



Viivi Virtanen, Liisa Postareff & Telle Hailikari

viivi.virtanen@helsinki.fi, liisa.postareff@helsinki.fi, telle.hailikari@helsinki.fi

Millainen arviointi tukee elinikäistä oppimista?

Kirjoituksessamme tuomme esiin, miten oppimisen arviointia tulisi kehittää, jotta se vastaisi paremmin arvioinnin uudellaisiin tavoitteisiin eli opiskelijan oppimisen ja elinikäisen oppimisen valmiuksien tukemiseen. Elinikäistä oppimista voidaan tukea arviointikäytännöillä, jotka harjaannuttavat arvioimaan omaa osaamista, ja tässä itse- ja vertaisarviointi nähdään olennaisina osina arvioinnin kokonaisuutta. Vertais- ja itsearviointien lisääminen luontevaksi osaksi arviointikäytäntöjä olisi tärkeää sekä opiskelijan osaamisen syventämisen että erilaisten työelämävalmiuksien ja elinikäisen oppimisen tukemiseksi. Lisäksi keskeisinä keinoina tuomme esiin kurssin lopussa tapahtuvan summatiivisen arvioinnin täydentämistä ennen kurssia tapahtuvan diagnostisen arvioinnin sekä kurssin aikana tapahtuvan formatiivisen arvioinnin avulla. Arviointikäytäntöjen muutos on hidasta ja edellyttää muutosta akateemisen yhteisön arviointikulttuurissa. Pohdimme kirjoituksessamme myös arvioinnin kehittämiseen liittyviä haasteita.

Asiasanat: arviointikulttuuri, elinikäinen oppiminen, formatiivinen arviointi, itsearviointi, kestävä arviointi, summatiivinen arviointi, vertaisarviointi

How to reform assessment practices for lifelong learning?

In the current paper, we draw attention to the assessment practices employed in higher education. Assessment practices are reportedly in need of change in order to equip students with skills that will enable learning throughout life. Lifelong learning is enhanced by assessment practices that increase student involvement in the assessment and, in this regard, self- and peer-assessment are viewed as potentially valuable approaches. In both of these activities, students are engaging with assessment criteria and standards. Thus, the use of self- and peer-assessment enhances the development of lifelong learning skills by requiring students to make sophisticated judgments about their own or their peers' learning. Recognition of the importance of the different purposes of assessment is also stressed in this paper by the recommendation that summative assessment (assessment of learning) should be complemented by formative and diagnostic assessment (assessment for learning). Achieving change in assessment practices is difficult. Development of sustainable assessment thinking and practice presumes the existence of change processes in the institutional culture of assessment. In this paper, we also discuss the challenges of particular concern in assessment change.

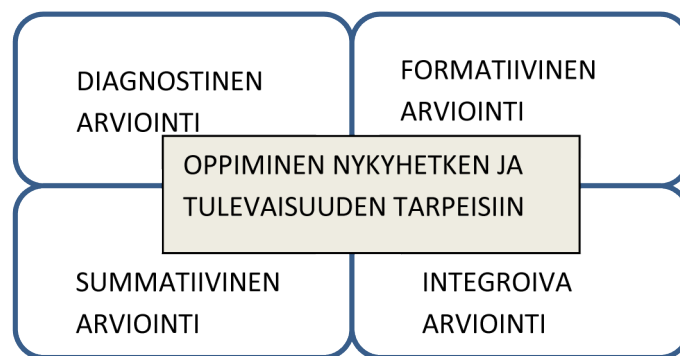
Keywords: culture of assessment, formative assessment, lifelong learning, peer assessment, self-assessment, sustainable assessment

Arviointikulttuurin muutoksen tarve

Elinikäisellä oppimisella tarkoitetaan ihmisen kykyä oppia ja kehittyä läpi koko elämänsä, myös yliopistokoulutuksen tai muun formaalin koulutuksen päätyttyä, sekä työelämässä että muissa yhteisöissä (Boud, 2000). Kearney (2013) käyttää elinikäisestä oppimisesta käsitettä ”kestävä oppiminen”, joka on välttämätön taito tällä vuosituhannella. Elinikäisen oppimisen vaatimus on korostunut yhteiskunnan muututtua ns. riskiyhteiskunnaksi, jossa työurat ovat aiempaa hajanaisempia ja satunnaisempia

ja kilpailu työmarkkinoilla kovempaa (Barnett, 1999). Ihmisen tulee uudistua ja oppia läpi koko elämänsä, jotta hänellä olisi kyky sopeutua jatkuvasti muutuvaan työelämään (Boud, 2000; Kearney, 2013).

Muuttuva ja kehittyvä maailma vaatii yksilöiltä uudenlaisia taitoja ja kykyjä. Opiskelijoilla tulee olla tietojen ja taitojen lisäksi kykyä ajatella kriittisesti, analysoida, ratkaista ongelmia ja toimia yhteistyössä muiden kanssa. Asiantuntijuuden vaatimukset muuttuvat yhteiskunnan ja työelämän muutoksen myötä, ja tämä asettaa myös arvioinnille uudenlaisia haasteita: miten tuottaa asian-



Kuvio 1. Arviointia voidaan kuvata neljän eri tavoitteen avulla (Crisp, 2012, 39)

tuntijoita uusilla työelämätaidoilla varustettuina (Tynjälä, Slotte, Nieminen, Lonka & Olkinuora, 2006, 76). Arviointikäytäntöjen tulisi kehittyä muuttuvan yhteiskunnan mukana vastaamaan näihin uudenlaisiin työelämätarpeisiin ja huomioida elinikäisen oppimisen tarpeet. Arviointitutkijoiden keskuudessa on jo pitkään painotettu arviointikulttuurin tarvetta muuttua mittaamisen kulttuurista oppimisen kehittämisen ja tukemisen kulttuuriin (Birenbaum, 1996; Black, Harrison, Lee, Marshall & William, 2004; Brown, Bull & Pendlebury, 1997). Arviointikulttuurilla tarkoitetaan niitä arvoja ja uskomuksia, jotka vaikuttavat siihen, millaisena akateeminen yhteisö näkee arvioinnin tavoitteet ja tehtävät, ja jotka ohjaavat yhteisön arviointikäytäntöjä (Fuller & Skidmore, 2014). Vahvan kontrolloivan otteen ja opettajajohtoisuuden sijaan tulisi suosia arviointikäytäntöjä, joissa opiskelija on aktiivinen toimija (Boud, 2000; Black ym., 2004; Bryan & Clegg, 2006; Kearney, 2013; Sluijsmans, Dochy & Moerkerke, 1999).

Yliopistokoulutuksen arviointikäytäntöjen kehittymisen on todettu olevan kauttaaltaan hidasta verrattuna muuhun opetuksen kehittämiseen, ja arvioinnin kehittämiseen liittyy paljon muutosvastarintaa (Deneen & Boud, 2014). Omat tutkimuksemme ovat osoittaneet, että arvioinnilla pyritään pikemminkin testaamaan opiskelijoiden osaamista kuin tukemaan heidän oppimistaan tai elinikäisen oppimisen valmiuksiaan (Asikainen, Parpala, Virtanen & Lindblom-Ylänne, 2013; Hailikari, Postareff, Tuononen, Räisänen & Lindblom-Ylänne, 2014; Halinen, Ruohoniemi, Katajavuori & Virtanen, 2013; Postareff, Virtanen, Katajavuori & Lindblom-Ylänne, 2012). Arvioinnin merkitystä oppimisprosessissa ei kuitenkaan tule sivuuttaa. Tutkimus on osoittanut, että arviointi ohjaa opiskelua ja oppimista enemmän kuin mikään muu tekijä oppimistilanteessa (Entwistle & Entwistle, 1992; Hodgson & Pang, 2012; Segers & Dochy, 2006; Struyven, Dochy & Janssens, 2005). Opiskelijat suuntaavat herkästi omaa oppimisprosessiaan arvioinnin mukaisesti, toisin sanoen he opiskelevat sen mukaan, miten heidän suorituksiaan kurssilla arvioidaan (Biggs & Tang, 2007; Brown ym., 1997). Näin ollen arvioinnin huolellinen suunnittelu on opettajalle varteenotettava mahdollisuus edistää opiskelijoiden elinikäisen oppimisen taitoja.

