilmiöistä, on vallalla ainakin kaksi suuntausta (esim. Rodwell 1996). Ensimmäisen näkökulman mukaan voimaantuminen lähtee ihmisestä itsestään – voimaa ei voi antaa toiselle (esim. Karl 1995, 14; Siitonen 1999; Fetterman 1996; Grimmert 1996; Heikkilä-Laakso & Heikkilä 1997; Neufelt & Grimmert 1994; Robinson 1994). Toisen näkökulman mukaan voimaa voidaan antaa yksilölle ulkoapäin (esim. Murrell 1985; Ruohotie 1998; Thomas & Velthouse 1990). Sen mukaan yksilö tai ryhmä voidaan voimaannuttaa tai valtaistaa siirtämällä voimaa yksilöltä tai ryhmältä toiselle yksilölle tai ryhmälle. Kolmas näkökulma on, että voimaantuminen voi tapahtua molempien edellisten yhteisestä vaikutuksesta – ulkoa- ja sisältäpäin ohjautuen (esim. Adams 1991; Pollari 2000, 68).

Loppujen lopuksi yksilön reaktioissa ja toiminnassa on aina kyse yksilön aivoissa, lihaksissa, hermoissa ja tahtotiloissa tapahtuvista reaktioista. Tärkeää voimaantumista ajatellen on, mikä aiheuttaa ja laukaisee kyseisen reaktion tai prosessin. Beachin (1999) ja Siitosen (1999) mukaan voimaantuminen ja motivaatio ovat hyvin lähellä toisiaan (ks. luku 4). Myös motivaatio on yksilön sisäinen prosessi. Se on psyykinen tila, joka määrää, millä vireydellä ja miten suuntautuneena yksilö kulloisessakin tilanteessa toimii. Yksilö voi olla motivoitunut, ja yksilöä voidaan motivoida toimintaan.

Voimaantumisessa ja motivoitumisessa ympäristön tarjoamalla edellytyksillä ja tuella on tärkeä merkitys. “Vahvaksi tuleminen” voi tapahtua yksilön sisäisten ja ulkoa tulevien ärsykkeiden vaikutuksesta. Voimaantumisessa on kyse erittäin monimutkaisesta ja yksilöä kokonaisvaltaisesti koskevasta psykofyysisestä prosessista, joka on riippuvainen vaihtuvista ympäristökonteksteista. Voimaantuminen saattaa olla josain tietyssä ympäristössä todennäköisempää kuin toisessa (Siitonen 1999) ja vaihdella voimakkuudeltaan ympäristön ja ajankohdan mukaan (Zimmerman 1995).

Siitonen (1999, 127) viittaa Lagerspetziin (1998) ja Fettermaniin (1994; 1996) käsitellessään ihmisen valinnanvapauden ja itsemääräämisoikeuden vahvistamista yhtenä voimaantumisenprosessin edellytyksenä. Siitosen mukaan ihmisen tulisi tarkastella kriittisesti oman valinnanvapautensa todellisuutta, sillä valinnat voivat tosiasiallisesti olla manipulatiivisen vallankäytön tulosta – vain näennäisesti omia. Ihmisen tulisi-kin arvioida kriittisesti, mistä ja kenen intresseistä voi olla kyse valintatilanteissa, joita hän kohtaa. Myös Lagerspetz (1998, 134–136) pitää tärkeänä ihmisen omaa kriittistä arviointia valinnanvapautensa todellisuudesta, vaikka katsookin vallankäytön arvioimisen olevan hyvin ongelmallista. Fettermanin (1994; 1996) mukaan valinnanvapauden arvioinnin yhteydessä ihmiset ja organisaatiot tarvitsevat tukea itsemääräytymisen edistämiseksi.

Bennisin ja Nanusin (1985) tutkimuksen pohjalta Siitonen (emt., 96) toteaa, että erityisesti hyvinvointia, avointa kommunikointia, kunnioitusta, arvostusta ja luottamusta tukevan kulttuurin edistämisen kautta voidaan ehkä auttaa ihmistä tunnista-

maan, luomaan ja kanavoimaan omat voimavaransa. Sisäisen voimantunteen saavuttaneista ihmisistä heijastuu Siitosen tutkimuksen mukaan (1999) myönteisyys ja positiivinen lataus, joka on yhteydessä hyväksyvään luottamukselliseen ilmapiiriin ja arvostuksen kokemiseen.

Sisäisen voimantunteen saavuttaneella ihmisellä on halu yrittää parhaansa ja ottaa vastuuta myös yhteisön toisten jäsenten hyvinvoinnista käyttämällä rohkeasti ja luottavaisesti toimintavapauttaan. Koettu vapaus tukee vastuun ottamista, ja molemmat ovat yhteydessä yhteisön turvalliseksi koettuun ilmapiiriin. Voimaantunut yksilö vaikuttaa ympäristöönsä, tuntee kuuluvansa ryhmään, toimii vaikuttavalla tavalla omaaloitteisesti ja ilmaisee tunteitaan monipuolisesti ja tarkoituksenmukaisesti (Kiikkala, Immonen, Sohlman, Ahonen, & Siitonen 2000).

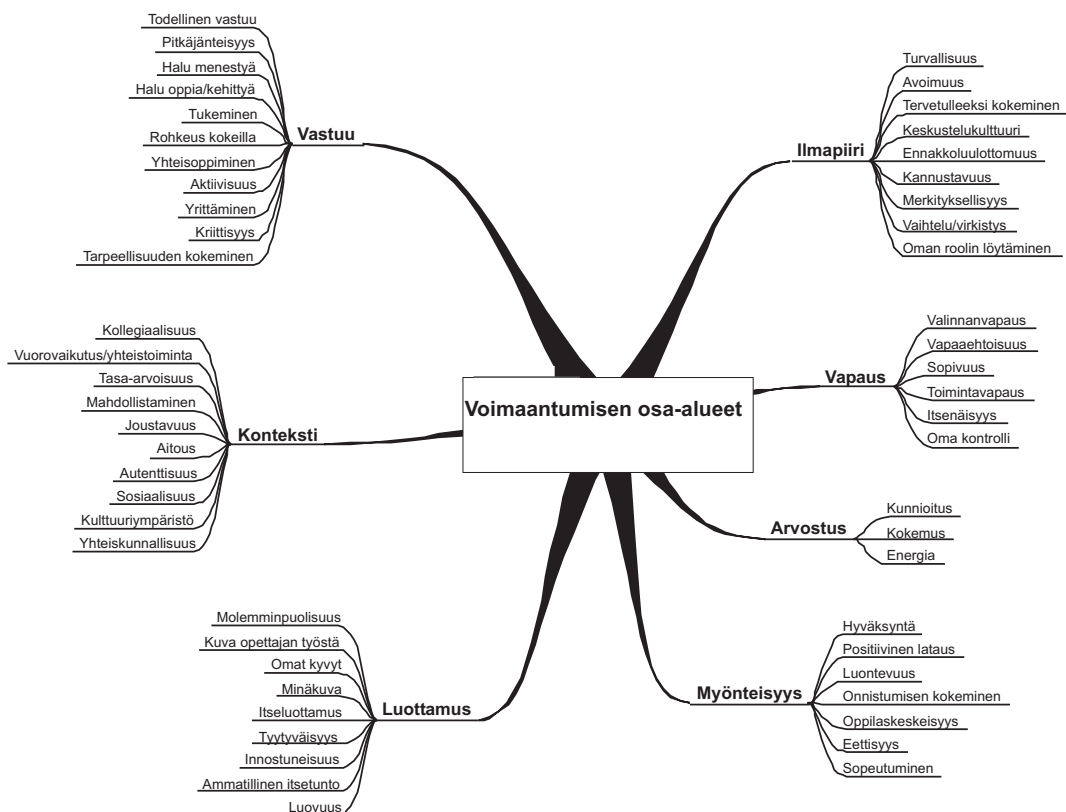
Monet edellä mainituista seikoista ovat juuri niitä ominaisuuksia, joita oppivassa organisaatiossa halutaan tietoisesti tukea ja edistää (ks. Nikkanen & Lyytinen 1996). Työyhteisön toimintakulttuuri on silloin energiaa vapauttava – ei sitä sitova.

Voimantunteen monitahoisuus

Sisäisen voimantunteen luokat eli kategoriat (vapaus, vastuu, arvostus, luottamus, konteksti, ilmapiiri, myönteisyys) ja niiden osa-alueet (ks. kuvio 1) ovat läheisesti kytkeytyneitä toisiinsa ja merkityssuhteessa keskenään. Siitosen (1999) tutkimuksen tuottamissa luokissa on eritasoisia ja osin teoria-alaltaan päällekkäisiä elementtejä. Tutkimusaineisto kerättiin neljältä opettajaksi opiskelevalta toisen vuoden opiskelijalta. Siitonen (emt. 61) väittää ilmeisesti tutkimustuntumansa perusteella, että ”ratkaisevat puutteet yhdessäkin osa-alueessa voivat heikentää tai jopa estää sisäisen voimantunteen rakentumista tai ylläpitämistä”. Väitteelle ei kuitenkaan löydy hänen tutkimuksestaan riittävää empiiristä todistusta. Toisaalta on helppo ajatella, että mainittua heikkenemistä saattaisi tapahtua.

Ydinkäsitteen löytyminen johti Siitosen mielenkiintoiseen havaintoon empiiristä aineistoa analysoidessaan: opiskelijan myönteinen asennoituminen ja sisäinen voimantunne johtavat sitoutumiseen ja ahkeraan työskentelyyn. Tästä syntyi hänelle ns. katalyytti-idea: voimaantuminen vaikuttaa katalyytin tavoin sitoutumiseen. Vahva katalyytti johtaa vahvaan sitoutumiseen ja heikko katalyytti heikkoon sitoutumiseen. Sitoutumisprosessi johtaa vastuulliseen ja aktiiviseen toimintaan, jota luonnehtii omistautuminen, ja joka voi parhaimmillaan johtaa jopa uudistavaan ajatteluun ja toimintaan. (Siitonen 1999, 79–80).

Kiinnostusta voimaantumista ja sen prosesseja kohtaan ovat voimakkaasti lisänneet erityisesti ongelmat, jotka ovat koskeneet työntekijöitä: työssä uupuminen ja työssä jaksaminen, liiallinen fyysinen ja psyykinen työkuormitus sekä työpaikkakiusaaminen. Työntekijän työuupumus on tulos pitkästä prosessista. Prosessiin sisältyy runsaasti inhimillistä kärsimystä, jota kokevat yksilön lisäksi myös kanssatyöntekijät.



Kuvio 1. Voimaantumisen osa-alueet (vrt. Siitonen 1999, 62)

Prosessin aikana yksilö on menettänyt voimavaransa työn laadukkaaseen tekemiseen ja myös kiinnostuksensa työhön (Suonsivu 2001). Edellä jo ilmaistiin, että oppivassa organisaatiossa jäsenet voivat voimaantua. Voimaantumista voidaan lisätä myös sisäisten ja ulkoisten arviointiprosessien avulla. Ulkoisen arvioitsijan tehtävänä on toimia valmentajana, kehittäjänä ja arvioivana “kriittisenä ystävänä” eikä tilivelvollisuutta korostavana kontrolloijana.

Arviointi tukemaan itsenäistä päätöksentekoa

Arviointi pyrkii määrittelemään tietyn kasvatuksen ja koulutuksen tuotteen, prosessin tai ohjelman arvon niin, että se tukee suunnittelua, päätöksentekoa ja kehittämistä (ks. Nikkanen & Lyytinen 1996, 93). Fettermanin (1996) mukaan voimaannuttava arviointi on innovatiivinen lähestymistapa arviointiin. Hän määrittelee voimaannutta-

van arvioinnin arviointikäsitteiden ja tekniikkojen sekä löydetyn tiedon käytöksi, jonka avulla voidaan edistää ja vahvistaa yksilön oikeutta päättää omista asioistaan. Sen avulla voidaan tukea itsemääräytymistä. Fetterman (1996; 1998; 2003) ja Levin (1996) katsovat voimaannuttavan arvioinnin vahvistavan työyhteisön toimintakapasiteettia.

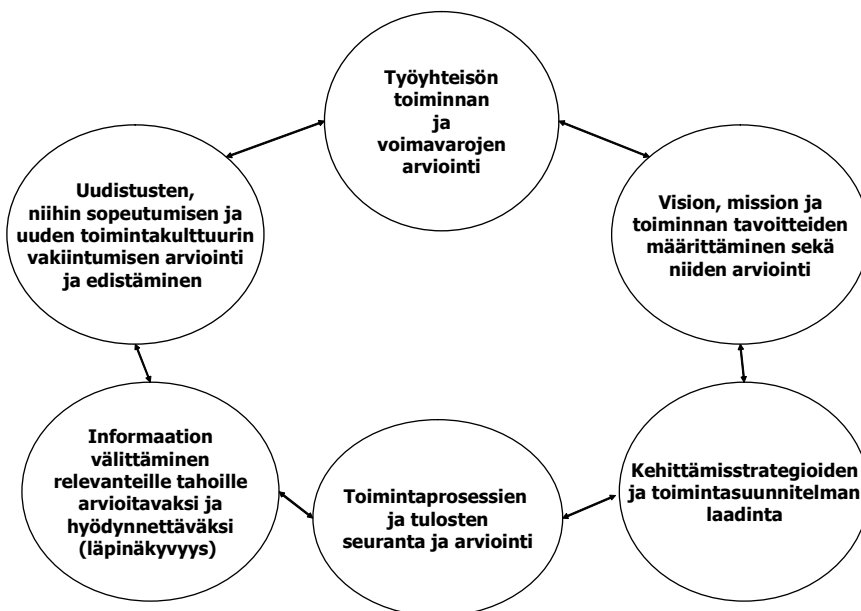
Fetterman on ollut mukana kehittämässä voimaannuttavan arvioinnin käsitettä ja ohjelmaa, jossa olennaista on auttaa ihmisiä auttamaan itseään. Heitä autetaan kehittämään toimintojaan itsearviointin ja reflektion avulla, ja näin edistetään ihmisten itsemääräytymistä. Fettermanin määritelmään sisältyy kaksi ajatusrakennelmaa: 1) Voimaantumisen edellyttää, että ne, joiden elämään tai toimintoihin vaikutetaan päätöksin, osallistuvat mielekkäällä tavalla päätöksentekoon. 2) Hyvät päätökset ovat tietoon perustuvia päätöksiä, joita voidaan parantaa systemaattisen arviointitiedon avulla (Fetterman 1996, 4–6; Levin 1996, 49–50). “Voimaannuttavassa arvioinnissa arviointisija ja yksilöt, jotka hyötyvät arvioinnista, ovat samalla tasolla, oppimassa toinen toiseltaan” (Fetterman 1996, 11). Ryhmätoimintaa, itsearviointia ja reflektointia korostava ohjelma perustui siihen, että arviointiin osallistuvat johtavat itse arviointiprosessia ja että ulkopuolinen arvioija toimii vain valmentavana henkilönä.

Siitosen ja Robinsonin (1998, 20) mukaan voimaannuttava arviointi ei erityisesti painota motivaatiota, vaikka se sisältääkin useita motivaatioteorioihin yhteneviä tekijöitä, kuten itsemääräytyvyys, vaikutuskyky, voimavarojen rakentaminen, näkemys kontekstista, luottamus, arvostus, vastuullisuus, tarpeet, kiinnostukset, kyvyt, kontrollin tunne ja vapaaehtoisuuden luonne. Huomio kiinnittyy pikemminkin kehitykseen, kehittämiseen ja elinikäiseen oppimiseen.

Arviointiprosessi voimaannuttajaksi

Voimaannuttava arviointi rakentuu useista vaiheista (ks. kuvio 2). Ne auttavat arviointisijoita jäsentämään ja arvioimaan toimintaansa ja luovat samalla “kommunikaatio-avaruuden” (Vanderplaats 1995) sekä helpottavat vuorovaikutuksen kohdentumista. (Habermas 1984; ks. Fetterman 2002, 63–75). Nämä vaiheet ovat: 1) työyhteisön toiminnan ja voimavarojen arviointi, 2) vision, mission ja toiminnan tavoitteiden määrittäminen, 3) kehittämisstrategioiden ja toimintasuunnitelman laadinta, 4) toimintaprosessien ja tulosten seuranta ja arviointi, 5) informaation välittäminen relevanteille tahoille (läpinäkyvyys) ja 6) uudistusten, niihin sopeutumisen ja uuden toimintakulttuurin vakiintumisen edistäminen.

Mainitut vaiheet ovat pitkälti yhdenmukaisia rakentavan ongelmanratkaisun vaiheiden kanssa. Ne kuuluvat tyypillisesti myös organisaation sisäiseen kehittämiseen. Kaikkia näitä vaiheita arvioitaessa voidaan käyttää myös valmiita kriteerejä ja strukturoituja mittareita (vrt. Lyytinen ym. 2002). Tärkeää on kiinnittää huomiota siihen, mitä arvioidaan, miten arvioidaan ja kuinka tulokset raportoidaan ja tulkitaan, jotta se



Kuvio 2. Voimaannuttavan arvioinnin prosessi

tukisi hyvää päätöksentekoa (Levin 1996, 49). Näin täytetään myös arvioinnin perusvaatimukset. Arviointi ei saa olla vain jäsentymätöntä toimintojen pohdintaa.

Ulkopuolisen arvioitsijan roolina edellä mainitussa arviointiprosessissa on toimia ryhmän tai organisaation tukijana ja avustajana sekä arviointitilanteen tasapainottajana prosessin eri vaiheissa. Hän auttaa arviointitehtävien suorittamisessa ja osallistujien kuulluksi tulemisessa. Arvioijan tulee olla myös analyttinen ja kriittinen. Hän pyrkii saamaan eri vaiheisiin liittyvien määrittelytehtävien avulla ryhmän jäsenet tietoisemmaksi omasta sekä yhteisön toiminnasta, toiminnan taustalla vaikuttavista arvoista, tavoitteista ja päämääristä, prioriteeteista, senhetkisestä todellisesta asiantilasta, omista ja yhteisön vahvuuksista ja heikkouksista sekä henkilösuhteista muihin jäseniin ja koko yhteisöön. Tehtävillä pyritään avaamaan keskustelua ja saamaan esille ryhmän jäsenten todelliset ajatukset käsiteltävistä asioista. Prosessin aikana edistetään itsemääräytymistä, luodaan syvempää tietoisuutta ja “vapaudutaan”. Tavoitteet ja tulospyrkimykset sovitetaan käytännön toteutuksen kannalta tarkoituksenmukaiselle kehitystasolle. (Fetterman 2003, 74.)

Avoimien ryhmätilanteiden tarkoituksena on 1) luoda demokraattinen tiedonkulku ja -vaihto, 2) vaikuttaa siihen ettei johto käyttäisi yhteisön ilmapiiriä heikentäviä rannaisutoimia tai manipulointia sekä 3) toimia niin, että mahdolliset “paljastukset” il-

maistaisiin diplomaattisesti ja rakentavalla tavalla. Vapaa keskustelu yhteisön arvoista ja normeista saattaa johtaa helposti epärealistisiin tuloksiin.

Yksi tärkeimmistä voimaannuttavan arvioinnin elementeistä on sen perustuminen vahvuuksiin heikkouksienkin osalta. Aina voidaan löytää jotakin positiivista, jota voidaan hyödyntää toiminnan vahvistamiseksi. (ks. Fetterman 2003, 63–75.) Arviointiin liittyy siis positiivinen orientaatio toiminnan vahvistamiseksi. Kehittämistarpeiden yksilöinti on kuitenkin tärkeää. Niiden pohjalta lähtevä kehittäminen avaa taas uusia mahdollisuuksia ja menestystekijöitä.

Yhteisöllisyydestä uutta mielekkyyttä arviointiin

Voimaannuttava arviointi on yhteistyöhön perustuvaa ryhmätoimintaa. Saadessaan toisiltaan tukea ja ohjausta ihmiset voimaantuvat. Voimaannuttavan arvioinnin avulla on mahdollista luoda ympäristö, joka edistää työyhteisön jäsenten voimaantumista ja itsemääräytymistä. Arviointiprosessi on perusluonteeltaan demokraattinen siinä mielessä, että se tarjoaa mahdollisuuden osallistua avoimessa foorumissa tarkastelemaan koko yhteisöä koskevia asioita. Kyseessä on jatkuva toimintojen kehittämisprosessi.

Fetterman (emt.) tähdentää, että yhteisön tai ryhmän saavutukset ja arvot eivät ole pysyviä, vaan ihmiset vaihtuvat ja samalla tavoitteet, tiedot ja arvot muuttuvat. Niin sisäisten kuin ulkoistenkin tekijöiden vaihtelu luo epätasapainoa. Näihin muutoksiin mukautumisessa auttaa Fettermanin mukaan dynaaminen ja responsiivinen arviointi, joka voidaan saavuttaa sisäistämällä ja vakiinnuttamalla interaktiivinen itsearviointikulttuuri. Tällöin arviointi reagoi niihin tarpeisiin, joita toimijalla on esim. toimintansa kehittämiseksi siten, että asetetut tavoitteet voitaisiin saavuttaa paremmin.

Itsearviointin ja reflektoinnin kohteina voivat olla toiminnan ja yhteisön arvot sekä kehittämissuunnitelmien laatiminen ulkoisen asiantuntija-arvioitsijan avulla. Osallistujat oppivat jatkuvasti arvioimaan kehitystään kohti itse määäämiään tavoitteita sekä muokkaamaan suunnitelmiaan ja strategioitaan arvioinnista saadun tiedon mukaan.

Fettermanin (2002, 74) mukaan arviointilogiikan ja -taitojen oppiminen yhdessä vahvistaa taitoa ja tahtoa jatkuvaan itsearviointiin sekä edistää ihmisten itsemääräytymistä. Patton (1997, 156) toteaa, että osallistuminen ja yhteistyö voivat johtaa pitkäkestoiseen sitoutumiseen arviointikeinojen käytössä, ja siten luoda yhteisöllistä oppimiskulttuuria ja yhteistä mielekkyyttä arviointiin.

Edellisissä luvuissa on jo tullut esille voimaannuttavan arvioinnin luonnehdintoja, jotka sopivat hyvin yhteen oppivan organisaation ominaispiirteiden kanssa. Mielenkiintoista on havaita, että Siitosen (1999, 62) esittämällä sisäisen voimantunteen osaluilla (kuvio 1) voitaisiin osin luonnehtia myös oppivaa organisaatioita. Stoll ja Fink (1996) ovat tiivistäneet seuraavasta taulukosta 1 ilmeneviä oppivan organisaation piirteitä eräänlaiseksi oppivan organisaation “seinätauluksi”:



Taulukko 1. *Oppivan organisaation "seinättaulu" (ks. Stoll & Fink 1996)*

PERIAATE	SISÄLTÖ
Yhdessä omaksi koetut tavoitteet	<i>Tiedämme, mihin olemme menossa!</i>
Vastuu menestymisestä	<i>Meidän täytyy onnistua!</i>
Kollegiaalisuus	<i>Me teemme tätä yhdessä!</i>
Jatkuva kehittyminen	<i>Pystymme vielä parempaan!</i>
Riskin ottaminen	<i>Opimme kokeilemalla uutta!</i>
Tuki	<i>On aina joku, joka tarvitsee tukeamme!</i>
Molemminpuolinen kunnioitus	<i>Jokaisella on jotakin tarjottavaa muille!</i>
Avoimuus	<i>Asioista voidaan keskustella!</i>
Juhliminen ja huumori	<i>On kiva olla yhdessä!</i>
Elinikäinen oppiminen	<i>Se kuuluu täällä jokaiselle!</i>

Oppivan työyhteisön kannustava ja rohkaiseva ilmapiiri vahvistaa ihmisten yhteenkuuluvuuden tunnetta ja avaa mahdollisuuksia uuden oppimiselle ja innovatiiviselle toiminnalle. Työyhteisön jäsen voi kokea edellä mainittujen periaatteiden sisällöt, jos hän tuntee kollektiivisen voiman, itsearvostuksen kohenemisen, työyhteisön myönteisen me-hengen sekä kokee olonsa turvalliseksi.

Yksilön ja yhteisön mielekkyyden kokemukset liittyvät motivaatioon: ihmiset tekevät mielellään asioita, jotka he kokevat mielekkäiksi. "Yksilön toiminnan määrän, laadun ja intensiteetin kannalta on tärkeää, miten mielekkääksi, hyödylliseksi ja merkitykselliseksi hän kokee mieltämänsä asian, kielellisen ilmaisun tai teon merkityksen ja tarkoituksen" (Olkinuora 1979, 53). Olkinuoran mukaan yksilöt ovat aikomussuhteissa elämänsä liittyvissä tavoitteissa ja päämäärissä, vuorovaikutussuhteissa muihin itselleen tärkeisiin ihmisiin, minän ilmaisusuhteissa toiminta- ja valintatilanteissa, oppimissuhteessa opittavina oleviin asioihin ja sisältöihin sekä havaintosuhteissa oppimisympäristöönsä ja yleensä vallitseviin olosuhteisiin. Oppivassa työyhteisössä nämä suhteet voidaan tulkita myös yhteisöllisiksi – ei vain yksilöä koskeviksi. Mielekkyyden kokemisen tunteisiin on yhteydessä suuri määrä erilaisia suhteita, joissa yksilö on ympäristönsä kanssa (Entwistle, Entwistle & Tait 1993). Mielekkäiksi koetut asiat ja suhteet voimaannuttavat, kun taas epämielikkäiksi koetut asiat lamauttavat, estävät energian suuntaamista toimintaan ja lisäävät halua vetäytyä pois.

Arvioimalla muutoksen hallintaan

Organisaation perimmäinen oikeutus lähtee siitä, että se kykenee tuottamaan jäsentensä osaamisen ylittävää osaamista. Arvioinnin yhteydessä jaettu osaaminen jalostaa ja kohottaa yhteisöllistä osaamista ja antaa siten paremmat edellytykset koko työyhteisön tuloksellisuudelle. Lisäksi arvioinnin yhteydessä jaetut kokemukset vahvistavat myös työyhteisöllistä muutoksen hallintaa ja henkistä lujuttua kestää muutoksessa muutos- ja kehitysvoimien väliset jännitteet. Tällöin arviointi on parhaimmillaan voimaannuttavaa kasvattaessaan työyhteisön “vääntömomenttia”.

Ketä arvioinnin tulisi sitten kaikessa laajuudessaan voimaannuttaa ja missä tarkoituksessa? Oppimisyhteisössä oppilaiden, opettajien ja vanhempien tulisi olla voimaantumisen ytimessä. Voimaantumisen tulisi palvella koko työyhteisön hyvinvointia ja perustehtävien mukaisia tuloksia sekä ennaltaehkäistä yhteisöllistä toimivuutta heikentäviä tekijöitä. Viimeksi mainittu merkitsee myös sitä, että tuloksellisuutta heikentäviä toimintoja tulisi myös oppia pois, jotta se toisi tilaa uuden oppimiselle ja aidolle itsemääräytymiselle ja -ohjautuvuudelle ja aiempaa syvemmälle ymmärrykselle.

Arviointiosaamisen kehittyessä ja tuottaessa aina parempia tuloksia syntyy tunne siitä, että asiat ovat hallinnassa ja syvemmän tietoisuuden ja näkemyksen pohjalta ohjattavissa. Voimaannuttavan arvioinnin tulisi vahvistaa osaamista, optimoida perustehtävän mukaisia tuloksia, edistää sopeutumista muutoksiin, vahvistaa uudistumista ja yhteisöllistä hyvinvointia.

Kun päätösvaltaa on siirretty oppilaitostasolle, professionaalinen itseohjautuvuus korostuu. Se taas edellyttää arvioinnin vahvistumista erityisesti voimaannuttavan arvioinnin näkökulmasta, jolloin arvioinnin tehtävänä on auttaa asiantuntijuutta auttamaan itseään. Tällöin arviointia ohjaa sitoutuminen totuuteen ja rehellisyyteen.

Miten sitten tulisi sovittaa yhteen perinteiset arvioinnin näkökulmat ja tavoitteet sekä voimaannuttavan arvioinnin idea? Kun voimaannuttavassa arvioinnissa korostetaan itsemääräytymistä ja kehittämistä, perinteinen arviointi tulkitaan systemaattiseksi arviointikohteen arvon ja ansiokkuuden tutkimiseksi. Miten sovitaan siis yhteen tuloksellisuus ja sen valvonta ja toimintojen kehittäminen sekä työyhteisön itsemääräytyminen? Yhteensovittaminen voi parhaimmillaan onnistua, kun nähdään arviointi keinoksi, jolla voidaan edistää tulosten arvioinnin ohella itsemääräytymistä ja kehittämiseen sitoutumista arvioinnin kaikissa vaiheissa. Tämä voidaan tehdä sortumatta sen enempiä liikaa heikkouksia ja vaikeuksia korostaviin analyysihin, itseruoskintaan, kuin liiallisen positiivisen identiteetin harhaankaan eli yleishyvän “hymistelyyn”. Positiivisen identiteetin harhalle leimallinen itseriittoisuus vääristää aitoa todellisuuskäsitystä ja heikentää aitoa voimaantumista. Kyseessä on siis paljolti uusi lähestymistapa ja asenne arviointiin, jonka menetelmät ovat paljolti samat kuin arvioinnissa yleensä. Vaikka arviointiin liittyy aina vallankäyttöä, sen tulisi voimaannuttavan arvioinnin yhteydessä saada legitimeettinsä yhteisistä kehittämisintresseistä eikä esi-



merkiksi työyhteisön jäseniin kohdennetusta vallankäytöstä.

Arvioinnin voimaantumista edistävä näkökulma tuo arviointiin innovatiivisen ulottuvuuden. Siinä korostuvat osallistavuus, yhteistoiminnallisuus sekä yksilölliset ja yhteisölliset vaikutusmahdollisuudet ongelman ratkaisuihin ja päätöksentekoon. Voimaannuttava arviointi voi parhaimmillaan vahvistaa yhteisölliselle kehitymiselle ja yhteisön energiataloudelle välttämätöntä sosiaalista ja tunnepääomaa sekä arvioinnin osaamispääomaa. Näin se luo suotuisan maaperän yksilön jatkuvalla uudistuvuudelle, uusille innovaatioille ja muutoksen hallinnalle yllättävästikin muuttuvassa toimintaympäristössä. Voimaannuttava arviointi korostaa itsemääräytymistä, kehittämisprosessia ja sen ymmärtämistä ja voi siten johtaa siihen, että toiminnan tuloksellisuuden arviointi voi jäädä vähemmälle huomiolle. Tähän on arviointituloksia tulkittaessa kiinnitettävä tietoisesti huomiota ja pyrittävä arviointiprosessissa molempien optimointiin. Tärkeää on myös tasapainottaa ulkoisen arvioitsijan ja yhteisön jäsenten intressit. Näin tulisi periaatteessa toteutua kansallisessa arvioinnissa myös ulkoisen arvioinnin ja paikallisen arvioinnin suhde.

Lähteet

- Adams, R. 1991. *Protests by pupils: Empowerment, schooling and the state*. Basingstoke: Falmer Press.
- Appelbaum, S., Herbert, D. & Leroux, D. 1999. Empowerment: power, culture and leadership – a strategy or fad for the millennium? *Journal of Workplace learning: Employee counselling Today*. II (7), 233–254.
- Barksdale-Ladd, M. & Thomas, K. 1996. The development of empowerment in reading instruction in eight elementary teachers. *Teaching and Teacher Education* 12, 161–178.
- Beach, A. 1996. Empowerment to the people: creating an atmosphere for growth. *Empowerment in Organizations* 4 (1), 29–33.
- Bennis, W. & Nanus, B. 1985. *Johtajat ja johtajuus*. Tampere: Weilin & Göös.
- Entwistle, N., Entwistle, A. & Tait, H. 1993. Academic understanding and contexts to enhance it: A perspective from research on student learning. Teoksessa T. M. Duffy, J. Lowyck & D. H. Jonassen (toim.) *Designing environments for constructive learning*. Berlin: Springer, 331–357.
- Fetterman, D. M. 1994. Steps of empowerment evaluation: From California to Cape Town. *Evaluation and Program Planning* 17 (3), 305–313.
- Fetterman, D. M. 1996. Empowerment evaluation: An introduction to theory and practice. Teoksessa D. Fetterman, S. Kaftarian & A. Wandersman (toim.) *Empowerment evaluation. Knowledge and tools for self-assessment & accountability*. London: SAGE Publications, 3–46.
- Fetterman, D. M. 1998. Empowerment evaluation and accreditation in higher education. Teoksessa E. Chelimsky & W. Shadish (toim.) *Evaluation for the 21st century: A handbook*. Thousand Oaks, CA: Sage, 381–395.
- Fetterman, D. M. 2003. Empowerment evaluation strikes a responsive cord. Teoksessa I. Steward Donaldson & M. Scriven (toim.) *Evaluation social programs and problems. Visions for the new millennium*. London: Lawrence Erlbaum, 63–76.
- Habermas, J. 1984. *The theory of communicative action*. Boston, MA: Beacon Press.

- Heikkilä-Laakso, K. & Heikkilä, J. 1997. Innovatiivisuutta etsimässä. Irtiottoa keskinkertaisuudesta. Turun yliopiston kasvatustieteiden tiedekunnan julkaisusarja B 57.
- Karl, M. 1995. Women and empowerment: Participation and decision making. Women & World Development Series. London & New Jersey: Zed Books.
- Kiikkala, I., Immonen, T., Sohlman, B., Ahonen, J. & Siitonen J. 2000. Sosiaali- ja terveydenhuollon työntekijä muutosten pyörteissä. Dialogi 6/2000, 1–8. Saatavana www-muodossa: <<http://www.stakes.fi/paasky/jutut/pyorre.html>>
- Koivunen, N. 2003. Leadership in symphony orchestras. Discursive and aesthetic practices. University of Tampere, School of Business Administration. Acta Electronica Universitatis Tamperensis 227. Academic dissertation.
- Lagerspetz, E. 1998. Itsemäärääminen ja valta. Teoksessa J. Pietarinen, V. Launis, J. Rääkkä, E. Lagerpetz, M. Rauhala & M. Oksanen (toim.) Oikeus määräämiseen. Helsinki: Edita, 97–141.
- Levin, H. 1996. Empowerment evaluation and accelerated schools. Teoksessa D. Fetterman, S. Kaftarian & A. Wandersman (toim.) Empowerment evaluation. Knowledge and tools for self-assessment & Accountability. London: Sage Publications, 49–78.
- Lyytinen, H.K., Vaso, J., Kangasoja, M. & Rönkä, A. 2002. Arvioi ja opi. Opi ja arvioi. Vantaa: Printel Oy.
- Neufeld, J. & Grimmett, P. 1994. The authenticity for struggle. Teoksessa P. Grimmett & J. Neufeld (toim.) Teacher development and the struggle for authenticity. Professional growth and restructuring in the context of change. New York: Teacher College Press, 205–232.
- Nikkanen, P. Lyytinen, H. K. 1996. Oppiva koulu ja itsearviointi. Jyväskylän yliopisto: Kasvatustieteiden tutkimuslaitos.
- Olkinuora, E. 1979. Oppimisen ja opiskelun mielekkyys. Katsaus kirjallisuuteen ja lähtökohtiin sekä tutkimusprojektin esittely. Jyväskylän yliopisto. Kasvatustieteiden tutkimuslaitoksen selosteita ja tiedotteita 121.
- Patton, M. 1997. Toward distinguishing empowerment evaluation and placing it in a larger context. Evaluation Practice 18 (2), 147–163. Saatavilla www-muodossa: <<http://www.stanford.edu/~davidf/patton.html>>
- Pollari, P. 2000. "This is my portfolio". Portfolios in upper secondary school English studies. University of Jyväskylä: Institute for Educational Research.
- Robinson, H. 1994. The ethnography of empowerment. London: Falmer Press.
- Rodwell, C. 1996. An analysis of the concept of empowerment. Journal of Advanced Nursing 23, 305–313.
- Ruohotie, P. 1998. Motivaatio, tahto ja oppiminen. Helsinki: Edita.
- Senge, P. 1990. The fifth discipline. The art and practice of the learning organization. New York: Doubleday/Currency.
- Siitonen, J. 1999. Voimaantumisteorian perusteiden hahmottelua. Oulun yliopisto. Oulun opettajakoulutuslaitos. Acta Universitatis Ouluensis. Scientiae rerum socialium. E 37.
- Siitonen, J. & Robinson, H.A. 1998. Empowerment: Links to teachers' professional growth. Teoksessa R. Erkkilä, A. Willman & L. Syrjälä (toim.) Promoting teachers' personal and professional growth. University of Oulu. Department of teacher education. Acta universitatis Ouluensis E 32.
- Stoll, L. & Fink, D. 1996. Changing our schools. Linking school effectiveness and school improvement. Buckingham: Open University Press.
- Suonsivu, K. 2001. "Tässä työpaikassa on ilo työskennellä kaikin puolin". Hoitohenkilöstön voimaantuminen (empowerment). Tehy ry. Sarja B: Selvityksiä 5.
- Thomas, K. & Velthouse, B. 1990. Cognitive elements of empowerment: an "Interpretive" model of intrinsic task motivation. Academy of Management Review 15 (4), 666–681.
- Vanderplaat, M. 1995. Beyond technique: Issues in evaluating for empowerment. Evaluation 1, 81–96.
- Zimmerman, M. 1995. Psychological empowerment: issues and illustrations. American Journal of Community Psychology 23(5), 581–599.



Arviointi opetuksen ja oppimisen tueksi

Artikkelissa tarkastellaan tuloksiltaan laadukkaiden oppimis–opetus-prosessien ja niiden ohjaamisen kannalta tärkeitä arvioinnin tehtäviä, välineitä ja ilmene-mismuotoja. Opetuksen laadun oleellisimmaksi kriteeriksi nähdään oppilaan monipuolinen inhimillinen kasvu. Tärkeäksi osoitetaan motivoiva, ohjaava ja kehittävä arviointi, joka hyödyntää monimenetelmäistä tiedonkeruuta. Ulkoiseen kvantitatiiviseen arviointitietoon ja indikaattoreihin otetaan jossain määrin kriit-tinen kanta opetusta tukevan arvioinnin näkökulmasta. Opetusinnovaatioiden toteutumisen kannalta ensisijaiseksi kuvataan opettajan pedagogista ajattelua ja rikasta pedagogista sisältötietoa kartuttava arviointi. Artikkelissa suositetaan, että arviointikäytäntöjen ja -välineiden kehittämisessä huomioidaan konstrukti-ivistinen oppimiskäsitys, optimistinen ihmis- ja kasvatuskäsitys sekä kommuni-katiivisen arvioinnin vaatimukset.

Asiasanat: oppimis–opetus-prosessi, kehittävä ja kommunikatiivinen arviointi, prosessiarviointi, opettajan pedagoginen ajattelu

Koulutuksen arviointineuvoston ohjelmassa on useita viittauksia kysymyksiin, jotka koskevat oppimis–opetus-tapahtumaa ja sen tuloksellista ohjaamista, eli ovat luon-teeltaan syvästi pedagogisia. Esimerkiksi perusopetuksessa korostetaan oppijan aktiivisuuden, itsenäisyyden ja toiminnallisuuden edistämistä, ja lukiossa nousevat esil-le jatko-opintojen kannalta tärkeät oppimaan oppimisen valmiudet ja kriittinen kansa-laisuus. Ammatillisessa koulutuksessa työssäoppimiseen on kohdistettu yhä enem-män odotuksia, ja aikuiskoulutuksen tulisi vastata valmiuksiltaan erilaisten oppijoi-den tarpeisiin. Tässä artikkelissa tarkastellaan, millainen arviointi voi edistää peda-gogisia prosesseja, jotka turvaavat yllä mainittujen tavoitteiden saavuttamisen.

Arviointi, opetustapahtuma ja opettajan ajattelu

Oppimisen ja opetuksen arviointia, eli *pedagogista arvon tai ansion määrittelyä*, on käsitteellistetty kirjallisuudessa lukuisin tavoin. On puhuttu diagnostisesta, formatiivisesta ja summatiivisesta arvioinnista tai eroteltu sanallinen ja laadullinen arviointi numeerisesta ja määrällisestä. Tuttua on myös väittely prosessin ja produktin ensisijaisuudesta tai siitä, milloin ja millaisessa suhteessa pitäisi painottaa kognitiivista tai sosioemotionaalista arviointia. Opetuksen kehittämisessä on samanaikaisesti vallalla osin keskenään ristiriitaisia pyrintöjä: toisaalta ihaillaan autenttista, prosessien laadullista arviointia ja toisaalta valtakunnalliset, määrällisesti painottuneet päättökokeet ja ulkoinen arviointi valtaavat koulutusmarkkinoita. Opetuksen ja oppimisen luontaisen monimuotoisuuden vuoksi kansallisen ja paikallisen arvioinnin kehittämisessä olisi vaalittava erilaisia metodologisia lähestymistapoja ja kartettava yksipuolisia trendejä.

Vaikka nykykäsityksen mukaisesti tähdennetään laadullista, kontekstuaalista ja prosessiin kohdistuvaa arviointia, sen pitää sitoutua aina tavoitteisiin: ellei päämäärä ole selvillä, arviointi ajautuu väärin uumiin. Ammattitaitoinen evaluointi edellyttää käsitteellistämistä, harkittua tiedonkeruuta, analyysia ja päättelyä, ts. *tutkimuksellista lähestymistä* arvioitavaan asiakokonaisuuteen. Oppimis- ja opetusprosesseja koskevan arvioinnin tulee kohdentua tavalla, joka mahdollistaa pätevän päättelyn esimerkiksi ongelmanratkaisukyvyyn, matemaattisen osaamisen tai sosiaalisten taitojen tilasta ja kehityksestä.

Opetuksen arviointi liittyy paitsi oppimiseen myös opettajan omaan kehittymiseen. Tutkiva opettaja on ongelmanratkaisija, joka kykenee pedagogisia prosesseja koskevan tiedon hankintaan ja sen kriittiseen arviointiin. Opettajankoulutuksen ja koko opettajauran tavoitteeksi voidaan perustellusti asettaa *pedagogisen ajattelun* edistäminen (Kansanen ym. 2000; Kompf & Denicolo 2003). Tämä tarkoittaa jatkuvasti tarkentuvaa havainnointi- ja päättelykykyä oppimis-opetus-prosessien ominaispiirteistä ja tavoista, joilla niihin voi vaikuttaa tuloksellisesti. Näin opettajaprofession näkökulma selvästi antaa suuntaa arviointikäytännöille, jotka voivat rikastaa opetusasian-tuntijuutta. Tämä puolestaan säteilee positiivisesti oppimisen ohjaamiseen.

