

JYX



JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

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

Author(s): Tarnanen, Mirja

Title: Arvioija valokeilassa: suomi toisena kielenä kirjoittamisen arviointia

Year: 2002

Version: Published version

Copyright: © Kirjoittaja & Kotikielen Seura, 2002

Rights: In Copyright

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

Please cite the original version:

Tarnanen, M. (2002). Arvioija valokeilassa: suomi toisena kielenä kirjoittamisen arviointia. *Virittäjä*, 106(3), 398-403.



ARVIOIJA VALOKEILASSA: SUOMI TOISENA KIELENÄ -KIRJOITTAMISEN ARVIOINTIA



MIRJA TARNANEN

VÄITÖKSENALKAJAISESITELMÄ
JYVÄSKYLÄN YLIOPISTOSSA
6. HUHTIKUUTA 2002

Arvoisa kustos, arvoisat vastaväittäjät ja arvoisa yleisö, pyydän teitä lektioni aluksi sulkemaan hetkeksi silmäne ja tarkastelemaan ensimmäiseksi mieleenne tulevaa arviointikokemusta.

Ehkä joidenkin teistä ensimmäiseksi mieleen tulleet kokemukset liittyvät arvioijana olemiseen, siihen, miten on tullut lähiaikoina arvioitua opiskelijoiden, oppilaiden tai työntekijöiden suoriutumista tai siihen, miten on tullut kommentoitua saatua palvelua asiointitilanteessa tai hetki sitten tämän tilaisuuden ensimmäisiä tunnelmia. Joidenkin teistä, ehkä useimpien kohdalla, ensimmäisenä mieleen nousevat kokemukset kuvastavat oppijan tuntemuksia, sitä, miten omaa osaamista on arvioitu. Erilaiset arviointikokemukset voivat olla omasta asemasta arviointitilanteessa ja tilanteen luonteesta riippuen positiivisia tai negatiivisia, motivoivia tai lamaanuttavia, oppimista ja toimintaa kehittäviä tai jopa ne pyssäyttäviä.

Arvioinnin vaikutuksista yksilölliset vaikutukset ovatkin kaikkein läpinäkyvimpiä, koska ne koskettavat meitä jokaista. Emme ehkä ole vaikutuksista aina kovin tietoisia, mutta käsityksemme juuret omasta osaamisestamme, esimerkiksi matematiikassa, käsitöissä tai kirjoittamisessa, johtavat usein lapsuuteen, kouluaikana saatuun palautteeseen ja numeroihin. Yksilön itse-tuntoon, minäkäsitykseen, eivät vaikuta vain koenumerot tai todistusarviot, vaan kaikki sanalliset ja sanattomat palautteet, joita hän saa osaamisestaan. Arvioinnin laatu, arviointitapa ja se, kuka arvioi, vaikuttavat siis merkittävästi yksilön käsityksiin itsestään oppijana ja taidon hallitsijana.

Poliittiset, sosiaaliset ja taloudelliset, osin koulutuksellisetkin, vaikutukset eivät välttämättä ole yhtä läpinäkyviä kuin yksilölliset, mutta niiden seuraukset voivat olla kauaskantoisempia sekä hyvässä että pa-

hassa. Se, miten yhteisö valitsee, luokittelee, jakaa ja arvioi tietoa, kertoo myös vallan jakaantumisesta ja sosiaalisesta kontrollista (Eggleston 1990: 59). Kenen tieto ja millainen tieto on hyväksyttävää ja kenellä on oikeus arvioida tietoa ja sen luotettavuutta? Arviointi on koulutuksellisesti ja poliittisesti keino kontrolloida yksilön pätevyyttä ja ohjata hänen valintojaan (ks. Norris 1995; Linnakylä 1998). Arvioinnin avulla voidaan seurata myös esimerkiksi koulujärjestelmän toimintaa ja sen tuloksellisuutta.

