



---

## Opettajien näkemyksiä matematiikan ja fysiikan päättöarvioinnin kriteereistä

*Laura Ketonen*

*Kasvatustieteiden ja psykologian tiedekunta, Jyväskylän yliopisto*



Peruskoulun arviointia on viime vuosikymmenen aikana uudistettu ja peruskouluun on luotu kriteeriperustainen päättöarviointi. Uudistuksen jälkeen on aika tutkia sen tuloksia. Toimivan kriteeriperustaisen arvioinnin edellytyksenä ovat toimivat arviointikriteerit. Suomessa opettajat vastaavat arvioinnin suunnittelusta ja toteutuksesta, ja siksi heidän näkemyksensä arviointikriteerien ominaisuuksista ovat tärkeitä. Tässä laadullisessa tutkimuksessa tarkastelen, millaisia ongelmia opettajat ( $N = 28$ ) näkevät fysiikan ja matematiikan arvioitavissa tavoitteissa ja arviointikriteereissä. Vastauksissa korostuvat tavoitteiden ja kriteerien epätarkkuus, epäjohdonmukaisuus kriteerien taustateoriassa ja arvosanojen vaatimustason kalibrointitarpeet. Tutkimuksessa tunnistettuja haasteita voidaan hyödyntää kriteerien suunnittelu- ja kehitystyössä niin peruskoulussa kuin laajemminkin.

**Asiasanat:** arviointi, päättöarviointi, peruskoulu, arviointikriteerit, opettajien näkemykset

Lähetetty: 22.1.2024  
Hyväksytty: 10.6.2024  
Vastuukirjoittaja: [laura.k.ketonen@jyu.fi](mailto:laura.k.ketonen@jyu.fi)  
DOI 10.23988/ad.142709

## Johdanto

Suomen perusopetuksessa on viimeisen vuosikymmenen aikana tehty suuri arviointiuudistus ja sen päätuotoksena peruskouluun on luotu kriteeriperustainen päättöarviointi. Päättöarviointi on nykyisen opetussuunnitelman perusteiden (Opetushallitus, 2014) aikana tarkentunut kahdella tavalla. Ensinnäkin vuoden 2014 opetussuunnitelman perusteissa (Opetushallitus, 2014) listattiin opetuksen tavoitteet kaikissa oppiaineissa, ja jokaiselle tavoitteelle esitettiin arvosanan 8 kriteerikuvaus. Edellisissä opetussuunnitelman perusteissa (Opetushallitus, 2004) oppimistavoitteita ja niiden arviointikriteereitä ei kytketty yhtä selvästi yhteen. Toisena seikkana lukuvuoden 2021 alusta otettiin arvosanan 8 kriteerien lisäksi käyttöön kriteerit arvosanoille 5, 7 ja 9 (Opetushallitus, 2020b).

Päättöarviointi vaikuttaa oppilaiden jatkokoulutusmahdollisuuksiin minkä vuoksi sen tulee kohdella oppilaita yhdenvertaisesti. Päättöarvioinnin kohteiden eli arvioitavien tavoitteiden valinta on tärkeä asia oppilaiden yhdenvertaisuuden kannalta. Se määrittää, minkälaisia tai *kenen* (Gipps, 1999) tietoja, taitoja ja asenteita arvostetaan. Suomen peruskoulussa oppiaineiden arvioinnin kohteet ovat laajat, sillä ne sisältävät arvoja, asenteita, taitoja ja tietoja, mukaan lukien työskentelyn tavoitteita (Opetushallitus, 2014).

Toinen yhdenvertaisuuden osatekijä on arvosanojen vertailukelpoisuus, eli se, että oppilaat saavat samantasoisella osaamisella saman arvosanan. Arvosanojen vertailukelpoisuudessa on kuitenkin todettu olevan ongelmia (Metsämuuronen, 2023; Ouakrim-Soivio, 2013). Arviointiuudistuksen ensisijaisena tavoitteena olikin, että kriteerit parantaisivat juuri arvosanojen vertailtavuutta (Opetushallitus, n.d.). Uudistuksen toimeenpanon jälkeen on perusteltua arvioida arviointikriteereitä ja uudistuksen seurauksia. Tässä tutkimuksessa aihetta lähestytään opettajien näkemysten kautta ja kohdennetaan tarkastelu päättöarvioinnin tavoitteisiin ja kriteereihin.

Opettajien näkemyksiä, käsityksiä ja uskomuksia arvioinnista on tutkittu varsin paljon (ks. esim., Barnes et al., 2014; Bonner, 2016). Näillä eri nimityksillä ei ole tarkkoja määritelmiä, vaan niitä käytetään keskenään ristiin (Barnes et al., 2014). Käytän tässä tutkimuksessa sanaa näkemys, sillä se kuvaa osuvimmin lähestymistapaa, jossa opettajien kertoman tiedon ajatellaan olevan merkityksellistä sellaisenaan. Perustelen valintaa sillä, että opettajat ovat päättöarvioinnin kriteerien ensisijaisia käyttäjiä ja tulkitsijoita, ja sen vuoksi heidän näkökulmansa on keskeinen. Opettajien arviointia koskevissa näkemyksissä voi painottua arvioinnin pedagoginen tai yhteiskunnallinen tehtävä (Remesal, 2011). Vaikka päättöarviointia itsessään tehdään juuri yhteiskunnallisista syistä, eli oppilaiden jaottelemiseksi jatko-opintoihin, opettajien oppilasvalintoja koskevissa näkemyksissä näkyy oppimisen näkökulma, niin kuin Remesal (2011) tutkimukseensa osoittaa.

Arvioinnin pedagogista ja yhteiskunnallista tehtävää ei voi erottaa toisistaan. Päättöarvioinnin toteutustavat, mukaan lukien päättöarvioinnin kriteerit, eivät vaikuta ainoastaan päättöarvosanoihin. Arviointi ohjaa oppimisprosessia sekä opettajan ja oppilaiden huomiota. Sillä on siksi väis-

tämättä seurauksia myös opetukselle ja oppimiselle. Arvioinnin kohteiden valinta vaikuttaa niin opettajan kuin oppilaiden toimintaan (Biggs & Tang, 2011; Silseth & Gilje, 2019), ja jopa arviointiasteikon valinnalla on merkitystä sille, miten oppilaat orientoituvat oppimiseen (Collins & Lundstedt, 2024). Oppimisen ja sen päätteeksi tehdyn arvioinnin näkemykset kietoutuvat siis väistämättä yhteen (Black & William, 2018).

Tämän artikkelin tavoitteena on tarkastella uudistunutta päättöarviointia ja tuottaa tietoa sen kehitystarpeista. Aihetta tarkastellaan opettajien näkökulmasta seuraavan tutkimuskysymyksen kautta: Minkälaisia ongelmia ja kehityskohtia opettajat näkevät matematiikan ja fysiikan arvioitavissa tavoitteissa ja arviointikriteereissä? Koska Suomessa arviointi on käytännössä täysin opettajan vastuulla ja opettajien autonomia on kansainvälisessä vertailussa suuri, opettajien näkemys arviointikriteereistä on keskeinen.

## **Kriteeriperusteinen arviointi tarvitsee hyvät arviointikriteerit**

Tämän tutkimuksen teoreettisena viitekehyksenä toimii kriteeriperustaisen arvioinnin tutkimus. Kriteeriperustainen arviointi tapahtuu vertaamalla oppilaan suoritusta ennalta laadittuihin arviointikriteereihin. Normiperustaisessa arvioinnissa, esimerkiksi ylioppilaskokeissa, oppilaan suoritus suhteutetaan vertailuryhmän arviointituloksiin. Peruskoulussa käytettiin ennen vuotta 1985 suhteellista arviointia, joka eroaa normiperustaisesta siten, että oppilaan suoritus voidaan suhteuttaa pienen joukon, usein yhden luokan tuloksiin (Ouakrim-Soivio, 2013).

Kriteeriperustaisen arvioinnin ajatellaan olevan hyvä tapa toteuttaa summatiivista arviointia (Brookhart & Chen, 2015). Sitä pidetään oikeudenmukaisempana kuin suhteellista tai normiperustaista arviointia, sillä aidosti kriteeriperustaisessa arvioinnissa toisten oppilaiden suoritus ei vaikuta yksittäisen oppilaan arvosanaan (Wallace & Ng, 2022), eikä se siksi kannusta oppilaita kilpailemaan toisiaan vastaan (Pitman & Dudley, 1985). Arviointikriteereistä on hyötyä myös oppimisprosessin aikana. Niitä voi käyttää pedagogisena työvälineenä oppimistavoitteiden konkretisoinnissa, oppimisprosessin mallintamisessa, itsearviointissa ja vertaisarviointissa (Andrade, 2005; Brookhart, 2018; Brookhart & Chen, 2015). Kriteerit parantavat oppimisen ja arvioinnin läpinäkyvyyttä, vähentävät oppilaiden ahdistusta, toimivat palautteen antamisen tukena, parantavat oppilaiden minäpystyvyyttä ja tukevat itsesäätelyä (Panadero & Jonsson, 2013). Andrade (2023) kuitenkin huomauttaa, että osaamisen tason määrittämistä varten laaditut kriteerit eivät useinkaan ole optimaalisia oppimisen tukemiseen. Tämä tarkoittaa sitä, että kriteereitä laatiessa täytyy valita, kumpi käyttötarkoitus on ensisijainen.