Pedagogisen arvioinnin tehtävät ja välineet

Arviointi voi olla luonteeltaan toteavaa, motivoivaa, ohjaavaa ja ennustavaa (Lahdes 1997; Helakorpi & Olkinuora 1997). Määrällinen poikkileikkaustyyppinen diagnosointi, jossa kartoitetaan, millaisia tavoitteita oppimiselle ja opetukselle on asetettu ja miten tavoitteisiin on tietyllä hetkellä ylletty, on tyypiltään *toteavaa*. Opetuksen tuen kannalta se on välttämätön lähtötilan ilmaisija, mutta opettajan ja oppilaiden kehittymisen eli tahtotilan synnyttämisen kannalta toteavan arvioinnin vaikutus on lyhytaikainen.

Samantapaisen luonnehdinnan voi esittää myös arvioinnin *ennustavasta* tehtävästä, vaikka se oikeaoppisesti tulkiten onkin merkittävä avarrus pelkkään toteamiseen verrattuna. Opetuksen ja kasvatuksen tulisi aina kurkottaa tulevaisuuteen, vaikka ihmisen oppimisen ja kasvun ennakointi ja mallintaminen on kyseisten ilmiöiden ominaisuuteen vuoksi ongelmallista. Monisäikeisyyttä ei tule mittaamis- ja tulosvastuun nolla unohtaa, jotta oppilasta tai oppilaitosta ei kategorisoidaisi hätiköiden ja kehitystä tyypistävällä tavalla.

Motivoiva ja ohjaava arviointi ovat oppimis-opetus-prosessien laadun turvaamisen avaimia. Ne koskettavat jokaista opettajaa ja oppilasta, ja ilman niitä laadukas opetus ja oppiminen eivät ole mahdollisia. Jotta oppijan aktiivisuus ja yhteistyötahto säilyy, tarvitaan jatkuvasti tietoa opiskelun kulusta, ja palautetta pitää saada muutenkin kuin opettajalta ja kirjallisista tuotoksista. Motivoiva arviointi on merkityksellistä myös opettajalle, ja yksi palkitseva tapa saada sitä on oppilaiden edistymisen kautta.

Kansallisen tai kansainvälisen *ulkoisen arvioinnin keinovalikoima* on tavallisesti *kvantitatiivinen* ja keskittyy kognitiiviseen tavoitealueeseen (esim. Jakku-Sihvonen & Heinonen 2001). Koulutusjärjestelyistä voidaan tuottaa erilaisia indikaattoreita, joihin tosin tulee aina suhtautua kriittisesti niiden metodologisten heikkouksien vuoksi (Hutmacher ym. 2001). Jos esimerkiksi evaluoidaan oppiainekohtaisia koulusaavutuksia, lukumäärältään suuret oppilasotokset voidaan tavoittaa vain kyselylomaketyyppisin ratkaisuin (esim. kansalliset koulukokeet tai asennemittaukset). Näiden vahvuutena on tutkimusotteisen toteutuksen ansiosta hyvä validiteetti ja reliabiliteetti. Tämä on osaamista, josta tulisi maassamme kantaa huolta kvalitatiivisen metodologian valtakaudellakin. Opetusprosessi jää kvantitatiivisia indikaattoreita käytettäessä tavallisesti mustaksi laatikoksi, mutta löydökset voivat – jos tulkinnassa säilytetään suhteellisuudentaju – lisätä palasensa arviointitietomosaikkiin.

Oppilaitoksissa on vähitellen opittu hyödyntämään konstruktivistisesti orientoituneen oppimisenäkemyksen mukaisesti vaikkapa erimuotoista *autenttista arviointia, jatkuvaa havainnointia, portfolioita, oppimispäiväkirjoja, aineisto- tai näyttökokeita sekä arviointikeskusteluja* (Darling-Hammond 1993; Niikko 2000; Uusikylä & Atjonen 2000; Virta 1999; Vuorinen 2000). Ne palvelevat oivallisesti produktin ja prosessin ymmärtämisessä niin opettajia, oppilaita kuin heidän muodostamiaan pienyhteisöitä, mutta trendien tai järjestelmien kuvaajiksi niistä ei ole. Vaikka mainittuihin evaluointityökaluihin liittyy subjektiivisuuden ongelmia, ne muistuttavat siitä perusasiasta, että vain ihminen voi tulkita itseään tai toisia ihmisiä. Tulkinta on myös syvällisempää kuin mikään koe tai testi.

Opetustapahtumaa ohjaava ja tukeva arviointitieto

Pedagogiselta kannalta arvioinnin tehtävä on huolehtia siitä, että opetus säilyy laadukkaana. Laadukkuuden perimmäisenä kriteerinä yleissivistävässä koulussa on *op-*

pilaiden monipuolinen inhimillinen kasvu. Kasvun uskottava todentaminen on sekä sisällöllisesti että metodologisesti haasteellista, mikä johtaa arvioijat helposti harha-teille: lavenevat arviointijärjestelmät alkavat elää omaa byrokraattista elämäänsä ja kadottavat sensitiivisyytensä opettajien ja oppilaiden arkitodellisuudelle. Arvioinnin varsinaisista subjekteista tulee objekteja, tavoitteet muuttuvat abstrakteiksi ja evaluointia tehdään varmuuden vuoksi.

Opetuksen arviointi on järkevää vain silloin, kun (Uusikylä & Atjonen 2000, 165)

- pohjana olevat arvot ja tavoitteet ilmaistaan selvästi
- kaikki osapuolet ymmärtävät ja hyväksyvät arvioinnin perusteet ja pitävät niitä oikeina ja tasapuolisina
- arviointi toteutetaan myös käytännössä siten kuin on sovittu, ja kaikki osapuolet ymmärtävät, mitä ideologiaa ja tavoitteita toimenpiteet palvelevat
- arviointimenetelmää tarkistetaan ja uudistetaan määräjain.

Edellä oleva luettelo auttaa vastaamaan kysymykseen, milloin arviointi ei tue opetusta. Jos arvioinnin päämäärä on hukassa, perusteita ei pidetä tasapuolisina, toteutus ja lupaukset eivät kohtaa toisiaan ja samat virheet toistetaan, palvelaan oppilaiden inhimillisen kasvun kannalta toisarvoisia, jopa epäeettisiä päämääriä. Mikäli arviointi lisäksi osoittelee vain virheitä, on metodologisesti epäluotettavaa, työllistää oppilaitoksia ylenpalttisesti ja jos tulokset julkistetaan vasta pitkän viiveen päästä ja asiaankuulumattomille tahoille, opettajien halu arviointiyhteistyöhön laantuu ymmärrettävästi.

Millaista tietoa opettaja sitten tarvitsee laadukkaana opetuksen toteuttamiseksi (Grossman 1990; Shulman 1987)? Ensisijaisinta on tieto oppimis-opetus-prosesseista, koska sitä ei voi korvata muilla tiedon lajeilla (Anderson 2003; Wragg 2001). Opettajalla tulee olla käsitys eri-ikäisten oppijoiden normaalista oppimisesta, sen keskeisistä ehdoista ja keinoista, joiden avulla voi tunnistaa poikkeavat kehitystarpeet ja tukea niitä (*prosessitieto, menetelmätieto*). Arviointia tapahtuu eniten ja luontevimmin juuri opetustapahtuman aikana, koska kaikki verbaalinen ja nonverbaalinen vuorovaikutus kartuttaa taito-tietoa siitä, miten opetusta tulisi kohdentaa (Brooks 2002, 59–63). Siksi prosessitarkastelulla on produktiarvioinnin rinnalla korvaamaton arvo. Sen tulisi myös toteutua tavalla, joka rakentaa oppilaiden positiivista kykykäsitystä (Kasanen 2003) eikä raunioita sitä.

Pelkkä *sisältötieto* (aineenhallinta) ja sitä koskeva arviointitieto ei riitä opettajalle, vaikkei sen merkitystä erityisesti koulutuksen toisella tai kolmannella asteella pidä väheksyä. Sisältötietoa tulee avartaa *opetussuunnitelmätiedolla*, joka auttaa oppimiskokonaisuuksien hahmottamisessa sekä virittää tavoitteiden ja filosofisten perusteiden pohdintaan. Kun laaditaan koulun tai oppilaitoksen opetussuunnitelmaa valtakunnallisten perusteiden pohjalta, vallitsevaa käytäntöä joudutaan tehokkaasti kyseenalaistamaan etusijalle nostettavien painalojen kartoittamiseksi. Myös *kontekstittietoon* tulisi kiinnittää arviointien toteutuksessa enenevästi huomiota. Tällöin esimer-

kiksi oppiainekohtaisista oppimistuloksista tai opiskeluvalmiuksista tehtyjen havaintojen pohjalta voitaisiin tehdä kestäviä praktisia päätelmiä.

Monet tutkijat puhuvat sisältö- ja menetelmätiedon yhdistävästä käsitteestä *pedagoginen sisältötieto* (Aaltonen 2003; Haring 2003). Jokainen opittavaksi tarkoitettu teema tulee organisoida, sovittaa oppijoiden tarpeisiin ja kiinnostuksen kohteisiin sekä “panna esille” (luento, työpaja, verkkokurssi, opintotehtävä, ryhmätyö) jollakin tavalla. Rakentaakseen itselleen rikkaan pedagogisen sisältötietämyksen opettaja – toimii hän sitten perusopetuksessa, lukiossa, ammatillisessa oppilaitoksessa tai aikuis-koulutuksessa – tarvitsee monenlaista arviointitietoa. Sitä hänen tulee keskeisin osin tuottaa ja tulkita itse käyttäen monia eri menetelmiä, sillä keskiarvo- ja trendikuvauksilla on taipumus juuttua etäälle pedagogisesta vuorovaikutuksesta.

Oppimiskäsityksen katalysoimia kehittämishaasteita

Perinteiset arviointimenetelmät eivät voi vastata kaikkiin niihin vaatimuksiin, joita nykyisestä oppimisnäkemyksestä heijastuu. Osaamista tulee käsitteellistää laaja-alaisesti, ja myös oppimisympäristöjen on mahdollistettava arviointi, joka tukee oppimisprosessia. Opettajan oman asiantuntijuuden kehityksen kannalta on ensisijaista *kehittävä arviointi*. Sen kautta omista pedagogisista ratkaisuista kertyy palautetta, mikä helpottaa vaihtelevien oppimis- ja opetusprosessien ohjaamista (Bereiter & Scardamalia 1993; Kauppi ym. 2003).

Oppimiskäsityksen muutos on johtanut Glaserin mukaan ainakin seuraaviin arviointiin kohdistuviin kehitysodotuksiin (Helakorpi & Olkinuora 1997):

- Oppimisen arviointiin on sisällytettävä laaja-alaisia tehtäviä ja kontekstuaalisten taitojen arviointia.
- Taitojen ja tietojen arvioinnissa on painotettava sekä oppimisprosessin että tulosten arviointia. Arviointikriteerien tulee myös motivoida ja suunnata oppimista ja opiskelua.
- Arvioinnin on opetettava itsearviointia. Tämä edellyttää opetuksen ja arvioinnin kiinteää sidosta.
- Oppimisen arvioinnin on tapahduttava entistä enemmän sosiaalisissa tilanteissa.
- Oppimisen tehokkuutta tulee arvioida entistä enemmän sen suhteen, miten se vaikuttaa opettamiseen.

Koulun ja paikallisen tason evaluoinnissa on mahdollista yltää ensimmäisen ehdon toetutumiseen. Sen sijaan esimerkiksi kansalliset päättökokeet taipuvat vain työläästi jo pelkästään *laaja-alaisuuden* vaatimukseen, *kontekstisensitiivisyydestä* puhumatta-kaan. Pedagogiselta kannalta katsoen tulevaisuudessa pitää pureutua vakavasti *prosessien validiin taltiointiin* tähtävien arviointikäytänteiden kehittämiseen. Tämä on

tarpeen siksi, että prosessien määrällinen ja laadullinen evaluointi vakauttaa oleellisesti opetuksen tuloksellisuutta koskevia päätelmiä (ks. Linn & Gronlund 2000; Oksanen 2003). Toisaalta yksinomaan prosessien punnitseminen ei riitä. Vasta produktiarviointi voi varmistaa tuloksellisuuden, joskaan se ei yksinään auta opetuksen kohentamisessa; viimeksi mainitusta ovat käyviä esimerkkejä koulujen rankinglistat tai puhtaasti kognitiiviset koulusaavutusmittaukset.

Itsearviointiin on kohdistettu paljon odotuksia päiväkotii-ikäisistä yliopisto-opiskelijoihin ja yksilöistä oppilaitoksiin saakka (Räisänen & Frisk 1996; Stipek ym. 1992). Suunta on oikea, sillä sekä opettajan että oppilaan tulee elämässä selviytyäkseen oppia itsenäisesti ankkuroimaan toimintansa tarkoituksenmukaisuutta. Samalla on tärkeää muistaa eri-ikäisten oppijoiden itsearvioinnin edellytykset, jotta niitä pystytään systemaattisesti kehittämään; itsearviointia ei opi itsekseen. Opettajan vastuuta opetus- ja kasvatustieteen asiantuntijana ei pidä ohentaa, puhuttiin sitten varhaiskasvatuksesta tai aikuiskoulutuksesta. Hänen tulee antaa palautetta ja toimia luotettavana reflektiopintana, jotta itsearvioinnin yli- ja alilyöntejä voidaan välttää. Koulutetulla aikuisella on useimmiten oppilaita enemmän tietoa ja kokemusta kasvun ja oppimisen toivottavista suunnista ja niihin kulkemisen kelvollisista reiteistä.

Monien jaetun asiantuntemuksen vaikuttavuutta osoittavien oppimistutkimusten mukaan oppimisen ja sen arvioinnin *sosiaalisen luonteen* tunnustaminen on niin ikään tavoiteltavaa (Hakkarainen ym. 2004). Työ- ja yksityiselämän vaatimukset haastavat työskentelemään yhdessä, ja tähän tulee saada harjoitusta jo koulussa. Silti yksilöllisiä eroja ei tule tyypistää arvioinnilla, koska kaikille ihmisille ei ole luontaista tai tuloksellista työskennellä ryhmässä.

Tämän artikkelin pääsanoman kannalta kiinnostava ja tärkeä on myös Glaserin luetelman viimeinen peruste: *evaluoinnin tulee* kaikissa muodoissaan *palvella* ainakin epäsuorasti *oppilaan ja opiskelijan kehittymistä*. Siinä *opetus on keskeinen välittävä aktiiviteetti*. Samaa perustetta on teroitettu myös opettajankoulutuksessa toisinpäin käännettynä: opetusta pitää arvioida yhä enemmän lähtien niistä vaikutuksista, joita sillä on oppimiseen. Lienee realistista tulkita vaikutus-sana tässä yhteydessä ei-kausaaliseksi ja määräaikaiseksi, vaikkei laajakantoisten vaikuttavuuskysymysten (outcome) tulisi hämärtyä lähitavoitteiden ja opetuksen arkirutiinien varjossa.

Oppijakeskeiseen oppimisen vaikuttavuuden arviointiin on pyritty löytämään menetelmiä, jotka yhdistyisivät *opiskelu-ympäristöön ja oppimisen tavoitteena olevan tuloksen käyttöympäristöön*. Tällöin merkittäviksi arviointikohteiksi nousevat esimerkiksi oppimateriaalit ja opetusvälineet sekä niiden kyky tukea halutunlaista oppimista ja opettamista. Yleisivistävässä koulussa edellä mainittu haaste on erityisen vaativa, eikä sitä ratkaista pelkästään tieto- ja viestintätekniikan pedagogisten sovellusten kehittämisellä (vrt. Jonassen & Land 2000). Ammatillisessa koulutuksessa, esimerkiksi työssäoppimisessa ja näyttökokeissa, on saavutettu jo merkittäviä edistysaskeleita opiskelu- ja käyttöympäristöjen integroinnin suuntaan, vaikkei ylioptimismiin tule sortua sielläkään (vrt. Varila & Rekola 2004).

Pedagogisen arvioinnin ihmiskäsitys

Nykyaikaisen pedagogisen arvioinnin taustalta voi tunnistaa humanistisen, optimistisen ihmiskäsityksen. Opettaja tai oppilas on osaava, kehittyvä ja aktiivinen oppimis-opetus-prosessin osapuoli, jota ei tule arvioinnilla kontrolloida ja alistaa vaan tukea ja motivoida. Näkemys haastaa suhtautumaan kriittisesti ulkoiseen arviointiin, joka helposti jakaa inhimillisen ihmisen näennäisesti mitattavissa oleviin osiin ja etäännyttää arviointitiedon sen loppukäyttäjistä. Evaluoinnin ”pohjoismaisen pehmeästä” otteesta, johon yhdistyy hyvin koulutettujen opettajien didaktinen vastuu, on syytä pitää kiinni tulostavastuu-, standardi- ja indikaattori-intoisen kansainvälisen arviointikulttuurin (ks. Raivola 2002) paineiden keskellä. Luottamus lisää luottamusta, virheiden osoittelu puolestaan mielipahaa, joka tukahduttaa kehittämisen.

Tutkijat raportoivat määräväleihin toistuviin kokeisiin ja testeihin luottavista maista (esimerkkejä näistä ks. Brooks 2002, 160–180) liian usein, ettei opetuksen innovatiivisia ratkaisuja pystytä toteuttamaan. Syynä on, että kokeet ja testit ovat ihmis- ja kasvatuskäsityksiltään kontrolloivia ja kohdettaan yhdenmukaistavia, koska ne keskittyvät tyypillisesti loppukäyttäjätymisen arviointiin. Mikäli kouluja ja opetusta halutaan aidosti profiloida oppilaiden, opettajien ja oppilaitosten tarpeiden ja osaamisen perusteella, kaikkia ei voida evaluoida samanlaisilla mittatikuilla. Opetuksen etävoitteeksi tulee nähdä uusien asioiden oppiminen eli viime kädessä innovaatioiden tuottaminen. Koska innovaatioiden kriteerit muuttuvat ajan myötä, opetuksen objektiivinen arviointi käy oikeastaan mahdottomaksi. Asia ei kuitenkaan korjaannu lisäämällä ja tiuhtamalla arviointeja. Sika ei siis liho siksi, että sitä punnitaan päivittäin.

Koulutuksen laatu on vaikuttavuutta, joka on subjektiivisesti koetun koulutustarpeen tyydyttämistä ja jossa jokaisen ihmisen tavoitteet muodostavat ainutlaatuisen kokonaisuuden. Etenkin ulkoisen evaluoinnin ongelmat kumpuavat siitä, että koulutuksen tulisi vastata myös yhteiskunnan tarpeisiin. Nämä tarpeet puolestaan ilmaistaan intressitaisteluiden muovaamissa koulutuspoliittisissa kompromissitavoitteissa. Tällöin koulutus menettää osan itseisarvostaan, ja arviointiperustaksi nousee oppimistulosten käyttö- ja välinearvo (Raivola ym. 2000). Oppilailta ja opiskelijoilta ei kuitenkaan voida kokeilla mitä tahansa; jokainen yritys vaatii aikansa ja jättää jälkensä, jota ei voida mitätöidä tai ”ottaa uusiksi”. Tässä konkretisoituu opettajien eettinen evaluointivastuu esimerkiksi tasa-arvon, oikeudenmukaisuuden ja avoimuuden ääntä kuunnellen.

Kommunikatiivisen evaluoinnin lupauksia

Opetusta tukevan arvioinnin luonnetta pohdittaessa on kommunikatiivisen evaluoinnin käsite hedelmällinen (Niemi 2000). Sen ensimmäinen tehtävä on *paljastaminen* (revelation), mikä ei tarkoita vain tavoitteiden saavuttamisen punnitsemista vaan

myös itse tavoitteiden kriittistä tarkastelua. Oppilaiden tekemä itsearviointi, opettajan, oppilaan ja vanhempien väliset arviointikeskustelut, erityyppiset summatiiviset tarkastelut ja opettajan pedagoginen reflektio valottavat opetuksen kehittämisen kannalta tärkeitä muutostarpeita. Taiten käytettyinä myös kansalliset ja kansainväliset poikkileikkausarviointit saattavat sysätä kehitystä innovatiiviseen suuntaan.

Toisena tehtävänä on *ennakointi* (anticipation), joka korostuu muuttuvissa oloissa (esim. koulutuksen siirtymävaiheet). Epävarmasta tulevaisuudesta on tunnetusti vaikea poimia yksiselitteisiä päämääriä, joten ratkaisevaa on epävarman tiedon arviointi monimuotoisten tietolähteiden varassa (vrt. triangulaatio tutkimusmetodologiassa). Merkittävimmän avun tähän antaa lopulta vain opettajan vankka pedagoginen asiantuntemus oppijan kehityksestä ja sen tukemisen keinoista. Opettajan vastuulla on tulkita eri keinoin tuotetun arviointitiedon muodostamaa kokonaisuutta ja sen implikoimaa tulevaisuutta.

Kolmas ominaispiirre on kenties oleellisin eli pyrkimys *kommunikaatioon ja yhteistyöhön* (aims for communication and partnership). Ilman niitä eivät edellä kuvatut kaksi tehtävää voi toteutua konstruktivistisen, optimistisen oppimis- ja ihmiskäsityksen hengessä. Moniammatillinen yhteistyö muodostaa opetusasiantuntijuuden perustan, kun tavoitteena on tulkita erilaista tietoa ja vaalia sen pedagogista hyödynnettävyyttä. Yhteistyötä tarvitaan myös oppilaiden kesken vertaisarvioinnin toteuttamiseksi; tärkein haaste on ehkä kuitenkin vertikaalisen yhteistyön kehittyminen opettajan ja oppilaan välisessä suhteessa. Opettajan auktoriteettia ja valtaa (Harjunen 2004) tulee käyttää niin, ettei oppilasta redusoida vain arvioinnin objektiksi.

Luottamusta ja tulevaisuuteen katsomista

Koulutuksen arviointineuvoston ohjelmassa todetaan heti sen toimintaperiaatteita kuvaavan luvun alussa, että arvioinnin vaikuttavuus edellyttää arvioitsijan ja arviointien kohteiden välistä *luottamusta*. Samassa luvussa puolletaan arvioinnin *kehittävää* luonnetta, eli arvioinnin tulee antaa relevanttia tietoa oppimista edistävästä prosesseista. Näiden kahden periaatteen toteutuminen myös arkisessa evaluointityössä, ei vain arviointineuvoston ponnisteluisissa, on ollut tämän artikkelin keskiössä. Samat periaatteet toimivat myös tulevan toiminnan toiveikkaina evästyksinä.

Laadukkaan opetuksen ja tuloksellisen oppimisen lähtökohtana tulee olla optimistinen ihmis- ja kasvatuskäsitys. Ainoastaan usko aiempaa *paremman tulevaisuuden tekemiseen* koulutuksen keinoin sekä luottamus yksilöiden ja heidän muodostamiensa yhteisöjen kehityspotentiaaliin voi kypsyttää opetus- ja kasvatustyössä pitkällä aikavälillä tuoreita hedelmiä. Valoisa katsantokanta velvoittaa myös arvioinnissa kommunikatiivisuuteen, kärsivällisyyteen ja tahdikkuuteen. Tämä siitä huolimatta, että ajassa saattaa liikkua mittaamisen, kiirehtimisen ja piittaamattomuuden valtavirtoja, jotka ovat ymmärrettäviä ei-pedagogisessa suhteessa.

Rakentavista lähtökohdista ponnistava arviointi ei tarkoita kritiikittömyyttä olemassa olevia pedagogisia käytänteitä kohtaan. Toisinaan huonostikin suunnitellun piikin pistäminen lihaan ulkopuolelta voi stimuloida oikeita innovaatioita, mutta se ei voi olla käytäntö, jonka varassa evaluointi lepää. Sekä ulkopuolisuuteen että piikin kerta-luonteisuuteen liittyy liikaa epäoleelliseen sortumisen vaaroja. Arvioinnin klassisia termejä käyttäen voi tiivistää, että *prosessien, produktien ja kontekstin tulee olla läsnä kaikessa arvioinnissa*, mikäli tavoitteena on oppimis- opetus-prosessien tukeminen aidosti oppilaan monipuolisen inhimillisen kasvun nimissä.

Lähteet

- Aaltonen, K. 2003. Opettajan pedagogisen ajattelun ja toiminnan suhde. Opetustaan integroivan opettajan tietoperusta lähihoitajakoulutuksessa. Joensuun yliopisto. Kasvatustieteellisiä julkaisuja 89.
- Anderson, L. 2003. Classroom assessment: Enhancing the quality of teacher decision-making. Mahwah: Erlbaum.
- Bereiter, C. & Scardamalia, M. 1993. Surpassing ourselves. An inquiry into the nature and implications of expertise. Chigaco: Open Court.
- Brooks, V. 2002. Assessment in secondary schools. The new teacher's guide to monitoring, assessment, recording, reporting and accountability. Buckingham: Open university press.
- Darling-Hammond, L. 1993. Authentic assessment in practice: A collection of portfolios, performance tasks, exhibitions and documentation. New York: National Center for Restructuring Education, Schools and Teaching.
- Grossman, P. 1990. The making of teacher knowledge and teacher education. New York: Teachers College.
- Hakkarainen, K., Lonka, K. & Lipponen, L. 2004. Tutkiva oppiminen. Järki, tunteet ja kulttuuri oppimisen syyttäjänä. Porvoo: WSOY.
- Haring, M. 2003. Esi- ja alkuopettajien pedagogisen ajattelun kohtaaminen. Joensuun yliopisto. Kasvatustieteellisiä julkaisuja 93.
- Harjunen, E. 2004. Miten opettaja rakentaa pedagogisen auktoriteetin? Otteita opettajan arjesta. Suomen Kasvatustieteellinen Seura. Kasvatusalan tutkimuksia 10.
- Helakorpi, S. & Olkinuora, H. 1997. Asiantuntijuutta oppimassa. Porvoo: WSOY.
- Hutmacher, W., Cochrane, D. & Bottani, R. 2001. In pursuit of equity in education: Using international indicators to compare equity policies. Dodrecht: Kluwer.
- Jakku-Sihvonen, R. & Heinonen, S. 2001. Johdatus koulutuksen uudistuvaan arviointikulttuuriin. Opetushallitus. Arviointi 2. Helsinki.
- Jonassen, D. & Land, S. (toim.) 2000. Theoretical foundations of learning environments. Mahwah: Lawrence Erlbaum.
- Kansanen, P., Tirri, K., Meri, M., Krokfors, L., Husu, J. & Jyrhämä, R. 2000. Teacher's pedagogical thinking. New York: Peter Lang.
- Kasanen, K. 2003. Lasten kykykäsitykset koulussa. Joensuun yliopisto. Yhteiskuntatieteellisiä julkaisuja 58.
- Kauppi, A., Hanski, K., Nousiainen, R. & Lahtiranta, K. 2003. Arviointi pedagogisen kehittämisen välineenä. Teoksessa H. Kotila (toim.) Ammattikorkeakoulupedagogiikka. Helsinki: Edita, 255–270.

- Kompf, M. & Denicolo, P. (toim.) 2003. *Teacher thinking twenty years on: Revisiting persisting problems and advantages in education*. Lisse: Swets & Zeitlinger.
- Lahdes, E. 1997. *Peruskoulun uusi didaktiikka*. Keuruu: Otava.
- Linn, R. & Gronlund, N. 2000. *Measurement and assessment in teaching*. Upple Saddle River: Merrill.
- Niemi, H. 2000. Opettajankoulutuksen vaikuttavuus. Teoksessa R. Raivola (toim.) *Vaikuttavuutta koulutukseen*. Suomen Akatemian julkaisuja 2, 169–193.
- Niikko, A. 2000. *Portfolio oppimisen avartajana*. Helsinki: Tammi.
- Oksanen, R. 2003. Laadun määrittely perusopetusta koskevassa kunnallisessa päätöksenteossa. Joensuun yliopisto. *Kasvatustieteellisiä julkaisuja* 88.
- Raivola, R. 2002. Tavoitetaanko laatu standardeihin yltämällä? Teoksessa E. Olkinuora, R. Jakkusihvonen & E. Mattila (toim.) *Koulutuksen arviointi – lähtökohtia, malleja ja tilannekatsauksia*. Turun yliopisto. *Kasvatustieteiden tiedekunnan julkaisuja B* 70, 15–30.
- Raivola, R., Valtonen, P. & Vuorensyrjä, M. 2000. Käsitteet, mallit ja indikaattorit koulutuksen tehokkuutta ja vaikuttavuutta arvioitaessa. Teoksessa R. Raivola (toim.) *Vaikuttavuutta koulutukseen*. Suomen Akatemian julkaisuja 2, 11–28.
- Räisänen, A. & Frisk, T. 1996. Silta uuteen opiskelija-arviointiin. Arviointia opiskelija-arvioinnista. Opetushallitus. *Arviointeja* 6. Helsinki.
- Shulman, L. 1987. *Knowledge and teaching: Foundations of the new reform*. *Harvard Educational Review* 57, 1–22.
- Stipek, D., Reccia, S. & McClintic, S. 1992. *Self-evaluation in young children*. University of Chicago.
- Varila, J. & Rekola, H. 2004. Mitä on työssä oppiminen. Teoreettisia ja empiirisiä näkökulmia työssä oppimiseen. Joensuun yliopisto. *Kasvatustieteiden tiedekunnan tutkimuksia* 83.
- Virta, A. 1999. Uudistuva oppimisen arviointi: Mahdollisuuksia ja varauksia. Turun yliopisto. *Kasvatustieteiden tiedekunnan selosteita B* 65.
- Vuorinen, J. (toim.) 2000. *Arviointi ja kehityskeskustelu. Koko kuva oppijasta*. Jyväskylä: PS-kustannus.
- Wragg, E. 2001. *Assessment and learning in the primary school*. London: Falmer.
- Uusikylä, K. & Atjonen, P. 2000. *Didaktiikan perusteet*. Juva: WSOY.



Utvärdering till stöd för lokalt beslutfattande

Den nationella utvärderingen av utbildningen har byggts upp till ett instrument för utveckling, som baseras på att utvärdering av resultat kopplas till beslutsberedning och vid behov till riktade åtgärder och särskilda prioriteringar. I den nu pågående omorganiseringen av utvärderingen eftersträvas bland annat en större bredd och förbättrade påverkansmöjligheter i konstruktionen av utvärderingarna. Undersökningar har dock påvisat att nyttjandegraden av de nationella utvärderingsresultaten har varit låg, när man kommer till användning av resultaten lokalt inom kommuner och enskilda skolor. Problematiken kring denna fråga belyses och frågan om uttolkningen och användningen av resultaten diskuteras. Artikeln ger förslag till utvecklade samarbetsformer kring användning av resultaten på lokalplanet.

Nyckelord: svenskspråkig utbildning, utvärdering, lokalt beslutfattande, utnyttjande av utvärderingsresultat

Utvärdering har på allvar introducerats inom den nationella utbildningsadministrationen under de senaste drygt tio åren. Ett digert arbete har lagts ner på att utveckla ett utvärderingssystem med en helhetssyn på metoderna och modellerna för genomförandet av utvärderingen. Samtidigt som den nationella utvärderingen småningom har kommit i gång har man också noterat att det finns signaler som antyder att utvärderingen inte till fullo lyckas inta den position som varit avsikten. Kan utvärderingen uppnå den position av central styrmekanism också inom de allmänbildande skolorna och på den lokala nivån så som man ursprungligen avsett? Kan utvärderingen svara på de förväntningar som man riktat på den? I den internationella diskussionen om evaluering rasar debatten kring de nya trenderna för utveckling av utvärderingen. Strävan

är att genom att ändra fokus inom utvärderingen eventuellt bidra till att påverka utvärderingens position.

Utvärdering på nationell nivå

På nationell nivå har syftet med utvärderingen varit att producera information om läget inom utbildningen för att påvisa utvecklingens riktning, för att påvisa samband med förändringar i insatser och strukturer och för att bedöma situationen i perspektiv av uppställda mål. De nationella utvärderingsresultaten har givits offentlighet och publicerats för att utgöra ett beslutsunderlag och för att utgöra en neutral, sakkunnig och analytisk beskrivning som beslutsfattarna kunnat ta del av. Offentliggörandet av resultaten har likaså erbjudit de inom utbildningen verksamma, de ansvariga upprätthållarna samt alla övriga intressenter möjlighet att ta del av rönen.

I det första skedet av utveckling av den nationella utvärderingen på 1990-talet var det särskilt organiseringen eller funktionsdugligheten, resultatskapandet och ekonomin som utgjorde kärnområden inom utvärderingsstudierna. Under samma tidsperiod växte samtidigt kraven på en tydligare NPM (New Public Management) -inriktad utvärdering. Tidigare betonades ett allmänt vetenskapligt förhållningssätt till de undersökta problemområdena och användningen av vetenskapligt beprövade metoder. Samtidigt sökte man nya former och ett bredare anslag i kontexten för utvärderingen, vilket bland annat tillämpades inom de tematiska utvärderingar som Utbildningsstyrelsen genomförde under 1990-talet, där också en intern ifrågasättande diskussion var en del av systemet.

I den samtida NPM -inriktningen utgjorde strävan att bygga upp ett stort mått standardiserade och kvantifierbara mätare en tydlig trend. Man önskade också genomföra en klar koppling mellan utvärderingsresultaten och resurstilldelningen. Målet var att presentera utbildningen som en service där medborgaren ges möjlighet att välja mellan alternativa tjänsteproducenter. Dragkampen mellan dessa två riktningar präglade diskussionerna på central nivå, men var inte en lika dominerande fråga på lokal nivå. På den lokala nivån har utmaningarna rört sig kring förutsättningarna för utvärdering och därtill kring intresset att ta till sig utvärderingsresultat.

I skolor och kommuner

Det förlopp som gjorde att utvärderingen fick en så snabb frammarsch under 1990-talet hade ett samband med behovet av ett passande verktyg i den nya decentraliserade och deregulerade styrmodellen, där en stor del av beslutsfattandet förts till den kommunala nivån och till de enskilda skolorna. De tidigare minutiösa regelverken har ersatts med informationsstyrning och strävan att servera feedback i form av utvärderingsdata.

Trots att avsikten varit att utvärderingsresultaten skall vara till nytta för det lokala

beslutfattandet likaväl som för de nationella utbildningspolitiska linjedragningarna (kulturutskottets ställningstagande 1997), har undersökning efter undersökning påvisat att detta fungerat synnerligen bristfälligt på lokal nivå. Inom yrkesutbildningen förefaller utvärderingen ha blivit rätt bra etablerad, medan de allmänbildande skolorna varit långsammare i att följa utvecklingen. (Rajanen 2000)

Under uppföljningar har man också kunnat konstatera att utvärderingsresultaten inte omsätts eller tas i användning i den utsträckning som förväntats samt att utvärdering som styrinstrument inte integrerats i det praktiska arbetet på regional och lokal nivå så som avsetts (Svedlin 2003; Karlsson 2003; Niemi 2004). Dessa brister är inte för Finland specifika problem, utan är gemensamma för en lång rad andra länder med jämförbar utvecklingshistoria inom utvärderingen.

Den ideala modellen, där utvärderingsresultaten tas till vara för förändringar under den kommande verksamhetsperioden i en ständigt återkommande kedja, har till en del visat sig vara en idealbild som inte alltid varit lätt att förverkliga.

Det finns också påståenden om att det på nationell nivå kan uppstå en tendens till glidning från de ursprungliga intentionerna. I sin analys av utvecklingen i vårt grannland Sverige väcker Lundahl (2003) frågan om administrationen egentligen fick för mycket utvärderingsinformation. Han undrar om beredskapen att ta till sig informationen, att reflektera kring denna och väva in den i den nationella utbildningspolitiska diskursen egentligen var alltför svagt utvecklad. Den ursprungliga föresatsen, där utvärderingsresultaten utgjorde ett centralt element i informationsstyrningen, blev inte fullt etablerad i Sverige. Lundahl anser därför att beslutet att omforma utvärderingsupplägget så att det snarare fick en prägel av kontroll än av en beskrivning av process och händelseförlopp, härrör sig från detta tillkortakommande. Denna omändring har genomförts på nationell nivå, utan större uppmärksamhet, trots att den innebär ett betydande perspektivskifte. I den nuvarande inriktningen av utvärderingen, eftersträvar man lätthanterlig, färdigt standardiserad information. Ett exempel är de kommunblad som man utvecklat vid Skolverket (Skolverket 2004). Kommer man att också i Finland gå från den öppna reflektiva diskussionen kring utvärderingen till en hårt styrd utvärderingsmall, som utan att ifrågasättas appliceras på skolorna?

Utvärderingen och skolan

Kombinationen av utvärderingsinformation och kravet på åtgärder gällande de missförhållanden som uppdragas, skapar behov av snabba reformer också i de situationer där man inte känner till hur denna intervenering bör genomföras. Under de senaste årens diskussioner om utvärderingens genomförare och genomförandereformer, har frågan om hur man faktiskt integrerar utvärderingen i samhällsutvecklingen, så att den får en befast och positiv funktion också på den lokala arenan, fått allt för litet utrymme. Kunskaperna om hur man tar itu med informationen, hur man kan bearbeta resultaten, skapa utvecklingsprojekt och bygga upp lärande miljöer borde utvecklas ytter-

ligare. För att komma djupare i analysen av problemställningen ingår här ett försök att skissera påverkande faktorer.

Situationerna och förutsättningarna varierar naturligtvis mellan kommunerna och de större städerna har en annan utgångspunkt i utvecklingsarbetet än de mindre kommunerna. Merparten av de finländska kommunerna är ändå små kommuner med färre än 15 000 invånare och en stor del av landets skolor är belägna i dessa. Strategin för att hantera de små kommunernas situation inom skolutvecklingsarbetet har baserats på strävan att bidra till och stöda uppkomsten av samarbetsnätverk, där de större orterna agerar primus motor och försöker få de mindre orterna engagerade och med i arbetet. I många projekt har denna modell fungerat rätt bra. De större orterna har kunnat bjuda på kunskaper, modeller och samarbetsmönster. Konstellationen har ändå inte varit problemfri genom att ledarkommunen kan uppfatta sitt ansvar och sin roll som för betungande och egentligen varken haft tillräckliga personresurser eller andra resurser för att möta behoven. Modeller som skapats för de större sammanhangen har inte heller alltid gått att föra över till de mindre kommunerna. Det kan uppstå lägen där man måste konstatera att det finns skillnader i prioriteringar och tidtabeller som gör att de gemensamma projekten inte alltid haft ett fruktbarande slutresultat. (Se till exempel Hyvin tehty, 2003)

Utvärderingens logiska konsekvenser

En realitet i de mindre kommunerna är att såväl personresurserna som de ekonomiska resurserna är mycket begränsade. På kommunnivån följer förändringarna inom utbildningssektorn i hög grad den allmänna samhällsekonomiska utvecklingen i kommunen. De kommunala förändringsvågorna och utvecklingen inom den offentliga sektorn är i mycket beroende av den allmänna ekonomiska utvecklingen. I detta utgör eventuella utvärderingsresultat ett mycket finstilt inlägg mellan de andra mer dominerande frågorna. Att upprätthålla diskursen mellan utveckling och utvärderingsresultat kräver därför särskild insats och mycket målmedvetenhet. Det är därtill inte ovanligt att det på lokal nivå redan hunnit inträffa flera förändringar inom det utvärderade området, varvid resultaten från data insamlade något år tidigare kan upplevas vara passé. Makronivåns analysresultat är inte omedelbart överförbara till mikronivåns verklighet.

Utförda analyser visar också att andelen lärarfortbildningsdagar är rätt begränsade inom den allmänbildande utbildningen och på hela taget betydligt lägre än inom yrkesutbildningen. Därtill är det vanligen de stora städerna som organiserat goda fortbildningsmöjligheter för sina lärare. Lärare på mindre orter har fått markant mindre fortbildning och de svenska skolornas lärare har signifikant färre fortbildningsdagar än sina finskspråkiga kolleger. (Jakku-Sihvonen & Rusanen 2000)

Det finns uppenbara risker för att mindre skolenheter inte klarar av att hålla sig ajour

med den allmänna pedagogiska och skoldidaktiska diskussionen. Detta understryker vikten av att man under den kommande utvecklingen av utvärderingen tar fasta på den faktiska situationen inom de mindre kommunerna, när man planerar fortsatta utvecklingsprojekt och informationsstyrning om och kring utvärdering.

Dagens diskussioner går i mycket ut på att omforma genomförandeprocessen till en mer "demokratisk" utvärdering, till exempel genom att utöka inslaget av "stakeholder" -delar. Genom att frånga den tidigare dominerande oligarkiska expertmodellen för utvärderingens värderingssekvens och ersätta dem med olika intressentgruppers möjligheter att påverka och ta del i konstruerandet av utvärderingen, förväntar man att intresset och nyttjandegraden av utvärderingsinformation kommer att öka. Det är naturligtvis möjligt att detta vitaliserar intresset, men ändå kan centrala frågor förbli olösta. I diskussioner om utvecklingen av regionala utvärderingsprojekt har det framkommit vissa frågor, som påverkar den lokala hanteringen av utvärderingsproblematiken likaväl som nyttjandet av utvärderingsresultaten.