Arvioinnin vaikutukset ovat negatiivisia, jos arviointi kapea-alaistaa tietokäsitystämme, jos se hierarkkisoii valtasuhteita ja jos se on epäoikeudenmukaista. Positiivisia ne taas ovat, jos arviointi motivoi ja tarjoaa työkaluja kehittymisen seuraamiseen ja tavoitteiden asettamiseen. Millaisin edellytyksin arviointi voi sitten olla positiivista, laadultaan hyvää arviointia?

Arviointikirjallisuudessa esitetään useita hyvän arvioinnin kriteereitä, mutta esitelen tässä kolme tärkeintä. Arvioinnilta edellytetään ensisijaisesti eettisyyttä. Arvioinnin on oltava eettisesti hyväksyttävää, oikeudenmukaista ja arvioitavan näkökulman huomioon ottavaa. Eettisyyttä voidaan pitää eräänlaisena arvioinnin yleiskriteerinä, josta kaikki alkaa ja johon kaikki palaa. Arviointi on eettistä, jos sen luotettavuus, eli validius ja reliabelius, perustuvat eettisesti hyväksyttäviin periaatteisiin (Messick 1994; Hamp-Lyons 1997; Kunnan 1997).

Arviointi on osuvaa eli validia, jos arviointitehtävät arvioivat sitä, mitä niiden on tarkoitus arvioida. Kirjoittamisen arvioinnissa tämä tarkoittaa sitä, missä määrin niin sanotut mittaamisvälineet eli kirjoitustehdävät, arviointiasteikot ja niiden soveltaminen arvioivat sellaista kirjoittamisen taitoa, mitä niiden oletetaan arvioivan.

Arvioinnin toistettavuus eli reliabelius

viittaa taas siihen, missä määrin tehtävät mittaavat taitoa samalla tavalla eri aikoina sekä siihen, miten johdonmukaisesti kukin arvioija arvioi ja miten yhdenmukaisia arvioiden antamat arviot ovat keskenään. Jos arviointitehtävät mittaavat tarkoituksesta eri taitoa eri arviointikerroilla ja jos arvioijat arvioivat suorituksia subjektiivisten mielitystensä mukaan, arviointia ei voida tietenkään pitää luotettavana toistettavuusmielessä.

Eettisyys, validius ja reliabelius eivät ole Messickin (1994: 13) mukaan vain mittaamisen tai laajemmin arvioinnin periaatteita, vaan ne ovat myös sosiaalisia arvoja. Ne eivät ole myöskään staattisia määritelmiä. Arviointi on dynaaminen prosessi, ja näkökulmiin arvioinnin laadukkuudesta vaikuttavat monenlaiset tekijät. Arvioinnin »hyvyys» voi riippua esimerkiksi arviointitilanteesta, käytetyistä arviointitavoista ja arviointiin liittyvistä yksilöistä.

Tässä tutkimuksessa valokeilassa ovat yksilöt, ja yksilöistä arvioijat. Heidän toimintaansa, vaikutustaan suomeksi toisena kielinä -kirjoittamisen arviointiin, tarkastellaan määrällisesti eli kvantitatiivisesti ja laadullisesti eli kvalitatiivisesti. Kvantitatiivisuus viittaa tässä arvioiden antamiin arvioihin eli siihen, miten johdonmukaisia ja yhdenmukaisia arvioiden antamat arviot ovat. Laadullisuus viittaa taas arvioiden käsityksiin kirjoittamisesta ja heidän tunteuksiinsa ja kokemuksiinsa arviointiprosessista.

Vastausten etsintä kvantitatiivisiin kysymyksiin edellyttää johdonmukaisuuden osalta sitä, että arvioijat arvioivat samoja suorituksia useammin kuin kerran ja yhdenmukaisuuden osalta sitä, että samoja suorituksia arvioi useampi kuin yksi arvioija. Tämän vuoksi tässä tutkimuksessa olikin kolme arviointikierrosta, joilla kaikilla arvioija arvioi samat suoritukset. Arvioijat käyttivät ensimmäisellä arviointikierroksella

▷

la kuusiportaista arviointiasteikkoa, jonka sisällön he määrittivät itse. Toisella ja kolmannella arviointikierröksellä arvioijat käyttivät Euroopan neuvoston viitekehysten perustuvaa holistista, myös kuusiportaista, kirjoittamisen asteikkoa.