Kriteeriperustaisen arvioinnin onnistumisen tärkein edellytys ovat toimivat arviointikriteerit (Brookhart, 2018). Arviointikriteeristöjä eli arviointimatriiseja on erilaisia. Arviointimatriisi voi olla sellainen, että se kuvaa ainoastaan hyväksytyyn suoritukseen tason, tai se voi eritellä eri arvosanojen vaatimukset (Gosling & Moon, 2002). Arviointimatriisit voidaan jakaa

yleisiin ja tehtäväkohtaisiin (Brookhart, 2018). Yleistä arviointimatriisia voi käyttää erilaisiin tehtäviin, kun taas tehtäväkohtaiset arviointimatriisit on laadittu yhden tietyn tehtävän tarpeisiin. Peruskoulun päättöarvioinnin oppiainekohtaiset arviointimatriisit ovat yleisiä, sillä yhden tehtävätyypin avulla ei voi opetussuunnitelman perusteiden mukaan edes arvioida kaikkia tavoitteita (Opetushallitus, 2020a). Toinen arviointimatriiseja määrittelevä jaottelu on, ovatko ne analyttisiä vai holistisia (Brookhart, 2018). Analyttisissä arviointimatriiseissa jokaiselle tavoitteelle on omat kriteerinsä, kun taas holistisessa arviointimatriisissa on vain yhdet kriteerit, joissa on huomioitu monta näkökulmaa. Perusopetuksen päättöarvioinnin arviointimatriisit ovat analyttisiä.

Arviointimatriisin tyypillä ei näytä olevan suoraviivaista yhteyttä annettujen arvosanojen vertailukelpoisuuteen (Brookhart, 2018), mutta on viitteitä, että matematiikan päättöarvioinnin tyyppisten tavoitteiden arviointi olisi hivenen yhdenmukaisempaa analyttistä arviointimatriisia käytettäessä (Jönsson et al., 2021). Kriteeriperustaisen arvioinnin toimivuuden tärkein edellytys on kuitenkin kriteerien laatu (Brookhart, 2018) ja kriteerien kehittäminen on tapa parantaa arvioinnin vertailtavuutta (Brookhart et al., 2016).

Opettajien näkemyksiä arviointikriteereistä on tutkittu yllättävän niukasti. Aihetta sivuavaa tutkimusta kuitenkin löytyy. O'Donovan ym. (2001) tutkivat opiskelijoiden näkemyksiä kriteereistä. He kertovat opiskelijoiden maininneen ongelmaksi yksityiskohtien puuttumisen ja sanoitusten epämääräisyyden. Opiskelijat esittivät huolensa siitä, että kyseisten ongelmien vuoksi arviointi on kriteereistä huolimatta hyvin subjektiivista ja opettajasta riippuvaa. Samantapaisia huomioita esittää Woolf (2004), joka hänkin tarkastelee kriteereitä korkeakoulutuksen puolella. Woolfin havaitsemia ongelmia olivat tulkinnanvaraiset sanoitukset, kuten ”*kypsä pohdinta*” ja ”*materiaalin asianmukainen laajuus*”, sillä tällaiset sanoitukset jäävät arvioijan subjektiivisiksi arvioiksi. Woolf mainitsee myös muita ongelmia. Toinen näistä on epäselvyys eri tavoitteiden painoarvosta suhteessa toisiinsa. Toinen on epäonnistumisen kuvaamisen kriteerissä, esimerkiksi ”*opiskelija yrittää osoittaa ymmärrystä aiheesta, mutta jotkin näkökulmat sekaantuvat*” (Woolf, 2004, s. 488).

Hopfenbeckin ym. (2012) tutkimus on lähempänä peruskoulun päättöarvioinnin kontekstia. Tutkimuksessa analysoitiin erilaisia Norjan opetushallituksen ja opettajien laatimia arviointikriteeristöjä sekä haastateltiin aiheesta opettajia. Tutkimustuloksena tunnistettiin kriteeristöissä seuraavia ongelmia: 1) Periaatteet arviointikriteerien taustalla olivat ristiriitaisia. Niissä oli esimerkiksi käytetty sekaisin Bloomin taksonomiaa, joka kuvaa erilaisia kognitiivisen prosessin tasoja, sekä oppilaan omatoimisuuden astetta, eli sitä kuinka itsenäisesti oppilas toimii. 2) Kriteeristöjen vaatimustasoissa oli eroja eli eri kriteeristöjen avulla annetut arvosanat eivät vastanneet toisiaan. 3) Kriteerit eivät erotelleet eri arvosanoja selkeästi. 4) Kriteerit eivät olleet linjassa opetussuunnitelman kanssa. Esimerkiksi matematiikan opetussuunnitelmassa tuotiin esille matematiikan sisältöjä, kun taas kriteerit keskittyivät taitoihin, kuten ongelmanratkaisuun ja kommunikointiin.

Hyvien kriteerien merkitys tunnustetaan siis yleisesti, mutta kirjallisuus tarjoaa edelleen niukasti tietoa siitä, minkälaiset sanotukset ovat hyviä tai ongelmallisia. Tämä tutkimus pyrkii osaltaan täyttämään kyseistä tutkimusaukkoa.

## Päätöarvioinnin perusteet fysiikassa ja matematiikassa

Opetussuunnitelman 2014 perusteissa oppiainekohtaiset kuvaukset koostuvat suurelta osin *opetuksen tavoitteista* ja tavoitteisiin liittyvistä *keskeisistä sisältöalueista*. Tarkennetuissa kriteereissä (Opetushallitus, 2020b) kuvataan opetuksen tavoitteen lisäksi *opetuksen tavoitteista johdettu oppimisen tavoite* (ks. taulukko 1). Kutsun näitä tässä artikkelissa yksinkertaisuuden vuoksi *tavoitteiksi* ja *sisällöiksi*. Päätöarvosanan muodostamisessa huomioidaan kaikki oppiaineen tavoitteet, joita ei ole erikseen rajattu päätöarvioinnin ulkopuolelle. Päätöarvosana kuvaa opetussuunnitelman perusteiden mukaan ”oppilaan osaamisen tasoa suhteessa kunkin oppiaineen oppimäärän tavoitteisiin ja päätöarvioinnin kriteereihin” (Opetushallitus, 2020a, s. 9). Jokaiselle päätöarvioinnissa huomioitavalle tavoitteelle on päätöarvioinnin kriteerit arvosanoille 5, 7, 8 ja 9 (ks. Opetushallitus, 2020b). Taulukossa 1 on esitetty esimerkki yhdestä tavoitteesta ja sen kriteereistä.

Taulukko 1. Esimerkki matematiikan arviointikriteereistä tavoitteelle T4.

Opetuksen tavoite	Opetuksen tavoitteista johdetut oppimisen tavoitteet	Osaamisen kuvaus arvosanalle 5
T4 kannustaa oppilasta harjaantumaan täsmälliseen matemaattiseen ilmaisuun suullisesti ja kirjallisesti	Oppilas ilmaisee matemaattista ajatteluaan täsmällisesti eri ilmaisukeinoja käyttäen.	Oppilas ilmaisee ohjattuna matemaattista ajatteluaan jollakin tavalla.
Osaamisen kuvaus arvosanalle 7	Osaamisen kuvaus arvosanalle 8	Osaamisen kuvaus arvosanalle 9
Oppilas ilmaisee matemaattista ajatteluaan joko suullisesti tai kirjallisesti.	Oppilas ilmaisee matemaattista ajatteluaan sekä suullisesti että kirjallisesti.	Oppilas ilmaisee perustellen matemaattista ajatteluaan.

Fysiikassa on yhteensä 15 tavoitetta ja ne on jaoteltu kolmeen ryhmään: 1) Merkitys, arvot ja asenteet, 2) Tutkimisen taidot sekä 3) Fysiikan tiedot ja niiden käyttäminen. Tavoitteet T3–T14 huomioidaan fysiikan päätöarvosanaa muodostettaessa. Matematiikassa on 20 tavoitetta, jotka on myös jaettu kolmeen ryhmään: 1) Merkitys, arvot ja asenteet, 2) Työskentelyn taidot ja 3) Käsitteelliset ja tiedonalakohtaiset tavoitteet. Ensimmäistä tavoitetta lukuun ottamatta kaikki huomioidaan päätöarvosanaa muodostettaessa. Oppiaineiden tavoiteryhmät ja tavoitteet on esitetty tiivistetysti taulukoissa 2 ja 3.

Fysiikan ja matematiikan tavoiteryhmien, tavoitteiden ja kriteerien ominaisuuksissa on eroja, osin johtuen oppiaineiden erilaisesta tiedon luonteesta, osin johtuen oppiainetyöryhmien tekemistä valinnoista. Esimerkki työryhmäkohtaisista valinnoista on tavoitteellisen työskentelyn arviointi (molemmissa oppiaineissa T2). Matematiikassa tavoite huomioidaan päättöarvioinnissa, fysiikassa ei. Matematiikan päättöarvioinnissa huomioitavat tavoitteet jakautuvat tavoitteellisen työskentelyn tavoitetta lukuun ottamatta kahteen ryhmään. *Työskentelyn taitojen* ryhmä (T3–T9) sisältää matemaattisen ajattelun ja työskentelytapojen taitoja. Työskentelyn taitoja ei ole sidottu mihinkään tiettyyn oppisisältöön, vaan niitä – esimerkiksi ongelmanratkaisua tai matemaattista ilmaisua – voi harjoitella ja arvioida kaikkien sisältöjen yhteydessä. Sen sijaan *Käsitteelliset ja tiedonalakohtaiset tavoitteet* (T10–T20) liittyvät tiettyyn oppisisältöön. Niitä siis harjoitellaan ja arvioidaan pääsääntöisesti tietyn sisällön kohdalla, esimerkiksi prosenttilaskennan tavoitetta prosenttilaskua harjoitellessa.