När enskilda utvärderingsprojekt genomförs, till exempel utvärdering av inlärningsresultat, eller tematiska, såsom utvärdering av specialundervisningen eller studiehandledningen, förutsätter uppbyggandet av utvärderingen en djup kunskap om substansfrågan, om datainsamlingsmetoder och om modeller för det konkreta organiserandet. Den inviterade "stakeholder" lokala representantens roll i detta konstruktionsbygge blir ofta att fungera som en del i en referensgrupp, att kommentera gjorda förslag, att pröva ut någon enskild detalj och att genom att beskriva de lokala förhållandena och arbetssituationerna bidra till en så välfungerande utvärderingsdesign som möjligt. Det låter sig sannolikt väl göras att förmedla den inom yrkeskåren existerande allmänna uppfattningen om situationen och sakernas tillstånd. Däremot är den lokala representantens beredskap att skapa en metaevalueringskunskap, att kunna göra bedömningar av konsekvenser av den valda designen och vald kontext för utfallet på regional och lokal nivå, sannolikt begränsad. Här finns sannolika förklaringar som till exempel de praktiska, vardagliga begränsningarna som ställer till med problem – tid att sätta ner på uppdraget vid sidan av det ordinarie värvet, möjligheter att fördjupa kunskaperna och att kumulativt bygga upp en sakkunskap. Frågan om användningen av utvärderingsresultaten på lokalt plan kommer inte att lösas genom en starkare integrering av på substansområdet verksamma enskilda personer, även om en sådan integrering kanske kan erbjuda andra fördelar.

Från utvärdering till utveckling

Utnyttjandegraden av utvärderingsinformation kan som fenomen jämföras med implementering av nya metoder eller tankemodeller på lokal nivå. Erfarenheter från nationella program för utveckling av undervisningsmetoder redan från 1980-talet beskriver en förändring av strategin från till enskilda lärare riktade fortbildningskurser och

informationskampanjer till mer omfattande helheter riktade till skolornas lärarteam. (Svedlin 2003) Det avgörande är sannolikt förmågan att kunna rikta sig till en grupp som har förmågan eller inom vilken man kan utveckla förmågan, att ta emot och bearbeta utvärderingsinformation. Men förutom detta bör man också hitta kanalerna för påverkan och åtgärder.

I många skolor inom den grundläggande utbildningen är skolans beredskap utvecklade, att som organisation ta emot central utvärderingsinformation och omsätta denna i planering, i åtgärder, i utvecklingsprogram som omfattar hela skolan. Färdigheterna att som ett kollektiv ta intryck, att ta till sig information som berör hela organisationen som helhet och att därpå sedan åtgärda enligt gemensamma uppgjorda strategier förutsätter en samlad, samordnad handling med ett långsiktigt perspektiv. Som rutin och handlingsmönster är nog detta fortfarande främmande för de flesta allmänbildande skolor.

Traditionen ger att lärarna gärna betonar friheten och självständigheten sitt arbete och inte önskar andras inblandning i sitt professionella yrkesutövande. Samtidigt ligger i denna frihet också en dimension av ensamhet och en känsla av uteblivet stöd från till exempel den kommunala skolledningens sida (Hansén 1997).

Denna bakgrund bidrar sannolikt till att informationen om till exempel elevers kritiska eller negativa attityder till ett ämne eller till någon verksamhet i skolan, är en svårbehandlad information för läraren. För en enskild lärare är det ett synnerligen svårt uppdrag att påverka attityderna inom en hel elevgrupp och de traditionella sätten att arbeta inom skolan har inte byggt på modeller för samordnad handling. Frågan om attityder är därtill något som kanske endast till en del har ett samband med företeelser i skolan. Attityderna har kanske skapats i hemmiljön, i kamratkretsen eller inom vissa subkamratgrupper. Det är givet att en enskild person inte kan rå på eller förändra dylika mycket sammansatta fenomen. Att åstadkomma en förändring förutsätter en samfällad överenskommelse och ett noggrant planerat åtgärdsprogram.

Resultaten från till exempel utvärdering av inlärningsfärdigheter i en rad ämnen visar på påtagliga utvecklingsbehov i de svenska skolorna (Svedlin 2000; Jakku-Sihvonen & Komulainen 2004). Vi har anledning att notera att det finns orsak att fundera på hur man kan bygga upp åtgärdsprogram inom de svenska skolorna.

Med beaktande av att utvärderingsinformationen och också en enskild skolas resultat beskriver en sammansatt helhet bestående av frågor om använda resurser, om undervisningsmetoder, mål för undervisningen, organisering av undervisningen, liksom också allmänna attityder till utbildningen och elevernas prestationsnivå m.m., skulle det vara viktigt att bearbetningen av informationen försiggår i en bred och ansvarsfull mottagargrupp, som därtill har kanaler att påverka de lokala besluten. En central del av skolornas utveckling handlar de facto om att organisera utvecklingsarbetet ända från den enskilda skolan till den kommunala skoladministrationen, så att det samtidigt finns en synlighet i genomförandet och en regelbundet återkommande diskussion gällande kommande prioriteringar. Skolorna kan knappast ensamma och inte obero-

ende av varandra skapa resultat. På den kommunala nivån handlar det dels om att centralt svara för och ge förutsättningar för utvecklingen inom skolorna. Här ingår att trygga resurstilldelningen, men också att stå för ett stödande ledarskap och ytterligare av att bland annat satsa på utveckling av substanskunskap. I praktiken kompliceras detta av att den kommunala nivån inte är, förenklat uttryckt, en endimensionell box i ett nationellt hierarkiskt bygge. En riktigare beskrivning är att det under ytan handlar om ett spänningsfält av möten mellan olika intressen, värderingar och behov. Intresser i detta är förutom de inom skolan anställda, även de som bär det förvaltningsmässiga ansvaret, samt de lokala beslutfattarna.

Mot denna bakgrund kan vi konstatera att ett mer utvecklat och bredare utarbetat samarbete mellan den nationella nivån och den lokala nivån för att stöda en fruktbar och på en positiv utveckling inriktad process, skulle vara nödvändig. En viktig insikt i detta är att ansvaret för utformandet av utvärderingen inte är avslutat i och med de offentliggjorda resultaten. Den efterföljande sekvensen, som handlar om hur resultaten landar, hur den används på den lokala nivån och de diskussioner utvärderingsinformationen föranleder, kan ge en ekoeffekt som förtar värdet av en aldrig så väl uppbyggd utvärderingsstudie. Den nuvarande informationsspridningen har nått dåligt fram och intresset för att ta del av informationen har varit oväntat svagt. Av en fungerande utvärdering kan man väl förvänta sig att den strävar att nå ända fram till de som upprätthåller och står för genomförandet av verksamheten?

Samarbetsmodeller i framtiden?

Eftersom de nationella utvärderingarna genom sina datainsamlingar får en översiktlig bild av situationen i olika landsändor och i olika skolor, torde en potentiell modell vara att efter resultatpublicering erbjuda ett mindre antal kommuner/skolor möjlighet att ta del av ett feedback program. Urvalet kan naturligtvis göras på olika grunder, i form av en samarbetsmodell med delad finansiering eller i form av särskilt utsedda skolor, representerande en variation i resultaten. Ett alternativ kunde vara att använda regionprogrammen som bas för samarbetsformerna. Kanske man i framtiden kan bygga upp en modell av reflekterande diskussion och "benchmarking" kring goda resultat inom och mellan regionerna.

I många fall uppfattas det som en självklarhet att det finns ett intresse och en ambition att ta till sig den utvärderingsinformation som presenteras i de nationella resultaten. I realiteten är frågan inte lika enkel. Man kan dels se detta som ett problem i frågan om jämförbarhet, i möjligheterna att överföra landsomfattande eller regionala utvärderingsresultat på den egna skolan. Men det berör också frågan om relevans, har slutsatserna i utvärderingsrapporterna drivits till för den lokala aktören relevanta slutsatser, ingår det i dessa konklusioner i vilka lärarens eller skolans andel i helheten synliggörs?

I förslagen till framtida modeller utgår man från att dessa borde baseras på “governing by participation”, där man i yrkesutövandet upplever att man själv ställer upp eller s.a.s. äger de mål som fastställts för verksamheten (Lundgren 2002). Detta förutsätter dock att man upplever sig kunna företräda rollen av en aktiv och central aktör, inte som en kugge i ett omfattande maskineri eller verksam utan möjligheter att förändra randvillkoren.

Vid sidan av de här nämnda mer organisatoriska frågorna har det visat sig att förutläggningarna för att ta till sig information också handlar om att bygga upp ett förtroende mellan utvärderare och utvärderade (Leeuw 2001). Man har i olika sammanhang talat om idén med en “critical friend”, den kritiska vännen, som innehar en förståelse av läget, men samtidigt kan ifrågasätta och ge kommentarer i egenskap av utomstående sakkunnig. Detta kan småningom leda till partnerskap i analys och tolkning av resultaten. Tyngdpunkten i utvärderingen skulle därmed sänkas från en toppstyrning till en bredsida vardagsförankring. En fungerande utvärdering medför således, så som här tidigare konstaterats, utmaningar som sträcker sig långt utöver själva planeringen och genomförandet av det enskilda utvärderingsprojektet. Strategin bör notera behovet av arbetsformer och samarbetsnätverk som utvecklar användningen och tillgodogörandet av utvärderingsresultaten.

Källor

- Hansén, S.E. 1997. “Jag är proffs på det här”: om lärarens arbete under en tid av förändring. Åbo Akademi. Rapporter från Pedagogiska fakulteten vid Åbo Akademi 16.
- Hyvin tehty! – meidän koulumme tulevaisuuden tekijänä 2003. Helsinki: Etelä-Suomen lääninhallituksen julkaisuja.
- Jakku-Sihvonen, R. & Komulainen, E. 2004. Alueellinen ja koulujen välinen vaihtelu perusopetuksen oppimistuloksissa. *Kasvatus* 35 (3), 272–285.
- Jakku-Sihvonen, R. & Rusanen, S. 2000. De svenskspråkiga lärarnas deltagande i fortbildning åren 1996-1998. I Renata Svedlin (red.) *Värderad utbildning – utvärderingsresultat 1997–1999*. Utbildningsstyrelsen. Utvärdering 5. Helsingfors.
- Karlsson, O. 2003. Evaluation politics in Europe: Trends and tendencies. *Studies in Educational Policy and Educational Philosophy*. E-tidskrift 2003:1. <<http://www.upi.artisan.se/docs/Doc171.pdf>> (3.II 2004)
- Kulturutskottets betänkande 3/1998 rd. KuUB3/1998 rd – RP 86/1997 rd.
- Kuntien opetustoimen itsearviointin kehittämisestä – kokemuksia lähikuntien arviointiyhteistyöstä. 2003. Helsinki: Etelä-Suomen lääninhallituksen julkaisuja.
- Leeuw, F.L. 2001. Reciprocity and the evaluation of educational quality. Assumptions and reality checks. EU Congress “The meaning of Quality in Education” Karlstadt, Sweden, 2–4 April 2001. Netherlands’ Inspectorate of Education & University of Utrecht, the Netherlands. <www.eu2001.se/education/eng/docs/Karlstad_leeuw.pdf> (26.II 2001).
- Lundahl, C. 2003. The formation of an assessment culture. *Studies in Educational Policy and Educational Philosophy*. E-tidskrift 2003:3. <www.upi.artisan.se/docs/Doc198.pdf> (3.II 2004).

- Lundgren, U.P. 2002. Political governing of the education sector: Reflections on change. *Studies in Educational Policy and Educational Philosophy*. E-tidskrift 2002:1.
<<http://www.upi.artisan.se/publish/docs/Doc138.pdf>> (3.II 2004).
- Mikkola, A. 2003 Täydennyskoulutus tukemaan opettajien työtä. Julkaisussa: Hyvin tehty! – meidän koulumme tulevaisuuden tekijänä. Etelä-Suomen lääninhallituksen julkaisuja.
- Niemi, E. K. 2004. Perusopetuksen oppimistulosten kansallinen arviointi ja tulosten hyödyntäminen koulutuspoliittisessa kontekstissa. Turun yliopiston julkaisuja. Sarja C 216.
- Rajanen, J. 2000. Selvitys koulutuksen paikallisen tason arvioinnin tilasta. Opetushallitus. Arviointi II. Helsinki.
- Skolverket. 2004. Kommunblad.
<<http://www.skolverket.se/fakta/statistik/kommunblad/04/index.shtml>> (3.II 2004).
- Svedlin, R. (red.) 2000. Värderad utbildning – utvärderingsresultat 1997–1999. Utbildningsstyrelsen. Utvärdering 5. Helsingfors.
- Svedlin, R. 2003. Utvärdering i skolan – mellan skolutveckling och styrpolitik. Helsingfors universitet. Pedagogiska institutionen, forskningsrapport 185.



Ennakoivaan arviointiin

Tässä artikkelissa tarkastellaan ennakoinnin ja arvioinnin rajapintaa. Artikkelin ensimmäisessä luvussa esitellään tulevaisuusajattelun ja ennakoinnin perusteet. Toisessa luvussa pohditaan ennakoinnin, sidosryhmätyöskentelyn ja tulevaisuuden tekemisen suhdetta. Kolmannessa luvussa tarkastellaan arvioinnin kehittämistä ennakoinnin menetelmillä.

Asiasanat: ennakointi, arviointi

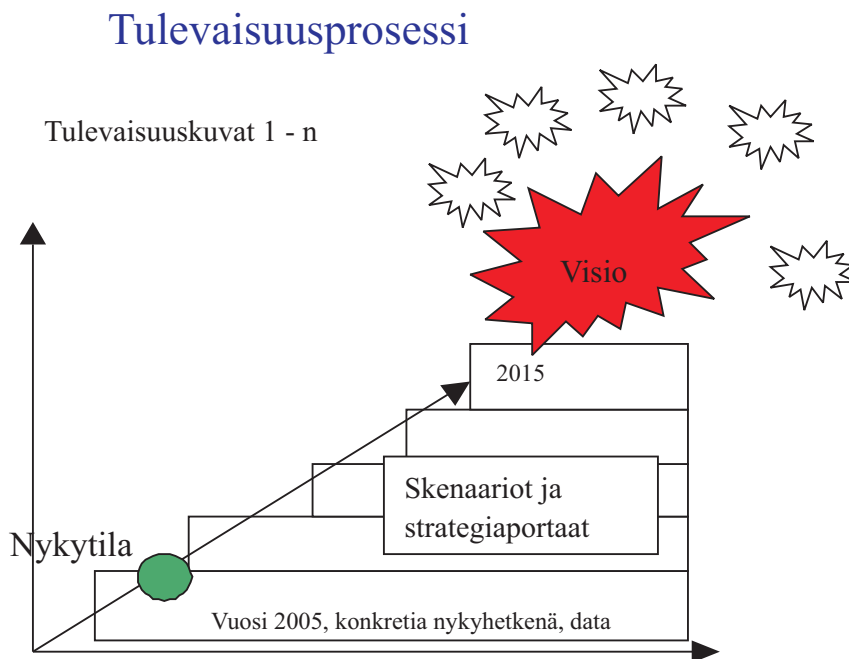
Ennakointi tulevaisuuden tekemisenä

Ennakoinnin tavoitteena on tuottaa tietoa tulevaisuudesta ja auttaa toimijaa tai toimijoita tällä tavalla varautumaan mahdollisiin muutoksiin.

Ennakoinnissa kerätään järjestelmällisesti tietoa toimintaympäristöstä sekä laaditaan tulevaisuuskuvia ja visioita keskipitkän ja pitkän aikavälin tulevaisuudesta. Teknisten, poliittisten, ekologisten, taloudellisten, sosiaalisten ja kulttuuristen murros- ja muutosprosessien ymmärtämisessä ja mallintamisessa voidaan hyödyntää muun muassa skenaario- ja systeemimenetelmiä sekä muita tulevaisuuden tutkimuksen perusmenetelmiä. Näitä ovat esimerkiksi Delfoi-asiantuntijapaneelitutkimukset, trendi- ja projektioanalyysi, ristivaikutusanalyysi ja heikkojen signaalien analyysi. (Ks. Godet 1994; Kaivo-oja, Kuusi & Koski 1997; Kuusi 1999; Keenan, Loveridge, Miles & Kaivo-oja 2003.)

Pelkkä olemassa olevan tiedon ja asiantuntijamielipiteiden yhteen kerääminen, systemaattinen analysointi ja todennäköisyyksien laskeminen ei kuitenkaan vielä varmistaa riittävää tulevaisuustiedon tasoa. Tehokas ennakointijärjestelmä on tiedon tuottamis- ja hallintaprosessi, joka sisältää datan lisäksi myös luovaa innovatiivisuutta,

strategista arvottamista sekä proaktiivista tulevaisuuden tekemistä (Deutsch 1997; Slaughter 1995; Kuusi 1996; Kaivo-oja 2003). Tällainen ennakoitijärjestelmä kertoo meille toimintaympäristön mahdollisten muutosten lisäksi myös, miten voimme vaikuttaa asioihin ja kenen kanssa yhteistyötä tekemällä voimme parhaiten saavuttaa sen tulevaisuuden tilan, jonka haluamme (kuvio 1). Tulevaisuusprosessi sisältää neljä työvaihetta (Ahokas 2003; Hietanen 2004):



Kuvio 1. Tulevaisuusprosessi koostuu nykytilan arvioinnista, mahdollisten, todennäköisten, haluttavien ja vältettävien tulevaisuuskuvioiden rakentamisesta, tavoiteltavan tulevaisuuden päättämisestä (visio) ja strategiaportaiden luomisesta (Ahokas 2003)

- Ennakoinnin, kuten muunkin kehittämisen, suunnittelun ja tutkimuksen lähtökohtana on tieteellinen tieto niistä ilmiöistä, rakenteista ja prosesseista, joista ennakoitava asia muodostuu. Monet ennakoitijärjestelmät pysähtyvät tähän. Tällöin kuitenkin menetetään tulevaisuustiedon hedelmällisin hyöty.
- Ennakointiprosessin toisessa vaiheessa tietoon lisätään mielikuviutus ja luovuus. Sen lisäksi, että ennakoitiprosessin on ymmärrettävä, miten asiat tulevaisuudessa todennäköisesti ovat, sen on myös pohdittava, miten ne voisivat olla. Inhimillistä toimintaa eivät yleensä rajoita taloudelliset tai tekniset rajoitukset, vaan hyvien ideoiden, luovuuden puute.

- Ennakointiprosessin kolmannessa työvaiheessa ennakointiin yhdistetään tiedon ja mielikuvituksen rinnalle myös arvot. Skenaariot ja tulevaisuuden kuvat jakaantuvat karkeasti määriteltynä mahdollisiin, todennäköisiin, toivottaviin ja vältettäviin tulevaisuuksiin. Näistä mahdolliset ja todennäköiset tulevaisuudet ovat melko lailla objektiivisia. Erilaisten mahdollisten tulevaisuuskuvioiden toivottavuus ja haluttavuus voi sen sijaan vaihdella toimijasta eli näkökulmasta riippuen. Tämän työvaiheen tuloksena ovat ensin skenaariot mahdollisista tulevaisuuksista. Työvaihe päättyy yhteisen vision valintaan, jolloin kaikista mahdollisista tulevaisuuksista valitaan yksi: tulevaisuus, jonka toimija haluaa toteutuvaksi.
- Ennakointijärjestelmän neljännessä vaiheessa rakennetaan strategiaporaatit nykyhetkestä haluttuun tulevaisuuden tilaan. Tämä vaihe on arvorationaalinen. Sen jälkeen kun on päätetty, minkälainen maailma halutaan, on tehokkaimman toteutustavan valinta melko lailla loogista ja objektiivista. Tässä työvaiheessa tulevaisuuden tutkimus, tulevaisuustieto ja ennakointi muuttuvat tulevaisuuden tekemiseksi.

Kun edellä mainitut neljä työvaihetta toistetaan säännöllisesti, niin ennakkoinnista syntyy tulevaisuusprosessi, joka on tehokas muutoksen hallinnan ja visionäärisen johtamisen työkalu.

Sidosryhmätyöskentely proaktiivisena ennakointijärjestelmänä

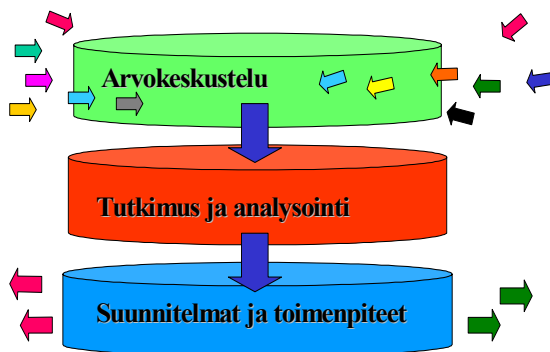
Ennakkoinnissa korostuvat usein myös osallistavat menetelmät, joilla pyritään edistämään muutoksen kannalta keskeisten avaintoimijoiden synergiaa ja verkostoitumista. Verkostomaisella toiminnalla lisätään verkoston jäsenten kollektiivista tietämystä tulevaisuuden haasteista ja mahdollisuuksista. Lisäksi edistetään ennakointijärjestelmän innovatiivisuutta ja proaktiivisuutta (Mattila ja Uusikylä 1999). Osallistavien menetelmien, esimerkiksi tulevaisuusverstaiden, avulla ennakointiprosessiin saadaan myös enemmän dataa ja useampia näkökulmia tutkittavan asian syistä ja seurauksista. (Pretty, Guijt, Thompson & Scoones 1995; Holman & Devane 1999.)

Verkostomaisen, sidosryhmätyöskentelyyn perustuvan ennakkoinnin onnistumisen edellytys on, että toimijat kykenevät löytämään ja määrittelemään yhteisen vision – toimijoiden täytyy siis olla yhtä mieltä halutusta tulevaisuuden tilasta. Yhteisen vision lisäksi avainsanoja ovat profilointi ja roolitus. Tavoitteena ei ole yksi ratkaisu kaikille, vaan jokaisella ennakointijärjestelmän toimijalla on omat polkunsa ja omat toimenpiteensä yhteisen vision toteuttamiseksi (Ahokas 2003; Hietanen 2004).

Ennakointijärjestelmän tehtävänä ei ole keksiä verkoston kaikille toimijoille yhteisiä uusia toimintatapoja, vaan kehittää kunkin toimijan erityisiä taitoja. Vaikka sidosryhmän visio (samat rusinat pullasta) onkin yhteinen, niin vision toteuttamisen keinot voivat olla hyvin erilaisia. Analogiana voidaan käyttää karttaa. Lähtökohtana on tilanne, jossa toimijat ovat eri pisteessä. Jos tavoitteena on päästä samaan pisteeseen, niin jokaisen on lähdettävä omaan suuntaansa; kaikki eivät voi lähteä oikealle sen paremmin kuin pohjoiseenkaan, vaan suunta on valittava sen mukaan, missä lähdön hetkellä sijaitsee. (Ahokas 2003.) Ennakointiin liittyy usein myös voimakas oppimisprosessi (Van der Heijden, Bradfield, Burt, Cairns, & Wright 2002).

Proaktiivisessa, sidosryhmätyöskentelyyn perustuvassa ennakointijärjestelmässä "tulevaisuussihdin" yläosaan kerätään moniäänistä ja monipuolista tietämystä ennakkoinnin kohteesta olevasta aiheesta. Mitä useampi toimija osallistuu tähän arvokeskusteluun sen parempi. Sihdin toisessa vaiheessa arvokeskustelua analysoidaan ja suhteutetaan olemassa olevaan tietoon. Tämän vaiheen toteuttavat asiantuntijat. Sihdin viimeisessä vaiheessa analysoitu arvokeskustelu – tulevaisuustieto – muutetaan toiminnaksi eli tulevaisuuden tekemiseksi (kuviot 2). Mitä useamman toimijan tavoitteita nämä toimenpiteet palvelevat sen paremmin ennakointi ja tulevaisuussihti toimii. (Hietanen, Kaskinen & Takala 2002.)

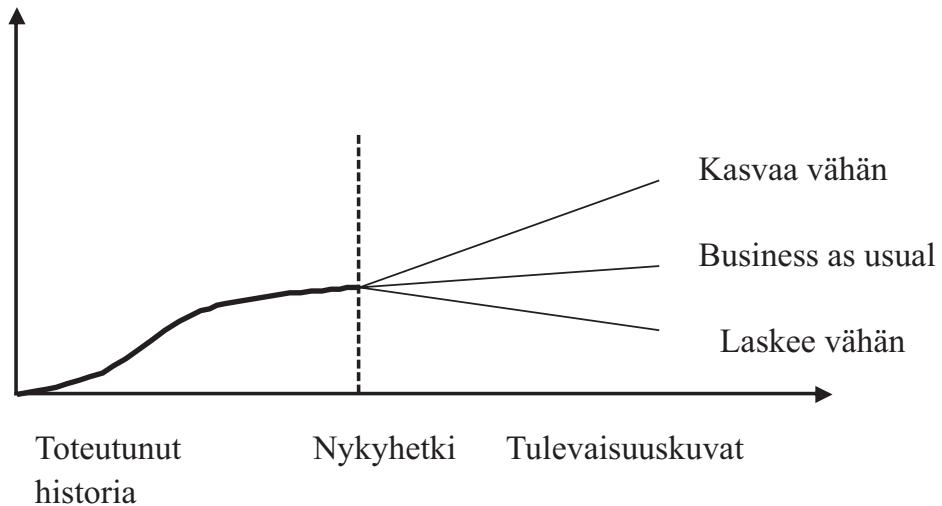
Proaktiivinen sidosryhmätyöskentely



Kuvio 2. Ennakointisihdissä moniääninen arvokeskustelu muokkaantuu proaktiiviseksi tulevaisuuden tekemiseksi (Ahokas 2003)

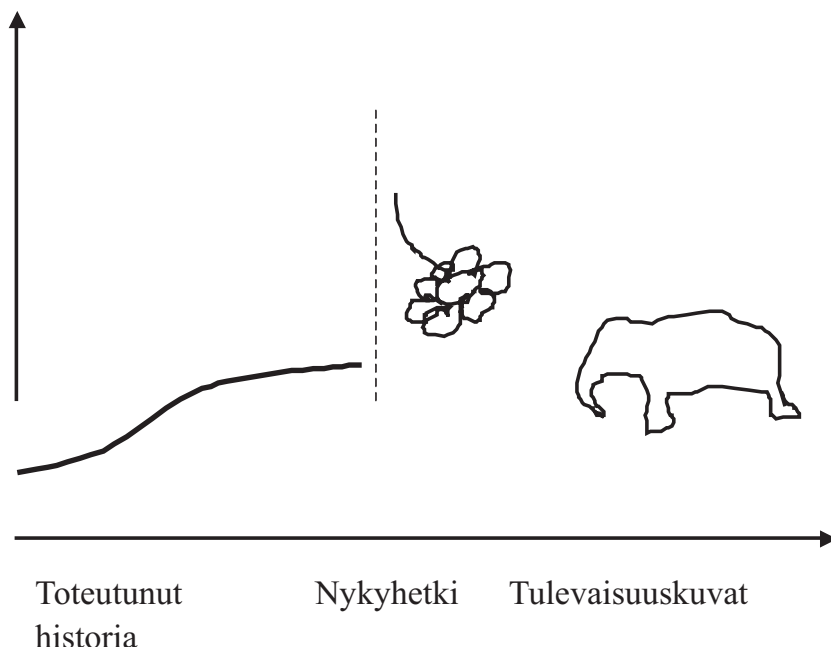
Todennäköiset ja luovat tulevaisuudet

Olellainen erottelu ennakkoinnin sisällä on, onko ennakointiprosessissa kyse vain ja ainoastaan aikasarjoihin perustuvasta trendiekstrapolaatiosta, jossa jo toteutuneen aikasarjan perusteella postuloidaan tulevaisuuskuvia tyyliin: jatkuu samanlaisena (business as usual), kasvaa vähän, laskee vähän (kuvio 3).



Kuvio 3. Trendiekstrapolaatioon, toteutuneisiin aikasarjoihin, perustuva ennakointi

Vaihtoehtona tälle todennäköisyyksiä etsivälle ennakkoinnille on heuristinen ote, jossa lähtökohtana on tahdonvapaus. Vaikka asiat olisivatkin aina ennen tapahtuneet tietyllä tavalla niin se ei tarkoita, että niiden pitäisi myös tulevaisuudessa tapahtua samalla tavalla. Ennakointiin samoin kuin arviointiin liittyy aina tietoa oppiminen historiasta (hindsight), nykyhetken ongelmista (insight) ja arvioiduista tulevaisuuksista (foresight) (Kaivo-oja, Seppälä & Katko 2004; Kaivo-oja 2004). Ihmisille on suotu tiedon lisäksi myös mielikuvitus – kyky keksiä uutta. Tulevaisuus muokkautuu aina monimutkaisen imitointi- ja innovointiprosessin kautta. Kun imitoidaan, seurataan voimakasta polkuriippuvuutta ja vallitsevia megatrendejä, trendejä ja parhaita käytäntöjä. Visionäärinen johtaminen on mahdollista, kun voidaan helposti tunnistaa tärkeät tulevat ilmiöt. Kun taas innovoidaan, ollaan herkkiä ja avoimia uusille ideoille ja innovaatioille, jolloin heikkojen signaalien analyysi nousee tärkeäksi tulevaisuuden tekemisen välineeksi. Yllättävät ja nopeat muutokset voivat edellyttää improvisoitua ja voimakasta operatiivista johtamista. (Mendonça, Cunha, Kaivo-oja ja Ruff 2004). Heuristinen ennakointi etsii ja jopa aktiivisesti luo epäjatkuvuuksia. Siksi heuristisessa ennakkoinnissa korostuvat myös arvot: keskustelu siitä, miten asiat voisivat olla ja miten asioiden pitää olla. (Kuvio 4)



Kuvio 4. Luovat tulevaisuuskuvat eivät perustu jatkumoille vaan epäjatkuvuuksille (Vrt. Kaivo-oja, Seppälä & Katko 2004)

Ennakoiva arviointi

Yllä esitelty todennäköisen ja luovan tulevaisuuden erottelu on myös hedelmällinen lähtökohta arvioinnin ja ennakkoinnin välisen rajapinnan ja synergian tarkasteluun. Arvioinnin tavoitteena voi olla muun muassa suunnitellun ja toteutuneen toiminnan vertaaminen tai esimerkiksi toteutuneen toiminnan vaikutusten ja vaikuttavuuden selvittäminen. Anu Räisänen esittelee toisaalla tässä julkaisussa nk. kehittävän arvioinnin, jolle Räisänen mukaan on luonteenomaista muun muassa osallistavuus ja eri tahojen välinen vuoropuhelu. Arviointitulosten ohella kehittävässä arvioinneissa painotetaan myös prosessia ja sen aikana syntyvää oppimista. Kehittävät arvioinnit johtavat myös – tulevaisuusprosessien tapaan – parantamisalueiden esille nostamiseen sekä suosituksiin ja kehittämisehdotuksiin. Näiden ominaisuuksiensa vuoksi erityisesti kehittävä arviointi on lähellä ennakkointia – ja siksi ennakkoinnilla voidaan parantaa erityisesti kehittävä arvioinnin vaikuttavuutta.

Edellisissä luvuissa esitellyt ennakkointimenetelmät, kuten tulevaisuusverstaat, soveltuvat myös kehittävä arvioinnin ryhmätyömenetelmäksi. Olennaisin ero ennakkoinnin ja kehittävä arvioinnin välillä on, että ennakkointi korostaa toiminnan tavoittei-

den arviointia: ovatko toiminnan tavoitteet oikein asetettuja? Ennakoivan arvioinnin kohteena voi olla muun muassa toiminnan tulevaisuuskuvioiden ja visioiden todennäköisyys ja haluttavuus sekä heuristisessa mielessä toiminnan innovatiivisuus ja “kaikeus”.

Ennakoivan arvioinnin menetelmänä voidaan käyttää esimerkiksi todennäköisiin kehityskulkuihin perustuva trendiekstrapolaatiota. Kehittävän arvioinnin oppimisprosessimaisuus hyötynee kuitenkin enemmän tulevaisuusprosessiin liittyvästä luovasta otteesta ja visionäärisestä muutoksen hallinnan näkökulmasta. Tällöin mahdollisia menetelmiä ovat muun muassa skenaariomenetelmät (kuten tulevaisuustaulukko), pehmeä systeemimetodologia, tulevaisuusverstasmenetelmät sekä Delfoi-haastattelu (tulevaisuuden tutkimuksen menetelmistä katso esim. Kamppinen, Kuusi ja Söderlund 2003).

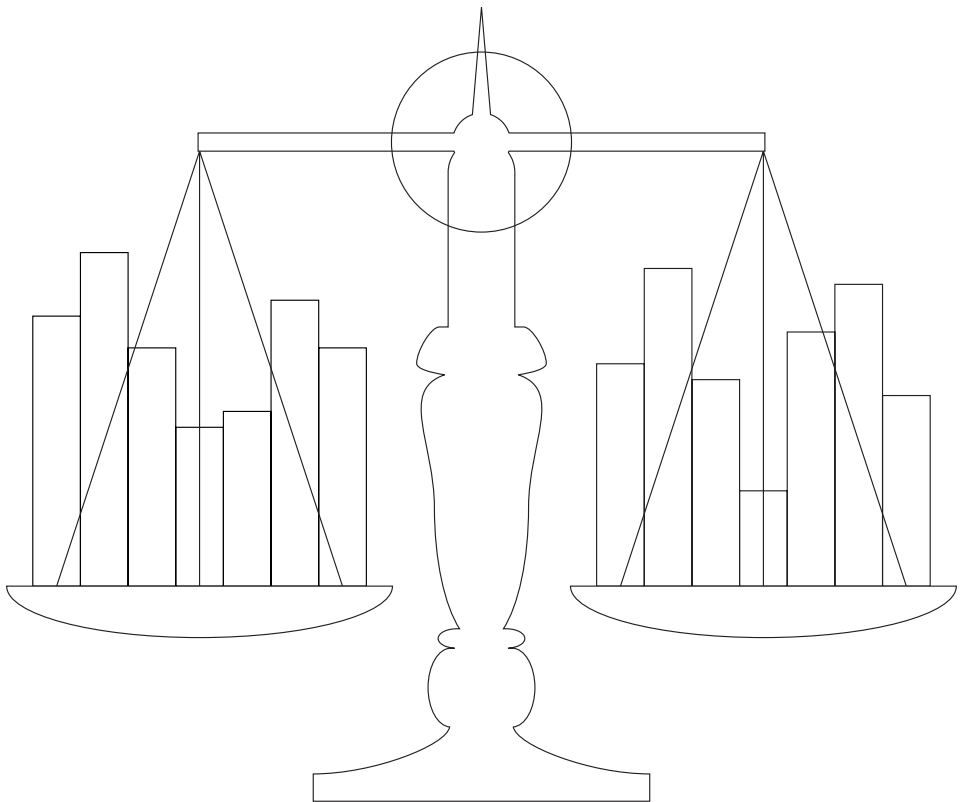
Lähteet

- Ahokas, I. 2003. Tulevaisuus ja turvallisuus: poliisin toimintaympäristö muutoksessa. Esiselvitys 2003. Länsi-Suomen lääninhallituksen julkaisusarja 13/2003. Saatavana www-muodossa: <<http://www.tukkk.fi/tutu/poltu.htm>>
- Deutsch, D. 1997. Todellisuuden rakenne. Vaasa: Terra Cognita.
- Godet, M. 1994. From anticipation to action. A handbook of strategic prospective. Paris: UNESCO.
- Hietanen, O. 2004. Future and security: Scanning the changing environment of the police. Presentation in Seminar on the Future of Police Management 8th–9th of June 2004, Nagu, Finland. Saatavana www-muodossa: <<http://www.tukkk.fi/tutu/poltu/>>
- Hietanen, O., Kaskinen, J. & Takala, A. 2002. KEKETU verkostoanalyysi, seudulliset strategiset verkostot innovaatiotekijänä ja sosiaalisena pääomana. Turun kauppakorkeakoulu, Tulevaisuuden tutkimuskeskus. Tutu-julkaisuja 1/2002. Saatavilla www-muodossa: <<http://www.tukkk.fi/tutu/etiето>>
- Holman, P. & Devane, T. 1999. The change handbook. Group methods for shaping the future. Berrett-Koehler Publishers.
- Kaivo-oja, J. 2003. Tulevaisuuden tekeminen strategisen ajattelun valossa. Teoksessa M. Kamppinen, O. Kuusi & S. Söderlund (toim.) 2003. Tulevaisuudentutkimus. Perusteet ja sovellutukset. Suomalaisen Kirjallisuuden Seuran toimituksia 896, Tampere, 226–249.
- Kaivo-oja, J. 2004. Analyses of historical and future problems of sustainable development: Research articles in spatial sustainability analysis, planning and evaluation. University of Tampere. Acta Universitatis Tamperensis 1038. Acta Electronica Universitatis Tamperensis 385. Väitöskirja.
- Kaivo-oja, J., Kuusi, O. & Koski, J.T. 1997. Sivistyksen tulevaisuusbarometri 1997. Tietoyhteiskunta ja elinikäinen oppiminen tulevaisuuden haasteina. Tulevaisuuden tutkimuskeskus. Julkaisuja 25. Helsinki: Opetusministeriö.
- Kaivo-oja, J., Seppälä, O. & Katko, T. 2004. Seeking for convergence between history and futures research. Futures. The Journal of Forecasting, Planning and Policy 36 (5), 527–547.
- Kamppinen, M., Kuusi, O. & Söderlund, S. (toim.) 2003. Tulevaisuuden tutkimus – menetelmät ja sovellukset. 2. korjattu painos. Helsinki: Suomalaisen kirjallisuuden seura.



- Keenan, M., Loveridge, D., Miles, I. & Kaivo-oja, J. 2003. Handbook of knowledge society foresight. Prepared by PREST and FFRC for European Foundation for the Improvement of Living and Working Conditions. Final Report. Dublin.
- Kuusi, O. (toim.) 1996. Innovation systems and competitiveness. VATT A22. ETLA B 125. Helsinki: VATT.
- Kuusi, O. 1999. Expertise in the future use of generic technologies. Epistemic and methodological considerations concerning the Delphi Studies. VATT. Research Report 59. Helsinki.
- Mattila, M. & Uusikylä, P. 1999. Verkostoyhteiskunta. Helsinki: Gaudeamus.
- Mendonça, S., Cunha, M. P., Kaivo-oja, J. & Ruff, F. 2004. Wild cards, weak signals and organisational improvisation, *Futures. The Journal of Forecasting, Planning and Policy* 36 (2), 201–218.
- Pretty, J. N., Guijt, I., Thompson, J. & Scoones, I. 1995. A trainer's guide for participatory learning and action. International Institute for Environment and Development. IIED Participatory Methodology Series. London, UK.
- Räisänen, A. 2005. Kehittävään arviointiin. Tässä teoksessa.
- Slaughter, R. 1995. The foresight principle. *Cultural Recovery in the 21st Century*. Adamantine Press.
- Van der Heijden, K., Bradfield, R., Burt, G., Cairns, G. & Wright, G. 2002. The sixth sense. Accelerating organisational learning with scenarios. Chichester, UK: John Wiley & Sons.

UUTEEN OPPIMISTULOSTEN ARVIOINTIIN



Kenties laajin mielenkiinto koulutuksen arvioinnissa kohdentuu oppimistulosten arviointiin. Se on yleisessä tietoisuudessa tutuinta ja omakohtaisesti koettua. Oppimistulosten arviointi mieltyy koulukokeiksi, testeiksi, tenteiksi ja tehtäväsarjoiksi, joiden tarkoituksena on saada tietoa oppimisesta edistymisestä. Käytännössä arviointi on tällöin suorituksen tason tai laadun numeerista tai sanallista määrittelyä. Nämä määrittelyt saavat ajattelussamme vahvan painoarvon. Kansallisella tasolla vertailulukuina esitettyinä ne saattavat vahvistaa kansallista identiteettiä ja ylpeyttä tai vastaavasti johtaa regressiivisiin reaktioihin ja itseruokintaan kuten kansainvälinen keskustelu OECD:n koulusaavutusten arviointiohjelman PISA:n tuloksista on osoittanut. Onpa PISA:n raportit ehditty karrikoida jo jälkimodernin yhteiskunnan 'pyhiksi kirjoiksi', joihin uskotaan kansallisen sivistystason mittana.

Oppimistulokset näyttävät siis vetoavan vahvasti tunteisiimme ja jopa kansalliseen identiteettiin. Oppimistulosten arvioinnin tärkeys ja seuraamukset asettavat myös paineita ja vaatimuksia niiden luotettavuudelle. Tämä edellyttää tutkimuksen ja arvioinnin kiinteää yhteyttä, jotta arviointimetodologiset ja -methodiset ratkaisut täyttäisivät myös tieteelliset kriteerit. Oppimistulosten arvioinnissa korostuvat myös testiteoreettiset vaatimukset koskien arviointitulosten yleistettävyyttä. Teoreettisten lähtökohtien pätevyys luo tukevan perustan myös tulkintojen ja johtopäätösten pätevyydelle. Arvioinnissa on oltava vahvoilla tiedollisesti ja taidollisesti, mutta myös eettisesti. Arviointi edellyttää siis vahvaa ja monipuolista professionaalista osaamista.

Arviointijärjestelmätasolla oppimistulosten arvioinnin vaatimukset koskevat laajaa kokonaisuutta yksilötason palautejärjestelmästä paikalliseen, alueelliseen, kansalliseen ja kansainväliseen tasoon saakka. Koulutusjärjestelmän tuloksellisuuden osoittajia ovat juuri oppimistulokset edettäessä aina yhteiskunnalliseen vaikuttavuuteen ja tulevaisuuden ennakointiin saakka. On pystyttävä arvioimaan sitä miten hyvin saavutetut tiedot ja taidot palvelevat myös tulevaisuuden tekemistä. Tämä näkökulmahan on sisään rakennettuna myös opetussuunnitelmassa, jonka toteutumista ja toimivuutta juuri oppimistulosten välityksellä arvioidaan.

Luodessaan edellytyksiä uudelle arviointijärjestelmälle eduskunta on lähtenyt siitä, että oppimistulosten arviointien yhteydessä tuotettaisiin tietoa myös siitä oppimisympäristöstä missä oppimistulokset syntyvät. Kehittämisen kannalta on erittäin oleellista saada tietoa oppimisympäristöstä, jotta kehittämiskohteet voitaisiin identifioida ja löytää ratkaisuja mihin toimiin tulisi ryhtyä nykytilanteen muuttamiseksi. Tavoitteena on siis tuottaa selityskykyisiä arviointituloksia.