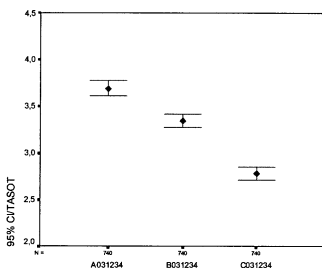
Jokaisen arviointikierroksen jälkeen arvioijat haastateltiin. Haastatteluissa puhuttiin muun muassa arvioijien itsensä määrittelemistä asteikoista, Euroopan neuvoston asteikon soveltamisesta, arviointiprosessista, tehtävistä ja niiden kirjoittajista. Ensimmäinen ja toinen haastattelu ovat puolistrukturoituja teema-haastatteluja. Kolmas haastattelu on konstruktiteoriaan (personal construct theory) perustuva valintaverkkohaastattelu (repertory grid), joka on strukturoitu haastattelumetodi. Valintaverkkohaastattelun avulla selvitetään arvioijien havaintoja arvioitavista teksteistä ja sitä, miten he arvottivat tekstejä tekemiensä havaintojen mukaan.

Ennen kuin jatkan eteenpäin, on syytä esitellä arvioitavien tekstien kirjoittajat ja

arvioijat. Tutkimuksessa on arvioitavien tekstien kirjoittajia kaikkiaan 248, jotka kaikki ovat aikuisia, äidinkielisiltä ja kulttuurisilta taustoiltaan erilaisia suomi toisena kielenä -oppijoita. Jokainen oppija kirjoitti kolme kirjoitustehtävää yhdeksän tehtävän sarjasta, mutta kuitenkin niin, että yksi tehtävä oli yhteinen kaikille kirjoittajille. Arvioijia on kaikkiaan 17. He ovat kaikki äidinkieleltään suomenkielisiä ja heillä kaikilla on kokemusta suomen opettamisesta toisena kielenä. Kukin arvioija arvioi vähintään sadan kirjoittajan tekstit eli yhteensä vähintään noin 300 tekstiä kullakin arviointikierröksellä.

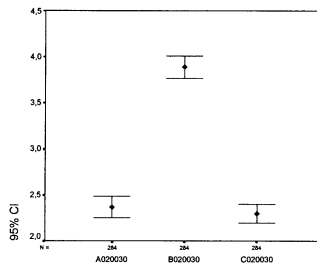
Mitä arvioijien antamat arviot kertovat arvioijien johdonmukaisuudesta ja keskinäisestä yhdenmukaisuudesta? Odotusten mukaan määritellyn asteikon käyttämisen olisi pitänyt johdonmukaistaa arvioijien arviointia toisella ja kolmannella arviointikierröksellä. Arvioijat voidaan jakaa heidän antamien arvioidensa keskiarvojen perusteella neljään tyyppiin: 1) ensimmäisen tyy-

TYYPPI 1
Arviointit
eroavat
toisistaan
keskimäärin
kaikilla
arviointi-
kerroilla
(N=7).



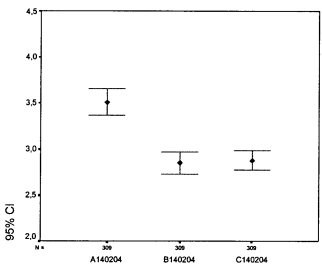
Kierros 1. Kierros 2. Kierros 3.

TYYPPI 2
Ensimmäi-
nen kierros
eroaa joko
toisesta tai
kolmannesta
kierroksesta
(N=5).



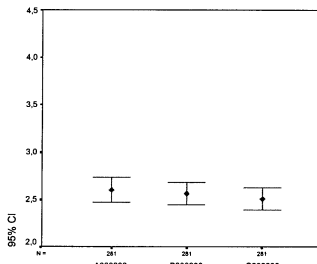
Kierros 1. Kierros 2. Kierros 3.