Matematiikan *Työskentelyn taitojen* tavoiteryhmää vastaa jossain määrin fysiikan *Tutkimisen taidot*, sillä fysiikan tieteenalalla työskentelyn taitojen voi ajatella koskevan tutkimuksen tekemistä. Taitoihin on fysiikassa sisällytetty TVT:n käyttö, samoin kuin matematiikassa. Kun matematiikan oppiaineessa ryhmässä toimimisen taidot on huomioitu tavoitteen T2 korkeimmissa kriteerikuvauksissa (arvosanat 8 ja 9), ne on fysiikassa sisällytetty projektityöskentelyyn osallistumisen tavoitteeseen T8.

Fysiikassa matematiikan kaltaisia puhtaita sisältötavoitteita ei ole yhtään. Oppisisältöjä on mainittu vain tavoitteessa T14, mutta siinäkin matematiikkaa yleisemmin, sillä kriteerikuvauksissa ei ole mainittu tarkkoja sisältövaatimuksia. *Fysiikan tiedot ja niiden käyttäminen* ei siis luettele tiettyjä käsitteitä, jotka pitäisi hallita, vaan oppilaan tavoitteena on esimerkiksi käyttää malleja ja käsitteitä. Tietotavoitteet on näin ollen määritelty matematiikassa tarkemmin kuin fysiikassa.

*Taulukko 2. Fysiikan tavoitteet tutkijan tiivistämänä. Yksityiskohtaiset muotoilut löytyvät Perusopetuksen opetussuunnitelman perusteista (2014).*

Tavoiteryhmä	Tavoite	Arvioinnin kohde
Merkitys, arvot ja asenteet	T1	Motivaatio fysiikan opiskelua kohtaan*
	T2	Tavoitteellinen työskentely*
	T3	Fysiikan merkityksen ymmärtäminen
	T4	Kestävä kehityksen tiedot ja taidot
Tutkimisen taidot	T5	Tutkimuskysymysten muodostaminen
	T6	Kokeellisen tutkimuksen toteuttaminen
	T7	Tutkimustulosten käsittely ja analysointi
	T8	Teknologiaprojektin tekeminen ryhmässä
	T9	TVT:n ja simulaatioiden käyttö

Fysiikan tiedot ja niiden käyttäminen	T10	Fysiikan käsitteiden käyttö ja jäsentyminen
	T11	Mallien käyttäminen
	T12	Tiedonhaku, tiedon arviointi ja argumentaatio
	T13	Luonnontieteellisen tiedon luonteen hahmottaminen
	T14	Jatko-opintovalmiudet mekaniikasta ja sähköstä
	T15	Fysiikan tietojen ja taitojen käyttäminen monialaisessa projektissa*

\*Ei huomioida päättöarvioinnissa

*Taulukko 3. Matematiikan tavoitteet tutkijan tiivistämänä. Yksityiskohdattaiset muotoilut löytyvät Perusopetuksen opetussuunnitelman perusteista (2014).*

Tavoiteryhmä	Tavoitteen numero	Arvioinnin kohde
Merkitys, arvot ja asenteet	T1	Motivaatio matematiikan opiskelua kohtaan*
	T2	Tavoitteellinen työskentely
Työskentelyn taidot	T3	Asioiden välisten yhteyksien ymmärtäminen
	T4	Matemaattinen ilmaisu
	T5	Ongelmanratkaisutaidot
	T6	Taito kehittää ja arvioida matemaattisia ratkaisuja
	T7	Matematiikan soveltaminen eri ympäristöissä
	T8	Tiedonhaku ja tiedon arviointi
	T9	TVT:n soveltaminen matematiikassa
Käsitteelliset ja tiedonalakohtaiset tavoitteet	T10	Päätely- ja päässälaskutaito
	T11	Peruslaskutoimitukset rationaaliluvuilla
	T12	Lukukäsitteen ymmärtäminen
	T13	Prosenttilaskennan osaaminen
	T14	Tuntemattoman käsitteen ymmärtäminen ja yhtälönratkaisutaidot
	T15	Muuttujan ja funktion ymmärtäminen
	T16	Tasogeometrian käsitteiden ymmärtäminen ja käyttö
	T17	Suorakulmaisen kolmion ja ympyrän ominaisuuksien hahmottaminen
	T18	Pinta-alojen ja tilavuuksien laskutaito
	T19	Tilastojen ja todennäköisyyksien osaaminen
T20	Ohjelmointitaidot	

\*Ei huomioida päättöarvioinnissa

Matematiikan ja fysiikan oppiaineiden tavoitteissa ja kriteereissä on eroja ja yhtäläisyyksiä. Yhtäläistä on kolme saman tyylistä tavoiteryhmää: asenteet, taidot ja tiedot. Merkittävin ero on, että matematiikassa noin puolet tavoitteista on sisältöihin sidottuja, jolloin niitä arvioidaan vain tiettyjen sisältöjen yhteydessä. Fysiikassa tavoitteet ovat yhtä lukuun ottamatta läpileikkaavia, eli niiden harjoittelu ja arviointi on mahdollista läpi kaikkien oppisisältöjen.

# Tutkimusmetodologiset lähtökohdat

## *Tutkimusaineisto*

Tutkimusaineisto koottiin osana Opetus- ja kulttuuriministeriön rahoittamaa Päätösarviointiuudistuksen tutkimusta (Parvi; ks. myös hankkeen raportti, Pulkkinen ym., 2024). Parvi-hankkeessa kerättiin valtakunnallisesti kattava kyselyaineisto yhteensä 226 koulun opettajilta. Kyselyn osallistujille tarjottiin mahdollisuus osallistua tutkimushaastatteluun. Haastatteluvien määrää täydennettiin ottamalla yhteyttä kaikkiin Suomen yhtenäis- ja yläkoulujen rehtoreihin ja pyytämällä heitä välittämään opettajilleen tiedon haastattelumahdollisuudesta. Haastatellut opettajat olivat eri ikäisiä ja eri vaiheissa työuraansa, heitä oli sekä suomen- että ruotsinkielisestä kouluista, ja eri puolilta Suomea. Osallistujissa oli monia arvioinnista kiinnostuneita opettajia, kuten arviointituutoreita ja arvioinnin alueellisessa kehitystyössä mukana olleita opettajia, mikä johtui todennäköisesti haastattelun vapaaehtoisuudesta. Opettajia haastateltiin lähtökohtaisesti yhden oppiaineen näkökulmasta, mutta muutama opettaja vastasi kattavasti kysymyksiin sekä fysiikan että matematiikan osalta. Haastateltavia oli 28, joista 15 vastasi ensisijaisesti matematiikan ja 13 fysiikan näkökulmasta.

Haastattelutilanteessa tutkija asemoi itsensä viestinviejäksi opettajien ja päätöksentekijöiden välillä, eikä arvottanut opettajien näkemyksiä tai toimia. Haastateltavat olivat motivoituneita kertomaan näkemyksistään ja puhuivat haastattelutilanteessa hyvin avoimesti näkemyksistään ja toimistaan. Puolistrukturoiduissa haastatteluissa opettajilta tiedusteltiin näkemyksiä arvioinnista yleensä ja sitten arviointiuudistuksesta, kokemuksia uudistuksen toimeenpanosta koululla ja uudistuksen vaikutuksesta käytännön työhön. Tämän jälkeen siirryttiin oppiainekohtaisiin kysymyksiin, joissa tukena käytettiin oppiaineen päätösarvioinnin kriteereitä. Opettajilta kysyttiin arviointikriteerien käyttötavoista ja käyttökelpoisuudesta, kriteerien toimivuudesta erilaisille oppilaille, tavoitteiden painottamisesta ja työskentelyn arvioinnista. Haastattelun lopussa opettajilta pyydettiin palautetta kriteereistä sekä kysyttiin mihin suuntaan he toivoisivat arviointia kehitettävän. Kaikki haastattelut tehtiin videotapaamisena, tallennettiin sekä litteroitiin muutoin sanasta sanaan, mutta täytesanat, myötäilyt ja toistot jätettiin pois. Haastattelujen keskipituus oli 51 minuuttia.

## *Analyysi*

Analyysi seurasi aineistolähtöisen sisällönanalyysin vaiheita (Schreier, 2012). Kirjoittaja teki kaikki haastattelut itse ja analysoi niitä myös muuhun tutkimuskäyttöön, joten hän tunsi aineiston hyvin. Analyysi aloitettiin tunnistamalla Atlas.ti -ohjelmassa haastatteluiden litteraateista otteita, joissa opettajat kuvailivat näkemyksiään tavoitteista ja kriteereistä. Tämän tuloksena löytyi 223 otetta, joiden keskimääräinen pituus oli 141 sanaa. Helpottaakseen otteiden käsittelyä, kirjoittaja tiivistä jokaisen otteen ydin-sanoman 1–3 lyhyeen lauseeseen, pyrkien käyttämään niissä puhujan omia sanavalintoja. Jos otteessa viitattiin oppiaineen tiettyyn tavoitteeseen tai



sen kriteeriin, kuvaukseen lisättiin tavoitteen numero. Esimerkiksi ote ”Ja tällöinen käsite, vaikka täällä on fysiikan tiedot ja niiden käyttäminen, tämä nyt on tässä mulla auki. Niin kun 5:ssä on, että oppilas selittää fysiikan ilmiöitä käyttäen joitakin fysiikan käsitteitä. Niin tähän on nyt taas hyvin ympäripyöreä. Ja jos on taas 7, ja siellä on käyttäen fysiikan keskeisiä, että se ero 7:ssä ja 5:ssä on se, että osaa käyttää jotakin, mutta 7:ssä sentään keskeisiä käsitteitä. Niin vähän ympäripyöreitä ilmauksia.” tiivistettiin muotoon ”FT10 Kriteerien muotoilu ympäripyöreä: jotain käsitteitä vs. keskeisiä käsitteitä”.