Oppimistulosten arvioinnissa on keskitytty perusopetukseen ja ammatilliseen koulutukseen. Näiden näkökulmat ja arviointistrategiat kuten myös arvioinnin kohteet ja metodiikka eroavat toisistaan. Siinä missä perusopetuksessa ollaan kiinnostuneita jonkin oppiaineen osaamisesta, ammatillisessa koulutuksessa ollaan kiinnostuneita erilaista osaamista yhtä aikaa vaatimasta ammatillisesta osaamisesta, jota vielä on voitava arvioida autenttisesti siellä, missä osaamista myös joudutaan käyttämään, eli muuttuvissa ja toisistaan eroavissa työelämän tilanteissa.

Organisatorisesti uudessa oppimistulosten arvioinnissa on lähdetty verkostoajattelusta monipuolisen kansallisen ja tarvittaessa aina kansainvälisenkin asiantuntemuksen hyödyntämiseksi. Asiantuntemuksen monipuolisuutta ja -tahoisuutta jo arvioitavan ilmiön kompleksisuudella. Asiaa on lähestyttävä samanaikaisesti monista näkökulmista ja monipuolisesti suhteuttaen.

Verkostona organisoitava koulutuksen arviointi perustuu lakiin. Lainsäädännössä arvioinnin tavoista säädetään koulutuksen tehtäväksi organisoida ulkopuoliset arvoinnit yliopistojen, Opetushallituksen ja muiden asiantuntijoiden verkostona. Säädöksissä on vielä Opetushallituksen asema yksilöity siten, että "Opetushallitus osallistuu osana koulutuksen arviointiasiantuntijoiden verkostoa opetussuunnitelmien ja tutkintojen perusteiden mukaisten oppimistulosten arviointiin". Opetushallituksella on siten itseoikeutettu asema osallistua oppimistulosten arviointiin osana yliopistojen ja muiden asiantuntijoiden verkostoa. Lainsäädäntö edellyttää siten oppimistulosten arvioinnin monitahoista organisointia mikä oppimistulosten arvioinnin luonteen ja merkityksen vuoksi on perusteltua.

Esko Korkeakosken artikkeli *Perusopetuksen oppimistulokset ohjausjärjestelmän osaksi* käsittelee oppimistulosten kansallisen arvioinnin kehittämistä järjestelmänä, arvioinnin tavoitteita ja edellytyksiä. Artikkelissa kuvataan nykyinen arviointijärjestelmä, arviointineuvoston oppimistulosten arviointijärjestelmää koskevat linjaukset ja tehdään ehdotuksia uuden arviointijärjestelmän kehittämiseksi.

Anu Räisänen ja **Mari Rökköläinen** tarkastelevat artikkelissaan *Oppimistulosten arviointi laadunvarmistuksena* ammatillisen koulutuksen oppimistulosten arviointijärjestelmää, sen historiaa, nykytilaa ja tulevaisuuden haasteita. Kirjoittajat tähdentävät oppimistulostenkin arvioinnissa vaikuttavuuden ja kehittävyden sekä laadun parantamisen merkitystä valvonnan ja tilivelvollisuuden ohella. Perustuessaan näyttöihin uusi arviointijärjestelmä nostaa myönteisinä puolina esiin arvioinnin avoimuuden, läpinäkyvyyden, jatkuvuuden ja taloudellisuuden, mutta samalla se luo haasteita arviointitiedon luotettavuuden, totuudellisuuden ja yleistettävyyden saavuttamiseksi.

Kathryn Ecclestonen artikkeli *Principles and processes in evaluation systems: an outsider looks in.....* käsittelee Suomen ja Ison-Britannian arviointijärjestelmiä minkä lisäksi artikkelin kirjoittaja tekee vertailevaa analyysiä näistä järjestelmistä. Artikkelin on siirtänyt ammatillisen koulutuksen oppimistulosten arviointijärjestelmään ja sen kehittämiseen mihin työhön kirjoittaja on osallistunut ulkopuolisena asiantuntijana usean vuoden ajan.

Jarkko Hautamäki ja **Jorma Kuusela** tarkastelevat artikkelissaan *Mitataan oppilaita, mutta päätellään koulusta* kriittisesti arvioivien johtopäätösten tekemistä koulusta oppilaiden suoritusten ja heidän asenteidensa ja hyvinvointinsa perusteella. Artikkelissa analysoidaan myös koulun vaikutuksiin liittyviä käsitteitä ja tulkintoja.

Kirjoittajien keskeinen sanoma on, ettei koulutuksen tuotoksia voida eristää muista lopputulokseen vaikuttavista tekijöistä. Kriittisen analyysin vastineeksi kirjoittajat esittävät ego-logisten tarkasteluiden periaatteita ja esittävät metodis-eettisiä suosituksia.



Perusopetuksen oppimistulokset ohjausjärjestelmän osaksi

Oppimistulosten kansallisen arvioinnin kehittämisen taustaksi kuvataan Koulutuksen tutkimuslaitoksen ja sen edeltäjien sekä Opetushallituksen tekemää oppimistulosten arviointia. Nykytilan kuvauksen ja Koulutuksen arviointineuvoston linjausten kautta syvennetään oppimistulosten arviointijärjestelmän kehittämistavoitteita ja niitä keinoja, joilla tavoitteisiin tulisi pyrkiä.

Oppimistulosten arviointijärjestelmän tavoitteena on tuottaa osuvaa, systemaattista, kattavaa, ajantasaista, monipuolista ja korkeatasoista pitkittäis- ja poikittaistietoa oppimistuloksista ja niihin yhteydessä olevista keskeisistä tekijöistä. Tietoa tuotetaan kansallisiin ja paikallisiin, sekä välittömiin että pitempiaikaisiin tarpeisiin. Kansainvälinen, kansallinen, alueellinen ja paikallinen tietotuotanto täydentää toistaan aukottomasti. Arvioinnit organisoidaan parhaiden asiantuntijoiden monitahoisena ja -tieteisenä mahdollisimman riippumattomana verkostona. Opetushenkilöstön perus- ja täydennyskoulutus, kansainvälinen yhteistyö, arviointitutkimus, arvioinnin arviointi sekä arvioijakoulutus tukevat järjestelmän kehittämistä.

Tavoitetilan saavuttaminen edellyttää yksiselitteistä säädöspohjaa, strategisia linjapäätöksiä, toimivia hallinnollisia rakenteita, arviointikulttuurin suunnitelmallista vakiinnuttamista sekä monitahoista, -ammattillista ja -tieteistä yhteistyötä. Lisäksi tärkeä edellytys on hallinnon, arvioijien ja kehittäjien yhdensuuntainen ja määrätietoinen toiminta.

Asiasanat: arviointijärjestelmä, oppimistulokset, opetussuunnitelma, kehittäminen

Oppimistulosten arviointijärjestelmän kehittäminen perusopetuksessa

Koulun toiminnan päätarkoitus on luoda edellytyksiä oppimiselle. Näin ollen on luonnollista, että eri tavoin arvioidaan niitä tuloksia, joihin koulu pyrkii. Kyse on ensi sijassa oppimistulosten arvioinnista. Järjestelmä-, tila- ja teema-arvioinnit tuottavat tuloksellisuustietoa, jota tulisi tarkastella yhdessä oppimistulosten kanssa. Valtaosa koulutuksen tuloksellisuustiedosta tuotetaan oppilaitoskohtaisesti. Paikallisesti ja kansallisesti tuotetun tiedon määrä on ollut Suomessa kasvussa noin kymmenen vuoden ajan. Kansainvälisiä koulusaavutustutkimuksia on tehty Suomessa 1960-luvun alusta lähtien.

Oleennaista oppimistulosten arvioinneissa on mitä, miten ja milloin tietoa tuotetaan sekä millaisen kokonaisuuden tämä tieto muodostaa. Tiedon tärkeys johdetaan perusopetuksen yhteiskunnallisista tehtävistä ja päätellään tuotetun tiedon soveltuvuudesta ja soveltamisesta kehittämistarpeisiin. Näiden tarpeiden kirjo ulottuu opetusprosessien kehittämisestä merkittävien kansallisten koulutusongelmien ratkaisemiseen. Kyse on tällöin arviointitoiminnasta, joka täyttää järjestelmälle asetettavat vaatimukset. Oppimistulosten arviointijärjestelmä on osa kansallista koulutuksen arviointijärjestelmää.

Formaalissa koulutuksessa oppimistuloksia on aina arvioitu. Tulokset ovat olleet merkityksellisiä arvioinnin kohteille, oppilaille ja heidän opettajilleen sekä oppilaiden vanhemmille. Koulukohtaisten tehtävien ja vastuun kasvaessa tuloksia on alettu tarkastella enemmän yksikkökohtaisesti. Perusopetuksen alkutaipaleella pidettiin yleisesti kuntakohtaisia kokeita ja kokeilukoulut osallistuiivat koetoimintaan valtakunnallisesti. Opetussuunnitelmien kansallinen yhtenäisyys tarjosi oivan tilaisuuden tarkastella tuloksia kansallisesti ja edistää peruskoulun toimeenpanoa maassa.

Kansalliset vaatimukset ja kansainväliset pyrkimykset ovat vetäneet koulutuksenkin mukaan keskusteluun tilivelvollisuudesta, tuloksellisuudesta, tuottavuudesta, laadusta, kehittämistarpeista ja kansalaisten yhdenvertaisuudesta palvelutarjonnan kannalta. Keskustelu koulutuksen vaikuttavuudesta on ulottunut myös keskusteluun opettajankoulutuksen ja opettajien täydennyskoulutuksen vaikuttavuudesta (esim. Kirkpatrick 1969; Hamblin 1974; Niemi 2000).

Perusopetuksessa vaikuttavuuskeskustelun päähuomio on toistaiseksi kohdistunut oppimistuloksiin ja niiden tasoon suhteessa kansallisesti asetettuihin tavoitteisiin. Kun tätä tietoa tuotetaan tavoitteisesti ja suunnitelmallisesti, voidaan järjestelmällisen tulosarvioinnin todeta alkaneen. Kun korkeatasoisen tiedon tuottamisen ohella kehitetään mekanismeja tulosten hyödyntämiseen, varmistetaan arviointitoiminnan laatu ja luodaan toimivat käytännöt arviointiosaamisen parantamiseen, voidaan jo puhua oppimistulosten arvioinnista järjestelmänä.

Oppimistulosten arvioinnin hyödyllisyyttä tulisi tarkastella yksilötasolta koulutuksen järjestelmätasolle ja aina yhteiskunnalliseen vaikuttavuuteen saakka. Arvioinnin

primääri tarkoitus on toimia kehittämisdynamiikan käynnistäjänä ja moottorina (esim. PL 21 §; Norris ym. 1996, 88). Koulutuksen perimmäinen tarkoitus on kuitenkin aina yksittäisen oppilaan tai opiskelijan oppimisen ja hyvinvoinnin edistäminen, vaikka oppimistuloksia tulee tarkastella myös koko koulutusjärjestelmän tuloksellisuuden näkökulmasta.

Koulutuksen arviointineuvoston ohjelmassa painottuu uuden oppimistulosten arviointijärjestelmän kehittäminen ja sen toimivuuden testaaminen (Koulutuksen arvioinnin uusi suunta 2004, 42). Perusopetuksen oppimistulosten arvioinnin kehittämistä tarkastellaan seuraavaksi 1) kansallisen arviointitoiminnan kehitysvaiheina 1950-luvun lopulta nykypäivään, 2) nykytilan kuvauksena, 3) strategisina kehittämistarpeina koulutuksen arviointijärjestelmän ja opetussuunnitelmajärjestelmän kontekstissa sekä 4) metodisina kehittämistarpeina.

Kansallisen arviointitoiminnan kehitysvaiheita

Oppivelvollisuuskoulumme arviointitoiminnassa pääpaino on aina ollut oppimistulosten arvioinnilla. Se on ollut oppilaan arvostelun ja oppilaan arvioinnin perusta. Se on ollut myös tausta, johon laajempi arviointinäkemys on kiinnittynyt. Koulun toiminnassa on ollut tarpeellista jollakin tavalla selvittää muitakin koulun aikaansaannoksia.

Kansakoulun tarkastajien työ keskittyi aikanaan toiminnan lainmukaisuuden valvontaan, vaikka tarkastuksiin sisältyi muutakin tiedonkeruuta. Tarkastuskertomuksen kohteet perustuivat todennäköisesti Uno Cygnaeuksen ja K. G. Leinbergin yhteistyössä kehittelemään kaavaan, joka vuonna 1872 sisälsi tietoja, jotka koskivat mm. opetusta, todistuksia, käytöstä, opetusmenetelmiä ja oppimistuloksia. (Länsi-Suomen lääninhallitus 2002, 21.) Varsinainen tarkastustoiminta alkoi kansakouluissa vuonna 1885.

Peruskoulun myötä korostui kansallisesti yhtenäisten tavoitteiden mukainen arviointi. Tämä tapahtui kansallisen opetussuunnitelman, oppimateriaalien, hallintovalvonnan (mm. opetussuunnitelman vuositarkisteen) ja oppilaan arvioinnin normien avulla. Oppimateriaalit tukivat nykyistä enemmän opetussuunnitelman mukaisten tavoitteiden arviointia, koska koulujen tuli käyttää opetuksessaan Kouluhallituksen hyväksymiä oppikirjoja. Oppimateriaalien tarkastustoiminta päättyi vuonna 1992 (Niemi 2004, 4) ja koulujen tarkastustoiminta päättyi vuonna 1991.

Vuoden 1985 opetussuunnitelman perusteet antoi jo selvästi tilaa paikallisille opetussuunnitelmille, mutta vasta vuoden 1994 opetussuunnitelman perusteet avasi mahdollisuuden koulukohtaiseen opetussuunnitelma-ajatteluun (Opetushallitus 1994, 9). Muutos lisäsi tarvetta myös paikallisen arvioinnin kehittämiseen.

Kasvatustieteiden Tutkimuskeskus (KTK) Jyväskylässä tarjosi palvelujaan koulun- uudistusten edistämistä sekä kasvatus- ja opetustyön kehittämistä varten vuodesta

1957 alkaen. Esimerkiksi vuonna 1967 oppimistulosten osalta oli valmistumassa koko ikäluokan oppimistulosten arviointi ruotsin kielessä ja useita muita tutkimuksia ensimmäisestä vuosiluokasta alkaen. Lisäksi oli käynnissä IEA:n kansainvälinen koulu-saavutustutkimus äidinkielessä, luonnontieteissä, yhteiskuntaopissa ja englannin kielessä. (Kasvatustieteiden Tutkimuskeskus 1967, 5–9.)

Suomi on ollut IEA-järjestön toiminnassa mukana 1950-luvun lopulta saakka, ensin kokeilevassa vaiheessa ja sen jälkeen viidessätoista yhteistutkimuksessa. Pitkäjänteinen ja todellinen kansainvälinen tutkimusyhteistyö avasi ikkunoita maailmaan, mutta käynnisti samalla koulutuksen seuranta- ja arviointitutkimusten sarjan, jolloin aika ajoin hankittiin vertailutietoa koulutustuotoksista (mm. lukutaito, matematiikka ja luonnontieteet) ja niiden taustalla olevista tekijöistä. IEA-yhteistyö on rakentanut suomalaisen koulun (erityisesti oppivelvollisuuskoulun) kehitysvaiheista empiirisen tietokannan, jonka tutkimusmerkitys on jatkuva. Tuotettu tieto on suunnattu erityisesti koulutuksen järjestelmätasolle. Tämä Kasvatustieteiden Tutkimuskeskuksessa alkanut työ on jatkunut nykyisen Koulutuksen tutkimuslaitoksen toimintaan saakka. (Leimu 2004.)

Kasvatustieteiden tutkimuslaitos jatkoi Kasvatustieteiden Tutkimuskeskuksen toimintaa vuonna 1968, jolloin oli erityinen tarve saada tutkimustietoa mm. peruskoulu-uudistuksen valmisteluun, erilaisten vaihtoehtojen suunnitteluun ja arvioimiseen sekä uudistusten toimeenpanon seurantaan. Laitoksen jatkaessa vuodesta 1996 alkaen Koulutuksen tutkimuslaitoksena arviointitoiminnan painopiste siirtyi kansainvälisten arviointitutkimusten kansalliseen toimeenpanoon. Laitoksen edeltäjineen voidaan todeta tehneen pitkään monipuolista kansallista arviointi- ja tutkimustyötä erityisesti perusopetuksen hyväksi. Tynjälän (1993) mukaan tutkimuksella on voitu todeta mm. se, että oppimistulokset ovat parantuneet lähes kaikilla tutkituilla alueilla rinnakkaiskoulujärjestelmään ja peruskoulun alkutaipaleeseen verrattuna.

Oppimistuloksia yleisempää koulutuksen tuloksellisuuskeskustelua lienee käyty maassamme vasta reilut 10 vuotta (Hämäläinen ym. 1993, 5). Uuden tuloksellisuusajattelun suuntaista arviointia edustivat jo Opetushallituksen raportit “Lukion tila 1994” ja “Toteuttaako peruskoulu tasa-arvoa?” (1996). Vuoden 1996 elokuun alusta Opetushallitus organisoi arviointitoimintansa ns. pääjohtajan esikuntana viiden opetusneuvoksen vahvuisena. Sen jälkeen henkilöresurssit kasvoivat vaiheittain erityisesti oppimistulosten arvioinnin alettua Opetushallituksessa vuonna 1997. Opetushallitus julkisti oppimistulosten arviointijärjestelmänsä vuonna 1998 (Opetushallitus 1998a), jolloin se toteutti myös ensimmäisen oppimistulosarviointinsa (Korhonen 1999). Vuoden 1999 alusta Opetushallituksen tehtäväksi tuli huolehtia mm. ulkopuolisten arviointien toimeenpanosta (PL 1998, 21 §).

Oppimistulosarviointeja on toimeenpantu Opetushallituksessa vuosina 1998–2004 kaikkiaan 22, joista viisi (modersmål) on rinnastettavissa osaksi kansallista äidinkielen ja kirjallisuuden arviointia. Oppimistulosten kansallisessa arvioinnissa yhtenä keskeisenä lähtökohtana on ollut seurata koulutuksellisen tasa-arvon toteutumista

perusopetuksessa (esim. Jakku-Sihvonen & Komulainen 2004). Varsinaisiin uusiin innovaatioihin koulutuspolitiikan tarkistamiseksi tai merkittävien koulutuksen ongelmien ratkaisemiseen arviointitulokset eivät ole johtaneet (Rönholm 2005, artikkeli tässä julkaisussa).

Oppilaitoslainsäädäntöä muutettiin vuonna 2003 siten, että opetusministeriön yhteyteen tuli Koulutuksen arviointineuvosto (PL 2003, 21 §). Sen tuli organisoida arviointitoiminta yliopistojen, Opetushallituksen ja muiden arviointiasiantuntijoiden verkostona. Valtioneuvoston asetuksessa koulutuksen arvioinnista (20.2.2003/150, 4 §) todetaan täsmäten, että Opetushallitus osallistuu opetussuunnitelmien perusteiden mukaisten oppimistulosten arviointiin osana koulutuksen arviointiasiantuntijoiden verkostoa. Opetushallitus antaa lisäksi määräyksen opetussuunnitelman perusteista, opetussuunnitelmien toimeenpanoon liittyvää täydennyskoulutusta ja normit oppilaan arvioinnista (esim. Opetushallitus 2004, 260; PL 2003, 22 §). Opetusministeriön ja Opetushallituksen välisen tulossopimuksen nojalla Opetushallitus arvioi edelleen oppimistuloksia. Koulutuksen arviointineuvostolle tai arviointiasiantuntijoiden verkostolle ei ole toistaiseksi osoitettu resursseja oppimistulosten arviointeihin.

Arvioinnin kehittämistarpeista

Kansainväliset arvioinnit perustuvat avaintaitoihin, jotka ovat tärkeitä jatko-opintojen, työelämän, aktiivisen kansalaisuuden ja elinikäisen oppimisen kannalta. Tästä johtuen arviointimenetelmissä suositaan elämänläheisiä, autenttisia tehtäviä, joiden ratkaisemisessa oppilas joutuu käyttämään tietojaan yli oppiainerajojen (Linnakylä 2002). Tämä tieto on välttämätön lisä kansalliseen oppimistulostietoon, koska oppimistulosten arvioinnit ovat keskittyneet oppiaineiden oppimistavoitteiden saavuttamiseen. Kansainväliset tulokset antavat myös viestin kansallisen arvioinnin tarkemmille kohdennuksille. Kansainvälinen tulostieto vakuuttaa erityisesti toistuessaan. 1970-luvulta peräisin oleva kuuden oppiaineen tutkimuksen yhteydessä tehty vertailu (IEA Six Subject Study) sekä vuosina 1991 ja 1995 tehdyt kansalliset kouluviihtyvyyden arvioinnit eivät antaneet kovin myönteistä kuvaa suomalaisten oppilaiden asennoitumisesta kouluun ja opiskeluun (Linnakylä 1993; Linnakylä ym. 1996). Kuva ei ole juuri muuttunut uudemmissa vertailuissakaan. Lisäksi kansallisessa arvioinnissa osoittautui, että vuosien 1997 ja 2001 välisenä aikana heikointa osaamista osoittaneiden yhdeksäsluokkalaisten poikien oppimisasenteet olivat muuttuneet oppimisen kannalta haitalliseen suuntaan (Hautamäki ym. 2003).

Kansallisia kokeita laadittaessa on kiinnitetty vain vähän huomiota siihen, millaista oppimiskäsitystä mittarit edustavat, miten sisältöosiot on poimittu tavoitekokonaisuudesta tai miten arvioijien subjektiiviset arvot vaikuttavat mittareiden tehtävävalintoihin. Osaamista arvioitaessa arvovalinnat ovat tärkeitä, koska valintojen avaruus ulottuu tiedonaloista (oppiaineet) kasvatuksen yleistavoitteisiin sekä oppiaineiden sisäl-

tökokonaisuuksista yksittäisiin tehtäviin saakka. Tehtävän osaamisella on toki oma arvonsa tuloksena, ja tehtävän ratkaisemisella oppimista edistävänä prosessina oman sa. Kontekstista riippuu, mitä näkökulmaa painotamme. Arvioinneissa valintojen arvoperusta opetussuunnitelma-analyyseineen on dokumentoitava näkyväksi. Ellei näin menetellä, voivat tuloksien yhteydessä esitettävät arvioivat päätelmät saada ristiriittaisen vastaanoton.

Oppimistulosten arvottamista ei varsinaisesti ole se, että tiedetään, kuinka suuri osa tehtävistä osataan, montako prosenttia tulos on maksimipistetasosta tai millä tasolla opetusryhmän tulos on kansalliseen keskiarvoon verrattuna. Havaintoaineiston monipuolinen kuvaus edeltää arvottamista. Arvottamisessa olennaista on tuloksen arvon esiin nostaminen. Arvottamisen painopiste on monipuolisen ja luotettavan tulostiedon jälkeen.

Asioiden tai tulosten arvoa ei voida määrittää objektiivisesti. Arvon antaminen perustuu subjektiivisiin arvostuksiin, ennalta sovittuihin arvottamisen kriteereihin tai yhteiskunnallisen päätöksenteon yhteydessä eri tavoin todettujen arvojen tulkintoihin. Vuoren (2004, 1) mukaan arviointitotuus on aina jossain määrin relativistista. Arviointistandarditkaan eivät ole arvoneutraaleja, koska arviointi on tiedon suhteuttamista arvoihin. Kansalaisten arvostuksissa voivat joskus olla vastakkain esimerkiksi koulutuksen antama lisäarvo tulonhankinnan kannalta sekä koulutuksessa kehittynyt eettisyyden ja esteettisyyden taju. Tuloksia arvoitettaessa keskeisiä ovat kuitenkin ne arvot, jotka nousevat laki-, asetus- ja säännöstekstien sekä opetussuunnitelman perusteiden tehtävämäärittelyistä ja päämääräasetteluista. Kyse on tulkinnallisesta analyysistä, jossa on mukana sekä määrällisiä että laadullisia kriteerejä. Arvottamisessa tehdään päätelmiä kohteen arvosta.

Oppimistuloksia arvioitaessa syntyy helposti katveita. Kohdealue voi olla vain esimerkiksi muutamien “kovien” oppiaineiden tiedollisten tavoitteiden saavuttamisen arviointi. Kapeimmillaan kontrolloidaan vain muistitiedon osaamista. Vaikuttavuutta tulisi arvioida kattavammin. Tällöin on esimerkiksi tarpeen kysyä aika ajoin, miten hyvän yleissivistyksen perusopetus antaa. Voimmeko arviointi- tai tutkimustiedon perusteella sanoa, miten hyvin perusopetus tukee oppilaiden kasvua ihmisyyteen ja eettisesti vastuulliseen yhteiskunnan jäsenyyteen tai millaisen tieto- ja taito- ja asennepohjan perusopetus antaa jatko-opintojen tai tulevan elämäntehtävän kannalta (vrt. PL 1998, 2 §)? Olennaisia oppimistuloksiin liittyviä tietoja ovat myös oppilaiden oppimisprosessit ja opiskelukäytännöt. Kyse on asioista, joiden tulisi näkyä jo silloin kun laaditaan opetussuunnitelmia, päätetään arviointikriteereistä, kehitetään koulun toimintaa, tuetaan yksilötason oppimisen perustaa ja laaditaan arviointiohjelmaa.

Oppimistulosten arviointijärjestelmässä olisi tärkeää a) saada tietoa koulutusjärjestelmän tekijöiden (koulutuspolitiikka, suunnittelu, toiminta, tulokset) keskinäisistä vuorovaikutusyhteyksistä, b) seurata systemaattisesti keskeisimpiä tulosindikaattoreita (esim. yleissivistyksen ydinalueet, yksilötason hyvinvointi, koulutuksellinen tasa-arvo, oppimaan oppimisvalmiudet), c) kytkeä kulloinkin paras kansallinen osaa-

minen tietotuotannon prosessin vaiheisiin (esim. yliopistot, tutkimuslaitokset, Opetushallitus, yksittäiset arviointiosaat, kansainvälinen asiantuntemus) ja d) tuottaa osuvaa, kattavaa ja ajantasaista tietoa perusopetuksen välittömiin ja pitempiäaikaisiin tarpeisiin. Tämä tarkoittaa mm. sitä, että arvioijat seuraavat intensiivisesti, mitä koulutuspolitiikassa ja koulutuksen kehittämisen saralla tapahtuu. Seuraava asetelma kuvaa oppimistulosten arvioinnin nykytilaa ja kehittämistarpeita ja esittää joitakin kehittämisstrategioita tavoitetilän saavuttamiseksi (kuvio 1).

NYKYTILA	KEHITTÄMISEN STRATEGIOITA	TAVOITETILA
<p>arviointitietoisuuden taso lisääntynyt</p> <p>parhaan ja monipuolisen asiantuntemuksen hyödyntäminen puutteellista</p> <p>verkostomainen toimintakulttuuri epäsystemaattista</p> <p>arvottamisen kriteeristö osin määrittelemätön</p> <p>oppimistulostieto usein erillään muusta arviointitiedosta</p> <p>arviointitiedon tuottaminen kapea-alaista ja resursseiltaan vähäistä</p> <p>tulosten hyödyntäminen vähäistä</p> <p>arviointitoiminnan vastuut osin selkiytymättömiä</p> <p>kansallinen ja paikallinen arviointi tukevat toisiaan riittävästi</p>	<p>säännöspohjan täsmentäminen ja vahvistaminen</p> <p>yhtenäisen käsitteellisen viitekehyksen vakiinnuttaminen ja kohdeanalyysien teoriaperustan vahvistaminen sekä verkostotietämyksen systematisointi</p> <p>vahvasti resurssoidun ja mahdollisimman riippumattomasti organisoidun tietotuotannon vahvistaminen</p> <p>pitkäkestoisen, kattavan ja koordinoitun arviointiohjelman vakiinnuttaminen</p> <p>kumulatiivisen tietotuotannon vahvistaminen (mm. pitkittäisarvioinnit monipuolisine taustatietoineen)</p> <p>tiedonkeruuta ja -välitystä tukevan järjestelmän luominen</p> <p>verkostomaisen asiantuntijuuden ja toimintakulttuurin edistäminen kaikilla tasoilla</p> <p>opetussuunnitelma-, oppimateriaali- ja arviointijärjestelmien kehittäminen rinnakkain</p> <p>arviointiosaamisen ja arviointitiedon kysynnän vahvistaminen sekä hyödyntämismekanismien kehittäminen järjestelmän eri tasoilla</p>	<p>Järjestelmä tuottaa osuvaa, systemaattista, kattavaa, ajantasaista, monipuolista sekä korkeatasoista pitkittäis- ja poikittaistietoa kehittämisen välittömiin ja pitkäaikaisiin tarpeisiin.</p>

Kuvio 1. *Oppimistulosten arviointijärjestelmän kehittämisen lähtökohtia*

Pelkän tulostiedon varassa on lähes mahdoton päätellä jotain koulujen paremmuudesta esim. perusopetuksen päättövaiheessa tai opetuksen laadusta tulosten suhteellisen aseman perusteella. Tätä päättelyä ei voida tehdä edes kahden mittauskerran perusteella (ks. Hautamäki & Kuusela 2005, artikkeli tässä julkaisussa). Esimerkiksi koulujen välillä todettujen erojen taustalla voivat vaikuttavat monenlaisten muuttujien kombinaatiot sekä oppilaiden taustojen että oppimisympäristöjen osalta. Oppiaineiden tulostietojen ohella on tärkeää arvioida vaikuttavuutta, jossa korostuu sellaisten yleisempien tietojen ja taitojen käyttö, jotka toimivat uuden oppimisen välineinä esim. kaikissa lukuaineissa (ks. Hautamäki ym. 2004, 7; Linnakylä 2002).

Kansalliset kokeet ovat olleet otosperustaisia. Opetushallitus toteaa kuitenkin, että kokeet voidaan toimeenpanna myös koko ikäluokan arviointeina (Opetushallitus 1998a, 9). Kokeita on pidetty lähinnä 9. vuosiluokan äidinkielen ja matematiikan oppiaineissa (ks. myös Opetushallitus 1998a, 11). Tuloksellisuuden käsitteistössä vaikuttavuus ymmärretään kuitenkin laaja-alaisempana (Opetushallitus 1998a, 6). Päättövaiheessa saatua palautetta oppimistuloksista ei voida juurikaan hyödyntää enää oppilastasolla, joskin palautteen tulkinnallinen itsearviointi antanee opettajille joitakin yleisiä opetuksellisia vihjeitä. Perusopetuksen vuosiluokkien 1–6 koulut ovat olleet lähes kokonaan toteutuneen järjestelmän ulkopuolella.

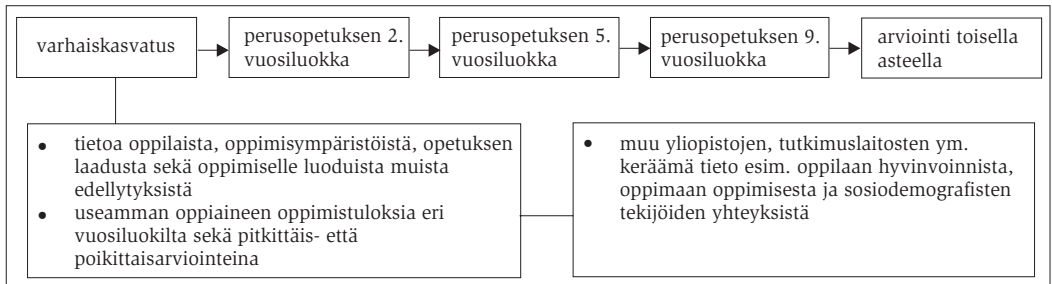
Kaikki koulutuksen järjestäjät eivät ole ostaneet tai eivät ole kyenneet ostamaan oppimistulosarviointeja kouluilleen. Näin ollen kehittävän arvioinnin näkökulmasta oppimistulosten arviointi ei vielä tue riittävästi oppilaan arvioinnin oikeudenmukaisuuden toteutumista. Se myös saattaa koulutuksen järjestäjät eriarvoiseen asemaan.

Arvioinnin kehittämisstrategioita

Oppimistulosten arvioinnin kehittämisessä tulisi nähdä yhteisiä tavoitteita yli koulutuksen sektorirajojen ja yli organisaatioiden intressien. Lisäksi tulisi tuntee eri tasojen pyrkimyksiä ja toimintaa sekä haluta aidosti kehittää oppimista arvioinnin keinoin koordinoituna kokonaisuutena, kansallisena järjestelmänä. Kansalliset kokeet perusopetuksen päättövaiheessa ja ylioppilaskirjoitukset lukiossa eivät täytä arviointijärjestelmän vaatimuksia. Oppimistulosten arviointia paikallisella tasolla voitaneen kehittää nykyistä enemmän osana kansallista vaikuttavuustiedon tuottamista, koska opetuksen perusta – ainakin virallisesti – on opetussuunnitelman perusteissa ja oppilaan arvioinnin ohjeissa (ks. Opetushallitus 2004, 8).

Oppimistulosten arviointi on kontekstiinsa sidoksissa oleva kompleksinen kokonaisuus. Kaikenlaisia asioita voidaan mitata: laajoja oppimisen valmiuksia, yksittäisiä muistitiedon sirpaleita, yhden tai useamman oppiaineen tai oppiainekokonaisuuden tuloksia sekä ajattelun taitoja. Toisaalta voidaan mitata myös lapsen ja nuoren kehitystä kokonaisuutena tai vaihtoehtoisesti vain syrjäytymiseen johtavia polkuja. Useamman oppiaineen samanaikainen arviointi, joka tehdään samoille oppilaille (ks.

Koulutuksen arviointineuvosto 2004, 45), lisää jo evidenssiä tehdä päätelmiä oppimistulosten tasosta koulussa yleisemminkin. Lisäksi tarvitaan pitkittäisseurantaa jopa varhaiskasvatuksesta alkaen aina toisen asteen oppilaitoksiin saakka, ts. tiettyjen oppilaskohorttien seuraamista esimerkiksi seuraavan kuvion 2 mukaisesti (ks. Korkeakoski 2003; Nikkanen & Lyytinen 1996, 149–150).



Kuvio 2. Esimerkki ikäkohortin seurannan ketjusta koulujärjestelmässä

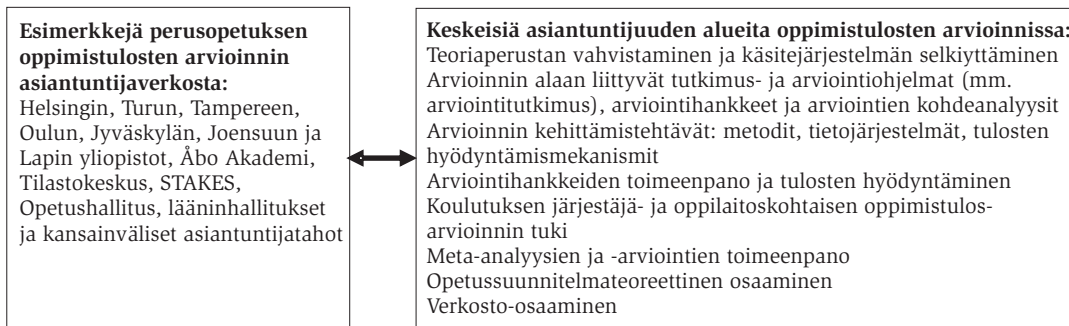
Pitkittäisseuranta antaa tietoa järjestelmän toimivuudesta eri vaiheissa eri tavoitteiden kautta. Tieto sisältää sekä tulostietoa että toimintaa kuvaavaa tietoa. Sen myötä vahvistuu myös koulutuksellisen tasa-arvon todentaminen vertailtaessa mm. alueellisia ja oppilaitoskohtaisia eroja.

Kun oppimistulosten yhteydessä kerätään tietoa oppilaan taustoista, kasvu- ja kehitysympäristöistä sekä koulun toiminnasta (esim. opetuksen laadusta), voidaan oppimistuloksista saatu tieto suhteuttaa paremmin kontekstiinsa. Tämä tarkoittaa mm. yhteyksien ja selitysmallien etsimistä hyvillä oppimistuloksille, jolloin saadaan myös vahvempaa evidenssiä “hyvien käytäntöjen” kuvauksille ja levittämiselle. Vain järjestelmällinen, metodisesti luotettava ja monipuolinen tiedonkeruu ja analysointi tuottaa kumuloituvaa tietoa, jota voidaan hyödyntää myöhemminkin arvioinneissa, tutkimuksissa ja kehittämisessä. Tämä tieto yhdistettynä kansainväliseen ja paikalliseen arviointitietoon – eräin edellytyksin tutkimustietoonkin – antaa evidenssiä arvioida myös koulutuksen yhteiskunnallista vaikuttavuutta.

Parhaimmillaan oppimistulosten arviointi on monitahoista, -ammattillista ja -tieteistä yhteistyötä; tiedonala-asiantuntijoiden, metodikkojen, opettajien sekä koulutuksen tutkijoiden ja arvioijien professionaalista vuorovaikutusta. Erityisesti oppimistulosten arvioinneissa arvioijien ja tutkijoiden metodinen vuoropuhelu korostuu, koska tarvitaan lisää käsitteellistä selkeyttä ja yhtenäisyyttä, kohdeilmiöiden teoreettisempaa haltuunottoa, otantamenetelmien, tiedonhankinnan ja -käsittelyn syvällistä tuntemusta. Jo oppimisilmiöiden sekä tiedonalojen moninaisuus ja kompleksisuus edel-

lyttää asiantuntijoiden systemaattista ja toiminnallisesti integroitua verkostoitumista.

Seuraavassa asetelmassa todetaan keskeiset oppimistulosten arviointiin liittyvät asiantuntijuuden alueet sekä esimerkkejä joistakin kansallisista toimijoista, joilla on osaamista yhdellä tai useammalla oppimistulosten arviointiin liittyvällä osa-alueella (Muistiot arviointineuvoston ja -sihteeristön verkostovierailuista 2004).



Kansainvälinen yhteistyö ja oppimistulosten arviointi tuovat mukanaan selvää lisäarvoa siihen, miten arviointitoimintaa tulisi kehittää metodisesti ja mitä kansallisia osaamisen alueita tulisi kehittää. Kansallisilla arvioinneilla voidaan sitten arvioida kohteita tarkemmin. Välijärven (2003, 4) mukaan esimerkiksi PISA:ssa ovat mukana kaikki Suomen kannalta tärkeimmät vertailumaat, jolloin suomalaisen peruskoulun tulokset sijoittuvat realistiseen ja monipuoliseen vertailukehykseen. Kansallista järjestelmää kehitettäessä olisi tunnettava paikallisen (oppilaitoksen), kansallisen ja kansainvälisen tason arviointien poikkeavat lähtökohdat ja sovitettava toimintakokonaisuudet yhteen näkökulmien ja verkottumisen kannalta.

Opetussuunnitelman perusteiden arvoista, tavoitteista ja sisällöistä sekä ohjaavista käsityksistä ovat johdettavissa ne priorisoinnit, luokittelut ja jäsenyykset, joita käytämme oppimistuloksia arvioitaessa. Ellei opetussuunnitelmaa ole avattu teoreettisten taustaolettamusten, ohjaavien käsitysten ja jäsenyyksien osalta, olemme pulmallisten arvovalintojen edessä silloin kun oppimistuloksia arvioidaan.

Kouluissa oppimistulosten arvioinnit tapahtuvat usein kokeilla, jotka edellyttävät opetusta ja opiskelua tietyn oppimateriaalikonaisuuden ohjaamana. Tämä on ongelmallista, koska ei ole luotettavaa tietoa siitä, miten hyvin oppimateriaalit vastaavat opetussuunnitelman tavoitteita. Siksi eri oppimateriaalien käyttö opetuksessa saattaa johtaa oppimistulosten vaihteluun, kun arvioidaan oppimistuloksia opetussuunnitelman perusteiden mukaan (esim. Niemi 2004). Oppikirja voi olla opettajalle jopa opetussuunnitelman asemassa (Kuusisto 1989). Oppimateriaalien arviointi olisi siis kan-

sallisesti välttämätöntä niiden keskeisen sisällöllisen, opetuksellisen ja koulun arviointeja määrittävän merkityksen vuoksi (Korkeakoski 1989, 138; Viiri 2003). On siis pääteltävissä, että oppimateriaalikomitean (Mietintö 1973; 1–2) luonnehdinta oppimateriaalien merkityksestä, “Oppimateriaalit muodostavat oppimisen rungon koulussa”, on yhä osa koulun käytäntöä. Edellä todetun perusteella opetussuunnitelma-, oppimateriaali- ja arviointijärjestelmät tulisi rakentaa pääosin samoille periaatteille. Yksittäisen opettajan toiminnassa tämä periaate toteutuukin, ainakin implisiittisesti.

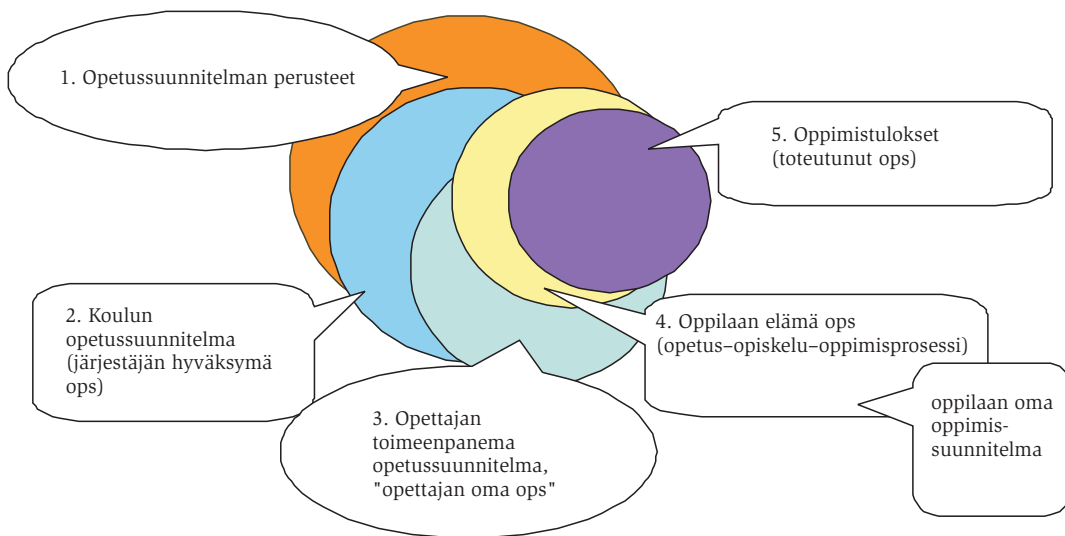
Arviointi opetussuunnitelmajärjestelmän kehityksessä

Opetussuunnitelman perusteet on keskeisin sisällöllinen kokonaiskuvaus kansallisesta sivistystehtävästä. Se on oppimistulosten kansallisen arvioinnin kannalta avainasiakirja. Vaikka Opetushallitus päättää opetussuunnitelman perusteista, sen voidaan ajatella edustavan yleisemminkin yhteiskunnallisia arvoja. Vuoden 2004 opetussuunnitelman perusteisiin pyydettiin lausunto 162 eri taholta ja lausuntoja annettiin 124, joista 30 oli koulutuksen ulkopuolisten tahojen antamia (Vitikka 2004). Se on myös kansallinen kehys, jonka pohjalta paikallinen, opetuksen järjestäjän hyväksymä opetussuunnitelma laaditaan.