TYYPPI 3
Ensimmäi-
nen kierros
eroaa sekä
toisesta että
kolmannesta
kierroksesta
(N=4).



Kierros 1. Kierros 2. Kierros 3.

TYYPPI 4
Eri arviointi-
kierrokset
eivät eroa
toisistaan
tilastollisesti
merkittävästi
(N=1).



Kierros 1. Kierros 2. Kierros 3.

Kuvio 1. Arvioijien johdonmukaisuustyytit.

pin arvioijiin, joiden antamat arviot eroavat keskimäärin kaikilla arviointikierroksilla, 2) toisen tyyppin arvioijiin, joiden ensimmäisellä arviointikierroksella antamat arviot eroavat joko toisen tai kolmannen arviointikierroksen arvioista, 3) kolmannen tyyppin arvioijiin, joiden ensimmäisen kierroksen arviot eroavat sekä toisen että kolmannen arviointikierroksen arvioista, ja 4) neljäs tyyppin arvioijaan, jonka eri arviointikierroksilla antamat arviot eivät eroa toisistaan.

Samansuuntaista tendenssiä osoittavat myös kierrostenväliset Pearsonin korrelaatiot ja samanmielisyysohjeet, joiden mukaan noin puolella arvioijista määritellyn asteikon käyttäminen johdonmukaisti arviointia ja noin toisella puolella sillä ei ollut vaikutusta johdonmukaisuuteen.

Arvioijien keskinäistä yhdenmukaisuutta voi tarkastella niiden arvioijien kesken, jotka arvioivat samojen kirjoittajien tekstejä eri arviointikierroksilla eli tehtäväsarjoittain. Tällaisia tehtäväsarjoja oli kaikkiaan neljä. Cronbachin ja mittauskertojen välisten korrelaatioiden mukaan erot arvioijien välisessä yhdenmukaisuudessa eri arviointikerroilla ovat niin pieniä, että arviointi näyttää olleen yhdenmukaista kaikilla kolmella arviointikierroksella asteikosta riippumatta.

Kaikkiaan kvantitatiiviset tulokset osoittavat, että vaikka arvioijien johdonmukaisuudessa oli jonkin verran yksilöllisiä eroja, he kuitenkin arvioivat tekstejä yhdenmukaisesti kaikilla kolmella arviointikierroksella. Miksi näin oli? Katsotaan, miten haastattelut vastaavat tähän kysymykseen.

Ennen kuin siirrymme haastatteluihin, pyytäisin teitä lukemaan seuraavan oppijan kirjoittaman mielipidetekstin ja saamaan siten vähän tuntumaa siihen, millaisia reaktioita oppijan kirjoittama teksti voi herättää ja millaisiin asioihin siinä voidaan kiinnittää huomiota.

Minua häiritsee hullujen muoti, miksi ihmiset käyvät päälaki paljana. Hyvä hattu, päähine pukee ihmistä. Päähineitä voi vaihtaa vaikka joka päivä. On olemassa hyviä kuosia, trendikkaita hattuja mutta ei. Mitä kannavat mallimuodissa sitä minakiin, vaikkapa läpi harmaan kiveen. Enitten muotia tule Afrikkasta, Afrikkassa on kuuma, he ovat ihan paljaana, mutta me asumme kylmässä vyöhykkeessä ja meidän täytyy käyttäytyä niin kuin sää vaatii. Hiljattain Helsinkiin lentokentällä, lentokoneesta tuli joukko miehiä, vanhoja juruttajia ja kaikki paltoissa, mutta päälaki tyhjä niin kuin päässä, aivoissa.

Osa teistä saattoi kiinnittää tekstiä luokiensa huomiota siihen, että se näyttää vieraalta, viehättävältä tai ehkä vain vaikealta ymmärtää. Joidenkin mielestä siinä saattoi olla outoja kielen muotoja tai sanoja. Huomiota saattoi kiinnittää myös itse aihe ja siihen liittyvä sisältö tai tapa käsitellä sitä. Vieraan kielen kirjoittamisen arvioinnissa arvioijat kiinnittävät tutkimusten mukaan huomiota yleensä kielen rakenteiden hallintaan, tekstin pituuteen, sanastoon, sisältöön, oikeinkirjoitukseen, jäsentelyyn ja käsialaan (esim. Vaughan 1991; Milanovic ym. 1996; DeRemer 1998; Smith 2000).