Sisällönanalyysia jatkettiin viemällä tiivistelmät digitaaliselle Miro-alustalle, kummankin oppiaineen tiivistelmät eri värisille post-it-lapuille. Lapuilla esitettyjä näkemyksiä alettiin oppiaine kerrallaan ryhmitellä sisältölähtöisesti. Jatkokäsittelyn helpottamiseksi aineisto-otteet ja niitä vastaavat tiivistelmät vietiin oppiaineittain Excel-tiedostoon, jota pidettiin ryhmittelyn aikana auki Miron rinnalla, jolloin tiivistelmiä pystyi vertaamaan alkuperäisiin aineistonäytteisiin, ja varmistamaan että näytteen koko sisältö huomioitiin. Ensin fysiikan näkemyksiä ryhmiteltiin niin kauan, että saatiin muodostettua koherentteja, toisistaan irrallisia ryhmiä, eikä ryhmiin tai yksittäisten lappujen sijoitteluun tullut enää muutoksia. Tämän jälkeen ryhmät nimettiin. Esimerkiksi edellä mainittu tiivistelmä löysi paikkansa ryhmästä, jonka nimi oli ”Tarvitaan konkretiaa (sisältöjä)”, jossa käsiteltiin konkretian puutetta fysiikan kriteereissä (ks. Kuva 1)

Tarvitaan konkretiaa (sisältöjä)					
T10 Kriteerien muotoilu ympäripyöreä: jotain käsitteitä vs. keskeisiä käsitteitä	FT10 Vitonen ja kutonen tosi hämäävä kun ei selviä mitä käsitteitä pitää olla.	Aika vähän konkretiaa kriteereissä, taivaanrantaa maaläilevät	FT14 Tarvitaan tarkemmat määrittelyt, mitä jatko-opintovalmiudet	Lisää konkretiaa eli tarkempia sisältöjä kriteereihin (ei olioita).	Sisältöä voisi mainita tarkemmin kriteereissä
Kriteerien muotoilussa ”joitakin” on epäselvä. Pitäisi listata mikä on minimi.	Kriteereistä puuttuu täsmäjutut, joista oppilas ymmärtäisi mitä vaaditaan	Fysiikassa sisältöä on mainittu vain 1-2 tavoitteessa. Aika vaikeaa fysiikan koulutuksella.	FT10 Käsitteet pitäisi mainita tarkemmin käsitetavoitteessa	FT14 Vain yksi tavoite puhuu sisällöistä, eikä sekään tarkasti.	Kriteereihin voisi lisätä konkretiaa. ”Joitakin käsitteitä” --> mainitaan käsitteet

Kuva 1. Esimerkki Miro-alustalle muodostetusta ryhmästä ja sen otsikosta.

Seuraavaksi matematiikan näkemykset ryhmiteltiin samalla periaatteella. Lopuksi fysiikan ja matematiikan ryhmiä vertailtiin keskenään. Havaittiin, että ryhmittelyissä oli samankaltaisuutta, mutta myös eroja. Eri oppiaineiden sisällöltään samankaltaiset ryhmät ryhmiteltiin Miro-alustalla vierekkäin ja lopuksi näille ryhmistä koostuville kokonaisuuksille annettiin niiden aihepiiriä kuvaavat yläotsikot. Kuvan 1 ryhmä päättyi kokonaisuuteen, jonka otsikko oli ”Konkretiasta laajasti”, ja sen vastinpari matematiikan puolelta oli ryhmä, jossa käsiteltiin ”ohjattuna”-termin käyttöä.” Tulosluku on jäsennetty näiden kokonaisuuksien mukaan.

## Tulokset

### *Opetuksen ja arvioinnin tarpeet ovat erilaisia*

Opettajien suhtautuminen fysiikan kriteereihin oli jännitteistä. Arvioitavien tavoitteiden kokonaisuutta pidettiin yleisesti hyvänä. Tavoitteiden monipuolisuutta arvostettiin ja sen nähtiin antavan oppilaille fysiikan tietojen lisäksi laajempaa tieteen ymmärrystä. Toisaalta oppilaan osaamistason määrittäminen kriteerien perusteella koettiin vaikeana. Opettajat tunsivat jännitteen ja kuvasivat sitä vastauksissaan esimerkiksi seuraavasti.

*Sitten siinä arviointityössä tai siinä, kun tuota kriteeristöä laaditaan, pitäisi miettiä tosi tarkkaan sitä, että mistä tavoitteista on edes mahdollista tehdä mitattavia näyttöjä. ... Tai sitten hyväksyä se, että näin ei, että se ei ole mahdollista. Mä olen itse hyväksynyt, että se ei ole mahdollista, ja noista läpileikkaavista ja jostain just varsinkin työskentelyn taidoista, niin sitä on käytännössä mahdotonta mitata millään järkevällä asteikolla. Niin jos tuota jotenkin halutaan kehittää, niin ainakin tiedostaa tämä haasteellisuus. Ei ole mitään varsinaista neuvoa siihen, koska tavoitteet itsessään on hyviä. Mutta sitten kun tuo järjestelmä pakottaa arvioimaan niitä tavoitteita, niin sitten ne ei enää toimi siihen. (V27)*

Esimerkkeinä tällaisista vaikeasti mitattavista taidoista mainittiin fysiikassa erityisesti tavoitteet T5, T12 ja T13. Näiden problematiikkaa arvioinnin osalta on kuvattu seuraavissa lainauksissa, ensimmäisessä tavoitteen T5 ja jälkimmäisessä tavoitteiden T12 ja T13 osalta.

*Ehkä yleisesti kysymyksiä kysyminen. Pedagogisesta näkökulmasta ymmärrän, miksi se on tärkeää. ... Olen kysymysten kysymistä aika paljon tehnyt ja testannut eri tavoilla, mutta se on silti mun mielestä arvioinnin kohteena hankala. Koska se on kuitenkin kauhean henkilökohtainen ominaisuus yläkouluikäisellä ihmisellä. ... Se on ehkä sellainen, missä aika paljon haastetta siinä, että mikä on oppilastodellisuus ja mikä on kriteeritodellisuus. (V6)*

*Tietolähteitten käyttäminen ja tuo fysiikan tai kemian luonnontieteellisen tiedon hahmottaminen on semmoisia vähän vaikeasti arvioitavia. ... Hyviä tavoitteita, mutta kun vaikea konkreettisesti sanoa, mikä se osaaminen on tuossa just tarkalleen. (V9)*

Matematiikassa tavoitteiden ja kriteerien kokonaisuus sai fysiikkaa positiivisempaa palautetta. Sekä tavoitteita, kriteereitä että niiden muodostama kokonaisuutta pidettiin vähintäänkin melko hyvänä myös osaamistason määrittämiseen. Seuraava lainaus kuvaa tällaista näkemystä:

*Mä itse näen, että matikan puolella tämä on toimiva homma ja vielä, kun tätä viedään eteenpäin tästä, ehkä tulee joskus vielä muuttumaankin ne kriteerit ja selkiytymäänkin, mutta matikan puolella ne on pääasiassa melko selkeät. Mulla on sellainen olo, että joku matemaatikko on ollut niitä laatimassa. (V23)*

Matematiikassa työskentelyn tavoitteiden ryhmässä (T3–T9) koettiin kuitenkin olevan saman tapaista problematiikkaa arvioinnin osalta kuin fysiikan puolella. Tästä esimerkkinä on seuraava lainaus, jossa opettaja puhuu tavoitteesta T3.

*Mä vähentäisin ylipäättänsä tätä, kun täällä on aika paljon näitä ”havaitsee ohjattuna opittavien asioiden välisiä yhteyksiä” ja tämmöisiä, koska näitä on aika vaikea sitten loppujen lopuksi mitata, että mitä tämä nyt sitten tarkoittaa, ja joka kertako mä nyt erikseen sitä kyselen häneltä, jokaiselta, ja painan vielä mieleen.” (V25)*

Toisaalta myös matematiikassa näkyi opetuksen ja arvioinnin tarpeiden välinen jännite, sillä opettajat toisaalta arvostivat näitä vaikeasti arviotavaksi näkemiään työskentelyn tavoitteita. Opettaja kuvaa näiden tavoitteiden merkitystä seuraavassa otteessa.

*Mä uskon, että just nuo T1-T9 ovat ne tärkeämmät. Joo, onhan se hyvä, että he osaavat piirtää jotain geometrista muotoa taikka prosenttilaskenta on aika tärkeä. Että joo, sitähan pitää vaatia, että voidaan testata heitä jollain tavalla. Mutta mä kyllä uskon, että nämä jutut, kun rohkaistaan oppilasta soveltamaan matematiikkaa muissakin oppiaineissa on paljon opiskelijan näkökulmasta hyödyllisempää kuin että päättätään jotain geometrista kaavaa. Mä pyrin kyllä siihen, että fokuroidaan T1:stä T9:iin sellaisella tavalla, että pyritään näyttämään, että tällä tavalla tätä matikkajuttua, esim. prosenttiyksikköä tai tilastotietoa voidaan käyttää. (V24)*

### ***Konkretian lisääminen on tärkeää***

Vaikka opettajat arvostivat laaja-alaisia ja vaikeammin arvioitavia tavoitteita opiskelun ja oppiaineen luonteen näkökulmista, he toivoivat päättöarvioinnin tarpeisiin lisää konkretiaa sekä matematiikassa että fysiikassa. Molemmissa toivottiin tarkennusta kriteereissä käytettyihin verbeihin. Opettajat pohtivat muun muassa, minkälaista toimintaa on ”osaaminen”, ”soveltaminen” ja ”käyttäminen”, ja epäilivät, että eri opettajien välillä verbien tulkinta saattaa erota paljon. Seuraava opettajan vastaus kuvaa verbien tulkinnanvaraisuuden problematiikkaa fysiikassa ja matematiikassa.

