

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Pollari, Pirjo

Title: "Mitä pitäisi kehittää ja miten?" : lukiolaisten kokemuksia arvioinnista ja sen menetelmistä englannin opinnoissa

Year: 2020

Version: Published version

Copyright: © 2020 Helsingin yliopisto

Rights: In Copyright

Rights url: <http://rightsstatements.org/page/InC/1.0/?language=en>

Please cite the original version:

Pollari, P. (2020). "Mitä pitäisi kehittää ja miten?" : lukiolaisten kokemuksia arvioinnista ja sen menetelmistä englannin opinnoissa . In R. Hilden, & K. Hahl (Eds.), *Kielididaktiikan katse tulevaisuuteen : haasteita, mahdollisuuksia ja uusia avauksia kielten opetukseen* (pp. 153-171). Helsingin yliopisto. Ainedidaktisia tutkimuksia, 17. <http://hdl.handle.net/10138/312321>

”Mitä pitäisi kehittää ja miten?” Lukiolaisten kokemuksia arvioinnista ja sen menetelmistä englannin opinnoissa

PIRJO POLLARI

pirjo.pollari@norssi.jyu.fi

Jyväskylän yliopisto, Normaalikoulu

Tiivistelmä

Opiskelija-arvioinnilla on merkittävä asema kaikkialla koulumaailmassa, niin myös suomalaisessa lukiossa. Vaikutusvaltaisesta roolistaan huolimatta arviointia ja sen käytännön toteutusta lukiossa tai lukion kieltenopetuksessa on tutkittu niukasti. Tämän tutkimuksen tavoitteena oli tutkia kyselylomakkeen avulla, kuinka lukiolaiset (146 lukiolaista yhdestä koulusta) itse kokivat arvioinnin osana englannin opintojaan, millaisia arviointimenetelmiä he olivat kohdanneet ja mitä niistä he pitivät hyödyllisinä. Tämän tutkimuksen perusteella opiskelijat arvostivat arviointia, jonka he kokivat osuvan varsin oikeaan ja olleen melko monipuolista. Arviointi ei kuitenkaan ohjannut heitä eteenpäin kovinkaan paljon. Lisäksi opiskelijoiden henkilökohtaiset kokemukset arvioinnista vaihtelivat huomattavasti. Jotta arviointi palvelisi vielä useampia opiskelijoita, sen tulisi olla entistä monipuolisempaa ja ohjaavampaa.

Avainsanat

lukion kieltenopetus, arviointi, arviointimenetelmät, formatiivinen arviointi

“What should be improved and how?” Students’ experiences of assessment and its methodology in their upper secondary English studies

Abstract

Student assessment has a powerful role in all educational contexts, also in the Finnish upper secondary school. Yet, in spite of its powerful role, there is rather little research on assessment and its practical implementation in foreign language teaching in upper secondary school. The aim of this study was to examine, with a questionnaire, how students (n = 146) in one upper secondary school experienced assessment as part of their English studies, what assessment methodology they had encountered and what they thought of the usefulness of these assessment methods. The article shows that most students appreciated assessment and found it accurate as well as quite varied in its methodology. However, assessment did not seem to feed forward very much. Furthermore, students’ individual experiences of assessment varied a great deal. In order for assessment to serve more students, it should be more multi-faceted and guiding.

Keywords

English as a foreign language in upper secondary school, student assessment, assessment methodology, formative assessment

Johdanto

Opiskelija-arviointi on käsitteenä hyvin monitahoinen ja laaja, sillä opiskelija-arviointi on eri koulukulttuureissa varsin erilaista. Niinpä opiskelija-arvioinnille on vaikea löytää yhteistä teoriaa: opiskelija-arvioinnin taustalla on kunkin maan opetussuunnitelma mutta myös esimerkiksi vallitsevat näkemykset arvioinnin tarkoituksesta, opettajan roolista ja luottamuksesta opettajan ammattitaitoon.

Arvioinnin tarkoitus

Yksi perinteisistä arvioinnin luokitteluperusteista on ollut jako formatiiviseen ja summatiiviseen arviointiin. Tämä jaottelu juontaa juurensa mm. Bloomin, Hastingsin ja Madausin (1971) teokseen *Handbook of formative and summative evaluation of student learning*, jossa jaottelu formatiiviseen ja summatiiviseen arviointiin tapahtui kolmen tekijän, arvioinnin tarkoituksen, ajoituksen ja yleistettävyyden, pohjalta. Sen jälkeen lukuisat tutkijat ovat tuoneet oman lisänsä näiden käsitteiden määrittelyyn. Konsensus nykyisin on, että jaottelun tärkein tekijä on arvioinnin tarkoitus. Formattiivisen arvioinnin tarkoituksena on edesauttaa oppimista muokkaamalla opiskelua ja opetusta, kun taas summatiivisen arvioinnin tehtävänä on arvioida oppimisen tuloksia. (Ks. esim. Black & William, 1998; Newton, 2007; Sadler, 1989, 1998.)

Formatiivisen ja summatiivisen arvioinnin lisäksi on puhuttu myös diagnostisesta arvioinnista. Sen tehtävänä on perinteisesti pidetty oppijan alkutaitojen kartoitusta joko opetuksen suunnittelun pohjaksi, oppijan taitojen ja tarpeiden mukaisiin opintoihin sijoittamiseksi tai oppimisvaikeuksien ja niiden syiden diagnosoimisen tueksi (ks. esim. Linnakylä & Välijärvi, 2005, ss. 26–27). Viime aikoina kieltenopiskelun saralla käsitettä diagnostinen arviointi on käytetty arvioinnista, jonka tarkoituksena on oppijan taitojen kartoittaminen niin vahvuuksien kuin heikkouksien mukaan. Sen avulla oppijalle voidaan tarjota sopivaa oppimista edistävää palautetta ja opetusta. (Alderson, 2005; Huhta, 2008; Lee, 2015.) Siten käsite on hyvin lähellä formatiivista arviointia.