Tässä tutkimuksessa arvioijien antamien arvioiden yhdenmukaisuutta voi selittää arvioijien keskittyminen kielen rakenteiden hallintaan, joka nousi kirjoittamisen eri kriteereistä yhtenäisimmin esiin sekä arvioijien itsensä määrittelemissä asteikoissa että valintaverkkohaastattelussa. Tämä ei kuitenkaan tarkoita sitä, että arvioijien kirjoittamiskäsitys olisi yhtä kuin kielen rakenteiden hallinta. Arvioijien käsitykset voidaan itse asiassa jakaa kahteen tasoon: yleiseen kirjoittamiskäsitykseen ja arvottavaan kirjoittamiskäsitykseen. Yleinen kirjoittamiskäsitys ottaa kirjoittamisen eri osat huomioon monipuolisesti. Sekä arvioijien asteikoissa että valintaverkkohaastattelussa esiin nousseita kriteereitä ovat esimerkiksi ymmärrettävyys, rakenteet, sa-

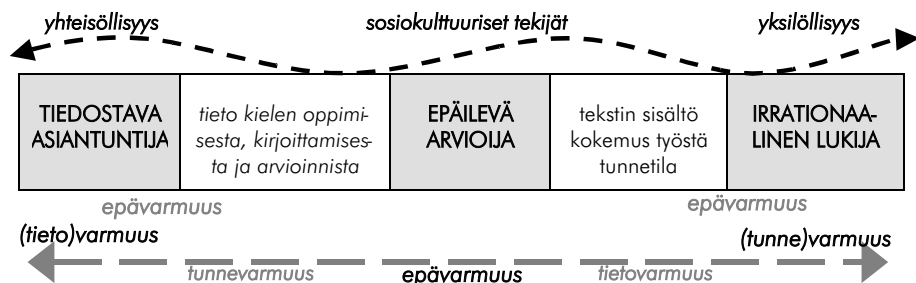
nasto, sisältö, tekstimäisyys, konventioiden hallinta, idiomaattisuus, tehtävänannon täyttyminen, tekstin pituus ja oikeinkirjoitus. Sen sijaan arvottavaa kirjoittamiskäsitystä ilmentävissä, arvioijien itsensä määrittelemissä asteikoissa, kuvataan systemaattisimmin rakenteiden etenemisen hallintaa. Samoin valintaverkkohaastattelussa tekstien arvottaminen rakenteiden ja sanaston mukaan korreloi vahvimmin tekstien yleistason mukaisen arvottamisen kanssa. Toisen kielen kirjoittamisen arvioinnissa tukeutuminen kielen rakenteisiin ei ole epätavallista, koska rakenteisiin on helppo ottaa kantaa niiden konkreettisuuden, hierarkisuuden ja opetuksessa olevan aseman vuoksi.

Arvioijien johdonmukaisuuden vaihtelua voi taas ymmärtää arviointiprosessin monitekijäisyyden kautta. Arviointitilanne on tekstin ja sen välityksellä kirjoittajan, asteikon ja arvioijan välinen vuorovaikutusprosessi. Arviointiprosessin keskiössä on arvioija, jonka kokemukset, vireystila, tunnetila, kulttuurinen viitekehys, tietoisuus ja tahtotila säätelevät sitä, millainen arviointiprosessi on luonteeltaan ja miten se etenee. Näine tekijöineen arviointiprosessi on dynaaminen ja jossakin määrin tekstikohteisesti ainutkertainen.

Johdonmukaisuuden vaihteluun on voinut vaikuttaa tässä tutkimuksessa myös määrittelyn asteikon huono sovellettavuus.