*Fysiikan tai kemian päättöarviointeja olen tehnyt tässä tänä vuonna, niin niissä ainakin tulee sama mikä näissäkin, että kun siellä on hienoja sanoja, että osaa käyttää ja tunnistaa ja ymmärtää, niin sitten on, että mitä se tarkoittaa. Jos sanotaan, että oppilas osaa käyttää yhtälöparia ongelmanratkaisussa. No minkälaisessa, kuinka vaikeassa ratkaisussa, miten sen pitää osata käyttää sitä? Tavallaan se, kun nämä on tietyllä tavalla ympäröityjä. Niin sen takia itselle on ehkä jäänyt enemmän se ajatus siitä, että tämmöisiin pitäisi keskittyä, mutta sitten juuri se, että miten niitä käytetään sitten oikeasti arvioinnissa, niin se on mun mielestä jotenkin haastavampi. (V4)*

Matematiikassa opettajien palaute kohdistui myös yksittäisiin kriteereihin, fysiikassa vain yleiselle tasolle. Opettajat olivat esimerkiksi havainneet,

että sieventämisen osaamistasoa ei ole kuvattu ylemmissä kriteereissä, ja jäi epäselväksi, pitääkö oppilaan osata sieventää esimerkiksi potenssilausekkeita. Toinen yksittäinen havainto oli epäsystemaattisuus eri tavoitteiden välillä. Moni ihmetteli, miksi oppilas saa joissain tavoitteissa osoittaa arvosanassa 5 kuvatun taitotason ohjattuna, mutta ei kuitenkaan tavoitteen T17 kohdalla. Seuraava lainaus on esimerkki tästä huomiosta.

*Muutenkin musta on niin kun mielenkiintoista, että täällä on aika paljon tämmöisiä että ”oppilas määrittää ohjattuna tyyppiärvon” monessa kohdassa lukee ohjattuna jotain, ja sitten osassa kohdassa ei ole tätä ohjattuna, niin musta on niin kun jännittävää että tässä hypotenuusan laskemisessa ei ole sitä ohjattuna, koska musta se olisi huomattavasti haastavampaa oppilaille kun esimerkiksi tilastollisissa tunnusluvuissa oppilas määrittää ohjattuna tyyppiärvon ja mediaanin. (V6)*

Fysiikan ja matematiikan kriteereistä nostettiin molemmista esille yksi omanlaisensa epäkohta, joka liitettiin kokonaan tai osittain konkretian puutteeseen. Fysiikassa se oli konkreettisten sisältöjen puuttuminen kriteerikuvauksista. Matematiikassa ongelmaksi koettiin ”ohjattuna” -termin käyttäminen, esimerkiksi ”oppilas jäsentää ohjattuna ongelmia”. Nämä on kuvattu seuraavissa alaluvuissa.

### *Sisältöjen puute fysiikassa*

Sisältöjen puute fysiikan kriteereissä herätti laajaa kritiikkiä. Valtaosa haastatelluista toivoi fysiikan kriteereihin lisää konkretiaa ja tarkoitti tällä erityisesti selkeästi nimettyjä sisältöjä. Konkretian puutteen nähtiin vaikeuttavan oppilaan vertailukelpoista arviointia ja kriteereistä viestimistä oppilaille. Opettajat kokivat epäkohdan merkittävänä. Esimerkiksi seuraavan otteen opettajan mielestä sisältöjen puute oli kaikkein päällimmäisin ajatus koko uudistuksesta:

*K: Mitkä on sulle ne päällimmäiset tunteet tästä arviointiuudistuksesta?*

*V: No nyt kun sen kanssa alkaa olla vähän paremmin sinut, niin ei se nyt ihan pelkkää pahaakaan ole, mutta se fysiikka on kyllä tosi outo. Siinä mä kyllä olen sitä mieltä, että se on kyllä tosi outo. Kun siinä on, arviointiin vaikuttaa 12 tavoitetta, ja niistä tosiaan yksi ainoa oikeastan puhuu fysiikan sisällöistä, ja sekään ei siis kerro, mitä niistä pitäisi osata. Mitä se on, se vitosen osaaminen fysiikassa tai kasin osaaminen, mitä on hyvä osaaminen fysiikassa? Vaan siinä sanotaan, että jatko-opintovalmiudet. (V12)*

Opettaja viittaa tässä fysiikan tavoitteeseen T14, eli tiedollisten jatko-opintovalmiuksien saavuttamiseen vuorovaikutuksesta ja liikkeestä sekä sähköstä, joka on hänen näkemyksensä mukaan, ja aivan perustellusti, ainoa tavoite, jossa mainitaan tiettyjä sisältöjä.

Lisää konkretiaa toivottiin tavoitteiden lisäksi myös kriteeritasolla. Nykyisiä ilmauksia ”joitain käsitteitä” ja ”keskeisiä käsitteitä”, joita käytetään useissa kriteereissä, ja jotka luovat joissain tavoitteissa eron arvosanojen

5 ja 7 osaamiseen, pidettiin liian epätarkkoina osaamisen määrittämiseen. Seuraavassa otteessa opettaja kuvaa asiaa tarkemmin:

*Ja tämmöinen käsite, vaikka täällä on fysiikan tiedot ja niiden käyttäminen, tämä nyt on tässä mulla auki. Niin kun vitosessa on, että oppilas selittää fysiikan ilmiöitä käyttäen joitakin fysiikan käsitteitä. Niin täähän on nyt taas hyvin ympäröityä. Ja jos on taas seiska, ja siellä on käyttäen fysiikan keskeisiä, että se ero seiskassa ja vitosessa on se, että osaa käyttää jotakin, mutta seiskassa sentään keskeisiä käsitteitä. Niin vähän ympäröityä ilmauksia. Mikä on toisaalta hyvä, koska siinä myös luotetaan tietyllä tavalla siihen opettajan aineosaamiseen ja tähän pedagogiseen hallintaan. Mutta ei ne ohjaa kyllä mihinkään selkeään arviointitulokseen. (V5)*

### *'Ohjattuna' matematiikassa*

Matematiikan kriteereissä runsaasti käytetty ohjattuna -termi koettiin huonoksi monestakin syystä. Ensimmäinen näistä liittyi konkretiaan, sillä ohjauksen sallittu taso jäi opettajille epäselväksi: Kuinka paljon ja millä tavoin oppilasta saa ohjata, jotta suoritus on vielä hyväksytty? Aiheesta oli myös kouluilla erilaisia tulkintoja yksittäisten opettajien ja opettajaryhmien, lähinnä aineenopettajien ja erityisopettajien, välillä. Opettajat myös kokivat, että ”ohjattu” -sanana käyttäminen kriteereissä vierittää vastuuta oppilaalta opettajalle. Hylätty suoritus voidaan pahimmillaan tulkita johtuvan riittämättömästä tai heikkotasoisesta ohjauksesta. Seuraavassa otteessa opettaja kuvaa näitä näkökulmia:

*Kuinka vahvasti voi ohjata, koska jokainen oppilas ylittää vitosen kriteereihin, jos se on ohjatusti, koska ohjaaminen voi jo olla sitä, että mä liikutan sitä kynää. [naurahtaa] Näin kärjistetysti, että mikä on se ohjaamisen raja, kuinka paljon ohjataan, koska sitten tässä tullaan siihen, että meillä on vitosia ja vitosia. Mikä on se vitosen alaraja? Itse mä kaipaaisin sitä, että siellä on selkeästi, että osaa, ei ohjattuna, vaan ne on ne taidot, mitä sen pitää osata. ... Sitten se on aina, jos se on ohjattuna, että miksi mä olen ohjannut niin huonosti. (V24)*

Ohjattuna -termin ongelmallisuuden ei nähty johtuvan ainoastaan epäselvästä määrittelystä. Opettajat eivät pitäneet oppilaiden elämän kannalta kovin arvokkaana ”ohjattuna osaamisen” taitoa. He pitivät tärkeämpänä joidenkin – vaikka sitten vain muutamien – taitojen ja tietojen itsenäistä osaamista. Tämä näkökulma näkyy seuraavassa vastauksessa:

*Selkeästi sitä arviointia, mä vielä kävisin ne ehkä suurennuslasilla läpi ja nimenomaan poistaisin ne ohjattuna -kohdat. Edelleen mä olen sitä mieltä, että mitä oppilas hyötyy siitä, että se osaa ohjattuna tehdä jotain, kun mun mielestä meidän pitäisi tähdätä nimenomaan siihen, että oppilaalla on itsenäisesti taidot ja valmiudet suoriutua arkipäivän haasteista. Miksi me kulutetaan aikaa siihen, että oppilas osaa ohjattuna tehdä jotain sen sijaan, että me panostettaisiin siihen, että se osaa oikeasti kunnolla tehdä jotain semmoista, mistä sillä äärimmäisen suurella todennäköisyydellä tulee olemaan elämässä hyötyä? (V24)*

Lisäksi pidettiin huonona, että kriteerit antoivat oppilaiden ymmärtää ohjauksen määrän vaikuttavan arvosanaan. Seuraavassa otteessa opettaja kuvaa, miten sanointu voi saada oppilaat vieroksumaan opettajan tukea:

*Ja mäkin itse asiassa sanoin, mä en ole lapsille puhunut tuosta yhtään mitään, tuosta, että tuo ”ohjattuna”-teksti vaikuttaa siihen arvosanaan, koska heikkojen oppilaiden kanssa käy niin, että jos mä sanon, että jos sua paljon ohjataan, sä saat vitosen. Ja jos sua ohjataan vähän, sä saat seiskan, niin ne ei ota vastaan sitä ohjausta, koska ne nyt jo saattaa, että ”mene pois”, että ohjaajalle sanoo, ei ne mulle sano, mutta joka tapauksessa, jos ne tietäisi sen, että se vaikuttaa siihen työskentelyn arviointiin, niin nehän kuvittelisi, että kun he ei ota ohjausta vastaan, niin he saa seiskan. Eli itse ajattelen sen niin, että sillä, että paljonko mää työskentelyn aikana heitä ohjaan, ei pitäisi olla mitään vaikutusta. (V25)*

Fysiikassa ei tällaista kritiikkiä herännyt. Tämä johtuu todennäköisesti siitä, että fysiikan kriteereissä vastaavaa kuvausta (ohjatusti) käytetään kolmesti, mutta matematiikassa 22 kertaa.