Formatiivisen ja summatiivisen arvioinnin käsitteitä on käytetty paljon. Koska ne ovat kovassa käytössä myös menettäneet käsitteellistä selkeyttään (Leahy &

Wiliam, 2012), niiden rinnalla on alettu käyttää englanninkielisessä arviointikirjallisuudessa termejä *assessment for learning* ja *assessment of learning*. Nimensä mukaisesti *assessment for learning* tarkoittaa arviointia, jonka tehtävänä on edesauttaa ja edistää oppimista. *Assessment of learning* taas kuvaa arviointia, jonka tehtävänä on mitata oppimisen tuloksia. (Ks. esim. Black, Harrison, Lee, Marshall & Wiliam, 2003; Wiliam, 2011.)

Monissa opetuskulttuureissa formatiivinen arviointi ja sen suunnittelu ja toteutus kuuluvat opettajan työhön, mutta summatiivinen arviointi on merkittävästi osin uskottu muihin käsiin. Niinpä se, joka päättää ja säätelee arviointia, on yksi määrittävistä tekijöistä: onko arviointi ulkopuolisten tahojen laatimaa (*external*) vai koulujen tai opettajien itsensä laatimaa (*internal assessment*) (ks. esim. Pickford & Brown, 2006, s. 4). Myös arvioinnin laajuus ja yhteismitallisuus eli se, onko arviointi esimerkiksi koko maassa tai osavaltiossa samanlaista (*large-scale*) vai yksittäisten koulujen tai kouluryhmien käyttämää (*small-scale assessment*), määrittää arviointia.

Lukioarviointi ja opetussuunnitelmien ohjeistukset

Tämän artikkelin aineiston keräämisen aikana olivat voimassa vuonna 2003 ilmestyneet Lukion opetussuunnitelman perusteet. Niiden yleisessä arviointia käsittelevässä osassa arvioinnin tehtävä, kohteet ja menetelmät määritellään Lukiolakiin (629/1998, 17§1mom) turvautuen seuraavasti:

Opiskelijan arvioinnilla pyritään ohjaamaan ja kannustamaan opiskelua sekä kehittämään opiskelijan edellytyksiä itsearviointiin. Opiskelijan oppimista ja työskentelyä tulee arvioida monipuolisesti. (Opetushallitus 2003, 220).

Lisäksi Lukion opetussuunnitelman perusteet (Opetushallitus, 2003, s. 221) ohjaavat keskustelemaan arviointiperusteista opiskelijoiden kanssa kunkin kurssin alussa:

Yleisten arviointiperusteiden lisäksi kunkin kurssin arviointiperusteet on selvitettävä opiskelijalle kurssin alussa, jolloin niistä keskustellaan opiskelijoiden kanssa.

Kukin kurssi arvioidaan sen päätyttyä, ja arvioinnin tehtävänä ”on antaa palautetta opiskelijalle kurssin tavoitteiden saavuttamisesta ja oppiaineen opiskelun etenemisestä” (Opetushallitus, 2003, s. 221). Kieliaineiden kohdalla Opetussuunnitelman perusteet vaativat vielä lisää monipuolisuutta arvioinnin sisältöjen suhteen: ”Oppiaineen arvioinnissa otetaan huomioon kielitaidon kaikki alueet kurssikuvausten painotusten mukaisesti.” (Opetushallitus, 2003, s. 100).

Nykyisin voimassa olevat Lukion opetussuunnitelman perusteet edellyttävät arviointia myös kurssin aikana, eikä vasta sen päätyttyä. Tämän kurssin aikaisen arvioinnin tehtävänä on edistää kurssin tavoitteiden saavuttamista. (Opetushallitus, 2015.) Tulevaisuudessa formatiivinen arviointi vahvistaa asemaansa, sillä vuonna 2021 voimaan tulevat uudet Lukion opetussuunnitelman perusteet vahvistavat ja selkeyttävät jakoa arvioinnin formatiivisen ja summatiivisen tehtävän suhteen. Uudet perusteet määrittävät arvioinnin ensimmäiseksi tehtäväksi oppimisen tukemisen ja ohjaamisen. Kyseistä ”tehtävää toteuttaa opintojakson aikana annettava palaute, formatiivinen arviointi. Palautteella kuvataan opiskelijan edistymistä suhteessa asetettuihin tavoitteisiin.” (Opetushallitus, 2019, s. 45.) Arvioinnin toinen tehtävä on ”osaamiselle ja oppimiselle asetettujen tavoitteiden saavuttamisen näkyväksi tekeminen. Tätä tehtävää toteuttaa osaamisen ja opitun arviointi, summatiivinen arviointi. Osaamisen arviointi perustuu todennettuun näyttöön siitä, kuinka hyvin ja missä määrin opiskelija on saavuttanut opintojaksolle asetetut tavoitteet.” (Opetushallitus, 2019, s. 45.)

Arviointi lukion kieltenopetuksessa tutkimuksen valossa

Tutkimusta arvioinnin käytänteistä suomalaislukioissa on ollut vähän. Emme siksi voi varmuudella tietää, miten arviointia suomalaisissa lukiossa ja niiden kieltenopetuksessa toteutetaan (Pollari, 2017; ks. myös Virta, 2002).

Kansallisen koulutuksen arviointikeskuksen laatima arviointijulkaisu oppimisen ja osaamisen arvioinnista perusopetuksessa ja lukiossa (Atjonen ym., 2019) on ensimmäinen laaja kartoitus arvioinnista ja sen käytänteistä suomalaisessa koulussa. Tämä selvitys käsittelee arviointia kaikissa oppiaineissa niin alakoulussa, perusopetuksessa kuin lukiossa. Tuloksia on julkaistu myös kouluaste-kohtaisesti, joten se valottaa myös lukioden arviointikäytänteitä jonkin verran. Opiskelija-arviointia on lisäksi sivuttu suuremmissa kartoituksissa kuten Lukiopedagogiikka-arviointihankkeessa (Väljärvi ym., 2009).