Tässä kirjoituksessa pohdimme yliopistokoulutuksen arviointikäytäntöjä elinikäisen oppimisen näkökulmasta sekä sitä, mihin suuntaan ja miten oppimisen arvioin-

tia tulisi kehittää. Kirjallisuuden perustan muodostavat arvioinnin elinikäisen oppimisen kysymyksiä käsittelevät artikkelit ja sitä täydentävät arvioinnin peruskäsitteitä selittävä kirjallisuus sekä omat suomalaisessa kontekstissa tehdyt tutkimuksemme. Erityisesti pyrimme vastaamaan tutkimuskirjallisuuden pohjalta seuraaviin kysymyksiin: Millaiset arviointikäytännöt tukevat yliopistokoulutuksen keskeistä tavoitetta – elinikäisen oppimisen taitoa? Miksi arvioinnin muuttaminen vaikuttaa olevan vaikeaa ja mitä edellytyksiä muutos vaatii?

Arvioinnin tavoitteet ja tehtävät

Arviointia voidaan jaotella sen perusteella, minkälainen tavoite tai tehtävä sillä on. Diagnostinen, formatiivinen ja summatiivinen ovat yleisesti käytettyjä termejä, joilla kuvataan arvioinnin tavoitteita (Brown ym., 1997; Crisp, 2012). Diagnostisella arvioinnilla pyritään arvioimaan osaamisen taso ennen koulutusta, eli sitä voidaan toteuttaa esimerkiksi ennakkotehtävillä. Formatiivisen arvioinnin keskeinen periaate on, että opiskelija saa oppimisestaan palautetta, jonka avulla hänellä on mahdollisuus kehittää omaa oppimistaan. Tähän tarkoitukseen sopii jatkuva, kurssin aikana tapahtuva arviointi esimerkiksi välitehtävien avulla. Summatiivisella arvioinnilla viitataan arviointiin, jonka tarkoituksena on arvioida, miten hyvin opiskelijat ovat oppineet sitä, mitä oli tarkoitus oppia. Kyse on tällöin loppuarvioinnista. Falchikov (2005) näkee arvioinnilla olevan kaksi keskeistä tehtävää. Kontrolloivan ja valikoivan arvioinnin tehtävänä on varmistaa, että opiskelijat saavuttavat tietyn osaamisen tason ennen siirtymistä opinnoissaan eteenpäin. Jos taas arvioinnin tehtävänä on oppimisen edistäminen, keskeiseksi nousee oppimisen ohjaaminen ja motivointi palautteen avulla ja tätä myötä opiskelijoiden itsearviointitaitojen kehittäminen.

Kaikki koulutuksen aikaiset arviointitapahtumat voitaisiin Crispin (2012) mukaan jakaa niiden tavoitteen mukaisesti neljään eri tyyppiin (kuvio 1). Termit diagnostinen, formatiivinen ja summatiivinen voitaisiin varata kuvaamaan totuttuja näkökulmia, joissa arvioinnin tavoitteena on vaikuttaa opiskelijoiden senhetkiseen oppimiseen. Termiä integroiva voisi käyttää spesifisesti kuvaamaan niitä arviointitapahtumia, joiden ensisijainen tavoite on vaikuttaa opiskelijoihin tulevaisuuden oppi-

joina. Ensimmäisen tavoitteen toteutuminen vaikuttaa käytännössä lähinnä siihen, miten opiskelijat selviytyvät ja kehittyvät summatiivisten arviointien suorituksissa, kun taas jälkimmäiseen kytkeytyvät opiskelijoiden kehityminen oman oppimisensa säätelyssä sekä elinikäinen oppiminen.

Crispin (2012) mallin taustalla on se, että elinikäisen oppimisen edistäminen nähdään yhtenä arvioinnin tehtävänä. Boud (2000, 151) on käyttänyt käsitettä ”kestävä arviointi” (sustainable assessment) arvioinnista, joka vastaa elinikäisen oppimisen tarpeisiin. Elinikäistä oppimista tukevan arvioinnin periaate on, että oppimisesta saadun palautteen tulisi hyödyttää opiskelijoita pitkäkestoisemmin myös tulevaisuuden oppimistilanteissa (Boud, 2014; Hounsell, 2007; Knight, 2002). Arvioinnin tulisi myös pidemmällä tähtäimellä antaa opiskelijoille mahdollisuuksia kehittää kykyään arvioida omaa edistymistään ja tavoitteiden saavuttamista sekä edistää opiskelijoiden luottamusta omiin kykyihinsä. Lisäksi keskeisenä nähdään se, että opiskelijoita tulisi osallistaa mukaan arviointiin, jotta arviointi ja palaute hyödyttäisivät heitä myös pitkäkestoisemmin (Boud & Falchikov, 2006). Arviointiin osallistamisen kautta opiskelijat oppivat arvioimaan omaa suoritustaan ja pohtimaan oman työnsä laatua, mistä on hyötyä valmistumisen jälkeen myös työelämässä (Falchikov, 2005).

Tavoitteesta ja tehtävästä riippumatta oppimisen arvioinnin tulisi olla luotettavaa ja oikeudenmukaista (Brown ym., 1997). Tähän liittyy keskeisesti kaksi käsitettä: arvioinnin reliabiliteetti ja arvioinnin validiteetti. Reliabiliteetilla viitataan siihen, miten tarkasti ja osuvasti arvioidaan juuri sitä mitä on tarkoitus arvioida ja toisaalta sitä, miten johdonmukaista arviointi on. Arvioinnin tulisi olla johdonmukaista ja vapaata sattumanvaraisuuksista ja muista ulkoisista muuttujista, kuten siitä, kuka arvioi ja missä tilanteessa. Esimerkiksi opettajan fyysinen kunto tai stressin aste ei saa vaikuttaa arvioinnin lopputulokseen eikä se, kuka arvioinnin tekee. Kahden eri arvioitsijan tulisi siis päätyä samankaltaiseen arvioon (Sadler, 2009). Arvioinnin validiteetilla puolestaan tarkoitetaan sitä, että arvioinnin tulisi mitata sitä osaamista, mitä sillä on tarkoitus mitata ja vastaavasti arvosanojen tulisi heijastaa luotettavasti osaamisen laatua, laajuutta ja tasoa (Sadler, 2009). Arvioinnin tulisi täyttää nämä keskeiset kriteerit ollakseen luotettavaa ja oikeudenmukaista.

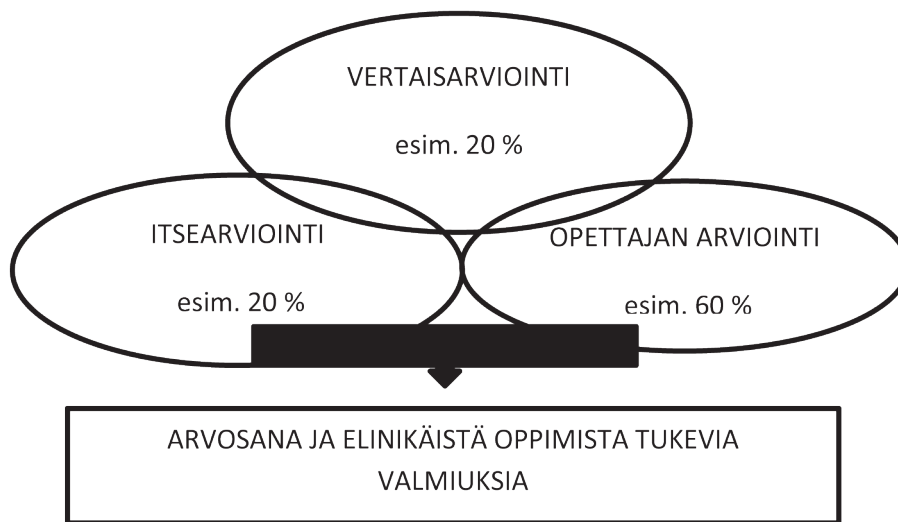
Luotettavuuden ja oikeudenmukaisuuden turvaamiseksi arvioinnin suositellaan perustuvan etukäteen asetettuihin kriteereihin (Brown ym., 1997). Kriteeriperusteisessa arvioinnissa opiskelijoiden osaamista arvioidaan suhteessa etukäteen laadittuihin kriteereihin. Niiden tulisi kuvastaa sitä, mitä opiskelijoiden tulisi osata kurssin jälkeen, eli olla yhteydessä kurssin tavoitteiden kanssa. Konstruktii-visesti linjakkaassa opetuksessa tavoitteita ei tulisi kuvata pelkästään määrittelemällä opittavia sisältöjä, vaan myös tavoiteltava ymmärryksen taso tulisi kuvailla (Biggs & Tang, 2007). Arviointikriteerien avoimuus ja läpinäkyvyys on tärkeää, jotta opiskelijat tietävät, millä perusteilla heitä tullaan arvioimaan. Opiskelijoiden tulisi saada myös tietää, mikä on arvioinnin tavoite ja tehtävä milloinkin (Crisp, 2012).