### ***Kriteerikuvaukset vaativat kalibrointia***

Kriteerien laadun lisäksi tärkeää on niiden skaalaus, eli yksittäisten kriteerien ja niiden kokonaisuuden vaatimustaso. Ongelmaksi koettiin lähinnä matalat ja korkeat arvosanat. Ne käsitellään seuraavissa alaluvuissa erikseen.

#### *Korkeat arvosanat*

Sekä matematiikassa että fysiikassa arvosanan 9 kriteereitä pidettiin yleisesti liian tiukkoina. Arvosanan 9 kriteerit koettiin niin vaativiksi, että oli epäselvää, mitä lisää kymppiin voisi enää vaatia. Opettajat ehdottivat, että arvosanalla 8 tai 8,5 pitäisi saada kiitettävä arvosana. Joku puolestaan tulkitse, että jos kaikki arvosanan 9 kriteerit täytyvät, todistusarvosana on 10.

Fysiikan kiitettävän arvosanan kriteerikuvauksia kuvailtiin haastaviksi myös opetuksen järjestämisen ja oppilaan oman toiminnan ohjaamisen näkökulmista. Opettajat kokivat, että oppilaan on mahdoton päästä arvosanan 9 kuvauksiin osallistumalla perinteiseen opettajajohtoiseen opetukseen, joten opettajan täytyy muuttaa pedagogiikkaansa juuri arvosanan 9 kriteerien vuoksi. Opettaja näkivät kiitettävien arvosanojen vaativan oppilailta fysiikan osaamisen vuoksi myös paljon omatoimisuutta. Eräs opettaja kuvasi asiaa seuraavasti:

*Meillä kuitenkin, jos ajatellaan vaikka tutkimustaitoja, meillä aika vähän on, kun on isot ryhmät, mahdollisuutta toteuttaa tosi avoimia tutkimuksia esimerkiksi, että suunnittele itse tutkimus ja kehitä kysymykset ja tee siitä koko juttu. Siellä aika paljon kuitenkin kymppin oppilaalta*

*esimerkiksi tai kiitettävän oppilaalta vaadittaisiin sitä, että se pystyisi täysin itsenäisesti suunnittelemaan ja toteuttamaan asioita. Mun mielestä hirveän paljon omatoimisuutta ja itsenäisyyttä, mitä ei ehkä kuitenkaan tavallisen peruskoululuokan kanssa hirveästi pysty välttämättä toteuttamaan. (V6)*

Arvosanalle 10 toivottiin omaa kriteeristöä, sillä kriteereistä ei käynyt ilmi, kuinka paljon arvosanan 9 vaatimusten tulisi ylittyä, jotta kympin voisi antaa. Asiaa pidettiin tärkeänä, sillä alueilla, joilla kilpailu lukioihin on kovaa, juuri arvosanojen 9 ja 10 erottelu nähtiin jatko-opintovalintojen kannalta tärkeimpänä rajanvetona.

### *Hyväksytyn ja hylätyn raja*

Arvosanan 5 kriteereitä pidettiin puolestaan helppoina. Sekä matematiikassa että fysiikassa useat opettajat mainitsivat, että vitosen kriteeri täyttyy, kun oppilas vain saapuu oppitunnille, ja että hylättyä arvosanaa on mahdoton antaa. Opettajien näkemykset jakautuivat sen suhteen, oliko alhainen vaatavuustaso hyvä vai huono asia. Osa opettajista koki, että rajan kiristäminen ei ole tarpeen, sillä se olisi hankalaa heikoille oppilaille tai aiheuttaisi haasteita esimerkiksi tukitoimien järjestämisen kanssa nykyisillä resursseilla. Alhainen vaatimustaso ei aina herättänyt suuria tunteita, esimerkiksi seuraavassa vastauksessa fysiikan opettaja suhtautuu asiaan varsin neutraalisti:

*Nuo arvosanojen tarkemmat määrittelyt. Se vitosen taso on se keskeisin. Siitä mä olen itse saanut sen välineen, että sellaiset oppilaat, jotka olisi muuten laitettu tähtimerkinnälle, niin mä olen katsonut, että ei, se ei ole opetussuunnitelman mukaista. Niin fysiikassa niitä sitten ei laiteta. Ja niitä ehkä on muissa aineissa laitettu, mutta fysiikassa ei tuosta johtuen ole laitettu [naurahtaa].”*

*[Haastattelija]: Sanoisitko sä, että tämä on jotenkin vaikuttanut työn kuormitukseen tai ajankäyttöön?*

*Jos jotenkin, niin helpottavasti. (V10)*

Opettaja siis kertoo, että arvosanan 5 taso on nyt fysiikassa niin alhainen, että jokainen oppilas pystyy pääsemään arvosanaan viisi, eikä yhdenkään oppilaan oppimäärää tarvitse yksilöllistää (tähtiarvosana). Hän pitää kuitenkin tilannetta hyvänä tai neutraalina.

Osa opettajista taas toivoi tiukempia vaatimuksia arvosanan 5 suhteen, jotta oppilailta voisi perustellusti vaatia työntekoa ja osaamista. Tällaisesta ajattelusta esimerkkinä on seuraava matematiikan opettajan näkemys:

*Mä koen että ei tällä hetkellä nelosta voi antaa. Kaikki saavuttaa, kaikki on paikalla, okei ehkä me sitten vain pakotetaan ne sinne luokkaan, me vaaditaan niiltä sanallisia vastauksia, mutta mulla on oppilaita, jotka ei suostu kirjoittamaan yhtäkään tehtävää. ...*

*[Haastattelija]: Minkälainen muutos auttaisi sua tuossa?*

*Siinäkin jotkut minimikriteerit, että nämä asiat on osattava ja check-lista, kyllä. Silloin voi päästä matematiikasta läpi peruskoulusta. (V27)*

Opettajan mukaan hylätyn arvosanan antaminen on matematiikan kriteerien perusteella käytännössä mahdotonta. Hän ehdottaa hyväksytyn rajalle tarkempia vaatimuksia, esimerkiksi yksilöityä listaa asioista, jotka oppilaan pitää osata.

## Pohdinta

Tämän tutkimuksen tavoitteena oli ymmärtää, millaisia ongelmia ja kehitystarpeita matematiikan ja fysiikan opettajat näkevät päättöarvioinnin tavoitteissa ja kriteereissä. Tarkoituksena ei ollut arvioida uudistuksen onnistumista tai opettajien perehtyneisyyttä, vaan nostaa esille näkökulmia kriteerien kehitystyön tueksi.

Opettajien näkemyksistä käy ilmi, että tavoitteiden ja kriteerien toimivuus opetuksen järjestämisessä on eri asia kuin niiden toimivuus päättöarvioinnissa. Samaa ristiriitaa sivuaa Andrade (2023) kun hän kuvaa, miten oppimista ohjaavat kriteerit voivat olla haastavia arvosanan antamisessa, ja arvosanan määrittämiseen soveltuvat kriteerit taas oppimisen tukemisessa. Matematiikassa opettajat kokivat tavoitteiden kokonaisuuden pääsääntöisesti toimivana niin opetuksessa kuin arvioinnissakin, joskin työskentelyn tavoitteet (T3-T9) nähtiin Andraden (2023) huomion suuntaisesti parempana oppimisen ohjaamiseen kuin arvosanan määrittämiseen. Fysiikassa kriteerien kokonaisuutta pidettiin hyvänä opetuksen järjestämiseen, mutta ongelmallisena päättöarviointiin. On mahdollista, että näkemykset käyttökelpoisuudesta ovat vaikuttaneet myös niiden käyttöön. Opettajat antoivat matematiikan kriteereistä yksityiskohtaista kriteerikohtaista palautetta, kun fysiikassa palaute jäi yleisemmälle tasolle. Tämä voi selittyä sillä, että toimivammaksi koettuja kriteereitä on käytetty aktiivisemmin, jolloin yksittäisten kriteerien ongelmat ovat nousseet esille.

Päättöarvioinnin yhdenmukaisuuden näkökulmasta opettajat toivoivat lisää tarkkuuta kriteereihin. He kaipasivat tarkemmin kuvattuja sisältöjä fysiikassa, yhteisen ymmärryksen rakentamista kriteereissä käytetyistä verbeistä ja yksittäisten ongelmalliseksi koettujen tavoitteiden ja kriteerien korjaamista. Samantapaisia kriteerien epätarkkuuden ongelmia on nostettu esille myös aiemmissa tutkimuksissa (Hopfenbeck ym., 2012; O'Donovan ym., 2001; Woolf, 2004), ja on selvää, että riittävän täsmällinen ilmaisu on tarpeellista. Aihe ei kuitenkaan ole yksioikoinen, sillä hyvin yksityiskohtaisten kriteerien käyttö voi olla kuormittavaa (O'Donovan ym., 2001), sekä kontrolloida ja kapeuttaa oppimista haitallisella tavoin (Bloxham et al., 2016); Cowie & Harrison, 2016).