Arviointia kieltenopetuksessa suomalaislukiossa on tutkittu kuitenkin varsin vähän: aiemmat tutkimukset ovat kohdistuneet joko yksittäisiin kurseihin tai arviointimenetelmiin kuten Eurooppalaiseen kielisalkkuun (Kohonen & Pajukanta, 2003; Lammi, 2002) tai portfolioarviointiin (Pollari, 2000). Viimeisin tutkimus paneutui lukiolaisten kokemuksiin arvioinnista englannin opinnoissa. Tämä tutkimus ei kuitenkaan perehtynyt niinkään arviointimenetelmiin ja -käytänteisiin vaan opiskelijoiden henkilökohtaisiin kokemuksiin siitä, auttoiko vai ahdistiko arviointi heitä opinnoissaan (Pollari, 2017).

Perusopetuksen puolella arviointia kieltenopetuksessa on tutkittu enemmän esimerkiksi osana tutkimusta, jossa kartoitettiin peruskoululaisten äidinkielen ja vieraiden kielten tekstikäytänteitä koulussa ja vapaa-ajalla (Luukka ym., 2008; Tarnanen & Huhta, 2011) tai osana kielten A-oppimäärien oppimistulosten kartoitusta perusopetuksen päättövaiheessa (ks. esim. Härmälä, Huhtanen & Puukko, 2014; Hildén & Rautopuro, 2014). Lisäksi oppilaiden käsityksiä suullisesta kielitaidosta ja sen arvioinnista on tutkittu (Ilola, 2018; ks. myös Ilola tässä julkaisussa).

Toisin kuin monissa muissa maissa, opettajalla on Suomessa arvioinnin ja sen toteutuksen suhteen täysi valta (ks. esim. Sahlberg, 2007; Vanttinen, 2011). Siten arviointi on – ainakin periaatteessa – opettajien suunnittelemaa ja toteuttamaa pienimuotoista luokkahuonearviointia (*small-scale internal classroom assessment*), jossa yksittäisillä kokeilla tai kurssiarvosanoilla ei ole kovin kriittistä painoarvoa (*low-stakes assessment*). Käytännössä kieltenopettajat kuitenkin turvautuvat arvioinnissa suurelta osin oppikirjasarjojen valmiisiin arviointimateriaaleihin. Lisäksi käytetään aiempien ylioppilaskirjoitusten tehtäviä (Atjonen ym., 2019). Ylioppilastutkintohan on koululaitoksemme ainoa laajamittainen ulkoinen tutkinto (ks. esim. Mehtäläinen & Välijärvi, 2013; Sahlberg, 2007) ja sillä on suuri merkitys opiskelijoiden tulevien opintojen kannalta (*external large-scale high-stakes examination*).

Aiempien suomalaistutkimusten valossa lukion opiskelija-arviointi – niin yleisesti kuin kieltenopetuksessakin – on vaikuttanut olevan varsin summatiivista ja numerokeskeistä arviointia, jonka keskeisenä tavoitteena on ollut kurssiarvosanan antaminen. Lukion kurssiarvioinnissa näytetään tukeutuvan varsin paljon kurssin summatiiviseen loppukokeeseen (Atjonen ym., 2019, ss. 145–146) tai

muihin erilaisiin kokeisiin. Arviointi ei siis ole menetelmiltään kovin monipuolista eikä itse- tai pariarvioinneilla ole merkittävää sijaa (Väljärvi ym., 2009; ks. myös Härmälä ym., 2014; Tarnanen & Huhta, 2011). Sen sijaan opiskelijat ovat hyvin tietoisia kurssien tavoitteista ja arviointiperusteista (Atjonen ym., 2019; Väljärvi ym., 2009) ja kokevat arvioinnin antaneen melko hyvän kuvan osaamisestaan (Väljärvi ym., 2009).

Arvioinnilla on myös monenlaisia vaikutuksia opiskelijoihin ja heidän hyvinvointiinsa. Väitöstutkimukseni (Pollari, 2017) mukaan arviointia pidettiin oikeudenmukaisena, mutta silti arviointi ahdisti ja aiheutti stressiä toisinaan monillekin opiskelijoille. Noin 15–20 % opiskelijoista koki arvioinnin tekijänä, joka ahdisti ja lannisti heitä huomattavasti ja siten haittasi opintoja ja opiskelumotivaatiota. Ahdistusta heille aiheuttivat varsinkin painoarvoltaan tärkeät kokeet ja koetilanteet. He myös kokivat, että käytetyt arviointimenetelmät eivät antaneet heille mahdollisuutta osoittaa kaikkea osaamistaan. Opiskelijat kaipaivat myös enemmän palautetta, joka tukee oppimista. Toisaalta noin 10 % opiskelijoista ei kokenut arvioinnin stressaavan, ahdistavan tai lannistavan yhtään. Suuri osa opiskelijoista sijoittui tähän väliin: useimmat olivat varsin tyytyväisiä arviointiin osana englannin opintojaan, mutta he ottivat sen koulutyöhön olennaisesti kuuluvana osana, joka ei todennäköisesti vaikuttanut heidän henkilökohtaiseen hyvinvointiinsa sen suuremmin. (Pollari, 2017.)

Tutkimusta, joka kartoittaisi nimenomaan lukion kieltenopetuksen arviointikäytänteitä ja -menetelmiä Suomessa, ei ole tietääkseni vielä tehty. Niinpä tässä tutkimuksessa pyrin vastaamaan seuraaviin kysymyksiin:

1. Arvostavatko opiskelijat arviointia? Auttaako arviointi oppimista ja opintoja heidän mielestään?
2. Millaisia arviointimenetelmiä englannin kursseilla on käytetty?
3. Miten opiskelijat itse arvioivat eri arviointimenetelmien hyödyllisyyden?