Elinikäistä oppimista tukevan arvioinnin suunnittelu

Yksittäisen opiskelijan osaamisen arviointi tietyn koulutuksen aikana koostuu lukuisista erilaisista tehtävistä, joiden tavoite ja tarkoitus vaihtelevat tilanteen mukaan. Arviointimenetelmällä tarkoitetaan sitä, millä keinoin ja millaisella tehtävällä osaamista arvioidaan (Brown ym., 1997). Se voi olla esimerkiksi essee, projekti, e-portfolio, tutkielma, monivalinta- tai avoimiin kysymyksiin perustuva kirjallinen kuulustelu, suullinen kuulustelu, esitys tai käytännön demonstraatio. Kyseessä voi olla yksilö-, pari- tai ryhmäsuoritus. Arviointimenetelmään voidaan myös ajatella kuuluvan sen, kuka toimii arvioijana, eli onko kyseessä itse- tai vertaisarviointi vai opettajan tekemä arviointi (Brown ym., 1997). Struyven ym. (2005) toteavat, että on mahdotonta löytää arviointimenetelmä, joka olisi jokaiselle opiskelijalle sopiva. Opiskelijat, joilla on huonot opiskelutaidot ja heikko luottamus omiin kykyihinsä, suosivat niin sanottuja perinteisiä arviointimenetelmiä, kun taas opiskelijat, joilla on kehittyneet opiskelutaidot, suosivat ns. uusia, vaihtoehtoisia arviointimenetelmiä (ks. myös Sambell, McDowell & Brown, 1997). Toisen tutkimuksen (Van De Watering, Gijbels, Dochy & van der Rijt, 2008) mukaan opiskelijat suosivat arviointimenetelmiä, jotka vähentävät heidän stressiään ja tenttiahdustustaan, ja näissä he myös suoriutuvat paremmin.

Arviointimenetelmän suunnittelun taustalla on oltava arvioinnin tarkoitus ja päämäärä: osaamisen diagnosointi, kontrollointi ja valikointi, oppimisen edistäminen tai elinikäisen oppimisen tukeminen (Boud, 2000; Crisp, 2012; Falchikov, 2005). Kysymys voi olla yksittäisen kurssin suunnittelusta tai laajemmasta kokonaisuudesta, mutta arviointimenetelmän tulee tukea oppimistavoitteita (Biggs & Tang, 2007). Arviointimenetelmän valintaan vaikuttavat myös opetuksen käytännön olosuhteet, kuten opiskelijamäärä ja resurssit. Opettajien kokemus kuormittuminen on yksi syy siihen, että oppimista edistäviä arviointimenetelmiä ei oteta käyttöön (Deneen & Boud, 2014) tai niistä luovutaan jonkin ajan päästä (Halinen ym., 2013). Arviointitutkija van der Vleuten (2014) esitti kiinnostavan yhteenvedon siitä, kuinka jokaisella yksittäisellä arviointimenetelmällä on vahvuutensa ja heikkoutensa, kun sitä tarkastelee valideetin, reliabiliteetin tai oppimisen edistämisen näkökulmista: mikään arviointimenetelmä ei sinänsä ole muita parempi, vaan olennaista on arvioinnin kokonaisuus. Arviointimenetelmät olisikin syytä suunnitella laajempien kokonaisuuksien kuin yhden kurssin puitteissa, jotta esimerkiksi opiskelijan etenemiseen vaikuttava summatiivinen arviointi olisi oikeudenmukaista tai jotta elinikäisen oppimisen taidot olisi huomioitu oppimistavoitteissa (van der Vleuten ym., 2012). Olennaista on myös, että opiskelija tietää, mihin tarkoitukseen arviointia milloinkin käytetään: onko arvioinnin tavoitteena diagnosoida senhetkistä osaamista tai kehittää taitoja tulevaisuuden työelämää varten (Crisp, 2012).

Itse- ja vertaisarviointi on arviointikirjallisuudessa nostettu esiin menetelminä, joilla voidaan sekä edistää opiskelijan oppimista että tukea elinikäisen oppimisen taitojen



Kuvio 2. Itse- ja vertaisarviointi osana elinikäistä oppimista tukevia arviointikäytäntöjä. Kuvio on muokattu Kearneyn (2013) artikkelin sekä siinä esitetyn kuvan 1 (s. 281) pohjalta

kehittymistä (Boud, 2000; Kearney, 2013). Taustalla on konstruktivistinen oppimiskäsitys, jossa opiskelijan osallistaminen ja aktivointi nähdään olennaisena oppimiseen vaikuttavana tekijänä. Itse- ja vertaisarvioinnissa opiskelija ei ole pelkästään arvioinnin kohteena, vaan hän osallistuu arvioinnin koko prosessiin ja samalla harjaantuu kehittämään työelämässä tarvittavaa oman osaamisensa arviointia (Boud & Falchikov, 2006; Kearney, 2013; Sadler & Good, 2006). Boud ja Molloy (2013) korostavat, että palautekäytännöt ovat osa arvioinnin prosessia, ja myös niiden tulisi tukea elinikäisen oppimisen taitoja. Opiskelijan tulisi olla palautekäytännöissään aktiivinen osallinen eikä pelkästään palautteen vastaanottaja. Australian yliopistokoulutuksen kehittämisprosessissa itse- ja vertaisarviointi on otettu elinikäistä oppimista tukevan arvioinnin osiksi opettajan tekemän arvioinnin rinnalla (Kearney, 2013; kuvio 2). Esitetyssä mallissa arviointi palvelee sekä summatiivista tarkoitusta että elinikäisen oppimisen tarpeita: opettajaopiskelijat pääsevät harjoittelemaan taitoja, joita he myöhemmin tarvitsevat toimiessaan opettajina. Kearney (2013, 281) kutsuu tätä mallia autenttiseksi, oppimista edistäväksi itse- ja vertaisarvioinniksi.

Vertaisarviointi arvosanan määrittäjänä ja palautteen antajana

Tutkimusten mukaan vertaisarviointi soveltuu hyvin sekä formatiiviseen että summatiiviseen tarkoitukseen (Falchikov & Goldfinch, 2000; Nulty, 2011; Topping, 1998). Tutkimukset eri tieteenaloilta ja sekä alku- että loppuvaiheen opinnoista ovat osoittaneet, että vertaisarvioinnissa opiskelija pääsee osalliseksi arviointiprosessista ja että opiskelijoiden aktiivisuuden kautta myös oppimisen laatu paranee (Asikainen, Postareff, Heino & Virtanen, 2014; Davey & Palmer, 2012; Gouli, 2008; McGarr & Glifford, 2013). Arvioimalla toisten opiskelijoiden suorituksia opiskelija samalla perehtyy oman tieteenalansa tiedon laadun kriteereihin ja oppii vertaisiltaan. Vertaisarviointi on nähty

keinona sisällyttää opintojaksoihin työelämätaitojen harjoittelua kuten palautteen antamista ja vastaanottamista, ja se auttaa opiskelijaa pohtimaan kriittisesti omia taitojaan, tietojaan ja asiantuntijuuttaan omalla tieteenalallaan. Näin ollen vertaisarviointi antaa valmiuksia elinikäiselle oppimiselle. Opettajan näkökulmasta vertaisarviointia voi pitää tehokkaana keinona antaa palautetta suurelle opiskelijajoukolle ja tehostaa yksittäisen opettajan ajankäyttöä (Asikainen ym., 2014; Davey & Palmer, 2012).