Kritiikki ohjattuna -termiä kohtaan matematiikan oppiaineessa ei liity ainoastaan termin epätarkkaan määrittelyyn. Kritiikin taustalla näkyy myös Hopfenbeckin ym. (2012) huomio kriteeristöjen ristiriitaisista taustateorioista. Päättöarvioinnin kriteerit on luotu Andersonin ja Krathwoh-



lin taksonomian (Anderson ym., 2014) avulla. Tämä tarkoittaa sitä, että kun arvosanat kasvavat, kognitiivisen prosessin, käsiteltävän tiedon tai molempien taso muuttuu vaativammaksi. Ohjauksen määrän kuvaus, joka esiintyy erityisesti alempien arvosanojen kriteereissä, ei ole Andersonin ja Krathwohlin taksonomiaan sisältyvä ulottuvuus eikä se sovi sen kanssa yhteen. Kaikki oppilaat voivat tarvita ohjausta, jos ovat nykyisen osaamisensa rajoilla. Tästä seuraa tulkintavaikeuksia kriteereitä käytettäessä.

Kriteeritasojen kalibrointi on haastavaa, kuten myös Hopfenbeck ym. (2021) havaitsivat. Tässä tutkimuksessa erityisesti alimpien ja ylimpien arvosanojen taso koettiin ongelmallisiksi. Ongelmat arvosanan 5 kohdalla juontanevat ainakin osin juurensa siitä, että kriteeristöä ei ensisijaisesti luotu hyväksytyin ja hylätyn rajan määrittämiseen, vaan ensin luotiin arvosanan 8 kuvaus ja sen jälkeen muut. Jos kriteerien laatiminen olisi aloitettu alimman hyväksytyin määrittelystä tekemällä, arvosanan 5 kriteerit saattaisivat näyttää hyvinkin erilaiselta. Jatkossa voisikin harkita oppiainneiden työryhmiin asiantuntijoita, joka keskittyisivät laatimaan Goslingin ja Moonin (2020) mainitsemia hyväksytyin suorituksen (arvosana 5) kriteereitä. Tällaisessa työskentelyssä voisi hyödyntää erityispedagogiikan asiantuntijoita. Heistä saattaisi olla apua myös matalan osaamistason määrittelyssä muuten kuin ohjaamisen määrän kautta.

Opettajien näkemyksiä pitää voida arvioida myös kriittisesti, ja pohtia, ovatko koetut ongelmat seurausta huonoista kriteereistä vai esimerkiksi puutteellisesta perehtymisestä. Tässä kohden on syytä huomioida, että tutkimukseen osallistuneet opettajat olivat valikoitunut otos. Koska haastatteluun ilmoittautuminen oli vapaaehtoista, arviointiin hyvin perehtyneitä opettajia oli osallistujajoukossa runsaasti. Kriittisiä huomioita tuli myös – ellei jopa erityisesti – heiltä. Osa tutkimuksessa esitetystä kritiikistä, esimerkiksi ”ohjattuna” -kuvauksen puuttuminen yhdestä matematiikan tavoitteesta, vaati tarkkaa perehtymistä kriteereihin. Täydennyskoulutusta tulee tarjota, mutta se voi tuskin yksinään tarjota ainakaan opetussuunnitelmasyklin aikajänteellä kaikille suomalaisille opettajille laajempaa ymmärrystä kuin asiantuntijaopettajilla oli. Siksi pidän kriteerien kehittämistä välttämättömänä.

Kriteerien kehitystyön rinnalla on syytä pitää yllä kriittistä pohdintaa siitä, kuinka tärkeä tavoite päättöarvosanojen vertailtavuus on ja kuinka suuri kompromisseja tavoitteen vuoksi ollaan valmiita tekemään. Opettajien näkemys tavoitteiden sopimisesta opetukseen hyvin mutta arviointiin kehnommin on esimerkki ongelmasta, jossa prioriteetteja pitää punnita tarkoin. On syytä pohtia, milloin kriteerien tulkinnanvaraisuus on väistämätöntä ja milloin sitä pystytään ilman liiallisia kompromisseja vähentämään. Opettajien kanta matematiikan ja fysiikan kriteerien suhteen on tässä selkeä: Tietty määrä sisältöihin sidottuja kriteereitä on hyvä asia. Opettajat kokivat sisältökriteerien olevan hyödyllisiä arvosanojen vertailtavuuden parantamiseksi. Samalla he kokivat arviointityönsä kevenevän, kun kaikkia kriteereitä ei tarvinnut kuljettaa mukana koko ajan.

Sisältötavoitteiden suosio johtuu todennäköisesti myös siitä, että arviointi on perinteisesti kohdistunut sisältöihin, ja esimerkiksi kustantajien arviointimateriaalit tarjoavat apua lähinnä sisältötavoitteiden arvioin-

tiin (Lyhty, 2020). Asenne- ja taitotavoitteiden arviointia voisikin tukea täydennyskoulutuksen ja arviointimateriaalien avulla. Näiden avulla opettajat voisivat saada ideoita vaikeaksi koettujen tavoitteiden arviointiin, sekä nähdä esimerkkejä, joihin omia arviointejaan verrata. Moni opettaja toivoikin Opetushallitukselta esimerkkejä monimuotoisista arvioitavista tehtävistä ja niiden arvioinneista.

Arvosana-asteikon kalibroinnissa on vaikea kuvitella opettajia parempia asiantuntijoita. Kriteerien tason määrittäminen on tärkeä osa kriteeriperustaisen arvioinnin kehittämistä, sillä arviointikriteeristö ei ole uskottava, jos sen erottelukyky on huono, tai jos sen avulla ei pysty antamaan kaikkia arvosanoja (Pitman & Dudley, 1985). Matematiikan ja fysiikan opettajien mukaan arvosanan viisi kriteerit ovat epätarkasti määritellyt sekä vaatimuksiltaan liian alhaiset, ja kiitettävän kriteerien vaatimukset ovat puolestaan liian korkeat. Nämä näkemykset kannattaa ottaa huomioon kriteerien kehitystyössä. Se että osa opettajista pitää hyvin matalaa hyväksytyin rajaa hyvänä asiana, kertoo nähdäkseni tuen järjestämisen haasteista. Moni opettaja ei nähnyt kiristyvien vaatimusten parantavan oppilaiden osaamista, vaan ajavan opettajan ja oppilaan mahdottomien vaatimusten eteen. Mikäli arvosanan 5 kriteerien vaatimuksia kasvatetaan, kuten suuri joukko vastaajista toivoi, pitää samalla pohtia, miten varmistetaan oppilaille tuki kohonneen vaatimustason saavuttamiseksi.

## Tutkimuksen rajoitusten pohdintaa

Tutkimuksella on omat rajoituksensa. Ensinnäkin sen osallistujajoukko on jossain määrin valikoitunutta. Vapaaehtoisiin haastatteluihin osallistui laaja kirjo erilaisia opettajia, mutta arviointiin perehtyneet opettajat olivat runsaasti edustettuna. Valikoitumisen negatiivinen vaikutus tuloksiin on se, että arviointikriteereihin vähemmän perehtyneiden opettajien kokemia haasteita ei välttämättä saatu kattavasti kuvattua. Valikoitumisen hyvä puoli on, että tulokset antavat ymmärrystä siitä, miten kriteerit toimivat, kun niihin on perehdytty. Kun tutkimuksen tavoitteena oli oppia tavoitteiden ja kriteerien toimivuudesta eikä opettajien osaamisesta, vahvojen arviointiosaajien keskimääräistä suurempi edustus on tarkoituksenmukaista.

Toisena näkökulmana on syytä huomioida, että tutkimukseen osallistuneet opettajat eivät ole yksiääninen joukko. Opettajilla oli aiheesta erilaisia, jopa omien näkemystensä kanssa vastakkaisia näkemyksiä. Kun kuvaan tutkimuksessa opettajien näkemyksiä, en tarkoita, että kaikki opettajat olisivat olleet ongelmista samaa mieltä. Tuloksissa esitettyjen ongelmien takana on kuitenkin monia opettajia. Olen tuloksia kirjoittaessa pohjittanut, teenkö vääryyttä tutkittaville, kun jätän jonkin yksittäisen tärkeän huolen kuvaamatta tai kun kuvaan yleistä tyytymättömyyttä jotain seikkaa kohtaan, joka yksittäisen opettajan mielestä oli oivallinen. Esimerkiksi ”ohjattuna” -termin käyttö matematiikassa herätti laajaa ärtymystä, mutta osa opettajista ei nostanut sen ongelmallisuutta esille. Eräs opettaja oli kehittänyt termille oman tulkintansa ja toteutti arviointia sen mukaan. Tuloksia tulee tulkita niin, että havainnot tavoitteiden ja kriteerien vahvuuk-

sista ja epäkohdista ovat aitoja ja merkittävän joukon, mutta ei välttämättä kaikkien opettajien jakamia.

## Johtopäätökset

Tutkimuksen tuloksena tunnistettiin useita opettajien ongelmalliseksi kokemaa tavoitteiden ja kriteerien ominaisuuksia. Löydöksiä voidaan hyödyttää arvioinnin kansallisessa kehitystyössä. Näkemykset koskevat matematiikkaa ja fysiikkaa, mutta osa niistä on sovellettavissa myös muihin oppiaineisiin sekä mahdollisesti myös muihin oppiasteisiin. Suosittelen Opetushallitukselle opettajien esittämien selkiytysten tekemistä, koska ne parantavat todennäköisesti niin kriteerien käyttökelpoisuutta kuin opettajien kokemusta käyttökelpoisuudesta. Käyttökelpoisuuden kokemus on puolestaan omiaan lisäämään motivaatiota kriteerien käyttöön. Hyvät arviointikriteerit eivät palvele ainoastaan arvosanojen vertailtavuutta, vaan myös oppimista (Panadero & Jonsson, 2013), joten perusteita kriteereihin panostamiseen riittää. Pidän järkevänä, että myös oppimisen näkökulma pidetään kehitystyössä mukana. Myös Perusopetuksen opetussuunnitelman perusteiden (2020a) mukaan oppimisen aikainen ja sen päätteeksi tehty arviointi tukevat toisiaan.