Menetelmä

Väitöstutkimukseni (Pollari, 2017) ja siten myös tämän artikkelin tutkimusaineisto kerättiin yhdessä lukiossa (omassa koulussani) maaliskuussa 2014 netti-

pohjaisella (MrInterview) kyselylomakkeella. Kyselylomakkeessa oli sekä Likert-tyyppisiä väittämiä (125 kpl) sekä avokysymyksiä (11 kpl). Taustakysymysten lisäksi kysymykset oli jaettu seitsemään arviointiin liittyvään osa-alueeseen. Ne olivat opiskelijoiden valta ja toimijuus arvioinnin suhteen, erilaisten arviointimenetelmien käyttö englannin kursseilla, eri arviointimenetelmien hyödyllisyys, arvioinnin oikeudenmukaisuus ja ohjaavuus, opiskelijoiden omat kokemukset ja näkemykset arvioinnista, ylioppilastutkinto sekä palaute. Vaikkakin kyselylomake sai tukea aiemmista tutkimuksista kuten Lukiopedagogiikka-tutkimuksesta (Väljærvi ym., 2009), suurin osa kyselylomakkeesta räätälöitiin tätä tutkimusta varten. (Ks. tarkemmin Pollari, 2017, s. 74–76).

Koulussamme oli tuolloin 199 2.–4. vuosikurssin opiskelijaa, ja heistä kyselyyn vastasi 146 (vastausprosentti 73,4 %). Toisen vuosikurssin opiskelijat vastasivat lomakkeeseen englannin oppitunnilla (79 vastaajaa), kolmannen ja neljännen vuosikurssin opiskelijat, jotka olivat jo aloittaneet valmistautumisen ylioppilaskokeisiin, vastasivat omalla ajallaan (67 vastaajaa). Vastaajista 86 ilmoitti olevansa naisia ja 60 miehiä. Heidän edellisen englannin kurssinsa arvosanojen keskiarvo oli 8,6 (arvosanat vaihtelivat välillä 6–10) ja takanaan heillä oli keskimäärin 6,7 englannin lukiokurssia (4–11). Näillä kursseilla heillä oli ollut 2–7 eri opettajaa (m=3,7). Vastaajat antoivat varsin kattavan kuvan koulumme opiskelijoista keväällä 2014.

Väitöstutkimukseni aineiston analysoinnin alkuvaiheessa kartoitin opiskelijoiden kokemuksia toteutuneesta arvioinnista sekä eri arviointimenetelmistä ja niiden hyödyllisyydestä. Analyysistä nousseiden havaintojen pohjalta keskityin opiskelijoiden henkilökohtaisiin kokemuksiin arvioinnista ja sen vallasta: auttaako arviointi opinnoissa vai ahdistaa se? Aineistoa analysoitiin monimenetelmäisesti, vaikkakin pääpaino oli kvantitatiivisissa menetelmissä (mm. pääkomponenttianalyysi ja askeltava regressioanalyysi).

Tässä artikkelissa painottuvat kuitenkin kuvailevat tilastolliset analyysimenetelmät, sillä keskityn aineistoon, jota en ole aiemmissa tutkimuksissani vielä raportoinut. Tämän artikkelin keskiössä ovat opiskelijoiden kokemukset arvioinnin oikeudenmukaisuudesta ja ohjaavuudesta, erilaisten arviointimenetelmien käyttö englannin kursseilla ja opiskelijoiden näkemykset näiden arviointimenetelmien hyödyllisyydestä, sillä tällaista tutkimusta Suomessa on edelleenkin hyvin vähän.

Tulokset

Opiskelijoiden näkemykset toteutuneesta arvioinnista

Päällisin puolin koulumme lukiolaiset vaikuttivat olevan hyvin tyytyväisiä arviointiin. Arviointi näytti myös noudattavan silloin voimassa olleiden Lukion opetussuunnitelmien perusteiden (Opetushallitus, 2003) vaatimuksia hyvin: kurssien tavoitteita, arviointiperusteita ja kriteereitä oli käsitelty käytännöllisesti katsottuna kaikkien kurssien alussa ($m=4,51$)³. Opettaja oli myös kertonut kunkin kurssin alussa, mistä kurssin arvosana koostuu ($m=4,52$). Opiskelijat olivat lisäksi voineet esittää toiveitaan ja näkemyksiään arvioinnin suhteen ainakin jollain kurseilla ($m=3,38$), mutta olivat tehneet niin itse melko harvoin.

Lukiolaiset olivat myös tyytyväisiä saamiinsa arvosanoihin: kurssiarvosanat olivat perustuneet niihin asioihin, joista kurssin alussa oli sovittu ($m=4,4$) ja opiskelijat kokivat saaneensa harvoin joko liian huonon ($m=1,68$) tai liian hyvän arvosanan ($m=1,44$).

Yleisvaikutelma tuntui olevan se, että opiskelijat olivat hyvin tyytyväisiä arviointiin ja sen oikeellisuuteen. Lukiolaiset eivät kuitenkaan olleet yhtä hyvin perillä konkreettista arviointiperusteista, esimerkiksi siitä, millainen koe tulisi olemaan ($m=3,53$), hyväksyttäisiinkö sanakokeessa synonyymeja ($m=3,12$) tai kuinka kokeet ja muut arviointitehtävät oli arvioitu ($m=3,19$).

Myös Lukion opetussuunnitelman perusteiden (Opetushallitus, 2003) vaatimukset arvioinnin monipuolisuudesta ja kielitaidon eri osa-alueiden huomiomisesta eivät tuntuneet aivan toteutuneen lukiolaisten mielestä kaikilla kursseilla. Vaikka arviointimenetelmät olivat olleet suhteellisen monipuolisia

³ Tässä osiossa Likert-asteikko oli viisiportainen:

5 = Jokaisella kurssilla

4 = Lähes jokaisella kurssilla

3 = Joillakin kursseilla

2 = 1–2 kurssilla

1 = Ei yhdelläkään kurssilla

($m=3,62$), kaikkia kielitaidon osa-alueita ei aina ollut otettu huomioon arvioinnissa joka kurssilla ($m=3,33$). Itsearviointiin oli toisinaan ollut mahdollisuus esimerkiksi ehdottamalla itselleen arvosanaa ($m=2,37$), minkä opiskelijat olivat tehneetkin ($m=2,35$).