Vertaisarvioinnin luotettavuus ja oikeudenmukaisuus

Vertaisarvioinnin yhteydessä herää välittömästi kysymys ja epäily sen luotettavuudesta ja oikeudenmukaisuudesta arviointimenetelmänä (mm. Patton, 2012). Sekä opettajia että opiskelijoita saattaa epäilyttää, onko vertainen pätevä arvioimaan toisen opiskelijan suoritusta vai pitäisikö arviointi jättää alan asiantuntijalle (Carvalho, 2013; Vu & Dall'Alba, 2007). Useat tutkimukset ovat kuitenkin osoittaneet, että opiskelijoiden antamat arvosanat eivät poikkea merkittävästi opettajien antamista arvosanoista (Asikainen ym., 2014; Davey & Palmer, 2012; Falchikov & Goldfinch, 2000; Nulty, 2011). Luotettavuuden mittarina on käytetty opiskelijoiden ja opettajien antamien arvosanojen yhtäläisyyttä, ja opettajan tekemä arvio on lähtökohtaisesti tulkittu luotettavaksi.

Esseiden vertaisarviointia käsittelevissä tutkimuksissa nousi esiin, että opiskelijoiden oli helpompi arvioida esseiden rakennetta ja muita kuin esseiden sisältöä (Lindblom-Ylänne, Pihlajamäki & Kotkas, 2006; Mostert & Snowball, 2013). Samantyyppinen tulos saatiin tutkimuksessa, jossa insinööriopiskelijat antoivat sanallisista tehtävistä keskimäärin korkeampia arvosanoja kuin opettajat, mutta numeeristen tehtävien kohdalla opiskelijoiden tekemä arviointi oli yhtenevä opettajien kanssa (Davey, 2011). Falchikovin ja Goldfinchin (2000) meta-analyysi sisälsi 48 kvantitatiivisen tutkimuksen tulokset. Yhteenvedossään he toteavat, että vertaisarvioinnin luotettavuus paranee, kun arviointi kohdistuu tuotoksiin

tai prosessiin, ei niinkään ammatillisiin käytänteisiin kuten opetustaitoon tai potilastyöhön. Myös selkeät arviointikriteerit ja niiden käsittely yhdessä opiskelijoiden kanssa parantavat vertaisarvioinnin luotettavuutta.

Opiskelijoiden kokemuksia vertaisarvioinnista

Opiskelijoiden kokemukset vertaisarvioinnista ovat tutkimusten mukaan pääosin myönteisiä ja opiskelijat arvostavat esimerkiksi sitä, miten paljon itse oppii arviointiin osallistumalla (Davey & Palmer, 2012; McGarr & Clifford, 2013; Vu & Dall'Alba, 2007). Vertaisarvioinnin avulla oppii ymmärtämään, mihin arviointi perustuu, millaista osaamista hyvä arvosana edellyttää ja miten hyvä vastaus tai muu suoritus rakentuu. Samalla saa palautetta osaamisestaan ja oppii itse arvioimaan oman osaamisensa tasoa suhteessa kriteereihin. Arviointia tehdessään, perehtymällä toisten opiskelijoiden vastauksiin tai muihin suorituksiin, opiskelija syventää omaa ymmärrystään käsitelystä aiheesta (Vu & Dall'Alba, 2007). Esimerkiksi Daveyn (2011) tutkimuksessa opiskelijat kokivat, että vertaisarviointi auttoi tehokkaasti oppimisessa ja herätti kiinnostusta sisältöön. Arviointikriteerit ja niihin tutustuminen olivat opiskelijoiden mielestä ehdoton edellytys arvioinnin sujumiselle. Vaikuttaa siltä, että arviointiprosessin ja arviointikriteerien käsittely yhdessä opiskelijoiden kanssa on olennaista siltä kannalta, miten positiivisena ja hyödyllisenä opiskelijat kokevat vertaisarvioinnin (Lindblom-Ylänne ym., 2006; Vu & Dall'Alba, 2007; Welsh, 2007).

Kaikissa opiskelijoiden kokemuksia käsittelevissä tutkimuksissa tulee kuitenkin esiin myös opiskelijoiden negatiivisia tuntemuksia. Opiskelijat saattavat epäillä etukäteen, onko vertaisarvioinnista hyötyä, onko vertainen pätevä ja oikeudenmukainen arvioimaan ja onko itse pätevä arvioimaan muita, mutta voivat arvostaa kokemusta jälkepäin (Vu & Dall'Alba, 2007). Lisäksi he saattavat kokea, että koko vertaisarviointiprosessi vie liikaa aikaa (Asikainen ym., 2014). Opiskelijat saattavat myös epäillä vertaisarvioinnin käyttöä loppuarvosanan määrittämisessä, vaikka arvostaisivat sitä muuten (Patton, 2012). Kysymys arvioinnin oikeudenmukaisuudesta on usein keskiössä opiskelijoiden negatiivisissa kokemuksissa. Carvalho (2013) tarkasteli yliopisto-opiskelijoiden kokemuksia tietyllä kauppatieteen kurssilla seitsemän perättäisen lukukauden ajan. Opiskelijoista (N = 120) 70 % piti vertaisarviointia myönteisenä ja 30 % negatiivisena. Negatiiviset kokemukset olivat yhteydessä siihen, että opiskelijat kokivat saamansa arvosanan epäoikeudenmukaiseksi. Daveyn ja Palmerin (2012) tutkimuksessa verrattiin kahden eri opiskelijajoukon antamien arvosanojen samankaltaisuutta opettajan antamien arvosanojen kanssa; toisessa ryhmässä opiskelijoilla oli käytössään mallivastaukset ja toisessa ei. Tulos osoitti, että vertaisarviointi oli yhtä luotettavaa molemmissa ryhmissä, mutta opiskelijaryhmät kokivat vertaisarvioinnin eri tavalla. Opiskelijat, joilla ei ollut mallivastauksia, kokivat arvioinnin haastavammaksi ja suhtautuivat aavisituksen varautuneemmin ajatukseen, että vertaisarviointia käytettäisiin jatkossa muilla kursseilla. Opiskelijoiden negatiiviset kokemukset eivät siis välttämättä vaikuta arvioinnin luotettavuuteen, mutta niillä on kuitenkin vaikutusta siihen, miten hyvin opiskelijat sitoutuvat arvioimaan toistensa suorituksia.

Vertaisarviointi käytännössä

Opintojen vaihe, oppimisympäristö tai ryhmän koko ei tutkimusten mukaan vaikuta vertaisarvioinnin luotettavuuteen (Asikainen ym. 2014; Davey, 2011; Mostert & Snowball, 2013; Nulty, 2011; Vickerman, 2009). Eroja ei ole myöskään havaittu, kun vertaisarvioinnin kohteena on ollut essee, loppuentti tai ryhmätyö (Davey, 2011; De Wever, Van Keer, Schellens & Valcke, 2011; Lindblom-Ylänne ym., 2006). Vertaisarvioinnin käytöstä luotettavana arviointimenetelmänä on näyttöä ainakin humanistisista tieteistä ja luonnontieteistä sekä fysioterapia-, insinööri-, opettaja- ja sairaanhoitajakoulutuksesta (Asikainen ym., 2014; Falchikov & Goldfinch, 2000; Davey & Palmer, 2012; McGarr & Clifford, 2013; Sluijsmans & Prins, 2006; Welsh, 2007).