Opetussuunnitelman perusteiden tavoitteet, sisällöt, ohjaavat käsitykset ja hyvän osaamisen kriteerit ovat arvovalintoja. Arvioijien arvovalinnoiksi jäävät arviointikohteet, tavoiteluokitukset, tehtävät ja niiden tasoryhmittelyt. Tällöin on vaara, että arvokkaat asiat nähdään suppeasti ja arvottaminen johdetaan muualta kuin opetussuunnitelmasta. Vaara voidaan osin eliminoida arviointiasiantuntijoiden monipuolisuudella ja verkostomaisella organisoinnilla.

Opetussuunnitelman perusteet on erittäin tärkeä arvioinnin kohde a) yhteiskunnallista kehitystä ja arvoja kuvaavana asiakirjana ja b) vertailulähteenä yhdessä kouluisa noudatettavien opetussuunnitelmien kanssa. Perusongelma tietysti on, miten opettajat noudattavat toimipaikkansa opetussuunnitelmia. Mittaamisen ja arvottamisen kannalta on kuitenkin olennaista, ovatko asiakirjan tavoitteet 1) täsmällisiä, 2) mitattavia, 3) saavutettavia, 4) merkittäviä ja 5) onko niillä ajallinen ulottuvuus. Jos opettajat eivät tunne tai arvosta opetussuunnitelmaa, sillä ei ole juuri käytännön merkitystä opetustyössä. Opetussuunnitelman tavoitteiden ja sisältöjen merkittävyys eri tahojen kannalta on yksi peruskysymys silloin, kun koulutuksen yhteiskunnallista vaikuttavuutta arvioidaan.

Oppimistuloksien taustalla on seuraava opetussuunnitelmallinen prosessiketju, joka on tunnettava oppimistuloksia arvotettaessa ja selitettäessä.



Kuvio 3. Oppimistulosten arviointi opetussuunnitelmajärjestelmän viitekehyydessä

Yhteyksien tarkastelut kuvatus ketjun pohjalta romuttuvat, jos kansalliset kokeet eivät perustukaan analyysiin opetussuunnitelman perusteista tai sen oppimisenäkemykseen, vaan yhteiskunnan kontrollin ulkopuolella olevat oppikirjat ja niihin liittyvät materiaalit ohjaavat opetusta (ks. Mikkilä-Erdman, Olkinuora & Mattila 1999; Niemi 2004). Törnroosin (2004, 218) tutkimuksen mukaan Opetussuunnitelman perusteiden 1994 pohjalta ei voida kuvailla esimerkiksi perusopetuksen 7. vuosiluokan oppilaiden matematiikan oppimismahdollisuuksia kovinkaan yksityiskohtaisesti, mutta analysoidut 5. ja 6. vuosiluokan matematiikan oppikirjat täyttivät kyllä hyvin opetussuunnitelman perusteissa asetetut tavoitteet. Lisäksi oppikirjojen erot olivat 7. vuosiluokalla selkeämpiä kuin vuosiluokkien 5–6 kirjoissa. Virallinen opetussuunnitelmaketju ohjausjärjestelmineen saattaa näin kätkeä hiljaa hyväksytyt ongelmanvyyhdin, joka vaikuttaa kansallisten oppimistulosten arviointiin. Arvioijien asiana on huolehtia siitä, että ohjausjärjestelmän ja opetussuunnitelmien ongelmat arviointien lähtökohtina nostetaan esiin.

Opetussuunnitelman perusteiden tavoitteet saavutetaan kansallisten oppimistulosten arviointien mukaan keskimäärin noin 60-prosenttisesti. Tasoon vaikuttaa myös, mille tasolle tehtävien vaikeustaso on ennakolta esikokein laadittu. Tason vaihtelua on esiintynyt oppiaineittain 45–68 %. (Korkeakoski 2002.) Kun opetussuunnitelman perusteiden tavoitteita ja niiden saavuttamista analysoidaan, voidaan eroja selittää kuvion 1 pohjalta puutteilla, joita esiintyy prosessin vaiheesta toiseen siirryttäessä. "Oikeiden" selitysten pitäisi löytyä pääsääntöisesti edellisen kuvion sisältä, eli tuloserot syntyvät siitä, että eroja on opetussuunnitelman perusteiden tavoitteiden toteuttamisessa käytännössä.

Olisi noloa, jos ensisijaiset selitykset tulosvaihteluille löytyisivät opetussuunnitelmallisen ohjausjärjestelmän ulkopuolelta silloinkin, kun tulosten vertailussa on jo otettu huomioon esimerkiksi sosioekonomiset taustatekijät. Kysymys on siitä, että opetussuunnitelmaa ei eri syistä noudateta. Tämä tilanne voi syntyä esimerkiksi silloin, kun oppimateriaalien tavoitteet ja sisällöt ohjaavat opetusta ja oppimista. Tämä selittäminen on osa opetussuunnitelmajärjestelmän arviointia oppimistulosten arvioinnin viitekehyksessä. Opetushallitushan lähtee siitä, että opettajan tulee noudattaa opetuksen järjestäjän vahvistamaa opetussuunnitelmaa (Opetushallitus 2004, 8).

Arvioinnin kohteiden valinnalla voidaan ohjata myös opetuksen kehittämistä. Se mitä arvioidaan, koetaan tärkeäksi myös opettaa ja oppia. Opetussuunnitelma on virallisesti opetuksen, opiskelun ja arvioinnin perustana, mutta opetusta ohjaavat käytännössä enemmän opettajien oppaat, oppikirjat ja niiden kirjalliset tehtävät sekä koe-paketit (esim. Leino 1985; Törnroos 2004, 219). Tämä on pedagoginen konteksti, jossa koulun arki usein toimii. Koulun pedagoginen todellisuus muotoutuu usein liikaa materiaalien edustamasta oppimisen perustasta käsin, mihin ei heijastu esim. konstruktivistinen oppimiskäsitys siten, että se auttaisi oppilasta rakentamaan mielekkään asiakokonaisuuden (Viiri 2003; Mikkilä-Erdman, Olkinuora & Mattila 1999).

Musiikkioppilaitoksien arvioinnissa havaittiin, että mitä aktiivisemmin ja sitoutuneemmin opettajakunta oli osallistunut opetussuunnitelman tekoon, sitä tutumpi ja käytäntöön integroituneempi suunnitelma tuntui olevan (Kurkela & Tawaststjerna 1999). Koska virallisen opetussuunnitelman tuki käytännön opetustyölle on ollut ja on yhä varsin vaatimaton (Korkeakoski 2001, 159–162), joutuvat opetussuunnitelman perusteiden mukaan arvioidut oppimistulokset oudon validiteettiongelman eteen: mitataan sitä mitä virallisesti pitääkin, mutta ei mitata sitä, mitä todella opetetaan ja opiskellaan.

Opetuksen ja opiskelun päämäärä on oppiminen. Tarkasteltaessa opetussuunnitelman perusteiden 1970, 1985, 1994 ja 2004 pedagogisia lähtökohtia ja painotuksia voidaan havaita, että pedagogiikan ja pedagogisen ohjausaineksen painoarvo ohenee kaiken aikaa. Lehrplan-tyyppinen tavoitteisiin ja sisältöihin keskittyvä suunnittelu on vallannut alaa, kun taas curriculum-tyyppinen, oppilaiden oppimiskokemuksiin pohjaava pedagogisempi linja on ohentunut. Virallista oppimiskäsitystä lukuun ottamatta ajatellaan, että pedagogiikka kuuluu opettajan profession ja vapauden piiriin. Pedagoginen vapaus on aina ollut Suomessa laaja, mikä vain vahvistaa sitä käsitystä, että oppimistulosten kannalta opettajan oma pedagoginen ajattelu on avainasia (ks. Kansanen 1991). Edellä kuvatusta taustasta johtuen opetuksen laadun arvottamisen perustaa on etsittävä muualta kuin opetussuunnitelman perusteista. Tätä aukkoa pyritään täyttämään hyvän perusopetuksen ja tuloksellisesti toimivan koulun laatusuosituksilla. Opetusministeri Haataisen (2004) mukaan laatusuositukset ovat parhaillaan laadittavana. Suositusten arvo tulee keskustelun kohteeksi sen jälkeen, kun niiden argumentaatioperusta on julkistettu.

Opetussuunnitelman perusteiden 1994 myötä tapahtui opetussuunnitelmallisen päätösvalan merkittävä delegointi paikalliselle tasolle, joten arvioinnin säädöspohjan uudistus (1998) oli luonnollinen seuraus valitulle koulutuspolitiikalle. Opettajien tulee olla opetussuunnitelma- ja arviointiprosessien valmistelun ja toimeenpanon keskiössä (esim. Dewey 1901; Alexander 1984, 2). Muutoin uudistukset jäävät käytännössä näennäisiksi ja ulkokohtaisiksi. Opettajat ovat myös kansallisen oppimistulosten arviointijärjestelmän kehittämisessä avainasiantuntijoita. He tuntevat koulun ja opetuksen todellisuuden.

Edellä todetun perusteella on ainakin jossain määrin harhaista väittää, että kansallisissa kokeissa mitataan sitä, mitä kouluissa opetetaan. On todennäköistä, että materiaalien luomat koulukokeiden normaalikäytännöt ovat ohjanneet vuoden 1996 jälkeen kansallisten kokeiden laadintaa. Olisi siis välttämätöntä, että oppimateriaali-, opetussuunnitelma- ja arviointijärjestelmiä kehitetään rinta rinnan toisiaan tukevaksi kokonaisuudeksi sekä paikallisesti että kansallisesti.

Menetelmällisiä kehittämiskohteita

Arviointitoiminnan menetelmälliset kehittämiskohteet liittyvät tiedon hankintaan, analysointiin, levittämiseen ja hyödyntämiseen. Järjestelmä itsessään voidaan ymmärtää myös menetelmäksi, jonka avulla on mahdollista kehittää päätöksentekoa, opetusta, osaamista ja arviointia (vrt. koulutuksen arvioinnin tavoitteet; asetus 20.2.2003/150). Menetelmällisiksi strategioiksi ovat luokiteltavissa esim. pitkittäisarvioinnit, aineistojen yhdistäminen kumulatiivisen tiedon tuottamisessa sekä opettajan ammatilliseen taustaan liittyvien tietojen, oppilaan arviointitietojen (esim. oppilaskorttien) ja itsearviointitietojen hyödyntäminen. Valmiiden tietorekisterien hyödyntäminen voi olla ongelmallista tietosuojasyistä.

Opetussuunnitelmien tavoiterakenne on tunnettava silloin, kun halutaan oppia tuntemaan olennaiset kysymykset eri tiedonalojen oppimisen kannalta. Sisältöluettelot ovat usein vain temaattisia ryhmittelyjä. Niihin ei sisälly tiedonalojen omaa sisäistä struktuuria tai oppimisen vaiheistamisen tai painottamisen kannalta perusteltua psykologista rakennetta. Toisaalta yleisellä tasolla lausutut tavoitteet voivat johtaa tulokinnalliseen epämääräisyyteen ja pirستاletiedon kasvuun materiaaleissa ja opetuksessa. Tieteessä käytetyt taksonomiat eivät aina sellaisenaan ole sovellettavissa eritasoisen oppimisen osoittamiseksi, mutta ne antavat hyviä eväitä tavoiteluokitteluille ja osaamisen rakenteellisille kuvauksille (ks. Korkeakoski 2003b, 145). Opettajien laadullisen arviointiosaamisen tarve kasvaa sitä mukaa kun numeroarviointi vähenee ja sanallinen arviointi lisääntyy sekä kun henkilökohtaiset opetuksen järjestämistä koskevat suunnitelmat (HOJKS) tai henkilökohtaiset oppimissuunnitelmat lisääntyvät opetuksessa. (Ks. Opetushallitus 2004, 20, 28.)

Kun arviointitietoa hankitaan esim. kokeilla, auditoinneilla tai portfolioilla, tarvitaan mittarien suunnitteluun ja kehittelyyn monipuolista asiantuntemusta. Auditoinnit voivat olla laadunvarmennusjärjestelmien arvioinnin ohella riippumatonta yleistä toiminnan arviointia (OPM 2004). Auditoinneilla voidaan todeta myös, vastaavatko toimintatavat organisaation itselleen asettamia tavoitteita (Opetushallitus 1998a). Auditoinnin tarkoitus voi olla myös koulun itsearviointeja täydentävä ulkopuolinen arviointi, joka saattaa sisältyä koulun toimeksiantona omiin kehittämistarpeisiin. Auditointien taustalla on silloinkin järjestelmällinen ennakolta laadittu systemaattinen tiedonkeruun suunnitelma. Menetelmällisesti auditoinnit voivat sisältää useita ratkaisuja. Arvottavien päätelmien teko ja palautteen antaminen kysyy tällöin poikkeuksellisen vahvaa oppilaitoksen toiminnan, opettajuuden, opetuksen ja oppimisen asiantuntijuutta. Menetelmällisesti auditointi edustaa laadullista lähestymistapaa.

Yliopistojen aktiivinen mukanaolo arviointijärjestelmän kehittämisessä on välttämätöntä siksi, että ne ovat pitkälti vastuussa opettajien perus- ja täydennyskoulutuksesta. Ne ovat lisäksi avainroolissa silloin, kun luodaan perinnettä empiiriselle tutkimukselle, joka kohdistuu koulujärjestelmään ja Suomessa vähän tutkittuihin muutoksen ongelmiin (esim. Norris ym. 1996, 89). Yliopistot vastaavat myös pitkälti kaikesta perusopetukseen kohdistuvasta tutkimuksesta ja arviointitutkimuksesta, joten on luonnollista, että menetelmien kehittämisessä ja niiden empiirisessä testaamisessa vuorovaikutus yliopistojen kanssa on systemaattista.

Korkeatasoisen tiedon tuottaminen edellyttää yliopistojen asiantuntemusta myös testiteoreettisissa kysymyksissä, otantaratkaisuissa, mittareiden laatimisessa ja niiden osioanalyysissä sekä aineistojen, erityisesti yhdistettyjen aineistojen meta-analyysissä. Koetulosten virheanalyysit ovat osa tätä prosessia. Luotettavan tiedon levittäminen käyttökelpoisessa muodossa on olennaista arviointien hyödyntämisen kannalta. Arviointiportaalit tiedon keruuta ja levittämistä varten ovat välttämättömiä. Oppimistulosten arvioinnit palvelevat parhaiten paikallista kehittämistä silloin, kun kaikki koulut ja koko ikäluokka voivat olla mukana arvioinneissa. Tarkoituksesta riippuen arvioinnit voivat vaihdella esimerkiksi nopeista ja kevyistä selvityksistä raskaampiin koko ikäluokan oppimistulosarviointeihin.

Oppimistulosten arviointijärjestelmän kehittämisessä avoin vuoropuhelu hallinnon (mm. päättäjät), tutkijoiden ja arvioijien välillä on välttämätöntä. Niiden, jotka hyödyntävät arviointituloksia, tulisi olla mukana järjestelmän kehittämisessä. Tutkijoiden kriittinen valppaus on tarpeellista erityisesti silloin, kun arviointien tulokset ja arvottamisen kriteerit tulkitaan liian tarkoitushakuisesti tai ne perustuvat vain praktiiseen ymmärrykseen. Erityisiä menetelmällisiä haasteita oppimistulosten arviointi tarjoaa silloin, kun on arvioitava kasvatustavoitteiden (esim. affektiivisten, sosiaalisten ja eettisten tavoitteiden) toteutumista, taide- ja taitoaineiden osaamista tai yleensäkin, kun mitataan yli oppiainerajojen ulottuvia oppimisen tavoitteita (ks. Uusikylä & Atjonen 2000, 169).

Oppimistulosten arviointijärjestelmän kehittämisen tulee olla jatkuvaa. Se edellyttää myös järjestelmän arviointia ja tehtyjen arviointien meta-arviointia aika ajoin. Olennaisia arviointikohteita ovat tällöin, miten järjestelmä palvelee 1) yksilötason oppimisprosesseja oppimistilanteista kouluaikeisen osaamisen kehitykseen saakka, 2) opetusta ja opiskelua välittöminä ja vaihteittaisina palautemekanismeina, 3) kansallisen ja paikallisen opetussuunnitelmallisen ohjauksen onnistumista, 4) opiskelun kehittämistä oppilaan tasolta aina koulujärjestelmän kehittämiseen saakka ja 5) opetussuunnitelmien ja oppimateriaalien kehittämistä yhteiskunnallisten arvojen kuvajana. Kun mainitut linjat nousevat esiin järjestelmän tuottamasta tiedosta, luodaan nykyistä vahvempi perusta koulutuksen yhteiskunnallisen vaikuttavuuden todentamiselle ja suomalaisen osaamisen entistä näkyvämmälle kansainväliselle kilpailukyvyille.

Lähteet

- Alexander, R. J. 1984. *Primary teaching*. London: Holt, Rinehart & Winston.
- AI50/2003. Valtioneuvoston asetus koulutuksen arvioinnista.
- Dewey, J. 1901. Are the schools doing what the people want them to do? *Educational Review* XXI, 459–474.
- Haatainen, T. 2004. Opetusministeri Tuula Haataisen puhe 1.12.2004 Yhdessä kasvamaan -projektin päätösseminaarissa Helsingissä.
- Hamblin, A. C. 1974. *Evaluation and control of training*. London: McGraw Hill.
- Hautamäki, J., Arinen, P., Hautamäki, A., Kupiainen, S., Lindholm, B., Mehtäläinen, J. & Niemivirta, M. 2003. Oppimaan oppiminen yläasteella 2. Tilanne vuonna 2001 ja muutokset vuodesta 1997. Opetushallitus. Oppimistulosten arviointi 6/03. Helsinki.
- Hautamäki, J., Kupiainen, S., Mehtäläinen, J., Rantanen, P. & Ruuth, M. 2004. Peruskoululaisten oppimaan oppimisen arviointi osana yleissivistävän koulutuksen arviointistrategiaa. Helsingin kaupungin opetusviraston julkaisusarja A2: 2004.
http://www.opetusministerio.fi/opm/uutiset/archive/2004/12/01_1.html
- Hämäläinen, K., Laukkanen, R. & Mikkola, A. 1993. *Koulun tuloksellisuuden arviointi*. Helsinki: VAPK.
- Jakku-Sihvonen, R., Lindström, A. & Lipsanen, S. (toim.) 1996. *Toteuttaako peruskoulu tasa-arvoa?* Helsinki: Opetushallitus.
- Jakku-Sihvonen, R. & Komulainen, E. 2004. Perusopetuksen oppimistuloksien meta-arviointi. Opetushallitus. Arviointi 1/2004. Helsinki.
- Kansanen, P. 1991. Pedagogical thinking: The basic problem of teacher education. *European Journal of Education* 26 (3), 251–260.
- Kasvatustieteiden Tutkimuskeskus 1967. *Kasvatustieteiden Tutkimuskeskuksen toimintakertomus vuodelta 1967*. Jyväskylän yliopisto.
- Kirkpatrick, D. L. 1998. *Evaluating training programs: The Four Levels*. 2nd ed. San Francisco (CA): Berrett-Koehler.
- Korhonen, H. 1999. Peruskoulun matematiikan oppimistulosten kansallinen arviointi 1998. Opetushallitus. Oppimistulosten arviointi 1/1999. Helsinki.
- Korkeakoski, E. 1989. Opetussuunnitelma opettajan näkökulmasta peruskoulun ala-asteella. Tampereen yliopisto. Kasvatustieteen laitos. Tutkimusraportti A45.

- Korkeakoski, E. 2003a. Opetuksen laatu ja oppimistulokset perusopetuksen tavoitteiden kannalta. Teoksessa P. Kansanen & K. Uusikylä (toim.) *Luovuutta, motivaatiota, tunteita. Opetuksen tutkimuksen uusia suuntia*. Jyväskylä: PS-kustannus.
- Korkeakoski, E. 2003b. Kirja-arvostelut: Perusopetuksen opetuskokeilussa lukuvuonna 2003–2004 noudatettavat opetussuunnitelman perusteet vuosiluokille 3–9 ja perusopetuksen opetussuunnitelman perusteet vuosiluokille 1–2. Helsingin yliopisto. Opettajankoulutuslaitos. *Didacta Varia* 8 (2), 144–147.
- Koulutuksen arvioinnin uusi suunta 2004. Arviointiohjelma 2004–2007. Koulutuksen arviointineuvoston julkaisuja I. Jyväskylä.
- Kurkela, K. & Tawaststjerna, E. T. 1999. Analyysi kuuden musiikkiopiston toiminnasta. Teoksessa T. Heino & M.-L. Ojala (toim.) *Musiikkioppilaitosten perusopetuksen arviointi 1998*. Opetushallitus. Arviointi I/1999. Helsinki.
- Kuusisto, J. 1989. Oppimateriaalit peruskoulun ala- ja yläasteella 1988. Jyväskylän yliopisto. Kasvatustieteiden tutkimuslaitoksen julkaisusarja A:26.
- Leimu, K. 2004. IEA-tutkimusyhteistyön taustasta ja luonteesta. Teoksessa K. Leimu (toim.) *Kansainväliset IEA-tutkimukset Suomi-kuvaa luomassa*. Jyväskylän yliopisto: Koulutuksen tutkimuslaitos, 43–76.
- Leino, J. 1985. Opetussuunnitelman toteutumisen esteet. Teoksessa E. Kangasniemi (toim.) *Opetussuunnitelma ja sen toimeenpano*. Jyväskylän yliopisto. Kasvatustieteiden tutkimuslaitos. *Selosteita ja tiedotteita* 268, 55–59.
- Linnakylä, P., Brunell, V. & Kankaanranta, M. 1996. Miten oppilaat kokevat koulun? Kouluelämän laatu ja sen kehittyminen 90-luvulla. Teoksessa R. Jakku-Sihvonen, A. Lindström & S. Lipsanen (toim.) *Toteuttaako peruskoulu tasa-arvoa?* Helsinki: Opetushallitus.
- Linnakylä, P. 1993. Miten oppilaat viihtyvät peruskoulun yläasteella? Teoksessa V. Brunell & P. Kupari (toim.) *Peruskoulun oppimisympäristönä*. Peruskoulun arviointi 90-tutkimuksen tuloksia. Kasvatustieteiden tutkimuslaitos: Jyväskylän yliopisto, 39–56.
- Linnakylä, P. 2002. Kansainvälisten ja kansallisten oppimistulosten arviointien välisestä suhteesta. Teoksessa E. Olkinuora, R. Jakku-Sihvonen & E. Mattila (toim.) *Koulutuksen arviointi: lähtökohtia, malleja ja tilannekatsauksia*. Turun yliopiston Kasvatustieteiden tiedekunnan julkaisuja B, 31–52. Länsi-Suomen Läänihallitus 2002. Kansakoulujen tarkastus kevätlukukaudella 1950. Moniste. Mietintö (1973). Oppimateriaalikomitean mietintö. Helsinki: VAPK.
- Mikkilä-Erdman, M., Olkinuora, E. & Mattila, E. 1999. Muuttuneet käsitykset oppimisesta ja opettamisesta – haaste oppikirjoille. *Kasvatus* 30 (5), 436–449.
- Niemi, E. 2004. Perusopetuksen oppimistulosten kansallinen arviointi ja tulosten hyödyntäminen koulutuspoliittisessa kontekstissa. Turun yliopiston julkaisuja 216.
- Niemi, H. 2000. Opettajankoulutuksen vaikuttavuus. Teoksessa R. Raivola (toim.) *Vaikuttavuutta koulutukseen*. Suomen Akatemian julkaisuja 2/00. Helsinki.
- Nikkanen, P. & Lyytinen, H.K. 1996. Oppiva koulu ja itsearviointi. Jyväskylän yliopisto: Koulutuksen tutkimuslaitos.
- Norris, N., Aspland, R., MacDonald, B., Schostak, J. & Zamorski, B. 1996. Arviointiraportti peruskoulun opetussuunnitelmaudistuksesta. Opetushallitus. Arviointi II/1996. Helsinki.
- Opetushallitus 1994. Peruskoulun opetussuunnitelman perusteet 1994. Helsinki.
- Opetushallitus 1998a. Kansallinen oppimistulosten arviointijärjestelmä. Arviointi 4/1998. Helsinki.
- Opetushallitus 1998b. Koulutuksen tuloksellisuuden arviointimalli. Arviointi 7/1998. Helsinki.
- Opetushallitus 2004. Perusopetuksen opetussuunnitelman perusteet 2004. Helsinki.
- Opetusministeriö 2004. Korkeakoulutuksen laadunvarmistus. Opetusministeriön työryhmämuistioita ja selvityksiä 2004:6. Koulutus- ja tiedepolitiikan osasto.
- Perusopetuslaki 21.8.1998/628 sekä perusopetuslain muutos (24.1.2003/32).

- Tynjälä, P. 1993. Oppiiko oppilas peruskoulussa? *Kasvatus* 24 (2), 194–196.
- Törnroos, J. 2004. Opetussuunnitelma, oppikirjat ja oppimistulokset. Seitsemännenn luokan matematiikan osaaminen arvioitavana. Jyväskylän yliopisto. Koulutuksen tutkimuslaitos. Tutkimuksia 13.
- Uusikylä, K. & Atjonen, P. 2000. Didaktiikan perusteet. Helsinki: WSOY.
- Viiri, J. 2003. Oppikirjojen asiasisällön tutkimuksesta; esimerkkinä vuorovesi-ilmiö. *Kasvatus* 34 (2), 139–149.
- Vitikka, E. 2004. Sähköpostitieto 18.10.2004. Opetushallitus.
- Vuori, J. 2004. Arviointi(tutkimus) on arvotonta ellei se arvota arviointiaan. Teoksessa: Hallinnon Tutkimus 23 (2), 1–3.
- Väljijärvi, J. 2003. PISA-tutkimus koulutuksen arvioinnissa. Teoksessa J. Väljijärvi & P. Linnakylä (toim.) Tulevaisuuden osaajat. PISA 2000 Suomessa, 1–8.



Oppimistulosten arviointi laadunvarmistuksena

Ammatillinen koulutus

Oppimistuloksia on ammatillisessa koulutuksessa arvioitu vuodesta 1995 lähtien. Arvioinnit on organisoitu alakohtaisina tai kaikkia aloja koskevin kansallisina kokeina. Yhtenäisistä kansallisista kokeista kuitenkin luovutaan ammatillisiin perustutkintoihin liittyvissä ammatillisissa opinnoissa ja sen sijaan siirytään näyttöihin perustuvaan arviointitapaan. Kaikille yhteisiin opintoihin (esimerkiksi äidinkieli, vieraat kielet, matematiikka) liittyvistä oppimistulosten arvioinneista ei ole tehty päätöksiä.

Artikkelissa tarkastellaan ammatillisen koulutuksen oppimistulosten arviointijärjestelmää ja sen perusteita. Painopiste on siinä, miten näyttöihin perustuvassa arvioinnissa ulkoisen arvioinnin edellyttämä arviointitieto tuotetaan opiskelijan arvioinnin yhteydessä ja mitä laadunvarmistuksella ja -kontrollilla tässä yhteydessä tarkoitetaan. Lisäksi nostetaan esille oppimistulosten arvioinnin kehittämisen kriittisiä kohtia ja arvioinnin organisoinnin haasteita.

Asiasanat: ammatillinen koulutus, oppimistulosten arviointi, näytöt

I Aiempi päättökokeisiin perustuva oppimistulosten arviointijärjestelmä

Ammatillisen koulutuksen oppimistulosten arviointi aloitettiin Opetushallituksen toimesta vuonna 1995 (Koulutuksen tuloksellisuuden...1995, Kansallinen oppimistulosten...1998). Arvioinnit perustuivat kansallisiin kokeisiin ja ne kohdistuivat **kaikille yhteisiin opintoihin, ammatillisiin opintoihin ja yleiskvalifikaatioihin. Kaikille yhteisten opintojen** arviointi kohdistui yleissivistykseen, muuttuvissa oloissa selviytymiseen, ammattitaidon kehittämiseen ja jatko-opintojen ja elämänhallinnan valmiuksiin. Kaikille yhteisten opintojen arviointi tehtiin oppiaineperustaisesti, jolloin eri alojen oppimistuloksia verrattiin toisiinsa. Arviointeja tehtiin matematiikassa, äidinkielessä, toisessa kotimaisessa kielessä, englannin kielessä ja fysiikassa ja kemiassa. Sen lisäksi kaikille yhteisiin opintoihin liittyvää osaamista arvioitiin alakohtaisten arviointien yhteydessä. Arviointiohjelmassa tavoitteeksi asetettiin yhden aineen arviointi kunakin vuonna.

Yleiskvalifikaatioista on arvioitu vain oppimaan oppimista (Hautamäki ym. 2002). Arviointiohjelman tavoitteiden mukaan arviointeja oli tarkoitus laajentaa myös muihin osaamisalueisiin (*ongelmanratkaisu- ja päätöksentekotaidot, arvo-osaaminen sekä vuorovaikutus- ja ryhmätyötaidot*), mutta voimavarojen vähäisyyden vuoksi arvioinnit rajautuivat vain oppimaan oppimiseen. Tavoitteena oli arvioida yksi osaamiskokonaisuus joka toinen vuosi. Kaikille yhteisten opintojen tapaan yleiskvalifikaatioiden arviointia liitettiin myös alakohtaisiin arviointeihin.

Ammatillista osaamista arvioitiin alakohtaisten arviointien osana. Arviointeja varten laadittiin kansalliset päättökokeet arviointiperusteineen ja -kriteereineen. Kokeet kestivät 3–5 päivää, ja niissä oli sekä kirjalliset että ammattityöhön liittyvät osuudet. Arvioinnit kohdistuivat työprosessin hallintaan (*työn tavoitteiden asettaminen, suunnittelutaidot, ongelmanratkaisu- ja päätöksentekotaidot, työmenetelmien hallinta, materiaalien käyttö ja valinta sekä joustava ja joutuisa toiminta ja toiminnan arviointi*). Ammatilliseen osaamiseen katsottiin kuuluvan myös kommunikaatio-, ryhmätyö- ja tiimityövalmiuksia, työturvallisuuden ja taloudellisuuden vaatimaa osaamista sekä luovuutta ja innovatiivisuutta.

Teoreettisen ja käsitteellisen perustan sekä käytännön taitojen edellytettiin integroituvan kokeissa toisiinsa. Tämä ilmeni toiminnan sujuvuutena, kykynä ratkaista ongelmia ja tehdä päätöksiä työtilanteissa. Osaamiseen nähtiin liittyvän myös kyky kriittisyyteen, oman osaamisen arviointiin ja uusien toimintakäytänteiden kehittämiseen sekä äänetön tieto ja oivallus.

Arviointiohjelman mukaan ammatillista osaamista tuli arvioida tutkinnoittain keran viidessä vuodessa (Kansallinen oppimistulosten...1998). Ohjelmaa ei voitu käytännössä toteuttaa määrärahojen vähäisyyden vuoksi. Osalla aloista ja suurimmassa osassa tutkinnoista ammatillista osaamista ei arvioitu lainkaan. Ammatilliseen osaamiseen liittyviä oppimistuloksia arvioitiin metsälalla, kone-, metalli- ja sähköalalla,

sosiaali- ja terveysalalla, kaupan ja hallinnon alalla, käsi- ja taideteollisuusalalla ja matkailu- ravitsemis- ja talousalalla (Kekkonen 1996; Vuorenmaa & Räisänen 1997; Räisänen 1999; Hollo ym. 2001; Nuotio ym. 2001; Koski ym. 2002). Arviointitulokset julkistettiin kansallisella tasolla joko erillisissä raporteissa tai alakohtaisten arviointiraporttien osana. Koulutuksen järjestäjille ja oppilaitoksille toimitettiin heitä itseään koskevat tulokset sekä kansalliset vertailutulokset kuukauden kuluttua päättökokeiden järjestämisestä. Kansallisessa raportissa tulokset kuvattiin ja analysoitiin yleisellä tasolla. Oppilaitoskohtaisia tuloksia ei julkistettu eikä järjestystä osoittavia luetteloita laadittu.

Arviointien vähäisyydestä johtuen oppimistulosten arviointijärjestelmän suurin ongelma oli vähäinen vaikuttavuus. Oppimistulosten kehityksestä ei arviointitietoa tuotettu lainkaan. Arviointien otantaperustaisuudesta johtuen osa koulutuksen järjestäjistä ei koskaan osallistunut arviointiin eikä siten saanut yksikkökohtaista palautetta edes kertaluonteisesti, elleivät ne hankkineet kokeita ostopalveluna. Vaikka arviointiin osallistuneet oppilaitokset olivat tyytyväisiä saadessaan niitä itseään koskevat tulokset ja kansalliset vertailutulokset, arvioinnit eivät satunnaisuudesta johtuen integroituneet kovin hyvin myöskään järjestäjien ja oppilaitosten itsearviointijärjestelmien osaksi.

II Uusi näyttöperusteinen oppimistulosten arviointijärjestelmä

Ammatillisiin perustutkintoihin liitettäviä näyttöjä kehitetään ammatillisen koulutuksen laadun ja työelämäyhteyksien varmistamiseksi. Oppimistulosten arvioinnin kannalta olennaista on, ettei uudessa järjestelmässä järjestetä erillisiä kansallisia kokeita, vaan arviointitieto tuotetaan näytöistä. Oppimistulosten arvioinnit integroidaan näyttöihin laadunvarmistuksen ja laatu kontrollin avulla.

Näyttöjärjestelmän vakiinnuttamista koskevat poliittiset linjaukset on tehty Valtioneuvoston vahvistamassa koulutuksen ja tutkimuksen kehittämissuunnitelmassa (2003) ja pääministeri Vanhasen hallitusohjelmassa (2004). Opetusministeriö on tehnyt myös päätöksen, joka koskee oppimistulosten arvioinnin liittämistä näyttöihin (212/430/98). Alustavien suunnitelmien mukaan uuteen näyttöjärjestelmään siirrytään vuonna 2006.

Mistä näytöissä on kyse?

Näytöillä tarkoitetaan koulutuksen järjestäjän ja työelämän yhdessä suunnittelemia, toteuttamia ja arvioimia työtilanteita tai työprosesseja. Näytöissä opiskelijat osoittavat, miten he ovat saavuttaneet opetussuunnitelmien perusteissa osoitetun työelämän edellyttämän ammattitaidon. Näytöissä osoitetaan keskeinen osaaminen.



Opetusministeriön kanta on ollut, että opetussuunnitelmien perusteiden lisäksi näyttöjä ohjataan näytön perusteilla, joissa määritetään näytettävä osaaminen, arvioinnin kohteet ja arviointikriteerit (Hallituksen esitys... 2004). Lisäksi ministeriö esittää, että näytön perusteissa voitaisiin antaa näytön suunnittelua, toteuttamista, arviointia, arvioinnin dokumentointia ja näyttöympäristöä koskevia ohjeita.

Ammatillisiin perustutkintoihin liittyvät näytöt eroavat aikuiskoulutuksen näyttötutkinnoista. Aikuisten näyttötutkinnoissa arvioidaan työtehtävissä tarvittavaa ammatitaitoa kiinnittämättä huomio siihen, miten osaaminen on hankittu. Ammatillisten perustutkintojen näytöt sijoittuvat koko koulutuksen ajalle, ja ne liitetään kaikkiin ammatillisiin opintokokonaisuuksiin. Ammatillisten perustutkintojen näyttöjen arviointi antaa mahdollisuuden kiinnittää huomion myös oppimista ja osaamista tuottaviin prosesseihin ja taustatekijöihin.

Näyttöjen avulla pyritään yhtenäistämään opiskelija-arviointia ja lisäämään työelämän tahojen osallisuutta arvioinnissa. Työelämäyhteistyötä lisätään muun muassa korostamalla työelämän ja koulutuksen järjestäjien yhteisvastuuta, vastuuttamalla paikallinen elin näyttöjen organisointiin ja korostamalla näyttöjen liittämistä työssäoppimisen osaksi. Näyttöjä voidaan kuitenkin antaa myös muilla työpaikoilla ja oppilaitoksissa.

Yksittäinen näyttö voi kattaa yhden tai useamman opintokokonaisuuden tai niiden osan. Näytöt voivat kohdistua myös ammatillisiin opintoihin liittyviin vapaasti valittaviin opintoihin. Yhteisiä opintoja ja muita vapaasti valittavia opintoja näytöt eivät koske. Näyttöjen kokonaismäärä vaihtelee tutkinnoittain. Näyttöjä on arvioitu olevan tutkinnoittain kolmesta neljään.

Miksi oppimistulosten arviointi liitetään näyttöihin?

Uusi oppimistulosten arviointijärjestelmä pyrkii ratkaisemaan aiemmassa järjestelmässä olleita arviointien vaikuttavuuteen ja kustannuksiin liittyviä ongelmia. Tarkoituksena on kehittää arviointijärjestelmä, joka on nivottu koulutuksen järjestäjien toiminnan osaksi ja joka tuottaa säännöllistä palautetietoa koulutuksen laadun parantamiseksi sekä paikallisella että kansallisella tasolla. Myös kustannusten on oltava realistiset.

Suurimmat odotukset liittyvät arvioinnin **vaikuttavuuden ja kehittävyys** lisäämiseen. Oppimistulosten arvioinnin **tavoitteena** on tukea opiskelijoiden oppimista, opetustoimen henkilöstön työtä, valtakunnallista ja paikallista päätöksentekoa ja koulutuksen kehittämistä oppimisen edellytysten turvaamiseksi (A150/2003). Oppimistulosten arvioinnin avulla voidaan myös varmistaa koulutuksen laatu erityisesti suhteessa työelämän odotuksiin ja vaatimuksiin.

Oppimistulosten arvioinnissa tuotetaan tietoa siitä, miten opetussuunnitelman perusteiden tavoitteet on saavutettu. Keskeistä on, että arviointiprosessi ja -tulokset **motivoivat** niin opettajia kuin opiskelijoitakin ja **kehittävät** opetusta, ohjausta ja koko

oppilaitoksen toimintakulttuuria ja kriittisyyteen perustuvaa arviointikulttuuria. Jotta arviointi olisi itsessään myös vaikuttavaa, sen tulee olla osuvaa, säännöllistä, ajantasaista ja luotettavaa.

Oppimistulosten arvioinnin ja näyttöperusteisen opiskelijan arvioinnin integroinnin tavoitteena on, että uudesta arviointijärjestelmästä muodostuisi ajan oloon näyttöjärjestelmän laatua jatkuvasti **kehittävä ja parantava mekanismi – todellinen laadunvarmennus**. Tätä arvioinnin kehittävää tehtävää voidaan tukea arvioinnin systemaattisuudella, osallistamalla opettajia ja kansallisia toimijoita arviointiprosessiin (vrt. moderation) ja kehittämällä näyttöperusteisen arvioinnin laadunvarmennusta.

Oppimistulosten arvioinnin integrointi näyttöihin **lisää myös arvioinnin tehokkuutta ja taloudellisuutta**, sillä uudessa järjestelmässä ei ole erillisiä kansallisia testejä. Koska arviointitieto kootaan jo muutoin muodostuvista aineistoista, määrärahoja ei tarvita erillisten kokeiden laadintaan ja niiden testaamiseen. On kuitenkin huomattava, että laadunvarmistus- ja laatuvalvontajärjestelmä edellyttää vastaavasti taloudellisia voimavaroja, sillä tiedetään, että oppimistulosten arvioinnin integrointi näyttöihin sisältää melko suuria luotettavuuteen liittyviä ongelmia, etenkin käyttöönoton alussa. Tämä voimavaratarve on erillistä koekäytäntöä vähäisempi. Laadunhallintaan suunnatut voimavarat tukevat myös näyttöjen käyttöönottoa ja yhdenmukaistavat opiskelija-arviointia jo järjestelmän kehittämisvaiheessa. Edelleen osaamisalueiden, arviointikriteereiden ja dokumentoinnin yhtenevyys kaikilla koulutusaloilla, kaikissa tutkinnoissa ja kaikissa näytöissä parantaa osaltaan taloudellisuutta ja tehokkuutta.

Oppimistulosten arvioinnin luonne muuttuu opiskelijan arvioinnin ja ulkoisen arvioinnin integroitessa toisiinsa. Koulutuksen päättövaiheeseen liittyvä oppimistulosten arviointi painotti tulosten arviointia ja oli luonteeltaan summatiivista. Koko koulutuksen ajalle ulottuva arviointi muuttaa arvioinnin luonnetta formatiivisen arvioinnin suuntaan. Arvioinnin muuttuminen prosessiarvioinniksi **antaa onnistuessaan mahdollisuuden muun muassa oppimistulosten kehityksen ja osaamisen kontekstista toiseen tapahtuvan siirtovaikutuksen (transfer) arviointiin (vrt. Räisänen & Vuorenmaa 2003)**.