Arvioinnin ei koettu olleen myöskään erityisen kannustavaa ($m=3,31$) kaikilla kursseilla. Arviointi ei vaikuttanut olleen ohjaavaa vaan ennemminkin toteavaa: noin puolet vastaajista oli sitä mieltä, että ”arviointi vain toteaa, se ei ohjaa tai auta oppimaan paremmin”. Noin 35 % vastaajista oli tästä eri mieltä⁴. Opettajien arviointilinjaukset erosivat opiskelijoiden mielestä jonkin verran: noin 45 % vastaajista oli sitä mieltä, että eri opettajat painottavat eri asioita arvioinnissa, kun taasen 30 % vastaajista oli eri mieltä.

Kaikista vastaajista noin 40 % koki, että arviointi ahdisti heitä ainakin ajoittain. Noin 15–20 % vastaajista – joista enemmistö tyttöjä – koki, että käytetyt arviointimenetelmät olivat lannistaneet heitä ja siten arviointi osaltaan oli myös vähentänyt heidän haluaan opiskella tai oppia englantia. Arviointi tuntui siis jakavan opiskelijoiden näkemyksiä ja kokemuksia.

Opiskelijoiden kokemukset ja näkemykset eri arviointimenetelmistä

Arviointimenetelmien monipuolisuus ja yleisyys

Aiemman tutkimuksen mukaan lukioarviointi on ollut varsin koekeskeistä. Lukiopedagogiikka-arviointihankkeessa puolet vastaajista koki, että kurssien arvostelussa kokeiden osuus painottui liikaa (Väljjarvi ym., 2009, s. 54). Myös Karvin arviointi antaa samankaltaisia viitteitä (Atjonen ym., 2019).

⁴ Tässä osiossa viisiportainen Likert-asteikko oli seuraava:

5 = Täysin samaa mieltä

4 = Jokseenkin samaa mieltä

3 = En osaa sanoa

2 = Jokseenkin eri mieltä

1 = Täysin eri mieltä

Kokeet nousivat kärkeen myös tämän aineiston englannin kurssien arviointimenetelmien joukossa. Käytännöllisesti katsoen joka kurssilla oli käytetty sanakokeita ($m=4,74$). Lisäksi melkein kaikilla kursseilla⁵ oli ollut jonkinlainen kurssikoe ($m=4,63$) sekä kuuntelukoe tai -kokeita ($m=4,51$). Neljantenä arviointimenetelmien frekvenssilistalla tuli kirjoitelma, joka oli useammin koulussa kuin kotona kirjoitettu. Jokin arvosanaan vaikuttava suullinen näyttö, joko koe ($m=2,63$) tai muunlainen näyttö kuten esitelmä ($m=2,63$), oli myös ollut melko yleinen.

Epätavallisemmät arviointikeinot kuten lunttilappukokeet ($m=1,81$), väittelyt tai ryhmäkeskustelut ($m=2,20$) sekä esitykset esimerkiksi video- tai äänitallenteena ($m=1,75$) olivat harvinaisempia, mutta niitäkin oli käytetty yksittäisillä kursseilla. Pistokokeet tai *open book* -kokeet (koe, jonne saa tuoda mukanaan joko kirjan tai muut muistiinpanot) olivat hyvin harvinaisia, samoin kuin tietokoneen avulla tai etänä tehdyt kokeet. Täytyy kuitenkin muistaa, että aineisto kerättiin ennen kuin ylioppilaskirjoitusten digitalisointi oli alkanut.

Arviointimenetelmissä korostui yksittäisen opiskelijan suoritus. Kokeet tai tehtävät, joissa tekijöiden osuutta ei eritelty, olivat melko harvinaisia ($m=1,47$). Sen sijaan itsearvioita oli tehty useammillakin kursseilla ($m=3,25$) ja pari- tai kaveriarvioitakin joillain kursseilla ($m=2,40$). Nämä eivät kuitenkaan pääsääntöisesti vaikuttaneet arvosanaan. Lisäksi formatiiviset ($m=1,58$) tai diagnostiset testit ($m=1,47$) vaikuttivat olevan harvinaisia. Näitä opiskelijat eivät juurikaan olleet havainneet käytettävän.

Vaikkakaan kyseessä ei ollut satunnaisotos, vastauksia tarkasteltiin myös yksisuuntaisen varianssianalyysin avulla (ANOVA). Vain yhden arviointitavan (pistokokeiden) kohdalla nais- ja miesopiskelijoiden vastaukset arviointitavan yleisyydestä erosivat tilastollisesti merkitsevästi. Siten vastausten yhteneväisyys menetelmien yleisyyden suhteen vahvistaa vastausten luotettavuutta.

⁵ Tässä kyselylomakkeen osiossa oli käytössä sama Likert-asteikko kuin aiemminkin, jossa 5 = jokaisella kurssilla, 1 = ei yhdelläkään kurssilla.

Arviointimenetelmien hyödyllisyys

Arviointitapojen hyödyllisyyden suhteen opiskelijat kokivat, että monipuolinen arviointinäyttö arvosanan pohjana oli esitetyistä vaihtoehdoista hyödyllisin⁶ (m=4,64). Noin 90 % vastaajista piti sanakokeita (m=4,30) ja noin 85 % erilaisia suullisen kielitaidon näyttöjä hyödyllisinä arviointikeinoina (m=4,15). Arvosanan perustumista pääsääntöisesti kurssikokeeseen piti hyödyllisenä puolet vastaajista, kun taasen noin 40 % vastaajista piti tällaista yhden kokeen painotusta hyödyttömänä (m=3,18). Sen sijaan lähes 60 % vastaajista piti jokseenkin tai täysin hyödyttömänä sitä, että kurssikoetta ei olisi lainkaan. Tähän asiaan 30 % ei osannut ottaa kantaa (m=2,38).