Useissa edellä mainituissa tutkimuksissa tuotiin esiin, että vertaisarviointi sopii parhaiten kvantitatiivisesti mitattavien suoritusten arviointiin, kun kyseessä on arvosanaan vaikuttava arviointi. Voisikin sanoa, että sen käyttöönotto on sekä opettajalle että opiskelijoille helpompaa eikä aiheuta ongelmia oikeudenmukaisuuden ja luotettavuuden suhteen, mikäli kyse on oikein-väärin-suoritusten arvioinnista (Asikainen ym., 2014; Davey, 2011; Falchikov & Goldfinch, 2000). Mikäli opiskelijoilla on käytössään selkeät kriteerit, kuten arviointimatriisi tai jonkinlaiset mallivastaukset, opiskelijoiden antamat arvosanat eivät poikkea merkittävästi opettajien antamista arvosanoista. Sen sijaan prosessien, työelämätaitojen tai laadullisten tuotosten loppuarvosanan arvioimisen suhteen vertaisarvioinnin onnistuminen vaatii opettajalta huolellisempaa suunnittelua (Vu & Dall'Alba, 2007). Opiskelijoiden perehdyttäminen ja osallistaminen arvioinnin tavoitteisiin ja toteuttamiseen on tällöin keskeistä (Bloxham & West, 2004; O'Donovan, Price & Rust, 2004).

Itsearviointin hyödyntäminen

Mikäli arvioinnin tehtäviin nähdään kuuluvan elinikäisen oppimisen tukeminen, itsearviointi on luonteva osa arviointikäytäntöjä: itsearviointi on aina oppimistilanne, jossa opiskelija itse arvioi oman osaamisensa tasoa ja usein myös reflektoi omaa oppimisprosessiaan (Andrade & Du, 2007; Falchikov & Boud, 1989). Kun opiskelija toimii itse arvioijana, hän saa käsityksen oman tieteenalansa osaamisen laadun standardeista ja oppii suhteuttamaan oman osaamisensa tason käytössä oleviin kriteereihin (Boud, 2000, 161). Parhaimmillaan itsearviointi kannustaa opiskelijaa omaksumaan syvällistä ymmärrystä tuottavan lähestymistavan oppimiseen ja edistää opiskelijan kykyä säädellä itse oppimistaan (Ozogul & Sullivan, 2007).

Kyky arvioida omaa osaamistaan ei kehity ilman harjoittelua, joten opiskelijoita tulee osallistaa arviointiprosessiin, käydä läpi arviointikriteereitä ja perustella, miksi itsearviointia tehdään (Kearney, 2013; Rust, Price & O'Donovan, 2003). Arviointia tulee myös harjoitella koko koulutuksen ajan (De Grez, Valcke & Roozen, 2012; Topping, 1998). Noviisi ei tunne tieteenalan käytäntöjä ja tapaa ajatella

eikä noviisilla ole asiantuntijan tietotasoa, joten oman osaamisen arviointi voi olla haastavaa. Meta-analyysi itsearviointia käsittelevistä tutkimuksista toikin esiin sen, että kokeneemmat opiskelijat olivat arvioinneissaan tarkempia kuin aloittelevat opiskelijat (Falchikov & Boud, 1989). Tutkimuksessa tarkasteltiin opiskelijoiden itsearviointia ja opettajan tekemän arvioinnin yhdenmukaisuutta ja havaittiin, että yhdenmukaisuutta vahvisti edellä mainittujen tekijöiden lisäksi arviointiprosessin huolellinen suunnittelu. Myös tieteenalojen välillä oli eroja siten, että luonnontieteissä itsearviointi oli yhdenmukaisempaa opettajan tekemän arvioinnin kanssa kuin muissa tieteissä. Omat tutkimuksemme farmasian ja biotieteen aloilta osoittivat, että itsearviointin käyttö ainakin tutkimuskohteena olleissa yhteisöissä oli erittäin vähäistä (Halinen ym., 2013; Postareff ym., 2012).

Arviointikriteerien merkitys

Arviointikriteerien ja opiskelijoiden arviointiin osallistamisen avulla on mahdollista suunnata opiskelijoiden oppimista haluttuun suuntaan (Handley & Williams, 2011; Rust ym., 2003). Keskeistä olisi keskustella opiskelijoiden kanssa arviointikriteereistä tai jopa osallistaa heitä kriteerien laatimiseen, jotta he todella ymmärtävät kriteerit ja voivat käyttää niitä myös itse- ja vertaisarviointin tukena. Arviointikriteerit helpottavat myös opettajan työtä opiskelijoiden arvioinnissa, sillä kriteerit helpottavat tulkintojen ja päätösten tekoa ja usein myös nopeuttavat arviointia (ks. esim. Hailikari ym., 2014; Sadler, 2005).

Arviointikriteerien käytöllä on kuitenkin myös kääntöpuolensa. Liian tarkat ja yksityiskohtaiset kriteerit saattavat estää opiskelijoiden kokonaisvaltaisen osaamisen arvioinnin. Aina ei myöskään ole mahdollista etukäteen määritellä, mitä opintojakson aikana opitaan. Tällöin ei ole mielekästä hyödyntää arvioinnissa tarkasti etukäteen määriteltäviä ja yksityiskohtaisia kriteereitä. Tällaisia kriteereitä on kritisoitu myös siitä, että ne eivät vastaa asiantuntijatyön käytäntöjä. Asiantuntijan työtä arvioidaan usein moniulotteisilla ja kokonaisvaltaisilla kriteereillä, jolloin mahdollistuu myös sellaisen osaamisen arviointi, jota ei ole etukäteen voitu määritellä (Sadler, 2009). Joka tapauksessa arviointikriteereistä keskusteleminen opiskelijoiden kanssa on keskeistä, jotta niitä voitaisiin mielekkäällä tavalla käyttää itse- tai vertaisarviointin tukena ja jotta opiskelija voisi peilata omaa osaamistaan suhteessa kriteereihin.

Arvioinnin kehittämisen keskeiset haasteet

Arviointikäytäntöjen muuttaminen on ilmeisen haasteellista, koska yhteisön arviointikulttuuri eli taustalla vaikuttavat käsitykset, arvot ja uskomukset vaikuttavat olennaisesti siihen, millaiset arviointikäytännöt menetelmienä ovat yleisesti hyväksytyjä (Biggs & Tang, 2007; Deneen & Boud, 2014). Arviointikulttuuri on vahvasti kytköksissä akateemisen yhteisön käsityksiin ja kokemuksiin arvioinnista. Tutkimuksen mukaan opettajien käsitykset arvioinnista vaikuttavat siihen, miten he arvioivat opiskelijoidensa oppimista (Postareff ym., 2012; Samuelowicz & Bain, 2002; Watkins, Dahlin & Ekholm, 2005), ja mitä tutumpi arvioin-

timenetelmä on opiskelijalle, sitä mielekkäämmäksi hänen sen kokee (Stryuven ym. 2005).

Deneen ja Boud (2014) ovat tutkineet tarkemmin arvioinnin kehittämiseen liittyvää muutosvastarintaa opetushenkilökunnan keskuudessa. He havaitsivat, että muutoksen vastustaminen johtuu usein siitä, että ei ymmärretä uudenlaisten arviointikäytäntöjen merkitystä. Toisaalta opettajat saattavat kokea, että he eivät voi omin päin lähteä muuttamaan arviointikäytäntöjään vaan että muutosta pitäisi johtaa ylemmältä tasolta. Kolmanneksi vastustusta aiheuttavat opettajien kiire ja suuri työmäärä, jolloin aikaa arvioinnin kehittämiseen ei tunnu löytyvän. Arvioinnin kehittäminen tulisikin nähdä koko yhteisön yhteisenä tehtävänä, aivan kuten opetuksen kehittämisen suhteen on monissa instituutioissa toimittu. Instituution arviointikulttuurin kehittäminen vaikuttaa myös siihen, miten myönteisesti opiskelijat suhtautuvat erilaisiin arviointikäytäntöihin (Segers & Tillema, 2011).