Arvioinnin prosessuaalisuudesta johtuen arvioinneissa voidaan tehdä päätelmiä opiskelijoiden ammattitaidon kehittymisestä ja osaamisen muutoksista, kun siirrytään lukuvuodesta ja oppimisympäristöstä toiseen (kumulatiivisuus). Tämä tosin edellyttää, että opetussuunnitelmissa ja opetus- ja arviointikäytänteissä otetaan oppimisen siirtovaikutus huomioon. Lisäksi voidaan tehdä päätelmiä eri alojen välisistä eroista. Uusi prosessiin sidottu arviointi mahdollistaa opintojen loppuun sijoittuvaa arviointia paremmin oppimisen ja sen edellytysten sekä oppimista säätelevien kontekstitekiöiden arvioinnin. Annettujen suositusten mukaisesti uusi järjestelmä mahdollistaa oppijan osallistumisen subjektina oman oppimisensa arviointiin.

Jos arviointijärjestelmään kehitetään yhtenäinen dokumentointi ja tietojen hallinnan verkkoperustainen palautejärjestelmä, niin se mahdollistaa **ajantasaisen tietojen koonnin, analysoinnin ja nopean ja vaivattoman palautteen saannin järjestäjille**

ja oppilaitoksille. Arvioinnin vaikuttavuuden kannalta on keskeistä, että koulutuksen järjestäjillä ja oppilaitoksilla on mahdollisuus saada tietoa omien oppimistulosten kehityksestä sekä mahdollisuus vertailla omia ala- ja tutkintokohtaisia tuloksia kansallisiin tai vertaisia kuvaaviin tuloksiin. Näin oppimistuloksia koskevaa vertailutietoa voitaisiin hyödyntää myös itsearvioinnissa ja oppimisen laadun jatkuvassa parantamisessa. Sama arviointitieto voidaan toimittaa myös vuosiraportteina tai muina raporteina kansallisesti. Tämä taas parantaa oppimistulosten huomioon ottamista päätöksenteossa, joka koskee tavoitteita ja kehittämishankkeita.

Uusi oppimistulosten arviointijärjestelmä helpottaa **arviointitutkimuksen integrointia oppimistulosten arviointiin.** Järjestelmä tuottaa itsessään paljon tutkimusaiheita ja -aineistoja erilaisten väitöskirjojen ja muiden tutkimusten tarpeisiin. Lisäksi yhtenäinen tietojärjestelmä ja dokumentointi mahdollistaa eri aloja ja tutkintoja yhdistävät ja vertailevat tutkimukset. Arviointitutkimuksen liittäminen oppimistulosten arviointiin voisi käytännössä merkitä esimerkiksi yhden teeman ottamista syväanalyysin kohteeksi vuosittain. Syväanalyysiin keskittyvän arviointitutkimuksen avulla voidaan tuottaa tietoa erityisesti oppimistulosten kehityksestä ja oppimistuloksiin liittyvistä eroista ja niiden syistä. Keskeistä on analysoida muiden selitystekijöiden lisäksi niin opiskelijoiden kuin oppimisympäristön kehystekijöiden yhteyttä oppimistuloksiin. Tällöin myös opetussuunnitelmaa, opetusta ja erilaisia oppimisympäristöjä koskevat arvioinnit sidostuvat oppimistuloksiin.

Huolimatta siitä, että suomalainen oppimistulosten arviointijärjestelmä poikkeaa eri maiden arviointijärjestelmistä, **se voi toimia hyvänä esimerkkinä kansainvälisissä yhteyksissä.** Vaikka eri maiden oppimistuloksia ei voitane suoraan verrata toisiinsa erilaisista arviointijärjestelmistä ja mittareista johtuvien syiden vuoksi, arviointijärjestelmien vertailujen avulla voidaan saada merkittävää apua järjestelmän kehittämiseksi.

Mitä arvioidaan?

Oppimistulosten arvioinnin perustana on opetussuunnitelmien perusteissa määritetty ammattikäsitys. Myös osaamista koskevat tulkinnat on sidottu yleisiin tavoitteisiin ja opintokokonaisuuksien ydintavoitteisiin ja osaamisalueisiin. Arvioinnin kohteet ovat kaikilla aloilla samat.

Oppimistulosten arviointi kohdistuu ammatilliseen ydinosaamiseen ja erityisosaamiseen sekä yhteisiin painotuksiin. Arviointi kohdistuu toiminnalliseen, tiedolliseen, sosiaaliseen ja reflektiiviseen osaamiseen. Arvioinnin kohteena on sekä osaaminen että osaamista tuottavat prosessit. **Toiminnallinen eli operatiivinen osaaminen** ilmentää työtehtävien, työtoiminnan, työmenetelmien sekä työprosessien hallintaa. **Kognitiiviset eli tiedolliset prosessit** osoittavat teoreettisen tiedon hallintaa ja kykyä soveltaa sitä, asioiden ja prosessien välisten yhteyksien ymmärtämistä sekä työn kokonaisuuden hahmottamista. **Sosiaaliset eli yhteisölliset prosessit** liittyvät kykyyn toimia sekä

itsenäisesti että vuorovaikutuksessa muiden kanssa ryhmän jäsenenä tai johtajana. **Reflektiiviset eli arvioivat ja kehittävät prosessit** liittyvät kykyyn oppia ja kykyyn arvioida ja kehittää sekä omaa että ympäristön toimintaa. Arvioinnissa seurataan myös **oppimisen siirtovaikutusta** eli tiedon ja osaamisen välittymistä tehtävästä, tilanteesta ja ympäristöstä toiseen koulutuksen aikana. (Vrt. Poikela 1998; Räisänen 1998; Kankaanpää 1997; Nuotio ym. 2001, Rökköläinen & Ecclestone 2005).

Edellä esitetyt osaamisen ulottuvuudet saavat sisältönsä kunkin perustutkinnon opetussuunnitelman perusteiden yleisistä tavoitteista sekä opintokokonaisuuksien keskeisistä tavoitteista ja osaamisalueista. Arviointitieto kootaan kaikista osaamisalueista, mikä antaa opiskelijalle mahdollisuuden osoittaa laaja-alaista ammatillista perusosaamista (ydinosaaminen), koulutusohjelmaopintoihin liittyvää erityisosaamista ja kaikilla aloilla tarvittavaa ammattisivistystä ja kansalaisvalmiuksia (yhteiset painotukset).

Kansalliset näyttöaineistot ohjaavat myös oppimistulosten arviointia. Näytöissä opiskelijan tulee yhdistää käytäntö- ja teorian tietoa ja soveltaa niitä työelämän toimintakokonaisuuksiin. Kansallisessa näyttöaineistossa määritetään perustutkinnossa arvioitava osaaminen opintokokonaisuuksittain yleisten arviointikohteiden ja kriteereiden pohjalta.

Oppimistulosten arvioinnin ja näyttöjen integrointi käytännössä

Oppimistulosten arviointijärjestelmä perustuu systemaattiseen laadunhallintaan, ja sen eri vaiheet ilmenevät kuviossa 1. Nämä oppimistulosten arviointiprosessin vaiheet kuvataan seuraavassa päälukittain. Näitä on testattu oppimistulosten arviointikokeilussa Opetushallituksessa (Rökköläinen & Ecclestone 2005).

1 Laatustandardit (quality standards)

- Monitahoinen asiantuntijaryhmä laatii eri alojen oppimistulosten arvioinnin perustaksi yhtenäiset laatustandardit, joissa määritetään arviointikohteet (osaamisalueet), kriteerit ja arviointiympäristöä ja arvioijia koskevat laatuvaatimukset.
- Asiantuntijaryhmä määrittää laatustandardeissa arvioitavan osaamisen **keskeisen sisällön** (ammatillinen osaaminen, ydinosaaminen, yhteiset painotukset) opintokokonaisuuksittain opetussuunnitelmien perusteissa määritettyjen tavoitteiden ja kansallisten näyttöaineistojen pohjalta. Kaikissa näytöissä arvioidaan työtehtävien (työmenetelmät, työvälineet ja työmateriaali), työprosessin sekä työn perustana olevan tiedon hallintaa ja työturvallisuuden hallintaa. Yhteinen ydinosaaminen (oppimaan oppimisen taidot, ongelmanratkaisutaidot, vuorovaikutus- ja viestintätaidot, yhteistyötaidot, eettiset ja esteettiset/emotionaali-



Standardit Quality standards	Laadunvarmistus Quality assurance	Arviointitiedon tuottaminen	Laatukontrolli Quality Control	Konsensus- ja vertaisarviointi - moderation	Itsearviointi ja ulkoinen arviointi
<ul style="list-style-type: none"> • ops-vaatimukset • osaaminen • mittarit • kriteerit • arvioijat 	<ul style="list-style-type: none"> • opetussuunnitelmat • näyttöaineisto • kertaluonteisuus • laatuleima • etukäteen 	<ul style="list-style-type: none"> • konsensus • ajankohta vapaa • kumulatiivisuus • dokumentointi • taustatiedot 	<ul style="list-style-type: none"> • rinnakkaisarviointi • arviointikäynnit • arviointitilanteet - ympäristö, - tehtävät, - kriteerit ja arvioijat • satunnainen • jälkikäteen 	<ul style="list-style-type: none"> sisäinen: <ul style="list-style-type: none"> • opettaja, opiskelija, työelämä • oppilaitokset verkostoina ulkoinen: <ul style="list-style-type: none"> • tulkinnallinen analyysi • kehittävä 	<ul style="list-style-type: none"> • itsearviointi • ulkoinen arviointi Kohteet: <ul style="list-style-type: none"> • oppimistulos • näyttöjärjestelmän toimivuus • opiskelijaarviointi

Kuvio 1. *Oppimistulosten arvioinnin vaiheet*

set taidot) ja yhteiset painotukset (kansainvälisyys, kestävä kehitys, teknologian ja tietotekniikan hyödyntäminen, yrittäjäyys, laadukas ja asiakaslähtöinen toiminta, palvelu- ja kuluttajaosaaminen, työsuojelusta ja terveydestä huolehtiminen) arvioidaan niissä opintokokonaisuuksissa, joiden tavoitteisiin ne liittyvät. Osaamisalueet tarkentuvat oppilaitosten omissa aineistoissa, erityisesti tehtävissä.

- Oppimistulosten arviointi on kriteeriperusteista. Kriteerit ovat kaikilla aloilla samat. Kriteeristö on kolmiportainen ja osaaminen kuvataan tasoilla tyydyttävä-hyvä-kiitettävä. Arviointiasteikko on 1–5. Tyydyttävä taso kuvaa ammatillisen osaamisen vähimmäisvaatimusta, josta seuraavien hyvän ja kiitettävän osaamisen tasojen vaatimukset johdetaan. Osaamisvaatimukset kumuloituvat tasolta toiselle siirryttäessä. Kansallisella tasolla laaditut yhtenäiset kriteerit tarkentuvat eri konteksteissa tehtävässä arvioinnissa oppilaitoksissa. Arvioinnissa opiskelijoita ei verrata toisiinsa eikä heitä aseteta järjestykseen (ei ranking-listoja).
- Koulutuksen järjestäjät tuottavat arviointitiedon organisoimistaan näytöistä opintokokonaisuuksittain koko koulutusprosessin (3 vuotta) ajalta. Näytöt organisoidaan siten, että kaikki osaamisen ulottuvuudet – toiminnallinen, tiedollinen, sosiaalinen ja reflektiivinen kyvykkyys – on mahdollista arvioida. Näyttöjen tulee olla riittävän kattavia, jotta opiskelijalla on mahdollista osoittaa laaja-alaista ammatillista perusosaamista ja suuntautumisopintoihin liittyvää erityisosaamista sekä yleisosaamista. Näytöissä opiskelija yhdistää käytäntö- ja teoretietoa ja soveltaa tietoa työelämän toimintakokonaisuuksiin.

Osanäytöt muodostavat yhdessä koko opintokokonaisuutta koskevan kokonaisuuden.

- Arvioijien arviointiosaaminen, erityisesti kriteereiden tulkinta, on oppimistulosten arvioinnin luotettavuuden kriittinen kohta. Tämä edellyttää arviointiosaamiseen liittyvien kehittämistarpeiden analysointia, koulutuksen suunnittelua ja järjestämistä sekä arviointiosaamisen kehittymistä tukevien näyttöaineistojen laadintaa. Vastuu opettajien ja työelämän arvioijien koulutuksen järjestämisestä on koulutuksen järjestäjillä. Oppilaitokset perehdyttävät myös opiskelijat arviointiin.

2 Laadunvarmistus (external quality assurance, internal quality assurance)

- Ulkopuoliset, riippumattomat asiantuntijat varmistavat kirjallisten aineistojen laadun etukäteen. Oppilaitokset varmistavat arviointiprosessin laadun tarkoituksenmukaiseksi katsomallaan tavalla. **Laadunvarmistus** kohdistuu prosessin kriittisiin kohtiin kuten opetussuunnitelmiin, näyttöaineistoihin ja arviointisijoiden osaamisen kehittämiseen. Tehtävien laatua ei ulkoisesti varmisteta, vaan niiden laadusta vastaavat opettajat. Tehtävät laaditaan ulkoisesti varmennettujen näyttöaineistojen pohjalta. Oppimistulosten arvioinnin edellyttämät mahdolliset korjaustoimenpiteet tehdään ennen arviointia.
- Laadunvarmistus on kertaluonteista, mistä syystä laadunvarmennettuja aineistoja (vrt. laatuleima) ja opetussuunnitelmia voidaan käyttää oppimistulosten arvioinnissa niin kauan kuin järjestelmä ei muutu tai kun aineistoja halutaan muuttaa sisällöllisistä tai muista syistä.

3 Oppimistulosten arviointitiedon tuottaminen

- Näytöt, joista oppimistulosten arviointitieto kootaan, arvioidaan konsensusperiaatteella. Arviointiin osallistuvat opettaja, työelämän edustaja ja opiskelija.
- Järjestäjät organisoivat näytöt parhaaksi katsomanaan ajankohtana, mikä lisää alueellista ja paikallista joustavuutta.
- Oppimistulosten arviointitiedon kokoaminen koko opiskeluajalta mahdollistaa opiskelijoiden kehittymisen ja kontekstista toiseen tapahtuvien siirtovaikutusten (transfer) arvioinnin. Eri vaiheissa koottuja tuloksia ei lasketa mekaanisesti yhteen eikä niitä painoteta laskennallisesti.



- Oppimistulosten arviointi mahdollistuu em. tavalla yhtenäisen dokumentoinnin avulla. Tietoverkkopohjainen järjestelmä mahdollistaa tulevaisuudessa siirtymisen paperittomaan arviointiympäristöön (arviointiportaali), jossa arviointitieto kulkee joustavasti molempiin suuntiin oppilaitosten ja kansallisten arvioijien välillä. Tämä mahdollistaa tulevaisuudessa myös sen, että järjestäjät ja oppilaitokset saavat omat tuloksensa ja vertailutulokset viipymättä itsearviointinsa perustaksi.
- Kontekstitekijöiden ja oppimistulosten yhteyksien arvioimiseksi kootaan sekä oppijaa, oppimisympäristöä ja näytön ajankohtaa (esim. lukuvuosi) koskevia taustatietoja kansallisella tasolla.

4 Laatukontrolli (quality control)

- Satunnaisen ulkoisen *laatukontrollin (quality control)* avulla lisätään erityisesti arvioinnin läpinäkyvyyttä ja mahdollistetaan arvioinnin luotettavuuden arviointi. Laatukontrollin avulla voidaan tuottaa tietoa tehtävistä, arviointiympäristöstä, arvioinnin organisoinnista ja arvioitsijoiden toiminnasta, kuten arviointiperusteiden ja kriteereiden tulkinnasta. Laatukontrolli voi sisältää myös rinnakkaisarvioiteja. Laatukontrollin tekevät ulkopuoliset riippumattomat asiantuntijat arviointikäyntien avulla. Laatukontrolleja voidaan suunnata erilaisiin oppilaitoksiin, kuten esim. oppimistulosten arvioinneissa hyvin ja huonosti menestyviin kouluihin.
- Laatukontrolleja voidaan suunnata erilaisiin oppilaitoksiin, esimerkiksi oppimistulosten arvioinneissa hyvin ja huonosti menestyviin kouluihin.

5 Moderation-metodi toisilta oppimisen ja jatkuvan reflektoinnin tukena

- Reflektiivistä toimintakulttuuria, toisilta oppimista ja järjestelmän laadun parantamista voidaan tukea oppilaitoskohtaisten (sisäinen) ja alueellisten konsensus- ja vertaisarviointien avulla (ulkoinen). Apuna voidaan käyttää moderation-menetelmää (esim. Radnor & Shaw 1996). Tämä metodi edellyttää systemaattisesti järjestettyjä ja ohjattuja, sekä järjestäjä- ja oppilaitoskohtaisia että usean oppilaitoksen yhteisiä istuntoja (esim. kerran tai kaksi kertaa vuodessa). Näissä istunnoissa opettajat tulkitsevat vertaisten ja mahdollisesti ulkopuolisten asiantuntijoiden kanssa yhdessä oppimistuloksiaan ja arviointiprosessejaan oppiakseen toisiltaan ja lähentääkseen oppimistulosten arviointiin liittyviä tulkintojaan.

6 Itsearviointi ja ulkoinen arviointi (self evaluation, external evaluation)

- Koulutuksen järjestäjät ja oppilaitokset voivat käyttää arviointitietoa vertailutietoineen itsearvioinnissa. Arviointitiedon tuottamisesta riippuen järjestäjät ja oppilaitokset voivat seurata tulostensa kehitystä ja verrata niitä toisten oppilaitosten tuloksiin ja valtakunnallisiin tuloksiin.
- Oppimistulosten arviointituloksista voidaan muodostaa indikaattoreita. Indikaattoritieto ja muu määrällinen tieto on mahdollista julkaista tiiviinä vuosiraporttina.
- Arviointiprosessin ja -tulosten läpinäkyvyys edellyttää arvioinnin monitahoisuutta. Oppimistuloksia koskevat tulkinnat tehdään monitahoisuutta edustavien arviointiryhmien toimesta. Tulevaisuudessa kansallisella tasolla voitaisiin laatia yhdestä kahteen arviointiraporttia vuosittain, minkä lisäksi voitaisiin julkaista artikkeleita. Näin julkistettavat tulokset voisivat liittyä tiettyyn alaan, osaamisalueeseen tai alojen välisiin eroihin. Arviointiraportteihin voitaisiin sisällyttää arviointiryhmän tulkintoja hyvin ja huonosti menestyneiden oppilaitosten/järjestäjien toiminnasta ja oppimistulosten erojen taustoista. Tämä edellyttää, että arviointiryhmä voisi tehdä myös arviointivierailuja oppilaitoksiin (esim. 5 parasta ja 5 heikointa oppilaitosta). Sen lisäksi, että tämä tapa mahdollistaisi hyvien käytäntöjen esille nostamisen, se mahdollistaisi myös palautteiden ja kehittämissuosituksen antamisen niille, joiden tulokset ovat muita huonommat. Heikosti menestyvien koulutuksen järjestäjien tukemisessa voisi Opetushallituksen rooli kehittämisvirastona olla keskeinen.
- Koulutuksen järjestäjät ja oppilaitokset sisällyttävät opiskelijan arvioinnin ja näyttöjärjestelmän ja oppimistulostensa arvioinnin itsearviointiinsa ja ulkoiseen laatuauditointiin.

III Oppimistulosten arviointikokeilu

Uutta arviointimallia on kehitetty ja osittain testattu Opetushallituksen toimesta valtakunnallisissa kokeiluissa yhteistyössä näyttöjärjestelmän kehittäjien, oppilaitosten, työnantaja- ja työntekijäjärjestöjen sekä tutkimuslaitosten ja kansainvälisten yhteistyökumppaneiden kanssa (Räkköläinen & Mäkihalvari 2003; Räkköläinen & Eccleston 2005). Euroopan sosiaalirahastosta rahoitettava hanke ajoittuu vuosille 2002–2006 (2007) ja siinä on kolme vaihetta: 1) *arviointimallin kehittäminen ja kokeilu*, 2) *arviointistrategia ja palautejärjestelmä* ja 3) *tulosten alakohtainen arviointi*. Arviointikokeilun ensimmäinen vaihe ajoittui vuosille 2002–2003 ja toinen ja kolmas vaihe

(2004–2007). Siirryttäessä ensimmäisestä vaiheesta toiseen vaiheeseen kokeilu laajeni koskemaan 14 tutkintoa.

Arviointimallin kehittämis- ja kokeiluvaiheessa on selvitetty ulkoisen arvioinnin ja näyttöjen yhdistämismahdollisuuksia, tehty vertailuja kansainvälisiin järjestelmiin, kehitetty ja testattu laadunvarmistusmekanismeja ja tuettu näyttöjen kehittämistä. Kokeiluun osallistui näyttöaineisto- ja pilottiprojektissa mukana olevia oppilaitoksia.

Vuosina 2002 ja 2003 testattiin arviointitiedon keruuta ja kehitettiin laadunvarmistuksen mekanismeja. Yhtenäisiä tehtäviä ei laadittu. Kokeilussa käytettiin kansallisia näyttöaineistoja, joissa on laadittu arvioitavan osaamisen kuvaukset sekä kolmiporaiset arviointikriteerit osaamisalueittain. Osaamisen arviointi kohdistui työtehtävien ja työprosessin hallintaan, ydinosaamiseen ja yhteisiin painotuksiin. Arvioinnin yleiset kohteet ja kriteerit ja tulosten dokumentointi olivat kaikilla aloilla samat. Arvioinnin tekivät opiskelijat, opettajat ja työelämän edustajat yhdessä.

Laadunvarmennus kohdistui koko prosessiin opetussuunnitelmista ja näyttöaineistoista näyttöjen suunnitteluun ja toteuttamiseen ja arviointiin. Laadunvarmennusprosessi sisälsi kirjallisen materiaalin arvioinnin, arviointikäynnit (vrt. laatukontrolli) oppilaitosten toteuttamiin näyttöihin sekä opettajien, työelämän arvioijien ja opiskelijoiden haastattelut. Näiden lisäksi järjestettiin vertaisarviointipalavereja (peer moderation) ja havainnoitiin näyttöjen arviointitilaisuuksia.

Arviointikokeiluun osallistui 190 sosiaali- ja terveystieteiden, talotekniikan ja rakennusalan perustutkintoihin opiskelevaa opiskelijaa seitsemästä oppilaitoksesta. Arviointitieto koottiin lukuvuoden aikana 376 näytöstä. Oppimistuloksia koskevan tiedon lisäksi opettajille, opiskelijoille ja työelämän edustajille tehtiin oppimistuloksia taustoittava kysely. Oppimistuloksista tuotettiin pikapalautteet ja oppilaitoskohtaiset raportit, minkä lisäksi laadittiin kansallinen raportti.

Kokeilun ensimmäisen vaiheen päätulosten mukaan:

- näytöistä saadut arvosanat olivat kaikilla aloilla erittäin hyvät ja arvosanojen väliset erot aloittain ja arviointikohteittain verrattuna olivat vähäiset
- kaikkien osapuolten mielestä näytöt ovat osuvia, oikeudenmukaisia ja kaikkia osapuolia motivoivia ja niiden tuottamaan arviointitietoon luotetaan
- noin puolet näytöistä toteutettiin työpaikoilla ja ne arvioitiin opettajien, työpaikkaohjaajien ja opiskelijoiden yhteisarvioiteina
- opiskelijat ja työelämän edustajat osallistuivat melko vähän näyttöjen suunnitteluun
- opiskelijat ja työelämän edustajat tunsivat puutteellisesti näytön arviointiperusteet
- opiskelijat arvioivat itseään vain satunnaisesti
- arviointikäytännöt vaihtelivat paljon aloittain, oppilaitoksittain ja saman oppilaitoksen sisälläkin eri tutkinnoissa
- kansallisen arvioinnin yhteensovittaminen opiskelijan arvioinnin kanssa aiheutti käytännön ongelmia ja päällekkäisyyttä ohjeissa ja dokumentoinnissa.

Kokeilu osoitti, että kansallinen arviointitieto voidaan koota näytöistä. Etenkin aidoissa työtilanteissa toteutettavat näytöt ovat osuvia, relevantteja ja autenttisia laaja-alaisen ammatillisen osaamisen arvioinnin välineitä (validiteetti). Näytöt tukevat oppilaitosten ja työpaikkojen välistä yhteistyötä, mahdollistavat yhteisarvioinnin ja tekevät arvioinnista tehokkaan kehittämisen välineen.

Ongelmana on näyttöjen ja niistä tehtävien tulkintojen yhdenmukaisuus ja vertailtavuus (reliabiliteetti). Näyttöjen toteuttaminen ja arviointikäytännöt vaihtelevat, arviointi ei ole johdonmukaista ja arvioijat tulkitsevat kriteereitä eri tavoin.

IV Oppimistulosten arvioinnin organisoinnin haasteet

Oppimistulosten arvioinnissa on kyse koulutuksen ulkoisesta arvioinnista, josta vastaavat opetusministeriö ja Koulutuksen arviointineuvosto yhdessä. Parlamentaarisen tahdon toteutuminen edellyttää, että arviointitoiminta organisoidaan arviointiasiantuntijoiden verkostona. Tämä on suuri haaste, jotta pystytään luomaan verkostomainen arviointikulttuuri myös oppimistulosten arviointiin. Opetushallituksen vastuu ja rooli kehittämisvirastona korostuu näyttöjärjestelmän ylläpidossa ja oppimistulosten perusteella tapahtuvassa kehittämisessä, minkä lisäksi sille on lainsäädännössä määritetty asema arviointiverkostossa.

Vaikka arviointijärjestelmän kehittämisessä on hyödynnetty kansainvälistä asiantuntemusta, kansainvälisen verkostoitumisen tarpeet ovat edelleen suuret. Laaja verkottuminen eri tahojen, erityisesti yliopistojen kanssa, mahdollistaisi myös arviointitutkimuksen integroinnin oppimistulosten arviointiin ja sen tarpeen ennakkoinnin nykyistä tehokkaammin. Arviointitiedon lisäksi myös arviointitutkimuksen tarve on suuri. Arviointitutkimuksen merkitys korostuu tulevaisuudessa, kun otetaan huomioon muun muassa oppimistuloksia selittävien oppimisympäristö- ja muiden kontekstitekijöiden arvioinnin vaatimukset.

Koulutuksen arviointineuvosto antoi (13.6.2004) lausunnon opetusministeriölle ammatillisen koulutuksen näyttöjärjestelmää koskevasta lakiehdotuksesta ja korosti siinä mm. kokeilujen merkitystä ja oppimistulosten arvioinnin organisointia laajan verkostopohjaisen toimintamallin mukaisesti. Lisäksi neuvosto teki opetusministeriölle erillisen aloitteen (20.4.2004) oppimistulosten arviointijaoston perustamisesta edistämään tämän arviointitoiminnan verkostoitumista. Oppimistulosten arviointijaoston tai muun vastaavan monitahoisuutta ja verkostomaisuutta edustavan toimintatavan avulla voitaisiin edelleen mallintaa oppimistulosten arviointijärjestelmää ja siihen liittyvää verkostoituvaa toimintakulttuuria. Monitahoinen verkosto voisi ensimmäisessä vaiheessa myös uusintaa ja selkeyttää arviointi- ja kehittämistoiminnan välistä yhteyttä, kehittää palaute- ja seurantajärjestelmiä (follow up), ratkoa arviointitiedon julkisuutta koskevaa problematiikkaa ja muita arvioinnin eettisiä kysymyksiä ja kehittää arviointimethodiikkaa.



V Uuden arviointijärjestelmän synnyttämää pohdintaa

Uusi ammatillisen koulutuksen oppimistulosten arviointijärjestelmä on ainutlaatuinen. Uuden järjestelmän keskeisin tunnuspiirre on, ettei yhtenäisiä kansallisia kokeita ole, vaan arviointitieto tuotetaan näytöistä. Välittävänä mekanismina on laadunvarmistus ja laatukontrolli. Uuteen järjestelmään liittyy uusia mahdollisuuksia, mutta myös ongelmia. Seuraavassa nostetaan tiivistetysti esille teemoittain **uuden järjestelmän vahvuuksia ja kriittisiä kohtia**.

- Uusi oppimistulosten arviointi sisältää **formatiivisen ja summatiivisen arvioinnin** elementtejä. Oppimistulosten arvioinnissa on summatiivisen arvioinnin vaatimus (tulostiedot) formatiivista arviointia suurempi, kun taas näytöissä korostuu myös formatiivisen arvioinnin merkitys. Tämä kaksisuuntainen vaatimus tuo jännitteitä arviointijärjestelmään, sen metodien ja kriteereiden kehittämiseen ja arvioinnin luotettavuuden varmistamiseen. Vain pitkäjänteinen kehittäminen, metodien testaaminen ja kokeilu voivat paljastaa, millä ehdoilla voidaan – samalla metodilla ylipäätään tuottaa tietoa sekä formatiivista että summatiivista arviointia varten.
- Uuteen oppimistulosten arviointimalliin sisältyy näkemys oppimis- ja toimintaympäristöstä toiseen tapahtuvan **siirtovaikutuksen (transfer)** ja oppimisen kumulatiivisuuden arvioinnista. Siirtovaikutuksen arviointi ei kuitenkaan ole mahdollista, jos opiskelijoiden osaaminen ja sen arviointi saa merkityksensä kunkin sisällöllisen kontekstin ja toimintaympäristön mukaan, jolloin kukin näyttö arvioidaan toisistaan riippumatta. Siirtovaikutuksen arvioinnin pohdinta on jo metodisen kehittämisen alussa nostanut esiin niin opetussuunnittelimen, arviointiperusteiden kuin arviointikäytäntöjen kriittisen analysoinnin ja arviointitutkimuksen arviointeihin liittämisen tarpeita.
- Oppimistulosten arviointiin liittyy aina keskustelu **tilivelvollisuudesta** ja tulosvastuusta (esim. Gipps 2002; Sirotnik 2004, 1–17). Kun tarkastellaan ulkoisen arvioinnin kontrollitehtävään liittyvää tilivelvollisuutta, vaatimukset metodien ja mittaamisen luotettavuudesta, standardoitujen kokeiden käytöstä, arviointiprosessin ja -ympäristön sekä arvioijien laatuvaatimuksista kasvavat. Standardoidut kokeet – vaihtoehtoina prosessiarvioinneille – ovat monissa maissa (esimerkiksi Hollanti, Englanti, USA, Kanada ja Uusi-Seelanti) merkinneet kokeiden sisällön, arviointiperusteiden ja -kriteereiden ja arviointiympäristön ja arvioitsijoiden pätevyyden yhtenevyyttä ja kontrolloimista vahvemmassa merkityksessä kuin näyttöperustaisessa arvioinnissa on aiottu.

Vaatimukset oppimistulosten *tilivelvollisuuteen* liittyvästä julkisuudesta – oppilaitosten välisiä eroja korostamalla – ovat painottuneet sekä *väestön että median odotuksissa* ja julkisuudessa käydyssä keskustelussa. Yleisen käsityksen mukaan tilivelvollisuusvaatimus sisältää kansalaisten oikeuden saada tietoa opetustoimesta, sen tuloksista ja voimavarojen käytöstä ja näihin tuloksiin perustuvan oikeuden vaikuttaa koulutuspalveluihin ja tehdä itseään koskevia päätöksiä. Tätä keskustelua ei ole vielä ammatillisen koulutuksen oppimistulosten arvioinnin yhteydessä käyty. Uudessa arviointijärjestelmässä korostetaan arviointiprosessien ja tulosten läpinäkyvyyttä mieluummin kuin julkisuuteen liittyviä pyrkimyksiä ja intressejä tuottaa tietoa koulujen välisistä eroista ranking-listoina.

- Uusi oppimistulosten arviointijärjestelmä sisältää käänteen kontrolloivasta sekä standardeihin ja mittaamiseen perustuvasta arvioinnista *kehittävään arviointiin ja laadun jatkuvaan parantamiseen*. Tällöin muuttuvat myös tilivelvollisuuden ja tulosvastuisuuden painotukset. Siinä missä tilivelvollisuutta painottava lähestymistapa perustuu luottamuksen puutteelle ja kontrollin tarpeelle, kehittävä arviointi perustuu avoimuudelle, luottamukselle, arvioinnin ohjaavuudelle ja jatkuvalla oppimiselle. Tämä käänne voi myös olla keino estää tilivelvollisuuden liiallisesta korostamisesta syntyneitä haittoja. Yksi suurimmista haitoista on, että opettajat alkavat pitää arviointia osoituksena opettajiin kohdistuvasta luottamuspulasta ja menettelynä, jolla pyritään horjuttamaan heidän käsitystään ammatillisesta pätevyystään, vähentämään luovuutta ja motivaatiota ja rajoittamaan heidän valinnanmahdollisuuksiaan ja vastuuta opetustyössä. (Vrt. Crooks 2002, 241 "244.)

Arvioinnin painopisteen muutos kontrolloivasta kehittävään arviointiin merkitsee myös *opettajien, työelämän ja opiskelijoiden välisen vuorovaikutuksen* ja yhteisen reflektoinnin painottumista. Tämä lisää myös arvioinnin avoimuutta ja läpinäkyvyyttä ja luottamusta. Arviointiin liittyvä luottamuksen ja oikeudenmukaisuuden kokemus on jo sinänsä merkki arviointitiedon luotettavuudesta. Luotettavuutta lisää myös se, että arviointitieto tuotetaan työelämässä. Lähtökohtaletuksena on, että työelämässä suoritettavat näytöt mittaavat parhaimmillaan työelämässä tarvittavaa osaamista paremmin kuin oppilaitoksissa keinokeisissa ympäristöissä järjestetyt kokeet ja näytöt.

- Uuden järjestelmän avulla pyritään lieventämään voimakkaan *riippumattomuuden* seurauksia. Uusi arviointijärjestelmä pyrkii lähentämään arvioijia ja arvioitavia toisiinsa vastakkainasettelun sijaan ja tuomaan arviointiin myös voimaannuttavaan arviointiin liittyviä elementtejä (vrt. Fetterman 1994, 1996). Fettermanin (emt.) mukaan voimakas riippumattomuuden korostaminen voi



synnyttää erillisiä ja irrallisia järjestelmiä ja etäännyttää arvioinnin arvioitavaa kohteesta niin, ettei arviointitieto ole hyödynnettävissä järjestäjien, oppilaitosten ja opettajien toiminnassa. Arvioinnin osuvuuden (validiteetti) heikkenemisestä seuraa myös, ettei ulkoinen arvioija saakaan luotettavaa ja todellista toimintaa koskevaa arviointitietoa.

- Oppimistulosten arvioinnin integrointi näyttöihin mahdollistuu näytöistä tuotettavien tietojen ja niihin liittyvän **yhtenäisen dokumentoinnin** (arviointidata) avulla. Tätä näytöistä kertyvää formatiivista arviointitietoa käytetään kansallisessa arvioinnissa osana tilivelvollisuutta. Kun formatiivista tietoa käytetään myös summatiiviseen tarkoitukseen, sisäisen, opettamiseen ja oppimiseen liittyvän arviointitiedon luonne muuttuu julkiseksi, yleiseksi arviointitiedoksi ja samalla muodollisemmaksi sekä virallisemmaksi. Tämä puolestaan nostaa esiin erilaisen arviointitiedon julkisuutta koskevat eettiset periaatteet ja arvioinnin haitat.
- Uuden oppimistulosten arviointijärjestelmän kehittämisessä on otettava kantaa, onko vertaileminen ylipäättänsä tarpeellista tai välttämätöntä. Uudessa järjestelmässä pidetään ulkoista vertaamista tärkeämpänä luottamukseen ja vuorovaikutukseen perustuvia oppilaitosten ja paikallisten toimijoiden **vertaisarviointeja, benchmarking-toimintaa ja oppilaitosten itsearviointia**. Oppimistuloksia koskevia tulkintoja tehdään myös paikallisten toimijoiden toimesta eikä vain niiden ulkopuolella. Tämä korostaa arviointikohteiden aktiivista ja vastuullista roolia arviointiprosessissa ja oppimistulosten arviointien sisällyttämistä myös koulutuksen järjestäjien ja oppilaitosten itsearviointeihin.
- Keskustelu näyttöjen ja oppimistulosten arvioinnin integrointivaatimuksesta ja arviointien yhteensovittamisongelmista nostaa esiin kysymyksen, pitäisikö ammatilliseen osaamiseen liittyvässä oppimistulosten arvioinnissa luopua perinteisestä käsityksestä **arviointitiedon luotettavuudesta, totuudellisuudesta ja yleislettävyydestä**. Tällöin yleistettävyyden sijaan voisi olla tarkoituksenmukaisempaa painottaa asioiden suhteellisuutta ja paikallisen tason ymmärryksen lisäämistä, jatkuvaa oppimista ja prosessinomaisuuteen liittyvää näkemystä laadunparantamisen ja oppimisen jatkuvuudesta ja loputtomuudesta. Tämänkaltaista arviointiprosessia ei voida täysin kontrolloida, vaan arvioijien ja arviointitulosten hyödyntäjien on siedettävä epävarmuutta (vrt. Guba ja Lincoln 1989).
- Uusi oppimistulosten arviointimalli merkitsee monista arvioinnin metodisista vaatimuksista tinkimistä. Tiedetään, että arvioinnin **luotettavuusongelmat** ovat etenkin järjestelmää luotaessa melko suuret. Oletuksena kuitenkin on, että ulkoinen laadunvarmistus ja -kontrolli takaavat riittävän luotettavuuden pitkällä

aikavälillä. Luotettavuutta parantavat muun muassa järjestelmään sisältyvä kriteeriperusteisuus ja jatkuvaa parantamista tukevat palautemekanismit. Luotettavuudesta tehtävien johtopäätösten tekeminen on mahdollista vasta metodien pitkäjänteisen kehittämisen ja testaamisen jälkeen. Tulevaisuudessa on löydettävä tasapaino validiteetti- ja reliabiliteettivaatimusten (vertailtavuus) välille. Vertailtavuuden ja yhdenmukaisuuden liiallinen korostus voi lisätä kontrollin ja yksityiskohtaisten standardien määrää, jolloin näytöt kaventuvat, tulevat mekaanisiksi ja menettävät merkitystään autenttisina ja valideina ammatillisen osaamisen arviointivälineinä.

- Arviointitoimintaa organisoidaan **verkostomaiseksi**. Verkostomainen toimintakulttuuri varmistaa monipuolisen ja parhaan mahdollisen asiantuntemuksen yhdistämisen ja hyväksikäytön myös oppimistulosten arvioinnissa. Tämä on erityisen välttämätöntä, kun luodaan metodisesti vaativaa, uutta ja eri tahojen kannalta katsottuna uskottavaa järjestelmää ammatillisen koulutuksen oppimistulosten arviointiin.
- Veloitteet näyttöjen ja oppimistulosten arviointien integrointiin antavat mahdollisuuden arvioida kriittisesti opetussuunnitelmien perusteiden ja koko opetussuunnitelmajärjestelmän toimivuutta sekä eri toimijoiden arviointiosaamista. Kokeilu osoitti, että kansallinen arviointi ja tulosten vertailu edellyttää selkeitä laatuvaatimuksia, laadunvarmistusta ja taustatietojen analysointia oppimistulosten ymmärtämiseksi. Lisäksi se edellyttää arvioijien koulutusta ja perehdyttämistä. Tärkeää on myös **oppimistulosten arvioinnin vaatimusten huomioiminen jo näyttöjärjestelmän kehittämisessä**. Tällöin voidaan päästä kokonaistaloudellisiin ratkaisuihin. Integroimalla oppimistulosten arvioinnin edellyttämä laadunvarmistus osaksi näyttöjen kehittämistä, vältetään päällekkäisiltä ja eri tahoja turhauttavilta sekä taloudellisia voimavaroja kuluttavilta työvaiheilta (kerralla oikein). Tällöin ulkoinen laadunvarmistus muodostuu näyttöjärjestelmän luonnolliseksi osaksi ilman erillistä oppimistulosten arvioinnin edellyttämää varmistusmenettelyä. Tämä on haaste kehittää uusia verkostomaisia työkäytäntöjä kehittäjien ja arvioijien ja muiden toimijoiden välille. Lisäksi haasteita riittää arviointitoiminnan mallintamisessa, verkostoituvassa toimintakulttuurissa, metodien kehittämisessä, palaute- ja seurantajärjestelmien luomisessa ja eettisten kysymysten ja arvioinnin vaikuttavuuden pohdinnassa.