Arviointimenetelmien tuttuus näkyi selkeästi vastauksissa. Harvinaisemmat menetelmät jakoivat näkemyksiä: esimerkiksi noin 45 % piti luntilappukoetta hyödyllisenä, 20 % ei osannut ottaa kantaa ja noin 35 % piti sitä hyödyttömänä menetelmänä (m=3,06). Tietokoneavusteinen koe oli näille opiskelijoille tuntematon, joten sen hyödyllisyydestä yli 55 % vastaajista ei osannut sanoa mitään ja reilu 30 % koki sen hyödyttömäksi: mies- ja naisopiskelijoiden vastaukset olivat lähes täysin yhtenevät. Formatiiiviset arviointimenetelmät vaikuttivat myös varsin tuntemattomilta: lähes kolmannes vastaajista ei osannut ottaa kantaa. Niitä kuitenkin kannatettiin enemmän kuin vastustettiin: noin 45 % piti niitä hyödyllisenä ja hieman alle neljännes opiskelijoista hyödyttöminä (m=3,20).

Itsearviot (m=2,96) ja pari- tai kaveriarviot (m=2,72) jakoivat myös opiskelijoiden näkemyksiä hyödyllisyyden suhteen. Itsearviota piti hyödyllisenä lähes 40 % vastaajista mutta hyödyttömänä lähes sama määrä (37,5 %). Pari- tai kaveriarviota hyödyllisenä arviointimenetelmänä kannatti 30 % vastaajista mutta

⁶ Tässä osiossa opiskelijoita pyydettiin vastaamaan, kuinka hyödyllisenä englannin kielien oppimiselle he pitivät mainittuja arviointitapoja. Likert-asteikko oli seuraava:

5 = Erittäin hyödyllinen

4 = Jokseenkin hyödyllinen

3 = En osaa sanoa

2 = Jokseenkin hyödytön

1 = Täysin hyödytön

vastusti lähes puolet. Arvosanan perustuminen lähinnä opiskelijan omaan itsearvioon sai tämän tutkimuksen vastaajilta lähes täystyrmäyksen: 85 % vastaajista piti sitä hyödyttömänä (m=1,76).

Arviointimenetelmien hyödyllisyyden suhteen yksisuuntainen varianssianalyysi (ANOVA) osoitti useampiakin tilastollisesti merkitseviä tai erittäin merkitseviä eroja mies- ja naisopiskelijoiden näkemysten välillä. Naisopiskelijat pitivät vähemmän koeahdistusta aiheuttavia arviointitapoja selkeästi hyödyllisempänä kuin miesopiskelijat. Näihin arviointitapoihin kuului portfolioarviointi tai lunttilappukoe sekä formatiiviset ja diagnostiset arviointimenetelmät.

Pohdinta

Tämän artikkelin ensimmäisenä tehtävänä oli selvittää, arvostavatko lukiolaiset arviointia osana englannin opintojaan. Tämän tutkimuksen perusteella opiskelijoista suurin osa oli hyvin tyytyväisiä arviointiin. Arvioinnin katsottiin antavan hyvä ja oikeudenmukainen kuva opiskelijan tiedoista ja taidoista ja opiskelijat olivat tietoisia kurssien tavoitteista ja arviointiperusteista. Näiltä osin tämän tutkimuksen tulokset ovat samansuuntaisia Lukiopedagogiikka-arvioinnin (Väljærvi ym., 2009) tulosten kanssa: opiskelijat kokivat summatiivisen arvioinnin toimineen varsin hyvin.

Seuraavaksi selvitin, millaisia arviointimenetelmiä englannin kursseilla oli käytetty ja mitä mieltä opiskelijat olivat eri arviointimenetelmien hyödyllisyydestä. Kieltenopetuksen ja sen arvioinnin perinteiset kulmakivet eli sanakokeet, kurssikokeet, kuuntelukokeet ja kirjoitelmat olivat käytetyimmät arviointimenetelmät. Siitä huolimatta vaikutti siltä, että arviointimenetelmien kokonaisuus oli hieman monipuolisempi kuin aiemmissa tutkimuksissa perusopetuksen puolella oli havaittu (Hildén & Rautopuro, 2013; Härmälä ym., 2014; Tarnanen & Huhta, 2011). Suullinen kielitaito otettiin huomioon ainakin muutamilla kursseilla, ja myös vähemmän perinteisiä menetelmiä kuten lunttilappukoetta tai ryhmäkeskusteluja oli toisinaan käytetty. Itsearviointit olivat melko yleisiä, pariarviointit eivät niinkään, mutta yleensä kumpikaan näistä ei vaikuttanut suoranaisesti arvosanaan.

Hyödyllisimpinä arviointimenetelminä opiskelijat pitivät monipuolista kokonaisuutta eli sitä, että arvosana koostui erilaisista arviointinäytöistä. Täten lukiolaiset itse selvästi kannattivat Lukion opetussuunnitelman perusteiden näkemystä monipuolisesta arvioinnista. Tutut menetelmät, kuten sanakokeet, koettiin hyödylliseksi, samoin suullisen kielitaidon arvioiminen. Sen sijaan harvinaisemmat arviointimenetelmät, kuten lunttilappukoe tai portfolioarviointi, jakoivat opiskelijoiden näkemyksiä. Pelkkään kurssikokeeseen perustuvaa arvosanaa tai kurssikokeen kokonaan puuttumista opiskelijat eivät suuresti kannataneet. Itsearvioon perustuva arvosana sai heiltä varsin selvän tyrmäyksen. Itseja pariarvioita opiskelijat eivät muutenkaan näyttäneet suuresti arvostavan. Myöskään tietokoneavusteinen koe ei heitä innostanut, sillä heillä ei ollut siihen vielä tarvetta. Ylioppilastutkinnon syksyllä 2016 alkanut asteittainen digitalisaatio ei koskettanut heitä.