Opettajat näkevät edelleen usein arvioinnin kurssin lopussa tapahtuvana irrallisena osana opetus-oppimisprosessia, jolloin arviointi ja palaute eivät linkity luontevaksi, oppimista ohjaavaksi ja edistäväksi toiminnaksi (Hailikari ym., 2014; Parpala & Lindblom-Ylänne, 2007; Postareff ym., 2012). Arvioinnin näkeminen irrallisena osana opetus-oppimisprosessia ja opittavan sisällön määrän painottaminen oppimisen laadun sijaan johtavat herkästi summatiivisten ja tietoa toistavien arviointimenetelmien käyttöön. Mikäli arviointi koetaan osaksi opetus-oppimisprosessia ja oppiminen nähdään tiedon rakentamisena ulkokohtaisen oppimisen sijaan, on todennäköistä, että opettajan arviointikäytännöissä korostuvat formatiivinen arviointi sekä tiedon soveltamista ja ongelmanratkaisua painottavat arviointikäytännöt (Postareff ym., 2012; Samuelowicz & Bain, 2002; Watkins ym., 2005). Opettajan käsitysten muuttaminen on haastavaa.

Myös opiskelijat pitää saada hyväksymään uudenlaiset arviointimenetelmät. Opiskelijoiden käsitykset arvioinnin oikeudenmukaisuudesta ovat yhteydessä kokemukseen siitä, vaikuttaako arviointi myönteisesti heidän oppimiseensa (Segers, Dochy & Cascallar, 2003; Sambell, McDowell & Brown, 1997). Usein uudet, vaihtoehtoiset arviointimenetelmät on nähty keinona vaikuttaa opiskelijoiden oppimiseen myönteisesti ja lisätä kokemusta arvioinnin oikeudenmukaisuudesta (Segers, Dochy & Gijbels, 2010). Tosin tutkimustulokset tähän liittyen ovat jotakuinkin ristiriitaisia. Sambell ym. (1997) osoittivat, että opiskelijat kokivat perinteisen, kurssin lopussa tehtävän arvioinnin olevan epäluotettava osaamisen mittari ja liian riippuvaisista ulkoisista tekijöistä, kuten fyysisestä vireystilasta tai tenttiahdistuksesta. Toisaalta Moge ym. (2007) osoittivat, että opiskelijat olivat huolissaan myös vaihtoehtoisten arviointimenetelmien oikeudenmukaisuudesta. Samansuuntaisia tutkimustuloksia saivat myös Segers ja Dochy (2001) itse- ja vertaisarviointin käytöstä arviointimenetelmänä. Yhteenvetona voidaan sanoa, että jokaiseen arviointimenetelmään liittyy tiettyjä haasteita ja on vaikeaa löytää arviointimenetelmää, jonka kaikki opiskelijat kokisivat luotettavaksi ja oikeudenmukaiseksi. Opiskelua ja oppimista voidaan siis suunnata arvioinnin avulla, mutta tämä ei kuitenkaan ole lainkaan yksiselitteistä: siihen, mi-

ten opiskelijat reagoivat tiettyyn arviointiin, vaikuttavat lukuisat tekijät, kuten sopiva haastavuustaso, muut kurssit ja niiden arviointimenetelmät sekä opiskelijoiden kokemus kuormitus ylipäättään (Cilliers ym., 2010; Gijbels, Segers & Struyf, 2008; Postareff, Lindblom-Ylänne & Parpala, 2014).

Haastavaa arvioinnin kehittämiseksi on myös se, että arviointikriteerejä ei välttämättä käytetä, vaan yliopisto-kontekstissa esiintyy edelleen normiperusteista arviointia: opettajat saattavat vertailla opiskelijoiden suorituksia toisiinsa tai muuttaa annettuja pisteitä jälkepäin tehdäkseen eroa muiden suorituksiin tai antaakseen tasaisesti eri arvosanoja opiskelijoille (Bloxham, Boyd & Orr, 2011; Hailikari ym., 2014). Arviointikriteerien puuttuessa arviointiin vaikuttavat herkemmin satunnaistekijät kuten esimerkiksi opettajan tai muiden arvioitsijoiden väsymys ja mielen-tila (Hailikari ym., 2014), ja opiskelijan suoriutumista verrataan herkästi muiden opiskelijoiden suoriutumiseen (Brown ym., 1997). Opettajalla saattaa olla arviointi-kriteerit, mutta ne perustuvat arvosanajakaumaan eli opiskelijoiden järjestämiseen suhteessa toisiinsa. Näin ollen normiperusteinen arviointi saattaa lisätä opiskelijoiden keskinäistä kilpailua ja estää kannustavan ja positiivisen ilmapiirin syntymistä (Sadler, 2005, 186–187). Normi-perusteinen arviointi ei myöskään tue elinikäisen oppi-misen edistämisen tavoitetta, sillä opiskelija ei voi peilata omaa osaamistaan suhteessa olemassa oleviin kriteereihin.

Johtopäätökset

Viimeaikainen tutkimus on nostanut esille kestävän arviointin merkityksen, jolloin arvioinnin tavoitteena on opiskelijoiden elinikäisen oppimisen valmiuksien ja itsearviointitaitojen tukeminen (Boud, 2014; Crisp, 2012). Taustalla on ollut huoli siitä, että nykyiset arviointikäytännöt eivät tue opiskelijoiden työelämävalmiuksien eli elinikäisen oppimisen keskeisten taitojen kehittymistä. Kirjallisuuden mukaan elinikäistä oppimista voidaan tukea arviointikäytännöllä, jotka harjaannuttavat arvioimaan omaa osaamista, ja tässä itse- ja vertaisarviointi nähdään olennaisina osina arvioinnin kokonaisuutta (Boud & Falchikov, 2006; Crisp, 2012; Kearney, 2013). Arvioimalla itse joko omaa osaamistaan tai vertaisensa osaamista saa käsityksen oman tieteenalansa osaamisen laadun standardeista ja oppii suhteuttamaan oman osaamisensa tason käytössä oleviin kriteereihin. Keskeisenä nähdään se, että opiskelijoita tulisi osallistaa mukaan koko arviointiprosessiin ja erityisesti palautekäytäntöjen tulisi kehittyä, jotta arviointi ja palaute hyödyttäisivät opiskelijoita pitkäkestoisemmin (Boud & Falchikov, 2006; Boud & Molloy 2013). Niinpä vertais- ja itsearviointin lisääminen luontevaksi osaksi arviointikäytäntöjä olisi tärkeää sekä opiskelijan osaamisen syventämisen että erilaisten työelämävalmiuksien ja elinikäisen oppimisen tukemiseksi.

Arvioinnin kehittämisen tarve ei kuitenkaan kohdistu yksittäisiin arviointimenetelmiin, vaan tarve olisi arviointikulttuurin muuttumiselle. Tutkijat ovat korostaneet arviointikulttuurin muutoksen tarvetta jo 1990-luvulta lähtien (mm. Brown ym., 1997). Vallitseva käsitys arviointitutkijoiden keskuudessa on, että osaamisen kehittämiseen perustuvan arviointikulttuurin avulla voidaan

vastata paremmin muuttuviin yhteiskunnan vaatimuksiin (Boud, 2000; Bryan & Clegg, 2006; Sluijsmans ym., 1999). Osaamisen kehittämiseen perustuvassa näkemyksessä arviointi nähdään pikemminkin mahdollisuutena edistää opiskelijoiden oppimista kuin mitata osaamista. Tällöin kysymys on erityisesti siitä, miten arviointi voidaan suunnitella siten, että se toimii opiskelijoille samalla oppimistilanteena ja oppimisen edistäjänä. Arvioinnin erilaisten tavoitteiden ja tehtävien ymmärtäminen on olennaista, jotta arviointikulttuuri voisi kehittyä elinikäistä oppimista tukevan arvioinnin suuntaan (Crisp, 2012).