Lähteet

- A 150/2003. Koulutuksen arvioinnista.
- Crooks, T. J. 2002. Educational assessment in New Zealand Schools. Profiles of educational assessment systems world-wide. *Assessment in Education* 9 (2), 237–253
- Fetterman, D. M. 1994. Steps of empowerment evaluation: From California to Cape Town. *Evaluation and program planning* 17 (3), 305–313.
- Fetterman, D. M. 1996. Empowerment evaluation. Teoksessa D. M. Fetterman, S. J. Kaftarian & A. Wandersman (toim.) *Empowerment evaluation and tools for self-assessment & accountability*. California: Sage, 3–39.
- Gipps, C. V. 2002. Educational accountability in England: The role of assessment. Seminaariesitys 22.4.2002, AERA-konferenssi, Chicago.
- Guba, G. E. & Lincoln, Y. S. 1989. *Fourth generation evaluation*. USA: Sage Publications.
- Hallituksen esitys Eduskunnalle laeiksi ammatillisesta koulutuksesta annetun lain ja ammatillisesta aikuiskoulutuksesta annetun lain 16 §:n muuttamisesta.
- Hautamäki, J., Arinen, P., Eronen, S., Hautamäki, A., Kupiainen, S., Lindblom, B., Niemivirta, M., Pakaslahti, L., Rantanen, P. & Scheinin, P. 2002. Assessing learning-to-learn. A framework. National Board of Education. *Evaluation* 4/2002. Helsinki.
- Hollo, M., Backman, H., Hakulinen, R., Katajisto, J., Koski, L. & Pernu, M-L. 2001. Käsi- ja taideteollisuusalan koulutuksen arviointi. Opetushallitus. Arviointi 6/2001. Helsinki.
- Kankaanpää, A. 1997. Ammatin kuvaus koulutuksen apuna. Ammattien kuvausjärjestelmän rakentamisen näkökulmia, ongelmia ja ehdotuksia. Helsinki: Opetushallitus.
- Kansallinen oppimistulosten arviointijärjestelmä. Opetushallitus. Arviointi 4/1998. Helsinki.
- Kekkonen, K. (toim.) 1996. Metsä vastasi. Metsä- ja puutalouden koulutuksen arviointi. Opetushallitus. Arviointi 9. Helsinki.
- Koski, L., Nyholm, N., Pernu, M-L., Pietilä, U. & Räisänen, O. 2002. Kaupan ja hallinnon alan koulutuksen arviointi. Opetushallitus. Arviointi 12. Helsinki.
- Koulutuksen arviointineuvoston aloite opetusministeriölle oppimistulosten arviointijaoston nimittämisestä 20.4.2004. Koulutuksen arviointineuvosto. Jyväskylä.
- Koulutuksen arviointineuvoston lausunto opetusministeriölle ammatillisen koulutuksen näyttöjärjestelmää koskevasta lakiehdotuksesta ja linjausluonnoksesta 13.6.2004. Koulutuksen arviointineuvosto. Jyväskylä.
- Koulutuksen tuloksellisuuden arviointimalli 1995. Opetushallitus. Arviointi 9. Helsinki.
- Koulutus ja tutkimus vuosina 2003–2008. Kehittämissuunnitelma 2003. Opetusministeriö. Koulutus- ja tiedepolitiikan osasto. Helsinki.
- Nuotio, P., Backman, H., Pernu, M. L. & Sisättö, M. 2001. Matkailu-, ravitsemis- ja talousalan koulutuksen arviointi. Opetushallitus. Arviointi 5/2001. Helsinki.
- OPM:n päätös (212/430/98) 23.6.1998.
- Poikela, E. 1998. Oppiminen, arviointi ja osaaminen. Teoksessa A. Räisänen (toim.) *Hallitaanko ammatti? Pätevyyden määrittelyä ammatin perustaksi*. Opetushallitus. Arviointi 2/1998. Helsinki, 35–46.
- Pääministeri Paavo Lipposen II hallituksen ohjelma 1999. Suomen hallitus: Valtioneuvoston kanslia.
- Pääministeri Matti Vanhasen hallituksen ohjelma 2003. Suomen hallitus: Valtioneuvoston kanslia.
- Radnor, H. & Shaw, K. 1996. Developing a collaborative approach to moderation. Teoksessa H. Torrance (toim.) *Evaluating authentic assessment*. Philadelphia: Open University Press, 124–143.
- Räisänen, A. 1998. Ammatillisen osaamisen arviointi. Teoksessa A. Räisänen (toim.) *Hallitaanko ammatti? Pätevyyden määrittelyä ammatin perustaksi*. Opetushallitus. Arviointi 2/1998. Helsinki, 9–20.

- Räisänen, A. (toim.) 1999. Komeetta. Kone- ja metallialan sekä sähköalan koulutuksen kokonaisarviointi. Opetushallitus. Arviointi 4/1999. Helsinki.
- Räisänen, A. & Vuorenmaa, M. 2003. Joustavuus, siirto-osaaminen ja liikkuvuus ammatillisessa koulutuksessa. Lähestymistapa ja arviointituloksia. Cost Action A II. Opetushallitus. Arviointi 3. Helsinki.
- Räikköläinen, M. & Ecclestone, K. 2005. The implications of using skills tests as basis for a national evaluation system. in Finland. Outcomes from a pilot Evaluation in 2002–2003 in Finland. Opetushallitus. Arviointi 1/2005. Helsinki.
- Räikköläinen, M. & Mäkihalvari, P. 2003. Oppimistulosten arviointijärjestelmän kehittäminen. Arviointikokeilu, LVI-ala, rakennusala ja sosiaali- ja terveysala. Saatavana www-muodossa: <www.oph.fi/ammattillinesr> tai <www.oph.fi/julkaisukaappi>
- Sirotnik, K. A. 2004. Critical concerns about accountability concepts and practices. Teoksessa K. A. Sirotnik (toim.) Holding accountability accountable. What ought to matter in public education. Teacher College. New York: Columbia University, 1–17.
- Vuorenmaa, M. & Räisänen, A. 1997. Lähihoitajat työelämään. Lähihoitajakoulutuksen käynnistysvaiheen arviointia. Opetushallitus. Arviointi 1/1997. Helsinki.



Principles and processes in evaluation systems: an outsider looks in.....

Introduction

In Finland Initial Vocational qualifications in future will incorporate a skills test based testimonial as proof of the student's attainment of the goals set in occupational studies. In parallel, a national system of evaluating learning results and new assessment practices to be used by the providers of education are being developed. The skills test based student assessment to be included in future vocational qualifications contains new challenges to be met and new questions to be answered in relation to how the Ministry of Education, the Finnish National Board of Education and other partners will work together to make the new evaluation system work effectively.

The decision by the Finnish Ministry to base national evaluation on skills tests not only changes routines and the ways of implementing the evaluation: it also changes the whole concept of evaluation and the evaluation culture and compels different organisations in Finland to re-evaluate the principles governing the reliability and practicality of evaluation and the aims of national evaluation.

In the light of these challenges, this article offers some perspectives by an external peer evaluator, Dr Kathryn Eccleston, from the University of Exeter in the UK who has been working on development work and a pilot evaluation project carried out by the National Board of Education (Räkköläinen & Eccleston 2005). This article outlines the main changes between old and new evaluation systems in the Finnish vocational system and argues that it is important to be clear about the principles, processes and purposes that underpin the new practices being proposed for evaluation in Finland's vocational system: without this clarity, an evaluation system might aim to promote certain values but end up

promoting different ones from those intended. Clarity also enables different groups and partners in the system to understand their respective roles and responsibilities in evaluation.

Descriptors: assessment, evaluation of learning outcomes, quality assurance

Changes to assessment and evaluation in Finland

Developing new assessment based on skills tests

Student assessment has been renewed so that vocational competence in future will be assessed by practical skills tests: this summative assessment will form the basis of a national evaluation system. A decision made by the Ministry of Education in 1998 required all vocational qualifications to be assessed by a practical skills test that demonstrated students' attainment of the aims of vocational training. The final transition to the new system of assessment and evaluation is scheduled for the autumn of 2006.

These proposals enabled tests to be arranged locally with the provider of education in charge of preparations. The transition to a skills test based assessment and evaluation system was supported at the national level with ESF (European Social Funding) funding through the production of skills test material and pilot tests in various sectors. Test materials were produced both nationally and by individual test organisers. However, there were no standard national tests. Instead, the provider of education was free to decide what material to use in the actual test. The skills tests are an integral part of the teaching and learning process, designed to unify the practices of providers whilst also diversifying student assessment to reflect local conditions and making the needs and desiderata of working life part of the student's vocational proficiency.

The purpose is to enhance the quality of vocational training and to ensure that the training meets the perspectives of working life. In addition to promoting better student assessment, the reform also strives to improve national evaluation. It is also crucial to evaluate whether national evaluation data can be extracted from the skills tests. A parallel aim is that the evaluation of national learning results further will be an evaluation of the effectiveness of training. The development project will reveal the strengths and weaknesses of these aims and suggest possible solutions.

Developing a new evaluation system

In parallel with developing a skills test-based assessment system, the Ministry of Education decided to abandon a system with separate national tests. Instead, the eval-

uation of learning results and the assessment by skills tests were to be integrated as effectively as possible. The decision as to whether one or more templates and strategies should be put to the trial was left to the judgment of experts.

In the past, national evaluation of learning results in Finland was based on nationally devised uniform tests and work assignments in the form of a final examination. The integration of national evaluation with skills tests has as its aim to abolish a system based on separate national tests and, instead, to gather the evaluation data straight from local tests flexibly administered by the institutions. This model has been tested in a national pilot project coordinated by the National Board of Education (NBE) in 2002–2003.

A report (based on the pilot project) to the Ministry of Education (Räkköläinen & Ecclestone, 2005) discusses the extent to which using the skills test simultaneously for assessing students' competence and for evaluating the effectiveness of training positively influences the concept of evaluation and the nature and quality of the evaluation data. It also addresses the tensions that have emerged between the parties involved in the Finnish evaluation system (e.g. control of and trust in evaluation, norms and flexibility of steering, fairness, objectivity and comparability of evaluation, and the difficult balance in all assessment and evaluation systems between validity and the requirement of reliability).

Objectives of the development project

The main objectives can be summarized as follows:

- To evaluate the strengths and weaknesses of a new evaluation system, taking into account the perspectives of different stakeholders.
- To describe key features of assessment and evaluation in vocational education and the values and principles that underlie them.
- To differentiate clearly between formative and summative assessment and evaluation and to define other important concepts (such as reliability, validity, standards, quality assurance and quality control).
- To analyze technical, professional, political and educational tensions within the existing system of evaluation and those that might emerge in an evaluation system based on skills tests.
- To base recommendations 1) on evidence from pilot projects that test different strategies for national and local evaluation and 2) on ideas from an impartial external expert.

- To pilot a new approach to evaluation through two (pilot) projects, in collaboration with representatives from colleges, vocational schools, employers, the social partners and government departments in the Finnish vocational education system.
- To offer advice to the National Board of Education and the Ministry of Education about the implications of using national skills tests for evaluating the Finnish vocational education system.

The project was launched by the Evaluation Unit at NBE in January 2002. Anu Räisänen, Counsellor of Education at NBE and Planning Specialist Mari Rökköläinen, manager of the project, devised the first project plan for the development project. Dr. Kathryn Eccleston from the University of Newcastle in the UK, now at the University of Exeter, was asked to act as peer evaluator and expert on assessment systems and practices in vocational education.

The role of a peer evaluator

An external evaluator can bring an outside perspective and expertise in assessment for vocational education to the project, based on clear objectives for the cooperation: this enables her to suggest ideas, advice and information at various points during the project. In this project, the peer evaluator met and interviewed various officials, teachers and social partners during six working periods in Finland (usually two each year) (see Rökköläinen & Eccleston, 2005: Appendix 1). Her general role was to offer an objective, impartial and supportive perspective, based on intensive experience of the project, an in-depth understanding of the Finnish vocational education and evaluation system and experience of clarifying principles, politics and practices in assessment for the English vocational system (see Eccleston, 2004). During working periods the expert familiarised herself with principles, values and features of different aspects of the assessment and evaluation system in vocational education in Finland and identified ideas about technical, practical and critical aspects of evaluation amongst different agencies and interested parties in Finland (all 7 work periods). This helped NBE officials understand other assessment and evaluation systems and the implications of their principles and features for Finland's system.

Developers from NBE also visited England and familiarized themselves with UK's education system in VET, focusing on quality assurance and quality control procedures and principles in assessment systems. Before these visits, they analysed the main problems and 'critical points' which might emerge in the new evaluation system. During the visit to the UK, officials from the NBE discussed ideas with those who evaluate and are evaluated, including the Qualification and Curriculum Authority (QCA), two large awarding bodies (City & Guilds and NCFE) and some vocational educa-

tion institutions. The aims of these visits was to understand the main features of evaluation in the UK's system and to explore the balance between external, summative evaluation required by government and internal evaluation undertaken by institutions. A list of working periods, interviews, visits and observation that have formed the basis for our cooperation are included in Råkköläinen & Ecclestone, 2005: Appendix 1.

Vocational education and national evaluation in Finland

Organisations and procedures

The Finnish Government defines the general objectives of vocational education and training, the general structure of the study programmes and the common studies. The Ministry of Education decides on the details and scope of the study programmes, while the Finnish National Board of Education (FNBE) issues the national core curricula that determine the objectives and core contents of the studies. The tripartite expert bodies and committees set up by the Ministry of Education participate to core curricula process. Based on these, each provider of education prepares his own curriculum. The providers also have tripartite expert bodies and consultative committees that participate in the planning and development of education at the local level.

The Ministry of Education grants licences (accreditation) to organise vocational education, determining the education providers' fields of study and total number of students, etc. Within the framework of the licence and the confirmed structure of the study programmes, the education providers may focus their education as they see fit, allowing for the needs of the local and regional economic and working life.

The Finnish National Board of Education is an expert body responsible for the development of educational objectives, contents and methods in basic, general upper secondary, vocational and adult education and training. The Board prepares and adopts the core curricula and works with providers themselves in order to evaluate the quality of learning results. It is also intended that it will work with the Evaluation Council responsible to the Ministry of Education but roles and responsibilities are still emerging. Responsibility for evaluation in polytechnics and universities rests with these institutions themselves. In this task they have the assistance of the Higher Education Evaluation Council, operating under the Ministry of Education.

Developing co-operation between education providers and the interest groups that comprise 'working life' in Finland, namely teachers, employers and trades unions, is regarded as desirable in order for education to better meet working life requirements. Such co-operation is necessary, if for no other reason, than the fact, that the Finnish education system is very institution-centred. Representatives of working life are involved in the advisory bodies of vocational education at both the central administration and the local level. Students' orientation to working life, on-the-job training or

fixed-term studies at workplaces have been included as a regular element of education in initial vocational and in polytechnic degree programmes.

There is no separate, external inspection system for schools or colleges in Finland. The steering of the education system is in the hands of the Government and the Ministry of Education. Many matters, however, have been entrusted to the providers of education. Their operations are steered through the core curricula and the objectives laid down in legislation. Feedback concerning the operations of the education system is collected by means of statistics and evaluations, and these provide information that by itself steers education.

Like most developed countries, Finland completed the basic infrastructure of education during the past decades and is now in a situation where the emphasis has shifted to the quality of education. One tool for quality assurance is the evaluation of educational outcomes. In recent years, powers of decision have been devolved to the local level, whereby evaluation has become a significant tool also for the steering of education. In legislation, evaluation duties have been assigned to both the education providers and the authorities.

Roles and responsibilities of different groups and organisations

It has become apparent during the course of the evaluation project that political bodies, providers of education and interest groups in any education system need to understand their own system, the roles and responsibilities of different organisations and the links between them. A figure showing organisations and stakeholders in the Finland vocational system is below: it does not include evaluation for schools or higher education.

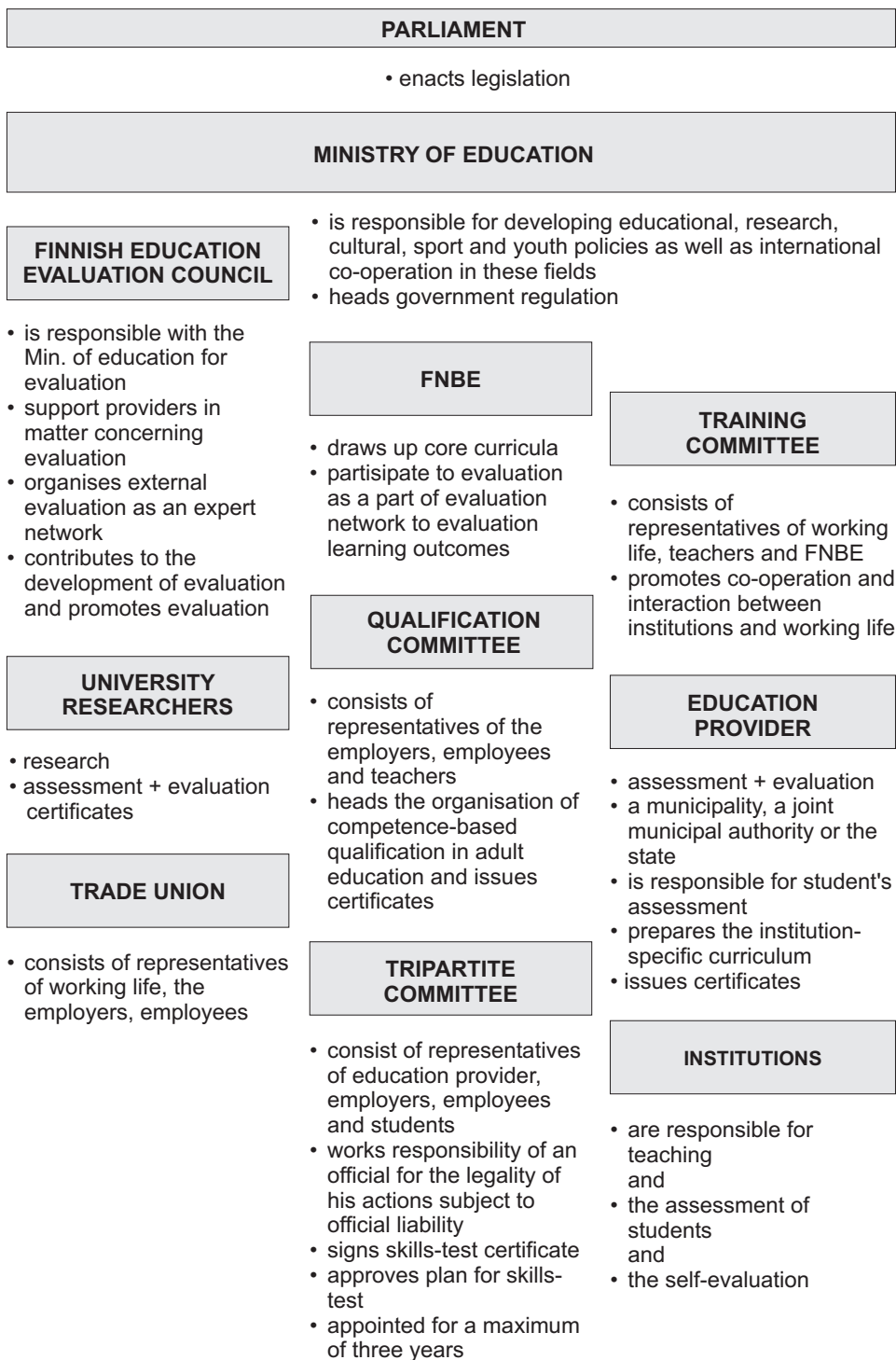


Figure 1. Organisations and stakeholders in the Finnish vocational system

Implications of integration

The integration of evaluation by learning results entails changes in the skills test-based system. However, one of the biggest challenges this poses for evaluation bodies in Finland is how evaluation based on skills tests can be reliable on a national level: the 4 skills tests were not designed originally for this new purpose. It was therefore clear from the outset that the decision for a new system of evaluation based on skills tests would create three main problems (Räkköläinen & Ecclestone 2005, 15):

- Comparison: skills tests and assessment practices vary a lot at the local level
- Time: it takes a long time to get the results and institutions implement their tests based on their own curricula
- Overlap: how to link and connect assessment at the local level with evaluation at the national level.

Principles and processes in evaluation

Definitions, terms and concepts

Organisations and providers in any vocational system need to understand how the system is organised and how different approaches to evaluation evolve over time. It is also important to be precise and clear about different concepts and principles because these affect roles and responsibilities in the system.

All European vocational systems show confusion amongst policy makers, evaluation and assessment bodies and teachers about the meaning of key concepts, purposes and processes in assessment and evaluation (Breur, 2002). Confusion is exacerbated because key terms are used interchangeably amongst these different interest groups: this problem is evident in Finland where there is no distinct word for 'assessment'. Despite these difficulties of definition and differences in the uses of terms and concepts, the project aimed to clarify from the outset, how different terms might be defined and then to use them more precisely than is currently the case in many assessment and evaluation systems (see Ecclestone, 2004). Key processes are:

- ***Assessment of learning outcomes:*** diagnostic, formative and summative, internal/external
- ***Evaluation of institutional provision:*** different levels such as macro (national), meso (institutional) and micro (classrooms and individual teachers); it is

therefore necessary to define *self-evaluation*, *peer review* and *peer evaluation* and *external audit*

- **Assessment and evaluation are inextricably linked** and the two terms are used interchangeably: the data gained from formative and summative assessments can therefore contribute to formative and summative evaluation
- **Learning results and learning outcomes** also need to be defined and differentiated: results has a narrower meaning (e.g. grades) than outcomes
- **Quality assurance**: internal processes built into qualifications and learning programmes from the outset, in order to increase reliability and validity of assessment
- **Quality control**: external processes to ensure that standards of reliability and validity are met
- **Standards**: norm-referenced standards are based on being able to measure students' results reliably and criterion-referenced standards are based on explicit measures of validity and authenticity.

The purposes of evaluation

There are different purposes of evaluation. It is therefore helpful to differentiate as precisely as possible which data are needed for which purposes and who should collect and use these data. Methods and instruments for collecting data, as well as the data themselves, can therefore serve each purpose separately or all of them together!

- External (formative) evaluation, for example carried out by advisers or people responsible for staff development and curriculum development
- External (summative) evaluation, for example carried out by inspectors appointed by the Ministry of Education, or peer reviewers appointed by an institution or a Ministry of Education
- Internal (formative and diagnostic) evaluation, for example using self and peer evaluation by colleagues, learning groups, quality circles
- Internal (summative) evaluation, for example carried out by peer reviewers evaluating courses within an institution.

Each purpose has a different ‘audience’ or ‘stakeholder’, although, of course, different stakeholders can be interested in all purposes! It is problematic and confusing for an evaluation system to try to achieve all these purposes with the same data and procedures/processes. This has been a problem in the UK where assessment data is used for many different evaluation purposes, both formative and summative. It is therefore crucial to differentiate between formative and summative purposes of evaluation because this helps to clarify more precisely which audiences and stakeholders have a legitimate interest in the processes and outcomes of evaluation.

Summative evaluation can be carried out by an institution or external agency for the following purposes: *accountability* for meeting government’s targets and/or for funding allocations; *system-wide standardization* of a) assessment practices, b) curricula, and c) learning results; *system-wide comparison of quality* against national criteria (reliability); *a lever for curriculum change* (e.g. to make vocational education more relevant to working life); *motivation of teachers and institutional managers* to work for improvements; *a basis for professional autonomy* (e.g. designing local assessments to suit local conditions).

Quality assurance

This aims to reduce sources of error caused by particular, local circumstances whilst preserving the validity of assessment: QA processes therefore develop quality and monitor progress. They include: accreditation – formal approval that institutions can carry out processes of quality assurance and quality control leading to formal recognition of achievement; validation (approval) of institutions to run courses and to carry out assessments; design of curriculum and certification (NBE and institutions); exemplars of students’ work at different levels of attainment; internal moderation of assessment results (standardisation – validity versus reliability); internal verification of assessment results to ensure that procedures have been followed; approved centres to run tests and curricula against centrally-designed and evaluated criteria (this is called validation in UK); exemplar materials of tests and learning materials; guidance for centres to follow.

Quality control

Quality control ensures public accountability for the results or outcomes of summative assessments carried out in local contexts. These processes might be both internal and external but their emphasis is on summative evaluation rather than formative evaluation. QC might involve the NBE (other government-appointed agencies) in processes to: *moderate assessors judgments* and decisions (e.g. external moderators acting on behalf of NBE; regional and national moderation meetings with teachers and evaluators; postal sampling, *in situ* sampling of assessment results by internal or

external evaluators; *verify that QA and QC procedures have been adhered to* within centres and by assessors acting on behalf of centres; *offer national quality awards* for best practice in quality assurance and quality control; *organise action research projects* by teachers into their own assessment practices (although these might be regarded largely as an informal form of QC); *collect and interpret data* on behalf of schools.

Comparing evaluation systems in Finland and the UK

Goals and values in the Finnish education system

A number of the features of the Finnish system indicate that democratic and humanist aims are prominent at all levels, from Ministry to teachers, social partners and students. These include for example: helping teachers and social partners understand the underlying principles and purposes of evaluation, rather than merely issuing top-down guidance and advice or prescriptive specifications to follow. Advice, support and trying to find ways of producing useful data replace inspection, ‘name and shame’, and the publication of league table results. There is a strong political antipathy to such measures espoused by officials at the Ministry of Education and FNBE. There is also a holistic approach that emphasizes local validity of assessment and curriculum design, and local and professional autonomy in deciding how to implement legislative obligations; this requires different layers and methods of evaluation.

From the Ministry to teachers, and social partners and students principles of evaluation in Finland appear to emphasize the following values:

- *A holistic approach* that emphasizes local validity of assessment and curriculum design and local and professional autonomy in deciding how to implement legislative obligations; this requires different layers and methods of evaluation
- *The prominence in official texts of students’ personal and social development*, rather than an emphasis on working life competence. From this perspective, the NBE states:

the standard of an individual provider of educational services or of the whole education system is determined by its capacity to support the students’ personal growth and educational aspirations as well as to prepare them for social and working life (NBE 1999, 12)
- *Helping teachers and social partners understand the underlying principles and purposes of evaluation* rather than merely issuing top-down guidance and advice or, as in the UK, prescriptive specifications to follow. Advice, support and trying to find ways of producing useful data replace inspection, “name and

shame” and publication of league table results. Officials at the Ministry of Education and NBE express a strong political antipathy to such measures

- A leading idea in the reformation of educational administration has been that the high quality of public educational services, which is necessary for the equality of citizens, cannot be achieved through strict uniformity, detailed regulations and state control but rather through improving the conditions for operation (NBE 1999, 11)
- *An expectation of direct employer involvement in educational activity*: if employers want more relevance to working life in vocational qualifications, they must take part in assessing students’ competence
- *An openness to admitting problems and tension* and to spending time on trying to resolve them, alongside social partners
- *A strong emphasis on evaluation for improvement of education* and the conditions for learning, together with a recognition that underlying conditions can act as profound barriers to good quality.

Evaluation in the UK

In contrast to the system of consensus between social partners in Finland, the UK’s vocational education and training system is strongly centralised, expensive and complex, with many organisations and interest groups playing a part in evaluation. The system has undergone many major upheavals and restructuring over the past 20 years, where funding, regulation and evaluation systems are reorganised simultaneously.

The figure below shows the complexity of the UK’s system. There are over 1 200 different qualifications accredited by over 400 awarding bodies, offered by 450 further education colleges, 150 universities and over 2 000 providers of work-place training (e.g. employers, private training companies). Providers have to be approved formally by awarding bodies to run these qualifications and, in turn, the qualifications must conform to national standards and criteria defined by the Qualifications and Curriculum Authority (QCA). The QCA regulates awarding bodies through a common code of practice that governs quality assurance and quality control in each qualification.

Outside the main qualification pathways (National Vocational Qualifications, general advanced level education and vocational advanced level education), awarding bodies offer institutions and providers a very wide range of courses and qualifications: some of these are designed locally to suit the particular needs of a group of students or an organisation. However, there is not the same degree of local autonomy and flex-



ibility in designing local curricula as in the Finnish system.

Governments in the UK place a strong emphasis on reliability in the results of summative assessment so that they can evaluate the overall norms of achievement each year in the form of ‘national standards’: this system is used in post-16 qualifications for Advanced general and vocational level. The work-based assessment of National Vocational Qualifications uses a form of skills-based (competence-based) tests: although attempts are made to ensure that assessment is as fair and reliable as possible, there is no attempt to collate the results of these individual tests for national evaluation.

Instead of using national tests to evaluate institutions, internal and external summative evaluation focuses on institutions and teachers by using a set of quantitative indicators such as achievement of formal qualifications, rates of retention and progression to employment or further education and qualitative indicators from external inspection that focus on the quality of teaching and learning, quality assurance and management of institutions and student achievement. Institutions and providers are graded 1–5 (5 is exceptionally poor and 1 is excellent) and results of inspections are published in the form of reports. Institutions and providers must undertake formative evaluation of the quality of their provision using the same criteria as inspectors and produce a self-evaluation report before summative evaluation in the form of external inspection. Many institutions invest heavily in external consultants to prepare for summative evaluation in order to maximise their grades.

In addition to summative external inspection, institutions must undertake quality assurance and quality control of assessment procedures and outcomes on behalf of the awarding bodies that offer qualifications in a particular college. For colleges offering a range of qualifications to their local community, this might mean procedures for 15 different awarding bodies. These procedures include internal moderation of assessors’ judgments in order to ensure consistency of grading, external moderation of a sample of students’ work by the awarding body and visits by moderators to institutions.

Finally, a number of bodies undertake research and development projects to implement government policy or more independently in response to various concerns such as widening participation in formal education. Some of these organisations receive money directly from the Department of Education and Skills or the Learning and Skills Council and others operate more independently in terms of funding and scope to decide what to research. These bodies are identified in the map of the UK’s system below.

The map indicates the main organisations involved in the UK’s post-compulsory system.

THE SYSTEM OF VOCATIONAL AND ADULT EDUCATION IN ENGLAND

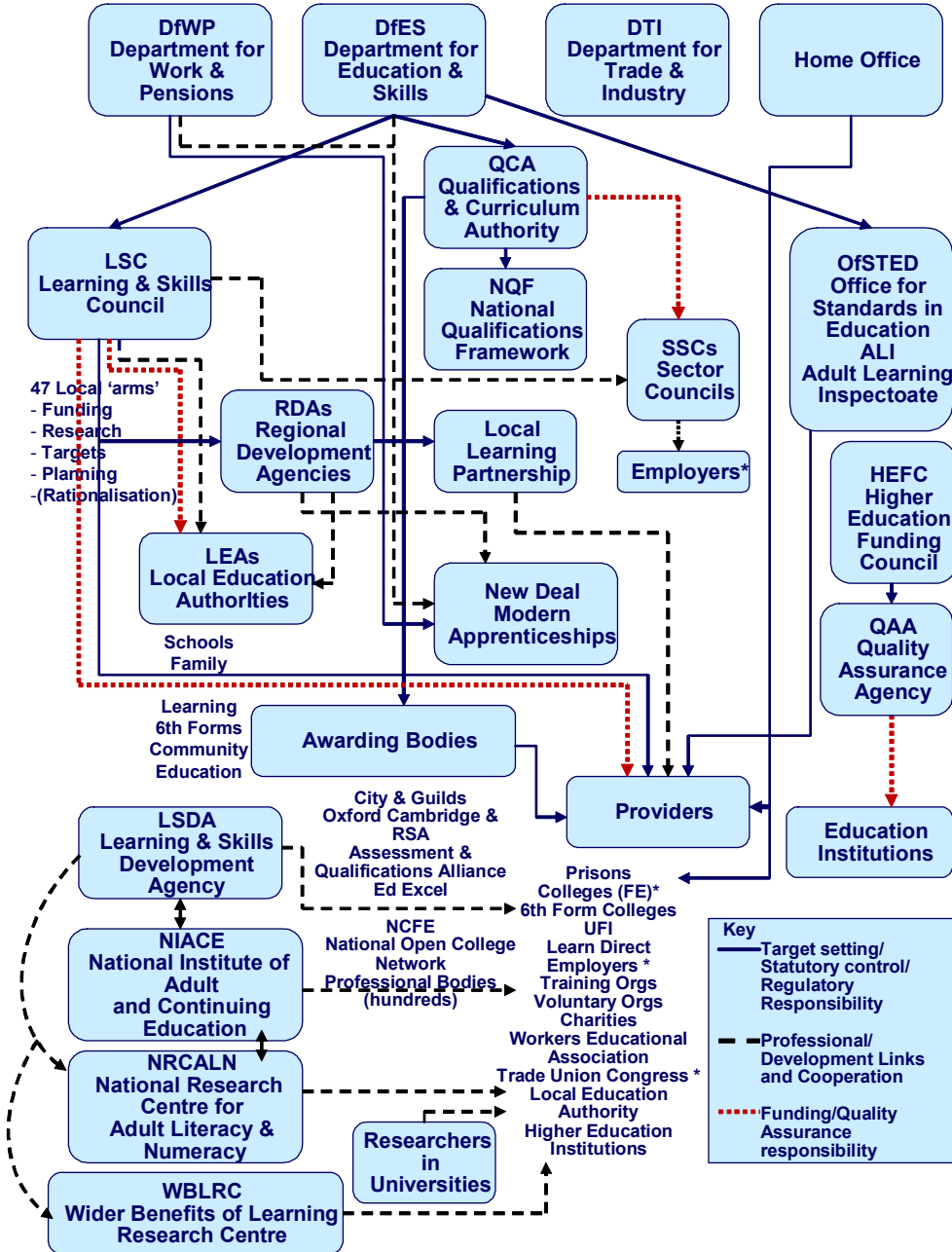


Figure 3. The system of vocational and adult education in the UK

Table 1. *The main differences between the two systems can be summarised as follows*

UK	Finland
<p>Complex, large and confusing with many different bodies and organisation involved in funding, regulation, quality assurance and quality control and evaluation.</p> <p>Major change and restructuring of the system at all levels, including funding, organisation, regulation and evaluation has occurred three times in the past 15 years.</p>	<p>Small and simple, with clear lines of responsibility between a small number of organisations.</p> <p>Change is much less frequent and tends to be introduced more slowly than in the UK.</p>
<p>Many different vocational programmes, assessment systems and qualifications (certification)</p>	<p>Few different programmes and forms of qualification or certification</p>
<p>Many awarding bodies with different approaches to quality assurance and quality control – institutions/providers must choose between them and work with different systems</p>	<p>No awarding bodies – institutions operate autonomously</p>
<p>Emphasis on national standards, national systems, comparison of providers</p>	<p>Emphasis on local standards, local agreements and a high level of local autonomy</p>
<p>National prescriptions for regulation, quality assurance and inspection and assessment systems</p>	<p>No prescription but, instead, guidance and advice and a strong emphasis on professional judgement and autonomy</p>

Summary

The goals and values of an education system are reflected both overtly and implicitly in its evaluation and assessment systems. Clarifying principles and processes, together with their underlying purposes in an education system enables those creating and implementing national evaluation systems to see what is being communicated to teachers, social partners and students. The role of an external evaluator in the FNBE's project is to help to illuminate how these principles, concepts, goals and values are operating in a system.

Clarifying these aspects of evaluation also enables different organisations and bodies in the system to understand their respective roles in areas such as quality assurance and quality control, and for training and education in evaluation for various groups such as teachers, institution managers etc.

Clarifying the principles and concepts also enables questions to be asked about the proposed changes to evaluation of vocational education in Finland:

- to what extent can skills tests that were designed originally to maximise validity and authentic assessment in local settings be used for a different purpose, to evaluate the learning results of institutions?
- what is the role of skills tests in relation to other procedures that aim to assure and control the quality of assessment in local contexts?
- what are the implications of using the skills tests for quality assurance and quality control at national and local levels?
- to what extent can traditions of local autonomy and professional judgment in assessment be maintained in the new system?
- what degree of national standardisation and regulation is necessary for evaluation?
- what are the respective roles of the different organisations in the Finnish evaluation system and how should they work together to make a coherent system?
- how far can the new system for evaluation reflect humanist and democratic values and principles in the Finnish education system?

References

- Eccleston, K. 2004. Principles, practices and politics: understanding assessment and qualifications in post-compulsory education (second edition) Leicester, National Institute of Adult and Continuing Education.
- Räkköläinen, M. & Eccleston, K. 2005. The implications of using skills tests as basis for a national evaluation system in Finland. Outcomes from a pilot evaluation in 2002-2003 in Finland. Arviointi 1/2005. Opetushallitus. Helsinki.



Mitataan oppilaita, mutta päätellään koulusta

Artikkelissa tarkastellaan tilannetta, jossa oppilastason mittauksista päätellyt koulukohtaiset tulokset tulkitaan suhteessa kouluun. Kirjoittajat toteavat, etteivät oppilastason tulokset ja instituution tulokset ole sama asia, eikä koulun arvioiminen ole sama kuin oppilaiden arvioiminen.

Koulutuksen arvioinnissa on suuriin otoksiin perustuvissa arvioinneissa eritelty kuitenkin oppilaiden suorituksia ja heidän asenteitaan tai hyvinvointiaan. Arvioinnin osuvuuden ja oikeutuksen näkökulmasta on tärkeää olla kriittinen sen suhteen, mitä näiden tulosten perusteella voidaan päätellä koulusta, ja johtavatko jotkin metodiset tai tulkintaan liittyvät seikat päätelyn hakoteille. Käsittelemme artikkelissa muutamia 'koulun vaikutuksen' käsitteeseen ja tulkintoihin liittyviä näkökohtia. Kun puhutaan koulutuksen tuottamista valmiuksista, oletetaan, että koulutuksen tuotokset voidaan eristää muista lopputulokseen vaikuttavista tekijöistä, esimerkiksi kodin ja lähiympäristön vaikutuksista tai oppilaiden tai opiskelijoiden valikoitumisesta johtuvista eroista. Näiden eri tekijöiden roolin erottelu on mahdollista, mutta se edellyttää koulutuksen tuottaman muutoksen tarkempaa mallintamista ja arvioimista.

Artikkelissa esitetään ekologisten tarkastelujen periaate, käsitellään ekologisten korrelaatioiden tulkinnallisia riskejä sekä esitetään eräitä metodis-eettisiä suosituksia.

Asiasanat: arviointi, aggregaattitasot, kontekstualisointi, metodiset kysymykset

Voidaanko koulutuksen arviointi suorittaa mittaamatta oppilaita?

Tämän artikkelin lähtökohta voidaan ilmaista erään kirjan nimen avulla: *Grading Teachers, Grading Schools. Is Student Achievement a Valid Evaluation Measure?* (Millman 1997). Millmanin kirjassa esitellään eräissä USA:n osavaltioissa käytössä olevia added-value arviointisysteemejä, joiden tarkoituksena on varmentaa, että oppilastason mittauksista päätellyt koulukohtaiset tulokset tulkittaisiin asianmukaisesti. Sellainen tulkinta edistää koulutuksen ja samalla yksittäisten koulujen pedagogista kehittämistä. Tuon kirjan ja tämän artikkelin tarkoituksena on selvittää koulutuksen arvioinnin tätä kysymystä.

Perusmalli koulutuksen arvioinnissa on se, että on saatu sovittua eräistä toiminnan indikaattoreista, joiden katsotaan kertovan jotain oleellista koulutuksesta ja sen instituutioiden toiminnasta. Yksinkertaisin malli on hyvin yleinen, mutta sen kertaaminen on paikallaan:

”Laajimmassa merkityksessään suoritusindikaattori on sellainen kokoava tilastollinen mittaus, joka kohdistuu instituutioon tai systeemiin, ja jonka on tarkoitus olla jossain suhteessa sen toiminnan ’laatuun’. Tällaiset mitat voivat koskea systeemin erilaisia puolia ja heijastaa erilaisia tavoitteita: panosindikaattoreita kuten oppilas/opettaja suhde ... käytetään arvioimaan instituutiolle käytettävissä olevia resursseja; prosessimitat kuten keskimääräinen opetusaike oppilasta kohden ... pyrkivät ilmentämään organisatorisia rakenteita, kun taas tulosmittoja kuten koulun tutkintotuloksia [kuten ylioppilaskoe Suomessa] ... on käytetty arvioitaessa instituutioon tuloksellisuutta). Vaikkakin [artikkelimme] keskustelu liittyy myös panos- ja prosessi-indikaattoreihin, keskitytään tulosindikaattoreihin ..., koska nämä ovat nousseet sosiaalisesti tärkeiksi ja niillä on keskeinen rooli poliittisissa keskusteluissa instituutioiden vastuusta.

On käynnissä hyvin tärkeä keskustelu siitä, millainen olisi paras ja validein indikaattorimittojen yhdistelmä instituutioiden tehokkuuden mittoina: asetamme kyseenalaiseksi sen, ilmentääkö tulosten vaihtelu, vaikka olisi otettu huomioon ulkoiset tekijät, oikealla tavalla instituutioiden ’laatu’. ... Uskomme myös, että kaikilla suoritusindikaattoreiden potentiaalisilla käyttäjillä tulee olla ainakin perustietoisuus näistä kysymyksistä olivatpa nämä käyttäjät, esimerkiksi, sairaaloiden hallintomiehiä tai koululaisten vanhempia.” (Goldstein & Spiegelhalter 1996).

Goldsteinin ja Spiegelhalterin keskustelu virittyy artikkelimme otsikon ristiriidasta. Tuotosindikaattorit vaihtelevat – oli sitten kysymys koko terveydenhuollon tai sivistyssysteemin tai niiden organisatoristen yksiköiden, sairaaloiden tai koulujen tarkasteluista – ja indikaattoreiden vaihtelun lähde on kovin helppo sijoittaa instituution

laitoksiin, mutta tässä päättelyssä on oleellisen suuria vaikeuksia.

Eräs käsitteellinen selvennys auttane edelleen täsmentämään sitä, että kyse on tärkeästä arvioinnin perusteisiin kuuluvasta asiasta. Goldstein & Spiegelhalter käsittelevät 'lisäarvon' käsitettä (*added value*), jota käytetään myös Suomessa koulutuksen arvioinnin keskusteluissa. 'Lisäarvon' käsitteessä on mukana eräs koulutuksen kannalta vakava ongelma. Käsite tulee taloustieteestä, missä panos ja tuotos voidaan mitata samoissa rahayksiköissä. Mutta koulutuksessa panos-, prosessi- tai tulosmitat eivät useinkaan ole samassa asteikossa tai mittayksikössä edes periaatteessa. Vaikka koulun "tulos" onkin ehkä ilmaistavissa ylioppilaskokeen puoltoäänien määränä tai vastaavana vertailukelpoiselta vaikuttavana suurena, jää koulujen toisistaan poikkeava alkutilanne, niiden 'panos' useimmiten erittelemättä. Tämä vertailuissa useimmiten huomiotta jäävä alkutila on ao. lukioon tulevien oppilaiden aiempi osaamisen taso, heidän ehkä jo ennen kyseiseen kouluun tuloa omaksuttu asenteensa koulutyötä kohtaan, heidän kotitaustansa, mahdollisesti jopa se, kuinka paljon tyttöjä koulussa on tai millaisia ainevalintoja he ovat lukiossa suosineet. Tämän alkutason merkityksen huomioon ottamiseksi Goldstein ja Spiegelhalter suosittelivat lisäarvon käsitteen sijaan käytettäväksi käsitettä 'sovitettu vertailu' (*adjusted comparison*). Instituution kuten koulutuksen (*schooling*) tehokkuuden ja merkityksen arvioiminen on kansakunnan koulutuspolitiikan tehtävä. Silloin arvioidaan sitä, miten koululaitos on kyennyt täyttämään sille asetetut sivistystehtävät ja luomaan edellytyksiä ammatilliselle osaa miselle työelämässä.