Formatiivinen arviointi oli tämän tutkimuksen opiskelijoille ainakin käsitteenä vielä varsin vieras. Erityisesti naisopiskelijat kuitenkin arvelivat sen olevan melko hyödyllistä. Lisäksi opiskelijat kokivat, että vaikka arviointi toteaa osaamisen ja sen tason hyvin, se ei yhtä hyvin ohjaa oppimista eteenpäin. Niinpä nykyisten ja tulevien Lukion opetussuunnitelman perusteiden (Opetushallitus, 2015, 2019) mainitsema formatiivinen, oppimista edistävä arviointi sai näiltä opiskelijoilta kannatusta, vaikka se ei heille ainakaan terminä vielä tuttua ollutkaan. Oppimista edistävä arviointi kaipaisi siis selkeästi lisää huomiota ainakin meidän lukiomme englannin opetuksessa.

Tämä aineisto kerättiin keväällä 2014, jolloin voimassa olivat vuoden 2003 Lukion opetussuunnitelman perusteet. Sikäli aineisto ei ole ajantasaista. Sen jälkeen vuonna 2016 käyttöön otetut Lukion opetussuunnitelmat (Opetushallitus, 2015) ovat olleet käytössä muutaman vuoden. Lisäksi Opetushallitus on jo julkaissut uudet Lukion opetussuunnitelman perusteet (Opetushallitus, 2019). Ylioppilastutkinnon digitalisaatio ei myöskään vaikuttanut tutkimuksen opiskelijoiden elämään. On myös huomioitava, että tämän tutkimuksen aineisto kerättiin koulussa, jonka opiskelija-aines on tiedoiltaan ja taidoiltaan selkeästi valtakunnallisen keskiarvon yläpuolella. Tämän tutkimuksen tulokset eivät täten ole suoraan yleistettävissä muihin suomalaisiin lukioihin. Siitä huolimatta tulokset antavat runsaasti uutta tietoa, sillä mikään aikaisempi tutkimus ei ole pa-

neutunut lukion kieltenopetuksessa käytettyihin arviointimenetelmiin ja opiskelijoiden kokemuksiin ja näkemyksiin niistä yhtä tarkasti ja kattavasti. Lisäksi tämän tutkimuksen tulokset antavat vertailupohjaa tuleville tutkimuksille kieltenopetuksen arviointikäytännöistä.

Mitä kieltenopetuksen arvioinnissa sitten pitäisi kehittää ja miten? Niin tämän kuin aiemman tutkimuksen pohjalta (esim. Atjonen ym., 2019; Välijärvi, 2009) suositellen, että arvioinnin tulisi olla mahdollisimman monipuolista niin menetelmiltään kuin sisällöiltään. Formattiiviseen, oppimista edistävään arviointiin tulisi kiinnittää entistä enemmän huomiota opintojen aikana. Lisäksi palautteen tulisi auttaa opiskelijaa opinnoissaan eteenpäin eikä vain todeta tämänhetkistä tilannetta. Erilaisia palautteenantomenetelmiä samoin kuin toimivia itsearviointimenetelmiä tulisi kehittää.

Yksi vaihtoehto olisi, että lukio-opintojen aikana eri kursseilla painotettaisiin opetussuunnitelman kurssikuvausten ja -tavoitteiden sekä käytännön mahdollisuuksien mukaan erilaisia arviointitapoja ja myös arviointikohteita. Siten lukioarvioinnin kokonaisuus antaisi monipuolisen ja kokonaisvaltaisen kuvan opiskelijan tiedoista ja taidoista. Nyt tilanne on, että kieliaineiden arviointi – jos se noudattaa vaatimusta kaikkien kielitaidon osa-alueiden huomioonottamisesta – on melkoinen urakka sekä opiskelijoille että opettajalle, sillä samoja taitoja arvioidaan jokaisella kurssilla. Toinen vaihtoehto olisi toteuttaa pidempiä, useamman opintopisteen opintokokonaisuuksia. Tulevien Lukion opetussuunnitelmien perusteiden (Opetushallitus, 2019) pohjalta tämä olisikin mahdollista. Tällöin opintojen aikaiselle arvioinnille syntyisi enemmän aikaa ja tilaa, mikä mahdollistaisi entistä paremmin myös runsaamman ohjaavan palautteenannon opintojaksojen aikana.

Opiskelijoiden mahdollisuutta osallistua arviointiin ja sen suunnitteluun tulisi myös lisätä (ks. myös Atjonen ym., 2019). Opettajien arviointiosaamiseen ja sen päivittämiseen on viime aikoina kiinnitetty huomiota ja tämän kehityksen soisi jatkuvan. Lisäksi toivon lisää tutkimusta arvioinnista, sillä arviointi on oppimisen, opintojen ja kouluelämän keskeinen osa-alue, jota on tutkittu varsin vähän. Olisi tärkeää, että arviointia tutkittaisiin laajemmin esimerkiksi tulevien opetussuunnitelmien perusteiden ja niiden arviointiohjeiden pohjaksi. Arviointi

on niin merkittävä tekijä, että sen ohjeistusten ja käytänteiden tulisi aidosti perustua tutkittuun tietoon niin yksittäisten opettajien kuin koko koulutusjärjestelmän tasolla.