Tutkimus ei ole vielä löytänyt kriteeriperustaiseen arviointiin viisasten kiveä, jonka avulla poistettaisiin arvioijan vaikutus arvosanaan (Bloxham et al., 2016). Arviointikriteerien kehittäminen on tärkein tapa yhtenäistää opettajan tekemää arviointia, mutta toisaalta liian yksityiskohtaiset kriteerit tekevät arvosanan antamisesta työlästä ja ohjaavat oppilaita kapean mallin mukaiseen suoritukseen (Bloxham et al., 2016; Cowie & Harrison, 2016). Arviointia kehitettäessä pitääkin pohtia, mitä kaikkea arvioinnilta halutaan, kuinka suurta päättöarvioinnin yhdenmukaisuutta Suomessa tavoitellaan, ja mitä sen eteen ollaan valmiita tekemään ja uhraamaan. Esimerkiksi aika, jonka opettaja käyttää päättöarviointiin, on pois muualta. Toisaalta tavoite- ja kriteerityössä tehty päätös huomioida oppiaineiden päättöarvioinnissa myös laaja-alaisia taitoja on kunnianhimoinen, ja mahdolliset onnistumiset kiinnostavat myös kansainvälistä yleisöä (Torrance, 2017).

## Lähteet

- Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, M. (2014). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's*. New York: Pearson Education Limited.
- Andrade, H. G. (2005). Teaching With Rubrics: The Good, the Bad, and the Ugly. *College Teaching*, 53(1), 27–31. <https://doi.org/10.3200/CTCH.53.1.27-31>
- Andrade, H. G. (2023). What is next for rubrics? A reflection on where we are and where to go from here. In C. Gonsalves, & J. Pearson (Eds.) *Improving learning through assessment rubrics: Student awareness of what and how they learn*. Hershey: IGI Global. <https://doi.org/10.4018/978-1-6684-6086-3.ch001>
- Barnes, N., Fives, H., & Dacey, C. M. (2014). Teachers' beliefs about assessment. In H. Fives & M. G. Gill (Eds.) *International handbook of research on teachers' beliefs*, (s. 284–300). New York: Routledge.

- Biggs, J. B., & Tang, C. S. (2011). *Teaching for quality learning at university: What the student does* (4th edition). New York: McGraw-Hill, Society for Research into Higher Education & Open University Press.
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>
- Bloxham, S., den-Outer, B., Hudson, J., & Price, M. (2016). Let's stop the pretence of consistent marking: Exploring the multiple limitations of assessment criteria. *Assessment & Evaluation in Higher Education*, 41(3), 466–481. <https://doi.org/10.1080/02602938.2015.1024607>
- Bonner, S. M. (2016). Teachers' Perceptions About Assessment: Competing Narratives. In G.T.L. Brown & L.R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment* (s. 21–39). New York: Routledge.
- Brookhart, S. M. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education*, 3. <https://www.frontiersin.org/articles/10.3389/feduc.2018.00022>
- Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, 67(3), 343–368. <https://doi.org/10.1080/00131911.2014.929565>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research*, 86(4), 803–848. <https://doi.org/10.3102/0034654316672069>
- Collins, M., & Lundstedt, J. (2024). The effects of more informative grading on student outcomes. *Journal of Economic Behavior & Organization*, 218, 514–549. <https://doi.org/10.1016/j.jebo.2023.12.001>
- Cowie, B., & Harrison, C. (2016). Classroom Processes That Support Effective Assessment. In G.T.L. Brown & L.R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment*, (s. 335–350). New York: Routledge.
- Gipps, C. (1999). Socio-Cultural Aspects of Assessment. *Review of Research in Education*, 24, 355–392. <https://doi.org/10.2307/1167274>
- Gosling, D. & Moon, J. (2002). How to use learning outcomes and assessment criteria. Southern England Consortium for Credit Accumulation and Transfer.
- Hopfenbeck, T. N., Throndsen, I., Lie, S., & Dale, E. L. (2012). Assessment with Distinctly Defined Criteria: A Research Study of a National Project. *Policy Futures in Education*, 10(4), 421–433. <https://doi.org/10.2304/pfie.2012.10.4.421>
- Jönsson, A., Balan, A., & Hartell, E. (2021). Analytic or holistic? A study about how to increase the agreement in teachers' grading. *Assessment in Education: Principles, Policy & Practice*, 28(3), 212–227. <https://doi.org/10.1080/0969594X.2021.1884041>
- Lyhty, J. (2020). Kuinka arvioida asetettuja tavoitteita: Alakoulun historian opetussuunnitelma ja opettajanoppaat. *Kasvatus & Aika*, 14(4), <https://doi.org/10.33350/ka.91153>
- Metsämuuronen, J. (2023). Matematiikkaa COVID-19-pandemian varjossa III - Syventäviä analyyseja matematiikan 9. Luokan arvioinnista keväällä 2021. Kansallinen koulutuksen arviointikeskus. [https://www.karvi.fi/sites/default/files/sites/default/files/documents/KARVI\\_3123.pdf](https://www.karvi.fi/sites/default/files/sites/default/files/documents/KARVI_3123.pdf)
- O'Donovan, B., Price, M., & Rust, C. (2001). The Student Experience of Criterion-Referenced Assessment (Through the Introduction of a Common Criteria Assessment Grid). *Innovations in Education and Teaching International*, 38(1), 74–85. <https://doi.org/10.1080/147032901300002873>
- Opetushallitus (2004). *Perusopetuksen opetussuunnitelman perusteet 2004*. [https://www.oph.fi/sites/default/files/documents/perusopetuksen-opetussuunnitelman-perusteet\\_2004.pdf](https://www.oph.fi/sites/default/files/documents/perusopetuksen-opetussuunnitelman-perusteet_2004.pdf)
- Opetushallitus (2014). *Perusopetuksen opetussuunnitelman perusteet 2014*. [https://www.oph.fi/sites/default/files/documents/perusopetuksen\\_opetussuunnitelman\\_perusteet\\_2014.pdf](https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_perusteet_2014.pdf)

- Opetushallitus (2020a). Oppilaan oppimisen ja osaamisen arviointi perusopetuksessa. [https://www.oph.fi/sites/default/files/documents/perusopetuksen-arviointi-luku-10-2-2020\\_1.pdf](https://www.oph.fi/sites/default/files/documents/perusopetuksen-arviointi-luku-10-2-2020_1.pdf)
- Opetushallitus (2020b). Perusopetuksen päättöarvioinnin kriteerit. [https://www.oph.fi/sites/default/files/documents/Perusopetuksen%20p%C3%A4%C3%A4tt%C3%B6arvioinnin%20kriteerit%2031.12.2020\\_1.pdf](https://www.oph.fi/sites/default/files/documents/Perusopetuksen%20p%C3%A4%C3%A4tt%C3%B6arvioinnin%20kriteerit%2031.12.2020_1.pdf)
- Opetushallitus (n.d.). Tietoa perusopetuksen arvioinnin uudistusprosessista. <https://www.oph.fi/fi/koulutus-ja-tutkinnot/tietoa-perusopetuksen-arvioinnin-uudistusprosessista>
- Ouakrim-Soivio, N. (2013). Toimivatko päättöarvioinnin kriteerit? : Oppilaiden saamat arvosanat ja Opetushallituksen oppimistulosten seuranta-arviointi koulujen välisten osaamiseröjen mittareina. <https://helda.helsinki.fi/handle/10138/41026>
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144. <https://doi.org/10.1016/j.edurev.2013.01.002>
- Pitman, J. A., & Dudley, R. P. (1985). *Criteria-Based Assessment: The Queensland experience*. 11th International Conference of the International Association for Educational Assessment, Oxford..
- Pulkkinen, J., Hildén, R., Ketonen, L., Khawaja, A., Kotila, J., Mäkipää, T., Nissinen, K., Puhakka, E., Rantala, J., Rautopuro, J., Sulkunen, S., Söyrinki, S., Tammelin-Laine, T. (2024). Arviointiuudistuksen alkutaipaleella. Perusopetuksen päättöarviointitutkimuksen loppuraportti. Opetus- ja kulttuuriministeriön julkaisuja 2024:6
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472–482. <https://doi.org/10.1016/j.tate.2010.09.017>
- Schreier, M. (2012). *Qualitative Content Analysis in Practice* (1st edition). London: SAGE Publications Ltd.
- Silseth, K., & Gilje, Ø. (2019). Multimodal composition and assessment: A sociocultural perspective. *Assessment in Education: Principles, Policy & Practice*, 26(1), 26–42. <https://doi.org/10.1080/0969594X.2017.1297292>
- Torrance, H. (2017). Blaming the victim: Assessment, examinations, and the responsibility of students and teachers in neo-liberal governance. *Discourse: Studies in the Cultural Politics of Education*, 38(1), 83–96. <https://doi.org/10.1080/01596306.2015.1104854>
- Wallace, M. P., & Ng, J. S. W. (2022). Fairness of Classroom Assessment Approach: Perceptions from EFL Students and Teachers. *English Teaching & Learning*. <https://doi.org/10.1007/s42321-022-00127-4>
- Wolf, H. (2004). Assessment criteria: Reflections on current practices. *Assessment & Evaluation in Higher Education*, 29(4), 479–493. <https://doi.org/10.1080/02602930310001689046>