Monet arvioijista pitivät asteikkoa liian ylimalkaisena, tekstityyppilähtöisenä ja moniselitteisenä. Asteikon puutteellisuus aiheutti myös epävarmuutta ja sai arvioinnin tuntuun liian subjektiiviselta. Se nosti myös esiin kysymyksen arvion absoluuttisuudesta, siitä, onko kullekin suoritukselle olemassa jokin yksi oikea arvio. Tämä ilmeni nimenomaan epävarmuutena arviointiprosessin aikana. Epävarmuutta herätti myös arvioijan rooli geneerisenä lukijana siinä mielessä, että arvioija voisi edustaa keitä tahansa lukijoita ja pystyisi tekemään luottavuudesta päätöksiä näidenkin puolesta.

Arvioijan roolia lukijana voisikin havainnollistaa jatkumolla, jossa vuorottelevat toistensa ääripäinä asiantuntijalukija ja irrationaalinen lukija (ks. kuviota 2). Arvioijan yksilöllisyyttä, eräänlaista yksilöllistä tulkintaa edustaa irrationaalinen lukija, joka oli etualalla silloin, kun arvioijan toiminta hänen kommenttinsa mukaan oli senhetkisen mielialan tai vireystilan mukaista, esimerkiksi kun hän kommentoi kirjoittajan taustaan liittyviä tekijöitä, tekstien yksitoikkoisuutta tai väsymistä. Näin ollen irrationaalisen lukijan sekä negatiiviset että positiiviset kommentit ovat tavallaan tunnevarmoja. Asiantuntijalukijan kommentit ovat taas enemmänkin tietovarmoja, koska ne perustuvat yhteisöllisestikin neuvoteltuihin käsityksiin, tietoon ja kokemuksiin kirjoittamisesta ja arvioinnista. Asiantuntija-



Kuvio 2. Arvioijan lukijatyyppiä.

lukijalle on tyypillistä omien valintojen tiedostaminen, toimintatapojen kyseenalaistaminen ja ohjaaminen. Haluan kuitenkin korostaa, että jatkumo ei ole sarjoittainen vaan eräänlainen katkeamattoman liikkeen virta, jossa irrationaalisen ja asiantuntijalukijan eriaisteiset tyypit ovat vaihtelevasti etualalla ja jossa vaihtelevat samoin varmuuden ja epävarmuuden tunteet.

Arvioijan toiminnan tarkasteleminen sekä numeerisesti että prosessia ymmärtäen osoittaa, että johdonmukaisuus ja yhdenmukaisuus liittyvät läheisesti validiuteen. Kirjoittamisen arvioinnissa arvioijan käsitys ohjaa sitä, miten hän oman käsityksensä kautta arvottaa tekstejä, tulkitsee ja soveltaa arviointikriteereitä. Arvioijien käsitykset ovat yksilöllisiä, mutta ne edustavat myös sosiokulttuurisia käsityksiä. Ne ovat osa opetuskulttuuria, sen käsityksiä, traditioita ja normeja. Johdonmukaisuus ja yhdenmukaisuus eivät siis liity arvioijiin vain yksilöinä eivätkä ne ilmene vain lukuina, vaan ne ovat osa laajempaa arviointikulttuuria ja edustavat sen arvostuksia. Aikaisemmin viitatus Messickin eettisyys, validius ja reliabelius ovat myös sosiaalisia arvoja.

Suomi toisena kielenä -kirjoittamisen arvioinnissa on kyse monikulttuurisista kirjoittajista, joilta suomalaisen tietoyhteiskunnan jäsenyys edellyttää kirjoittamisen taitoa. Mikä on se kirjoittamisen taidon taso, joka riittää tyydyttävään selviytymiseen ja millaisten tehtävien, esimerkiksi työtehtävien, hoitamiseen tällainen tyydyttävä kirjoitustaito riittää? Miten syvälle luotava ja monipuolinen, esimerkiksi kulttuuriset tekijät huomioon ottava, kirjoituskäsityksemme on? Miten suvaitsevasti ja ymmärtävästi suhtaudumme oppijoiden teksteihin? Miten tottuneita olemme yleensäkin lukemaan niitä? Mitä kirjoittamisesta tulisi opettaa ja miten sitä tulisi arvioida? Entä miten opetuksen ja arvioijan odotuk-

set kirjoittamisen arvioinnista vastaavat oppijan odotuksia? Miten paljon ja missä määrin olemme valmiita siirtämään arviointivastuuta oppijalle itselleen ja millaista käsitys- ja asennemuutosta se edellyttää poliittisilta, taloudellisilta ja koulutuksellisilta tahoilta?