Lähteet

- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- Atjonen, P., Laivamaa, H., Levonen, A., Orell, S., Saari, M., Sulonen, K., Tamm, M., Kampi, P., Rumpu, N., Hietala, R. & Immonen, J. (2019). "Etä tietää missä on menossa" – Oppimisen ja osaamisen arviointi perusopetuksessa ja lukiokoulutuksessa. Helsinki: Kansallinen koulutuksen arviointikeskus. Julkaisut, 7:2019. <https://karvi.fi/publication/etta-tietaa-missa-on-menossa-oppimisen-ja-osaamisen-arviointi-perusopetuksessa-ja-lukiokoulutuksessa/>
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead, UK: Open University Press.
- Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148.
- Bloom, B. S., Hastings, J. T. & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.
- Hildén, R. & Rautopuro, J. (2014). *Ruotsin kielen A-oppimäärän oppimistulokset perusopetuksen päättövaiheessa 2013*. Helsinki: Opetushallitus.
- Huhta, A. (2008). Diagnostic and formative assessment. Teoksessa B. Spolsky & F. M. Hult (Toim.), *The Handbook of Educational Linguistics* (ss. 469–482). Malden, MA: Blackwell.
- Härmälä, M., Huhtanen, M. & Puukko, M. (2014). *Englannin kielen A-oppimäärän oppimistulokset perusopetuksen päättövaiheessa 2013*. Helsinki: Karvi.
- Ilola, M. (2018). Oppilaiden käsityksiä englannin suullisesta kielitaidosta, sen oppimisesta ja arvioinnista. JYU dissertations, (15). Jyväskylä: Jyväskylän yliopisto. <https://jyx.jyu.fi/handle/123456789/59522>

- Kohonen, V. & Pajukanta, U. (Toim.). (2003). *Eurooppalainen kielisalkku 2–EKS-projektin päätösvaiheen tuloksia*. Tampereen yliopiston opettajankoulutuslaitoksen julkaisuja, 28.
- Lammi, K. (2002). *Kielisalkku lukion ruotsin kielen opiskeluun motivoinnissa: Opiskelijoiden ja opettajan kokemuksia*. Jyväskylä: Jyväskylän yliopisto, soveltavan kielentutkimuksen keskus.
- Leahy, S. & Wiliam, D. (2012). From teachers to schools: Scaling up professional development for formative assessment. Teoksessa J. Gardner (Toim.), *Assessment and learning* (2. painos, ss. 49–71). London: Sage.
- Lee, Y. (2015). Diagnosing diagnostic language assessment. *Language Testing*, 32(3), 299–316. DOI:10.1177/0265532214565387
- Linnakylä, P. & Välijärvi, J. (2005). *Arvon mekin ansaitsemme: Kansainvälinen arviointi suomalaisen koulun kehittämiseksi*. Jyväskylä: PS-Kustannus.
- Luukka, M., Pöyhönen, S., Huhta, A., Taalas, P., Tarnanen, M. & Keränen, A. (2008). *Maailma muuttuu - mitä tekee koulu? Äidinkielen ja vieraiden kielten tekstikäytänteet koulussa ja vapaa-ajalla*. Jyväskylä: Jyväskylän yliopisto, soveltavan kielentutkimuksen keskus.
- Mehtäläinen, J. & Välijärvi, J. (2013). *Ylioppilaskokeiden arvosanojen vertailtavuus eri aineissa vuosina 2007–2011*. Jyväskylä: Koulutuksen tutkimuslaitos.
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 14(2), 149–170. DOI:10.1080/09695940701478321
- Opetushallitus. (2003) *Lukion opetussuunnitelman perusteet 2003*. Nuorille tarkoitettun lukiokoulutuksen opetussuunnitelman perusteet. Helsinki: Opetushallitus. https://www.oph.fi/sites/default/files/documents/47345_lukion_opetussuunnitelman_perusteet_2003.pdf
- Opetushallitus. (2014). *Perusopetuksen opetussuunnitelman perusteet 2014*. Määräykset ja ohjeet 2014:96. Helsinki: Opetushallitus. https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_perusteet_2014.pdf

- Opetushallitus. (2015). *Lukion opetussuunnitelman perusteet*. Määräykset ja ohjeet 2015:48. Helsinki: Opetushallitus. https://www.oph.fi/sites/default/files/documents/172124_lukion_opetussuunnitelman_perusteet_2015.pdf
- Opetushallitus. (2019). *Lukion opetussuunnitelman perusteet 2019*. Määräykset ja ohjeet 2019:2a. Helsinki: Opetushallitus. https://www.oph.fi/sites/default/files/documents/lukion_opetussuunnitelman_perusteet_2019.pdf
- Pickford, R. & Brown, S. (2006). *Assessing skills and practice*. Abingdon: Routledge.
- Pollari, P. (2000). *"This is my portfolio": Portfolios in upper secondary school English studies*. Jyväskylä: Koulutuksen tutkimuslaitos.
- Pollari, P. (2017). *(Dis)empowering assessment? Assessment as experienced by students in their upper secondary school EFL studies*. Jyväskylä studies in humanities 329. Jyväskylä: Jyväskylän yliopisto. <https://jyx.jyu.fi/handle/123456789/55478>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), 77–84. DOI: 10.1080/0969595980050104
- Sahlberg, P. (2007). Education policies for raising student learning: The Finnish approach. *Journal of Education Policy*, 22(2), 147–171.
- Tarnanen, M. & Huhta, A. (2011). Foreign language assessment and feedback practices in Finland. Teoksessa D. Tsagari & I. Csépes (Toim.), *Classroom-based language assessment*. Language testing and evaluation Vol. 25 (ss. 129–146). Frankfurt am Main: Peter Lang.
- Virta, A. (2002). Arviointi oppimisen ja opetuksen punaisena lankana. Teoksessa E. Lehtinen & T. Hiltunen (Toim.), *Oppiminen ja opettajuus* (ss. 63–86). Turku: Turun yliopisto. Kasvatustieteiden tiedekunnan julkaisuja B: 71.
- Väljjarvi, J., Huotari, N., Iivonen, P., Kulp, M., Lehtonen, T., Rönholm, H., Knubb-Manninen, G., Mehtäläinen, J. & Ohranen, S. (2009). *Lukiopedagogiikka*. Koulutuksen arviointineuvoston julkaisuja 40. Jyväskylä: Koulutuksen arviointineuvosto.

- Vänttinen, M. (2011). *Oikeasti hyvä numero: Oppilaiden arvioinnin totuudet ja totuustuotanto rinnakkaiskoulusta yhtenäiskouluun*. Joensuu: Itä-Suomen yliopisto.
- Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3–14.