Olemme tuoneet esiin, että omien tutkimustemme mukaan vallitsevat arviointikäytänteet ovat vielä melko perinteisiä, ja niiden valossa elinikäistä oppimista edistävän arvioinnin käytännöissä on vielä runsaasti kehittämisen varaa (Halinen ym., 2013; Postareff ym., 2012). On kuitenkin huomattava, että tutkimuskonteksti on melko rajattu eikä Suomessa ole tehty kattavaa tutkimusta yliopistokoulutuksen arviointikäytännöistä ja siitä, miten ne tukevat elinikäistä oppimista. Tuloksemme ovat kuitenkin samansuuntaisia muun arviointitutkimuksen kanssa: arviointi on muuttunut melko vähän verrattuna siihen, miten yliopistojen opetusta on kehitetty (ks. esim. Deneen & Boud, 2014, 589). Arviointikulttuuri heijastuu opettajien käsityksiin arvioinnista, ja nämä käsitykset puolestaan ohjaavat voimakkaasti arviointikäytäntöjä (Postareff ym., 2012; Samuelowicz & Bain, 2002; Watkins ym., 2005). Näin ollen syntyy herkästi ”perinteisen arvioinnin kehä”, jossa yhteisön arviointikulttuuri ilmenee opettajien arviointikäsitteissä ja sitä myöden perinteiset arviointikäytänteet juurtuvat myös käytännön tasolle. Nämä käytänteet puolestaan entisestään vahvistavat perinteistä arviointikulttuuria.

Arvioinnin kehittäminen ja pysyvät muutokset arviointikäytännöissä edellyttävätkin koko yhteisön avointa keskustelua arvioinnin tehtävistä ja tavoitteista (Deneen & Boud, 2014; Segers & Tillema, 2011). Arviointia tulisi siis suunnitella ja kehittää esimerkiksi koulutusohjelmatasolla, jotta koko yhteisö saadaan sitoutettua arvioinnin kehittämiseen (van der Vleuten ym., 2012). Muutosvastarintaa arvioinnin kehittämisen suhteen ilmenee usein juuri sen takia, että opettajat kokevat, etteivät he voi yksinään tehdä muutoksia arviointiin, ja muutosta tulisikin johtaa ylemmältä tasolta (Deneen & Boud, 2014). Toinen keskeinen muutosvastarinnan syy liittyy siihen, että opettajat eivät ole välttämättä tietoisia arvioinnin uusista tavoitteista ja niitä tukevista arviointikäytännöistä. Kolmas haaste liittyy ajan ja resurssien puutteeseen. Vaikuttaakin siltä, että opettajien pedagogisen tietoisuuden lisääminen sekä koulutuksen johtaminen ovat arviointikäytäntöjen kehittämisen kannalta avainasemassa. Myös opiskelijoiden osallistaminen arvioinnin kehittämiseen olisi olennaista, jotta opiskelijat saataisiin ymmärtämään arvioinnin tarkoitus ja hyväksymään uudet arviointikäytännöt.

Systemaattinen ja suunnitelmallinen lähestymistapa arviointiin, jota toteutetaan esimerkiksi joissain lääketieteen koulutusohjelmissa (ks. van der Vleuten ym. 2012), vaikuttaa kaiken kaikkiaan olevan harvinaista laajemmassa mittakaavassa. Tarve olisi nykyisten arviointikäytäntöjen kattavammalle kartoitukselle, jotta saataisiin tutkimus-

tietoa siitä, missä määrin elinikäistä oppimista tukevia arviointikäytäntöjä on jo nyt käytössä. Jatkossa olisi kiinnostavaa tutkia suomalaisten yliopistojen kontekstissa sitä, saataisiinko opiskelijoiden elinikäisen oppimisen taitoja kehitettyä, mikäli arvioinnin tavoitteet ja arviointikäytännöt suunniteltaisiin osana laajemman opetus suunnitelman kokonaisuutta.

Kiitokset

Kiitämme rakentavista kommentteista Helsingin yliopiston valtiotieteellisessä tiedekunnassa keväällä 2014 järjestetyn YP4 Oppimisen arviointi ja palautteen antaminen -kursille osallistuneita kurssilaisia. Kiitämme myös dosentti, opetusdekaani Mirja Ruohoniemeä suuresta avusta käsi- kirjoituksen työstämisessä.

Viivi Virtanen on pedagoginen yliopistonlehtori Helsingin yliopiston bio- ja ympäristötieteellisessä tiedekunnassa. Liisa Postareff ja Telle Hailikari ovat yliopistonlehtoreita Helsingin yliopiston käyttäytymistieteiden laitoksen YTY- yksikössä (yliopistopedagogiikan tutkimus- ja kehittämisyksikkö).

LÄHTEET

- Andrade, H. & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32 (2), 159–181.
- Asikainen, H., Parpala, A., Virtanen, V. & Lindblom-Ylänne, S. (2013). The relationship between student learning process, study success and the nature of assessment. A qualitative study. *Studies in Educational Evaluation*, 39 (4), 211–217.
- Asikainen, H., Postareff, L., Heino, P. & Virtanen, V. (2014). Peer assessment in a large introductory class of gene technology. *Studies in Continuing Education*, 43, 197–205.
- Barnett, R. (1999). Learning to work and working to learn. Teoksessa D. Boud & J. Garrick (toim.), *Understanding learning at work* (s. 29–44). London: Routledge.
- Biggs, J. B. & Tang, C. (2007). *Teaching for quality learning at university: What the student does*. Buckingham: The Society for Research into Higher Education and Open University Press.
- Birenbaum, M. (1996). Assessment 2000: Towards a pluralistic approach to assessment. Teoksessa M. Birenbaum & F. J. R. C. Dochy (toim.), *Alternatives in assessment of achievement, learning processes and prior knowledge* (s. 3–30). Boston: Kluwer.
- Black, P., Harrison, C., Lee, C., Marshall, B. & William, D. (2004). Working inside the Black Box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86 (1), 9–21.
- Bloxham, S., Boyd, P. & Orr, S. (2011). Mark my words: The role of assessment criteria in UK higher education grading practices. *Studies in Higher Education*, 36 (6), 655–670.
- Bloxham, S. & West, A. (2004). Understanding the rules of the game: Marking peer assessment as a medium for developing students' conceptions of assessment. *Assessment & Evaluation in Higher Education*, 29 (6), 721–733.
- Boud, D. (2000). Sustainable assessment. Rethinking assessment for the learning society. *Studies in Continuing Education*, 22 (2), 151–167.
- Boud, D. (2014). Shifting views of assessment: From secret teacher's business to sustaining learning. Teoksessa C. Kreber, C. Anderson, N. Entwistle & J. McArthur (toim.), *Advances and innovations in university assessment and feedback* (s. 13–31). Edinburgh: Edinburgh University Press.
- Boud, D. & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31 (4), 399–413.
- Boud, D. & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38 (6), 698–712.
- Brown, G., Bull, J. & Pendlebury, M. (1997). *Assessing student learning in higher education*. London: Routledge.
- Bryan, C. & Clegg, K. (2006). *Innovative assessment in higher education*. New York and London: Routledge.
- Carvalho, A. (2013). Students' perceptions of fairness in peer assessment: Evidence from a problem-based learning course. *Teaching in Higher Education*, 18 (5), 491–505.
- Cilliers, F. J., Lambert, W., Schuwirth, H., Adendorff, J., Herman, N. & van der Vleuten, C. P. (2010). The mechanism of impact of summative assessment on medical students' learning. *Advances in Health Sciences Education*, 15 (5), 695–715.
- Crisp, G. T. (2012). Integrative assessment: Reframing assessment practice for current and future learning. *Assessment & Evaluation*, 37 (1), 33–43.
- Davey, K. R. (2011). Student peer assessment: Research findings from a case study in a master of chemical engineering coursework-program. *Education for Chemical Engineers*, 6 (4), 122–131.
- Davey, K. R. & Palmer, E. (2012). Student peer assessment: A research study in a level III core course of the bachelor chemical engineering program. *Education for Chemical Engineers*, 7, 85–104.
- De Grez, L., Valcke, M. & Roozen, I. (2012). How effective are self- and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education*, 13 (2), 129–142.
- Deneen, C. & Boud, D. (2014). Patterns of resistance in managing assessment change. *Assessment & Evaluation in Higher Education*, 39 (5), 577–591.
- De Wever, B., Van Keer, H., Schellens, T. & Valcke, M. (2011). Assessing collaboration in a wiki: The reliability of university students' peer assessment. *Internet and Higher Education*, 14, 201–206.
- Entwistle, A. & Entwistle, N. (1992). Experiences of understanding in revising for degree examinations. *Learning and Instruction*, 2, 1–22.
- Falchikov, N. (2005). *Improving assessment through student involvement*. London and New York: Routledge.
- Falchikov, N. & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59 (4), 395–430.
- Falchikov, N. & Goldfinch, J. (2000). Student peer assessment in higher education. A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70 (3), 287–322.
- Fuller, M. B. & Skidmore, S. T. (2014). An exploration of factors influencing institutional cultures of assessment. *International Journal of Educational Research*, 65, 9–21.
- Gijbels, D., Segers, M. & Struyf, E. (2008). Constructivist learning environments and the (im)possibility to change students' perceptions of assessment demands and approaches to learning. *Instructional Science*, 36, 431–443.
- Gouli, E., Gogoulou, A. & Grigoriadou, M. (2008). Supporting self-, peer-, and collaborative-assessment in e-Learning: The case of the PEer and Collaborative ASSESSMENT Environment (PECASSE). *Journal of Interactive Learning Research*, 19 (4), 615–647.
- Hailikari, T., Postareff, L., Tuononen, T., Räisänen, M. & Lindblom-Ylänne, S. (2014). Students' and teacher's perceptions of fairness in assessment. Teoksessa C. Kreber, C. Anderson, N. Entwistle & J. McArthur (toim.), *Advances and innovations in university assessment and feedback* (s.99–113). Edinburgh: University Press.
- Halinen, K., Ruohoniemi, M., Katajavuori, N. & Virtanen, V. (2013). Life science teachers' discourse on assessment: A valuable insight into the variable conceptions of assessment in higher education. *Journal of Biological Education*, 48 (1), 16–22.
- Handley, K., & Williams, L. (2011). From copying to learning: Using exemplars to engage students with assessment criteria and feedback. *Assessment & Evaluation in Higher Education*, 36 (1), 95–108.
- Hodgson, P. & Pang, M. Y. C. (2012). Effective formative e-assessment of student learning: A study on a statistics course. *Assessment & Evaluation in Higher Education*, 37, 215–225.
- Hounsell, D. (2007). Towards more sustainable feedback to students. Teoksessa D. Boud & N. Falchikov (toim.), *Rethinking assessment in higher education* (s. 101–113). London: Routledge.
- Kearney, S. (2013). Improving engagement: The use of 'Authentic self-and peer-assessment for learning' to enhance the student learning experience. *Assessment & Evaluation in Higher Education*, 38 (7), 875–891.
- Knight, P. T. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education*, 27 (3), 275–286.
- Lindblom-Ylänne, S., Pihlajamäki, H. & Kotkas, T. (2006). Self-, peer- and teacher assessment of student essays. *Active Learning in Higher Education*, 7 (1), 51–62.
- McGarr, O. & Gifford, A. M. (2013). 'Just enough to make you take it seriously': Exploring students' attitudes towards peer assessment. *Higher Education*, 65, 677–693.
- Mogey, N., Sarab, G., Haywood, J., van Heyningen, S., Dewhurst, D., Hounsell, D. & Neilson, R. (2007). The end of handwriting? Using computers in traditional essay examinations. *Journal of Computer Assisted Learning*, 24, 39–46.
- Mostert, M. & Snowball, M. D. (2013). Where angels fear to tread: Online peer-assessment in a large first-year class. *Assessment & Evaluation in Higher Education*, 38 (6), 674–686.
- Nulty, D. N. (2011). Peer and self-assessment in the first year of university. *Assessment & Evaluation in Higher Education*, 36 (5), 493–507.