Tutkimus siis herättää monia sellaisia kysymyksiä, joihin se ei tarjoa suoraan vastausta. Se kuitenkin osoittaa, että arvioinnin vaikutuksellisuus, monikerroksisuus ja inhimillisyys edellyttävät arviointiprosessin ymmärtämistä ja sen tarkastelua osana laajempaa sosiokulttuurista viitekehystä, jotta arviointi olisi eettisesti hyväksyttävää, luotettavaa ja uskottavaa. Arviointi on aina jossakin määrin subjektiivista, mutta hyvä ja eettinen arviointi edellyttää pyrkimystä johdonmukaisuuteen ja yhdenmukaisuuteen. ■

LÄHTEET

- DEREMER, MARY L. 1998: Writing assessment: raters' elaboration of the rating task. – *Assessing Writing* 5, 1 s. 7–29.
- EGGLESTON, JOHN 1990: School examinations — some sociological issues. – Tim Horton (toim.), *Assessment debates* s. 57–67. London: Hodder and Stoughton.
- HAMP-LYONS, LIZ 1997: Ethics in language testing. – Caroline Clapham & David Corson (toim.), *Language testing and assessment 7. The encyclopedia of language and education* s. 223–333. Dordrecht: Kluwer Academic.
- KUNNAN, ANTHONY JOHN 1997: Connecting fairness with validation in language assessment. – Ari Huhta, Viljo Kohonen, Liisa Kurki-Suonio & Sari Luoma (toim.), *Current development and alternatives in language assessment. Proceedings of LTRC 96*. Jyväskylä: ▷

- University of Jyväskylä.
- LINNAKYLÄ, PIRJO 1998: Koulutuksen kansallisen ja kansainvälisen arvioinnin haasteita. – *Kasvatus* 2 s. 175–189.
- MESSICK, SAMUEL 1994: The interplay of evidence and consequences in the validation of performance assessments. – *Educational Research* 23, 2 s. 13–23.
- MILANOVIC, MICHAEL – SAVILLE, NICK – SHUHONG, SHEN 1996: A study of the decision-making behavior of composition markers. – Michael Milanovic & Nick Saville (toim.), *Performance testing, cognition and assessment: selected papers from the 15th language testing research colloquium, Cambridge and Arnhem* s. 92–114. Cambridge: Cambridge University Press.
- NORRIS, NIGEL 1995: Koulutusohjelmien arviointi. – Sauli Takala (toim.), *Arviointi ja koulutuksen laadun kehittäminen* s. 33–38. Jyväskylä: Kasvatus-tieteiden tutkimuslaitos.
- SMITH, DAVID 2000: Rater judgement in the direct assessment of competency-based second language writing ability. – Geoff Brindley (toim.), *Studies in immigrant English language assessment* 1 s. 159–189. Research series 11. Sydney: Macquarie University.
- VAUGHAN, CAROLINE 1991: Holistic assessment: what goes on in the rater's mind? – Liz Hamp-Lyons (toim.), *Assessing second language writing in academic contexts* s. 111–126. New Jersey: Ablex Publishing Corporation.

MIRJA TARNANEN *Arvioija valokeilassa: suomi toisena kielenä -kirjoittamisen arviointia*. Jyväskylä: Jyväskylän yliopiston soveltavan kielentutkimuksen keskus 2002.

Soveltavan kielentutkimuksen keskus, PL 35, 40014 Jyväskylän yliopisto
Sähköposti: tarnanen@jyu.fi