Kanadalainen David Olson (2003) esittää, että koululaitoksen uudistuminen ei ole vain sama kuin pedagogiikan tai opetussuunnitelman perusteiden muutos. Koululaitoksen arviointi edellyttää institutionaalista erittelyä suhteessa muihin oleellisiin instituutioihin, poliittisiin, taloudellisiin, uskonnollisiin ja akateemisiin instituutioihin. Tämän osana voidaan arvioida resurssien määrää ja käyttöä, suunnittelujärjestelmää ja yleistä seurantajärjestelmää, mutta kyse ei ole sen tutkimisesta, mitä todella tapahtuu jossain pedagogisessa prosessissa tai siinä, jossa jokin todellinen ihmisyksilö suoriutuu tai ei suoriudu saamastaan kouluun liittyvästä todellisesta haasteesta tai tehtävästä. Olson toteaa:

"Kansalliset ja kansainväliset arvoinnit ovat jonkin verran hyödyllisiä viestissään tavoitteista ja standardeista ja välittäessään mitan, jonka avulla opettajat voivat arvioida omia arviointikäytäntöjään. Mutta näillä kaikilla on vain rajallinen merkitys pyrittäessä ymmärtämään tai ohjaamaan pedagogista käytäntöä, joka ... perustuu yksilöiden subjektiivisiin tiloihin, tavoitteisiin, toiveisiin ja uskomuksiin" (s. 281).

"Koulun arvioiminen ei ole sama kuin oppilaiden arvioiminen. Koulut arvioidaan epidemiologisesti, siis käyttäen käsitteitä, joilla voidaan kuvata populaatioita, ja ehdotetut ratkaisut ovat sellaisia, jotka sopivat populaatioihin.



Niin kuin meille edellä selvisi [viittaa aikaisempaan jaksoon, jossa on selvitetty mitä tarkoittaa populaation kuvaus], verotusasteen nosto pienentää populaation tasolla kulutusta, mutta ei välttämättä jonkun yksilön kohdalla. Vastaavasti, koulut saattavat päättää muuttaa arviointitutkimuksen perusteella oppilaiden ja opettajien lukumäärän suhdetta, vaikka tutkimustulos ei ehkä olisikaan suoraan sovitettavissa mihin tahansa tiettyyn luokkaan, opettajaan tai oppilaaseen. Tulosten tilannekohtaisen relevanssin arviointi tulee jättää paikallisten asiantuntijoiden harkittavaksi siten, että otetaan huomioon juuri kyseinen opettaja, kyseiset lapset, kyseinen koulu, sen resurssit, lukujärjestykset ja muut aikataulut sekä opetettavat oppiaineet” (s. 283).

Koulutuksen arvioinnissa on suuriin otoksiin perustuvissa arvioinneissa eritelty kuitenkin poikkeuksetta oppilaiden suorituksia ja heidän asenteitaan tai hyvinvointiaan. Esimerkkejä tällaisista tutkimuksista ovat kansainväliset ja kansalliset oppimistulosten arvioinnit, oppimaan/ oppimisvalmiuksien arvioinnit sekä kouluterveyskyselyt. Arvioinnin osuvuuden ja oikeutuksen näkökulmasta on tärkeää olla kriittinen sen suhteen, mitä näiden tulosten perusteella voidaan päätellä koulusta, ja johtavatko jotkin metodiset tai tulkintaan liittyvät seikat päättelyn helposti hakoteille. Käsittelemme tässä luvussa muutamia 'koulun vaikutuksen' käsitteeseen – ja tulkintoihin – liittyviä perustavanlaatuisia ja yleisesti tunnettuja pulmia ja päättelyn sudenkuoppia.

Mitä tulokset kertovat koulun ”hyvyydestä”?

Aluksi arviointi on sijoitettava yleisempään viitekehykseen. Koulutuksen arviointia koskevassa lainsäädännössä koulutuksen järjestäjiä veloitetaan arvioimaan järjestämänsä koulutusta ja sen vaikuttavuutta. Vaikuttavaksi koulutus on määritelty esimerkiksi silloin, kun *sen tuottamat* valmiudet laadullisesti ja määrällisesti edistävät yksilön henkistä kasvua sekä yhteiskunnan, kulttuurin ja työelämän kehitystä (Opetushallitus 1998, 21; kursivointi kirjoittajien). Esimerkiksi Opetushallituksen arviointimallissa vaikuttavuus on eritelty 1) koulutustarpeen ja tarjonnan vastaavuudeksi, 2) oppimistulosten saavuttamiseksi, 3) oppimaan/ oppimisvalmiuksiksi, 4) kommunikaatiovalmiuksiksi sekä 5) elinikäiseksi oppimiseksi (emt. 33–39).

Ongelmana ei ole se, miten vaikuttavuus jaetaan komponentteihinsa, sillä määritelmä voi muuttua, ja kun yhteiskunnan tarpeet ja odotukset muuttuvat, määritelmiäkin on syytä tarkistaa. Sen sijaan tämän artikkelin kannalta on oleellista se, että määritelmistä riippumatta on eräitä seikkoja, joita voidaan eritellä ja jotka liittyvät siihen mitä järjestelmän tasoa koskevia johtopäätöksiä arvioinnin perusteella tehdään.

Kansainvälisten vertailevien arviointien tuloksien voidaan katsoa kuvastavan koulutuksen nykytilaa, mutta niissä on epäilemättä myös sukupolvien kumuloituvan kulttuurin aineksia. Tulokset hyväksytään sellaisinaan eritoten silloin, kun ne ovat mairit-

televat, niin kuin Suomen PISA-tulosten kohdalla on käynyt. Pulmallisemmaksi tilanne muuttuu, jos poikkileikkausaineistojen perusteella halutaan tehdä johtopäätöksiä yksittäisten koulujen ja oppilaitosten tai maan koulutusjärjestelmän, sen opettajan-koulutuksen tai muun vastaavan hyvin yleisen tason systeemin vaikuttavuudesta. Tässä tulee olla varovainen.

Goldstein ja Spiegelhalter (1996) toteavat esimerkiksi OECD:n PISA-arvioinneista, että tuloskirjan taulukoissa esitetään maittain korjaamattomia tunnuslukuja, joiden arvellaan kertovan (metahuomautus: Goldstein on kriittinen kaikkiin arviointeihin ja esittää siksi tämän PISA asian tällä tavalla. Tämän artikkelin kirjoittajien käsitys on se, että kyllä ne ilmeisesti kertovat) jotain oleellista koulutusysteemien suhteellisesta suoriutumisesta. Ongelmaa voisi liioitellen tai ironisoiden selventää, jos sijaluvun sijaan käytetään käsitettä suoriutuminen, tai vielä ylemmäksi ja mahdollisesti virhe-suuntaan kivuten, suoriutumisen käsitteen sijaan käytettäisiin termiä koulutuksellinen suorituskyky. Lisäksi artikkelimme kannalta mielenkiintoinen on sekä Goldsteinin huomautus, että maiden välisen vertailun kannalta olisi perustellumpaa esittää PISA-tutkimuksen eri alueilta useampia tunnuslukuja kuin vain yksi. Esimerkiksi matematiikan tulokset voitaisiin eritellä monivalintaosioiden ja vapaiden vastausten osioiden keskimääräisinä tasoina. Maiden järjestys muuttuu kun otetaan toinen osaamisen dimensio vertailun perustaksi, koska maiden välillä on eroa siinä, miten niiden oppilaat suoriutuvat erityyppisistä tehtävistä (Goldstein 2004). Tämä kritiikki yllä sopii koulujen vertailuunkin, vaikka Goldstein kirjoittaa kansallisten ”koululaitosten” vertailusta. Eräs seuraus on yksinkertaisesti vain se, että tuloksia kannattaisi selostaa moniulotteisesti niin kansainvälisissä kuin kansallisissa arvioinneissa.

Jotkut arviointituloksia lukevat oikaisevat rohkeasti ja päättelevät, että jos päättövaiheen arvioinnin keskimääräiset tulokset ovat hyviä, koulu on hyvä, jos alle keskitason, koulukin on huono. Esimerkkejä tarjoavat erilaiset koulujen ranking-listaukset, jotka perustuvat ylioppilaskirjoitusten tuloksiin. Äärimmäisyyteen vietyä tulkin-taa edustaa seuraava sitaatti: ”Tutkimuksen mukaan menestyneimmät koulut löytyvät kaupungeista. Kaupunkikoulujen taso on korkeampi kuin taajamissa tai maaseudulla olevien. Pääkaupunkiseudun parhaiden koulujen oppimistulokset ovat myös selvästi parempia kuin muualla maassa.” (Metro 8.9.2004).

Olemme huomanneet, että ei riitä, että lukioiden ”hyvyyttä” arvioidaan virheellisesti vain päättövaiheen oppimistulosten perusteella, vaan lukiot luokitellaan huonoiksi tai hyväiksi suhteellisen aseman perusteella, vaikka lukio olisikin saanut itse asiassa riittävän hyviä oppimistuloksia.

Kun puhutaan koulutuksen tuottamista valmiuksista, oletetaan, että koulutuksen tuotokset voidaan eristää muista lopputulokseen vaikuttavista tekijöistä, esimerkiksi kodin ja lähiympäristön vaikutuksista tai oppilaiden tai opiskelijoiden valikoitumisesta johtuvista eroista. Näiden eri tekijöiden roolin erottelu on mahdollista, mutta se edellyttää koulutuksen tuottaman muutoksen tarkempaa mallintamista ja arvioimista.

Muutosta ei voida arvioida yhden mittauksen perusteella, vaan käytännössä siihen tarvittaisiin vähintään kolme mittaukset. Kahden ensimmäisen avulla arvioidaan sekä keskimääräinen taso tarkemmin että tason muutos, ja kolmas tarvitaan mittaamaan muutoksen pysyvyyttä. Näin ei ole tehty suomalaisissa arvioinneissa ennen kaikkea siksi, ettei koulujen vertailu tai paremmuuslistojen laatiminen ole ollut eikä ole arvioinnin tavoite (Opetushallitus 1998; Jakku-Sihvonen & Kuusela 2002). Goldstein ja Spiegelhalter (1996, p. s. 394) osoittavat, että jos yhdistetään sekä keskiarvotieto että muutostieto, niin koulujen listoittaminen 'oikealla' tavalla on lähes mahdotonta, koska koulujen erilaiset järjestykset vaihtelevat niin paljon: sama koulu ei yleensä sijoitu samalle sijalle kaikissa ulottuvuuksissa ja indikaattoreissa.

Koska koulujen keskimääräiseen tasoon ja niiden eroihin on kiinnitettävissä perusteltua koulutuspoliittista huomiota, on käynyt tarpeelliseksi korvata puuttuva seuranta erilaisin ennusteilla. Lukioiden ylioppilaskirjoitustuloksia on ehdotettu korjattavaksi suhteuttamalla lukioiden tulokset niiden opiskelijoiden perusopetuksen päättötodistusten tasoon.

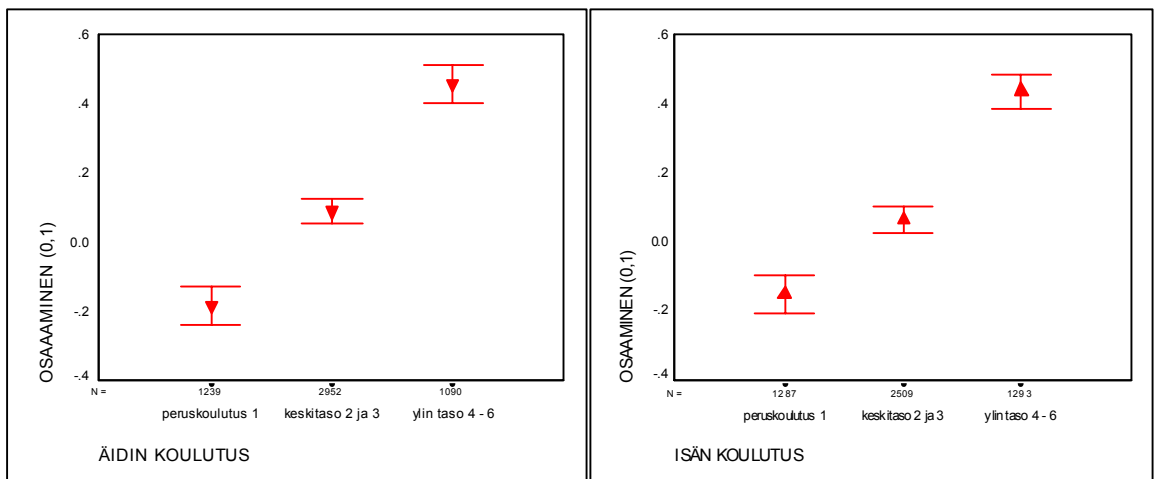
Perusopetuksen kohdalla vertailuissa on sen sijaan pyritty liittämään arvioinnin tulokset oppilaan kotitaustaan käyttämällä apuna erilaista demograafista taustatietoa, jonka käyttö perustuu oletukseen ja aiempaan näyttöön koulutuksen kulttuurisesta periytyvyydestä sekä hyvinvoinnin kumuloitumisesta. Taustatiedot on koottu joko oppilastasolla (esim. oppimaan/ oppimisvalmiuksien arviointi ja kouluterveyskyselyt) tai koulun tasolla käyttäen hyväksi erilaisia koulujen sijaintialueiden ominaisuuksia koskevia aggregaattitietoja (esim. oppimistulosten arvioinnit ja niiden meta-analyysit). Perustuivatpa ennusteet millä tavoin laskettuihin tietoihin tahansa, koulun vaikuttavuus tai "hyvyys" tulee määritellyksi suoran mittauksen sijaan poikkeamaksi ennusteesta. Ennusteisiin liittyvistä epävarmuustekijöistä huolimatta menettely antaa paremmin todellisuutta vastaavan kuvan kuin ainoastaan yhteen mittaukseen perustuvan tuloksen kontekstualisoimaton tarkastelu.

Ekologisten ennusteiden luonteesta

Aggregaattiaineistoja kutsutaan myös ekologisiksi aineistoiksi. Ekologisten yhteyksien yleinen piirre on, että ne nousevat huomattavasti korkeammiksi kuin yksilöitä koskevat ennusteet. Yhteyksien voimistumiselle on tilastollinen selitys, joka perustuu siihen, että kaikki keskiarvosta poikkeava yksilöllinen vaihtelu tulkitaan virhevarianssiksi. Kun virhevarianssi jakautuu satunnaisesti ennusteen molemmiin puoliin, poikkeamat kumoavat toisensa ja esimerkiksi oppimistulokset tiivistyvät keskiarvoiksi. Ylempien tasojen keskiarvo-estimaatit siirtyvät kohti populaation keskiarvoa ja näin lisäävät ennusteen tarkkuutta (Goldstein & Spiegelhalter 1996, p.s. 390). Taustalla on mielestämme järkevä oletus, että instituutiot olisivat periaatteessa samanlaisia, kun kaikki oleelliset tekijät oppilas- ja aluetasolla on otettu huomioon. Oletus liittyy

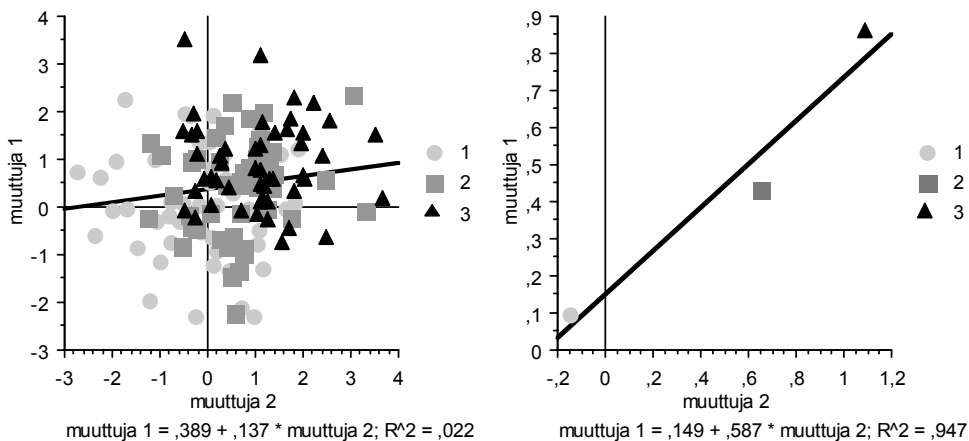
siihen, että suomalaiskoululaisten osaamisen taso on osoittautunut kaikissa kansainvälisissä vertailuissa hyvin tasalaatuisiksi, koulutuksen resurssit on jaettu poikkeuksellisen tasaisesti maan eri osissa, eikä ole perustetta olettaa, että opettajankoulutuksen taso vaihtelisi merkittävästi maan eri opettajankoulutusyksiköissä.

Keskiarvot puolestaan kuvaavat trendejä. Trendit kuvaavat todellisia yhteiskunnassa vallitsevia tilastollisia säännönmukaisuuksia, mutta niiden voimakkuutta yksilöiden elämänkulkua säätelevinä tekijöinä ei voi päätellä pelkän trendiä koskevan tiedon varassa. Havainnollinen esimerkki saadaan vaikkapa raportista ”Oppimaan oppiminen yläasteella 2” (Hautamäki ym. 2003), jonka sivulla 85 on kuvattu äidin ja isän koulutustason yhteys oppilaan testissä osoittamaan osaamiseen. Kolmen luokkaan jaetun vanhempien koulutustason ja oppilaan osaamisen välinen yhteys on lähes täydellisen lineaarinen, mutta vanhempien koulutuksen selitysosuus (eta²) on molemmissa tapauksissa vain vajaat kuusi prosenttia. Keskiarvoihin perustuva kuvaus paljastaa kuitenkin selkeästi trendin: mitä korkeammin koulutettuja vanhemmat ovat, sen paremmin heidän lapsensa ovat keskimäärin selviytyneet testitehtävistä. Tämä on kuitenkin populaatiotason ilmiö, josta ei voida päätellä, että tilanne olisi juuri näin verrattaessa kahden yksittäisen oppilaan tuloksia.



Kuvio 1. Äidin koulutuksen ja isän koulutuksen yhteys osoitettuun osaamiseen (Oppimaan oppimisen arviointi 9. vuosiluokilla) (95 % luottamusvälit)

Esitettäköön selvennykseksi seuraava simulaatio. Kuviossa 2 vallitsee kahden muuttujan välillä heikko positiivinen korrelaatio (vasemmanpuoleinen kuva).



Kuvio 2. Yksilö- ja aggregaattitason ilmiöiden selventäminen simuloidulla aineistolla

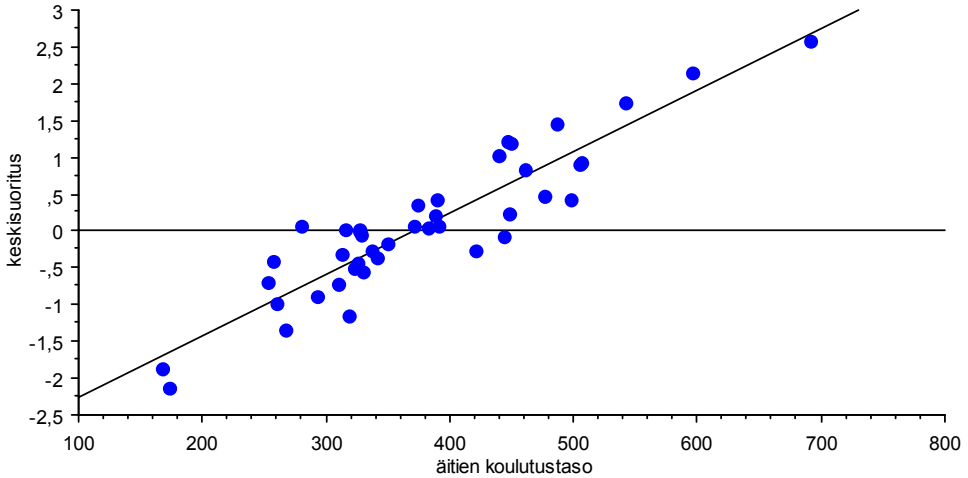
Korrelaatio on heikko, koska kysymys on tietyn menettelytavan mukaan generoiduista satunnaisluvuista. Ensimmäiselle 50 tapaukselle on arvottu kaksi sarjaa normaali-jakautuneita satunnaislukuja, joiden keskiarvo on 0, seuraaville 50 tapaukselle kaksi sarjaa satunnaislukuja, joiden keskiarvo on 0,5 ja kolmannelle kaksi sarjaa, joiden keskiarvo on 1. Kaikkien hajonta on 1. Ryhmien sisällä satunnaisluvut ovat täysin riippumattomia toisistaan, mutta keskiarvojen vaihtelun takia korrelaatioksi tulee 0,15 ja sitä vastaavaksi tilastolliseksi selityosuudeksi runsaat 2 %.

Vaikka yksittäisten tapausten korrelaatio on matalahko, koulujen keskiarvoja tarkasteltaessa satunnaislukujen sisään rakennettu keskeinen tendenssi tulee tietenkin esiin, ja keskiarvojen korrelaatioksi saadaan lähes yksi (oikeanpuoleinen kuva). Vain satunnaisvaihtelu aiheuttaa poikkeamat täydellisestä lineaarisesta korrelaatiosta. Kumpikaan kuvista ei yksinään ole kiinnostava, mutta parina ne havainnollistavat ekologisten yhteyksien luonnetta.

Kuvioparissa ei sinänsä ole mitään mystistä: se kuvaa ainoastaan sitä, mitä se on rakennettukin kuvaamaan. Mystiseksi asia muuttuu vasta, jos tilastolliselle ilmiölle aletaan keksiä pedagogisia selityksiä tai seurauksia. Paradoksaaliseksi tilanteen tekee se, että kuviopari pitää paikkansa, jos mitään ulkoisia häiriötekijöitä ei oleteta olevan, vaan keskiarvoihin vaikuttava satunnaisvaihtelu on ainoa vaihtelun lähde. Koulujen maailmaan sovellettuna erot koulujen opetuksen laadussa tai tasossa olisivat juuri tällainen ”ulkoinen häiriötekijä”, joka rikkoisi keskiarvojen yhteyden.

Kysymys ei ole ainoastaan simulaatioihin perustuvasta efektistä, vaan esimerkiksi Helsingin kouluista saadaan edelliselle teoreettiselle kuviolle empiirinen vastine.

Empiirisessä aineistossa luvut ovat aitojen säännönmukaisten todennäköisyyksien tuotteita. Parhaimmillaan tai pahimmillaan ne käyttäytyvät edellä kuvatun kaltaisesti



Kuvio 3. Äitien keskimääräisen koulutustason ja Helsingin yläasteiden koulujen keskimääräisten oppimistulosten yhteys (4 arviointia)

(Kuusela, 2003). Esimerkiksi vanhempien koulutustason ja lasten oppimistulosten välinen tilastollinen korrelaatio on oppilastasolla yleensä alle 0,3, eikä se voi nousta kovin korkeaksi jo koulutustason karkean luokittelun takia. Sen sijaan koulun tasolla laskettuna vastaava korrelaatio voi kivuta huomattavan korkeiksi ja olla jopa yli 0,9. Jos taustalla on kohtalainen yksilötason tilastollinen yhteys, ekologiset korrelaatiot nousevat sitä korkeammaksi, mitä eriytyneempiä koulut ovat selittävien taustamuuttujien suhteen ja mitä pienemmäksi jäävät opetuksesta aiheutuneet erot. Koulujen tasolla keskiarvojen korrelaatio olisi lähellä nollaa, jos oppilaat olisivat jakautuneet kouluihin satunnaisesti, vaikka yksilötason korrelaatio olisi korkeakin. Koulujen keskiarvojen korrelaatio pienenesi myös olennaisesti, jos koulujen toiminnasta johtuvat erot olisivat merkittävä vaihtelun lähde.

Ekologisen korrelaation tulkinnallisista riskeistä

Kaikkeen korrelatiiviseen tutkimukseen liittyy riski tilastollisten yhteyksien ylitulkinnasta tai kausaalitulkinnoista. Tilastollinen yhteys ei läheskään aina kuvaa syy-yhteyttä. Jos löydämme negatiivisen korrelaation esimerkiksi koulujen oppilaaksiottoalueiden toimeentulotukea saavien lapsiperheiden osuuden ja koulujen keskimääräisten oppimistulosten välillä, emme kuvittelekaan voivamme vaikuttaa oppimistuloksiin manipuloimalla toimeentulotuen myöntämisperusteita. Syy-yhteyden aiheuttavat välittävät tekijät, jotka puolestaan ovat pikemminkin todennäköisyyksiä kuin deterministisiä. Esimerkiksi vanhempien korkea koulutus lisää lapsen koulunkäyntiä tukevi-

en vuorovaikutusmallien todennäköisyyttä, mutta todennäköisyyksien ketju voi myös katketa. Vanhempien korkea koulutustaso ei aina ole koulua tukevien käyttäytymismallien edellytys tai tae, ja oppilas puolestaan saattaa menestyä tai olla menestymättä kodin vaikutuksesta huolimatta. (Keskiarvojen takana olevista jakaumista esimerkiksi Kuusela 2004, s. 19.)

Ylitulkinnan ”jalostuneempi” muoto saattaa olla kysymyksessä silloin, kun tehdään tilastollisin perustein johtopäätöksiä eliittikoulujen tuottamasta lisäarvosta vähemmän koulutettujen vanhempien lapsille. Aito lisäarvon toteaminen edellyttäisi kehityksen seuranta ja kehityksellisen muutoksen toteamista. Kuitenkin käsitykset hyvätaisoisten koulujen tuottamasta lisäarvosta perustetaan lähes yksinomaan korrelatiivisiin tutkimuksiin, jossa selittävänä muuttujana käytetään vanhempien koulutustasoa tai muita sosioekonomista asemaa kuvaavia tunnuslukuja. Ollakseen oikeutettu, päätelmä edellyttäisi joko determinististä oletusta vanhempien koulutustason ja kodissa vallitsevan koulutuksen arvostuksen välillä tai oletusta, että hyvämaineisen koulun lapsilleen valinneet vanhemmat olisivat satunnainen joukko matalasti koulutetuista vanhemmista. Kumpikaan oletus tuskin pitää paikkaansa.

Ekologisiin aineistoihin liittyvänä erityispiirteenä on ekologinen virhepäätelmä, jossa ryhmätason yhteydet yleistetään koskemaan yksilöitä. Empiirinen esimerkki voi havainnollistaa asiaa. Esimerkiksi Helsingin suomenkielisten yläasteen koulujen keskimääräiset oppimistulokset ja yksinhuoltajien osuus koulun oppilaaksiottoalueella korreloivat heikohkon negatiivisesti ($-0,30$). Jos jälkimmäinen olisi ainoa koulujen taustaa koskeva tieto, virheellinen yleistys olisi, että yksinhuoltajien lapset menestyvät keskimäärin muita heikommin. Kun ekologiseen malliin otetaan mukaan matalasti koulutettujen ja toimeentulotukea saavien osuus, yksinhuoltajien osuus saakin positiivisen painokertoimen ($\beta = +0,12$). Siis keskimääräiset oppimistulokset ovat jonkin verran sitä korkeammat, mitä enemmän alueella on yksinhuoltajaperheitä, *kun muut tekijät otetaan huomioon*. Tosiasiassa tällaisen aineiston perusteella ei voida sanoa mitään yksinhuoltajaperheiden lasten koulumenestyksestä kahden huoltajan perheiden lapsiin verrattuna, vaan aineisto kuvaa ainoastaan alueellisia yhteyksiä, ja kuten todettiin, kuvausmuuttujien yhteyden suunta (negatiivinen tai positiivinen yhteys) voi vaihdella siitä riippuen, mitä muuttujia on käytettävissä. Ainoa tapa selvittää edes tilastollinen yhteys olisi tarkastella sitä yksilötasolla tai aitona keskiarvovertailuna.

Uudet tilastolliset menetelmät, esimerkiksi monitasoanalyysi, tarjoavat entistä parempia mahdollisuuksia oppilaiden, luokkien ja koulujen tulosten tarkasteluun ja tarkastelun eri tasojen erottelemiseen toisistaan (Yang, Goldstein, Rath & Hill 1999). Monitasoanalyysijä on Suomessa käytetty eräissä Opetushallituksen arvioinneissa, oppimaanoppimisen arvioinneissa sekä PISA 2000 ja PISA 2003 raporteissa. Tässä emme ole keskittyneet niihin, sillä hienostuneimmatkaan analyysimenetelmät eivät tarjoa täydellistä suojaa päättelyvirheitä vastaan.

Jos halutaan arvioida yksittäisten koulujen vaikuttavuutta, arviointiasetelma on ra-

kennettävä siten, että koulun tuottamat vaikutukset, siis muutokset, todella saadaan esiin. Jos joudutaan turvautumaan ekologisiin tai muihin oppilaiden sosiaalista taustaa kuvaaviin malleihin, tulosten tulkinnoissa on syytä varovaisuuteen. Kolmantena mahdollisuutena on unohtaa tyystin epävarmat johtopäätökset oppimistuloksista koulujen paremmuuden mittarina, jolle poliittinen tahto muutu ja ala edellyttää sellaista koulujen julkista vertailua, jonka aiheuttamia haittoja monissa maissa (esim. Englannissa) parhaillaankin yritetään (toistaiseksi huonoin tuloksin) vähentää. On ilmeistä, että koulujen laaja-alaisen ”hyvyyden” arviointi on alue, jolla kvantitatiivista tutkimusta on täydennettävä monenlaisilla muilla aineistoilla. Silloin myös suuriin otoksiin perustuvat arviointitutkimukset ovat vain osa laajempaa ja monipuolisempaa tarkastelua, joissa pyritään vastaamaan siihen, vastaako kansakunnan koululaitos kansakunnan toiveita ja tulevaisuuden tarpeita.

Metodis-eettisiä periaatteita arvioinnissa: suosituksia

Kate Myers ja Harvey Goldstein (1997) esittävät seuraavat eettiset periaatteet oppilastason mittauksiin perustuvien, mutta koulutusinstituutioiden toimintaan kohdistuvien suoritusindikaattoreiden laatimiselle, hankkimiselle, esittämiselle ja tulkinnalle. Niiden avulla voimme luontevasti päättää lukumme. Periaatteet tukevat ja kokoavat edellä esitettyä, lähinnä periaate 6 kuitenkin uutena, mutta kuuluu kokonaisuuteen.

1. Perusteettoman haitan ehkäisemisen periaate

Indikaattoreiden avulla saatujen tulosten julkaiseminen ei saa aiheuttaa perustelematonta ja tarkoittamatonta haittaa niille kouluille, jotka mahdollisesti voidaan tuloksista yksilöidä. Perustelemattoman tai tarkoittamatottoman haitan käsitettä käytetään, koska joissain tilanteissa voi olla oikeutettua käsitellä jotain aidosti kehnoa tulosta tai tilannetta, jossa jokin yhteiskunnan tukea käyttävä laitos on toiminut velvollisuutensa laiminlyöden.

2. Periaate oikeudesta informaatioon

Jos on olemassa hyväksyttävän tarkkaa ja harhatonta tietoa julkisten instituutioiden suorituksesta, tulee tämän informaation olla saatavilla, edellyttäen että 1. periaate on sopivasti otettu huomioon.

3. Kontekstualisoinnin periaate

Kaikkien suoritusindikaattoreiden tulee pyrkiä antamaan tietoa, jonka avulla instituutioita voidaan arvioida reilulla tavalla. Ei tulisi käyttää sellaisia indikaattoreita, joiden tiedetään olevan riippuvia arvioitavan instituution ulkopuolisista tekijöistä. Yhtään indikaattoria ei tulisi julkistaa ilman, että selostetaan miten se liittyy ekologisiin muuttujiin, ilman että ne on kontekstualisoitu, ja



miten tätä liittämistä ja huomioonottamista voidaan parantaa. Tämä on selostettava ymmärrettävällä ja lukijoille selvästi havaittavalla tavalla (katso seuraava periaate).

4. Epävarmuuden arvioimisen periaate

Kaikki suoritusindikaattorit tulisi selostaa siten, että mukana ovat epävarmuuden tilastolliset arvioinnit. Näiden tulisi ilmentää otantavaihtelua sekä mittareihin ja tilastollisiin ratkaisuihin liittyvää mittavirhettä tai epävarmuutta. Taulukoihin ja kuvioihin tulisi siis liittää aina tilastolliset luottamusvälit.

5. Useiden indikaattoreiden periaate

Yhden indikaattorin rinnalla tai sijaan tulisi mahdollisuuksien mukaan käyttää useita suoriutumismittoja. Näin siksi, että arvioitavien yksiköiden järjestys saattaa vaihdella melkoisestikin, jos ja kun ilmiötä tarkastellaan eri näkökulmasta.

6. Instituution vastineen periaate

Arvioiduilla instituutioilla tulisi olla mahdollisuus esittää oma näkökantansa esitetyn tiedon perusteella ja perusteista. Eritoten tämä koskee yksilöityjä ja siten tunnistettavissa olevia instituutioita koskevia arviointituloksia. Jotta tämä olisi mahdollista, tulee kouluilla tai muilla arvioiduilla laitoksilla olla oikeus siihen tilastoaineistoon, johon tulokset perustuvat. Tietosuoja tulee säilyttää, mutta instituutiolla tulee olla oikeus itsekin tutkia primääriaineistoa.

7. Arviointitietoa julkaisevien toimijoiden vastuullisuuden periaate

Arviointia tekevien laitosten tai yksiköiden tulee kantaa vastuu kaikesta jakamastaan aineistosta ja tuloksista. Lisäksi niiden tulee selostaa metodiset periaatteet ja muut seikat, jotka voivat vaikuttaa tuloksiin ja niiden tulkinnaan.

Alun peruskysymystä – koulun arvioimista oppilaiden kautta – olemme eritelleet tavalla, jota voidaan kutsua kontekstualisoimisen vaatimukseksi. On tilanteita, joissa koulun tulos sellaisenaan on tärkeä. Tuntuu selvältä, että tulosten perusteella vanhemmat saattavat tehdä myös kouluvalintoja: pyrkiessään huolehtimaan parhaalla tavalla perheensä ja lapsensa eduista. Valintoja tehdään myös 7. vuosiluokalle ja lukioon ymmärsä. Poliitikot ovat usein verrattavissa vanhempiin, eivät koulutuksen arvioinnin asiantuntijoihin: uskomusten sopuleita. Kuitenkin kaikkien on syytä tuntea yleisimmät koulujen tuloksiin vaikuttavat tekijät, jotta niiden selostamista ja huomioonottamista vaadittaisiin. Jos vanhemmat, poliitikot ja lehdistö alkavat vaatia korjattuja vertailulukuja, tulee arvioinnin suorittajien niitä antaa. Se johtaa monipuolisempaan aineiston keruuseen. Tällainen olisi kaikkien etu.

Lähteet

- Goldstein, H. 2003. International comparisons of student attainment: some issues arising from the PISA study. Submitted for publication ([www.braintrack.com / Europe / Great Britain / University of London / Institute of Education / Goldstein](http://www.braintrack.com/Europe/Great Britain/University of London/Institute of Education/Goldstein))
- Goldstein, H. & Spiegelhalter, D. 1996. League tables and their limitations: statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society, A* 159, 385–443.
- Hautamäki, J., Arinen, P., Hautamäki, A., Kupiainen, S., Lindblom, Mehtäläinen, J., Niemivirta, M., Rantanen, P., Ruuth, M. & Scheinin, P. 2003. Oppimaan oppiminen yläasteella 2. Tilanne vuonna 2001 ja muutokset vuodesta 1997. Opetushallitus. Oppimistulosten arviointi 6/2003. Helsinki.
- Kuusela, J. 2003. Äidinkielen oppimistulokset Helsingin perusopetusta antavissa kouluissa. Helsingin opetusvirastolle tehty yhteenvetoraportti.
- Kuusela J. 2004. Raportti Helsingin oppimistulosten arviointien ja kouluterveyskyselyjen yhdistämisestä. Helsingin kaupungin opetusviraston julkaisusarja A3:2004.
- Millman, J. (toim.) 1997. Grading teachers, grading schools. Is student achievement a valid evaluation measure? Thousand Oaks: Corwin Press.
- Myers, K. & Goldstein, H. 1997. Failing schools in a falling system. Teoksessa A. Hargreaves (toim.) Rethinking educational change with heart and mind. ASCD 1997 Yearbook. Alexandria, VA: Association for Supervision and Curriculum Development.
- Olson, D. 2003. Psychological theory and educational reform. How school remakes mind and society. Cambridge University Press.
- Opetushallitus 1998. Koulutuksen tuloksellisuuden arviointimalli. Arviointi 7/1998. Helsinki.
- Yang, M., Goldstein, H., Rath, T. & Hill, N. 1999. The use of assessment data for school improvement. *Oxford Review of Education* 25 (4), 469–483.

Kirjoittajaluettelo

Päivi Atjonen, professori, Kasvatustieteen laitos, Joensuun yliopisto,
paivi.atjonen@joensuu.fi

Kathryn Ecclestone, Dr., School of Education, University of Exeter,
K.Ecclestone@exeter.ac.uk

Sven-Erik Hansén, professor, Pedagogiska fakulteten vid Åbo Akademi,
sven-erik.hansen@abo.fi

Jarkko Hautamäki, professori, Koulutuksen arviointikeskus, Helsingin yliopisto,
jarkko.hautamaki@helsinki.fi

Olli Hietanen, aluepäällikkö, Tulevaisuuden tutkimuskeskus,
Turun kauppakorkeakoulu, olli.hietanen@tukkk.fi

Kauko Hämäläinen, johtaja, Koulutus- ja kehittämiskeskus Palmenia, Helsingin
yliopisto, kauko.hamalainen@helsinki.fi

Jari Kaivo-oja, tutkimusjohtaja, Tulevaisuuden tutkimuskeskus,
Turun kauppakorkeakoulu, jari.kaivo-oja@tukkk.fi

Gunnel Knubb-Manninen, huvudplanerare, Sekreteriatet för utbildningsutvärdering,
Jyväskylän yliopisto, gunnel.knubb-manninen@eval.jyu.fi

Esko Korkeakoski, pääsuunnittelija, Koulutuksen arviointisihteeristö,
esko.korkeakoski@eval.jyu.fi

Jorma Kuusela, erikoistutkija, Opetushallitus, jorma.kuusela@oph.fi

Pirjo Linnakylä, professori, Koulutuksen tutkimuslaitos, Jyväskylän yliopisto,
pirjo.linnakyla@ktl.jyu.fi

Maisa Lovio, erityisasiantuntija, Kuntaliitto, maisa.lovio@kuntaliitto.fi

Heikki K. Lyytinen, pääsihteeri, Koulutuksen arviointineuvosto,
hekalyyt@eval.jyu.fi

Jouko Mehtäläinen, pääsuunnittelija, Koulutuksen arviointisihteeristö,
jouko.mehtalainen@eval.jyu.fi

Pentti Nikkanen, erikoistutkija, Koulutuksen tutkimuslaitos, Jyväskylän yliopisto,
pentti.nikkanen@ktl.jyu.fi

Anu Räisänen, pääsuunnittelija, Koulutuksen arviointisihteeristö,
anu.raisanen@eval.jyu.fi

Mari Räcköläinen, projektipäällikkö, Opetushallitus, mari.rakkolainen@oph.fi

Harri Rönholm, suunnittelija, Koulutuksen arviointisihteeristö,
harri.ronholm@eval.jyu.fi

Juha Sihvonen, rehtori, Valkeakoski-opisto, juha.sihvonen@vlk.fi

Heikki Silvennoinen, pääsuunnittelija, Koulutuksen arviointisihteeristö,
heikki.silvennoinen@eval.jyu.fi

Renata Svedlin, bildningschef, Karis stad, renata.svedlin@karis.fi

Jouni Välijärvi, johtaja, professori, Koulutuksen tutkimuslaitos,
Jyväskylän yliopisto, jouni.valijarvi@ktl.jyu.fi

Koulutuksen arviointineuvoston julkaisuja

Nro 1

Koulutuksen arvioinnin uusi suunta

Arviointiohjelma 2004–2007

Nro 2

Utbildningsutvärderingens nya inriktning

Utvärderingsprogram för perioden 2004–2007

Nro 3

New Directions in Educational Evaluation

Evaluation Programme 2004-2007

Nro 4

Bertel Ståhle

Toisen asteen koulu Pohjoismaissa

Toisen asteen koulujen pohjoismainen vertailu

“Pohjoismainen ISUSS-raportti”

Nro 5

Anu Räisänen

EFQM-arviointimalli ammatillisen koulutuksen

järjestäjien arvioinnin tukena

Nro 6

Heikki K. Lyytinen & Anu Räisänen (toim.)

Kehittämissuuntaa arvioinnista

Nro 7

Harri Rönholm & Anu Räisänen (toim.)

Arviointi tukee kehittymistä – miten arvioinnin kehittymistä tuetaan?

Koulutuksen järjestäjien tukeminen arviointiin liittyvissä asioissa

Tilaukset ja tiedustelut:

Osastosihteeri Maritta Leinonen

Koulutuksen arviointisihteeristö

s-posti: maritta.leinonen@eval.jyu.fi

puh. (014) 260 4685

os. PL 35, 40014 Jyväskylän yliopisto

Osa julkaisuista löytyy myös verkosta: www.edev.fi/portal/julkaisu



Koulutuksen
arviointineuvosto

Julkaisussa syvennetään Koulutuksen arviointineuvoston ohjelmaa vuosille 2004–2007. Kehittämissuuntaa tarkastellaan kolmesta näkökulmasta. Uuden suunnan askeleet johdattelevat kansallisen, paikallisen ja kansainvälisen arviointitoiminnan organisointiin ja sovittamiseen yhteen. Arvioinnin uskotaan muuttuvan kehittäväksi uusilla lähestymistapa- ja menetelmävalinnoilla ja korostamalla arvioinnin merkitystä ja hyödyntämistä koulutuksen ydintoiminnan eli opettamisen ja oppimisen kannalta. Tätä arvioinnin kehittävää luonnetta tarkastellaan osassa kaksi. Kolmannessa osassa puolestaan paneudutaan koulutuksen vaikuttavuuden kannalta keskeisimpään eli oppimistulosten arviointiin.

Julkaisun tavoitteena on jäsentää arvioinnin uutta suuntaa ja luoda perustaa arviointitoiminnan suunnan kehittymiselle arviointiasiantuntijoiden verkostona. Julkaisu on tarkoitettu kaikille koulutuksen arvioinnista kiinnostuneille. Siitä on hyötyä erityisesti niille, jotka osallistuvat aktiivisesti arviointitoimintaan joko arvioijina tai työryhmien jäseninä. Julkaisu on myös tarkoitettu päätöksentekijöille sekä koulutuksen kehittäjille.