- O'Donovan, B., Price, M. & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, 9 (3), 325–335.
- Ozogul, G. & Sullivan, H. (2007). Student performance and attitudes under formative evaluation by teacher, self- and peer-evaluators. *Education Technology Research and Development*, 57 (3), 393–410.
- Parpala, A. & Lindblom-Ylänne, S. (2007). University teachers' conceptions of good teaching in the units of high-quality education. *Studies in Educational Evaluation*, 33, 355–370.
- Patton, C. (2012). 'Some kind of weird, evil experiment': Student perceptions of peer assessment. *Assessment & Evaluation in Higher Education*, 37 (6), 719–731.
- Postareff, L., Lindblom-Ylänne, S. & Parpala, A. (2014). Explaining university students' strong commitment to understand through individual and contextual elements. *Frontline Learning Research*, 3, 31–49.
- Postareff, L., Virtanen, V., Katajavuori, N. & Lindblom-Ylänne, S. (2012). Academics' conceptions of the purpose assessment and their assessment practices. *Studies in Educational Evaluation*, 38 (3–4), 84–92.
- Rust, C., Price, M. & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education*, 28 (2), 147–164.
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30 (2), 175–194.
- Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 34 (2), 157–179.
- Sadler, P. & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, 11 (1), 1–31.
- Sambell, K., McDowell, L. & Brown, S. (1997). "But is it fair?" An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23, 349–371.
- Samuelowicz, K. & Bain, J. D. (2002). Identifying academics' orientations to assessment practice. *Higher Education*, 4, 173–201.
- Segers, M. & Dochy, F. (2001). New assessment forms in problem-based learning: The value-added of the students' perspective. *Studies in Higher Education*, 26 (1), 328–433.
- Segers, M. & Dochy, F. (2006). Enhancing student learning through assessment: Alignment between levels of assessment and different effects on learning. *Studies in Educational Evaluation*, 32 (3), 171–179.
- Segers, M., Dochy, F. & Cascallar, E. (toim.) (2003). *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Segers, M., Dochy, F. & Gijbels, D. (2010). Impact of assessment on students' learning strategies and implications for judging assessment quality. Teoksessa P. Peterson, E. Baker & B. McGaw (toim.), *International encyclopedia of education* (s. 196–201). Oxford: Elsevier.
- Segers, M. & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, 37 (1), 49–54.
- Sluismans, D., Dochy, F. & Moerkerke, G. (1999). Creating a learning environment by using self-, peer-, and co-assessment. *Learning Environments Research*, 1 (3), 293–319.
- Sluismans, D. & Prins, F. (2006). A conceptual framework for integrating peer assessment in teacher education. *Studies in Educational Evaluation*, 32 (1), 6–22.
- Struyven, K., Dochy, F. & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education*, 30, 325–341.
- Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68 (3), 249–276.
- Tynjälä, P., Slotte, V., Nieminen, J., Lonka, K. & Olkinuora, E. (2006). From university to working life: Graduates' workplace skills in practice. Teoksessa P. Tynjälä, J. Välimaa & G. Boulton-Lewis (toim.), *Higher education and working life – Collaborations, confrontations and challenges* (s. 73–88). Amsterdam: Elsevier.
- Van De Watering, G., Gijbels, D., Dochy, F. & van der Rijt, J. (2008). Students' assessment preferences, perceptions of assessment and their relationships to study results. *Higher Education*, 56, 645–658.
- van der Vleuten, C. (2014). 'Towards a future of programmatic assessment'. Keynote speech presented at the Earli Assessment & Evaluation Sig 1 Conference. 27–29 August 2014, Madrid.
- van der Vleuten, C. P. M., Schuwirth, L. W. T., Driessen, E., Dijkstra, J., Tigelaar, D., Baartman, L. K. J. & van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher*, 34, 205–214.
- Vickerman, P. (2009). Students' perspectives on formative peer assessment on colleges and universities: An attempt to deepen learning? *Assessment & Evaluation in Higher Education*, 34 (2), 221–230.
- Vu, T. T. & Dall'Alba, G. (2007). Students' experience of peer assessment in a professional course. *Assessment & Evaluation in Higher Education*, 32 (5), 541–556.
- Watkins, D., Dahlin B., & Ekholm, M. (2005). Awareness of the backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lecturers. *Instructional Science*, 33, 283–309.
- Welsh, M. M. (2007). Engaging with peer assessment in post-registration nurse education. *Nurse Education in Practice*, 7, 75–81.