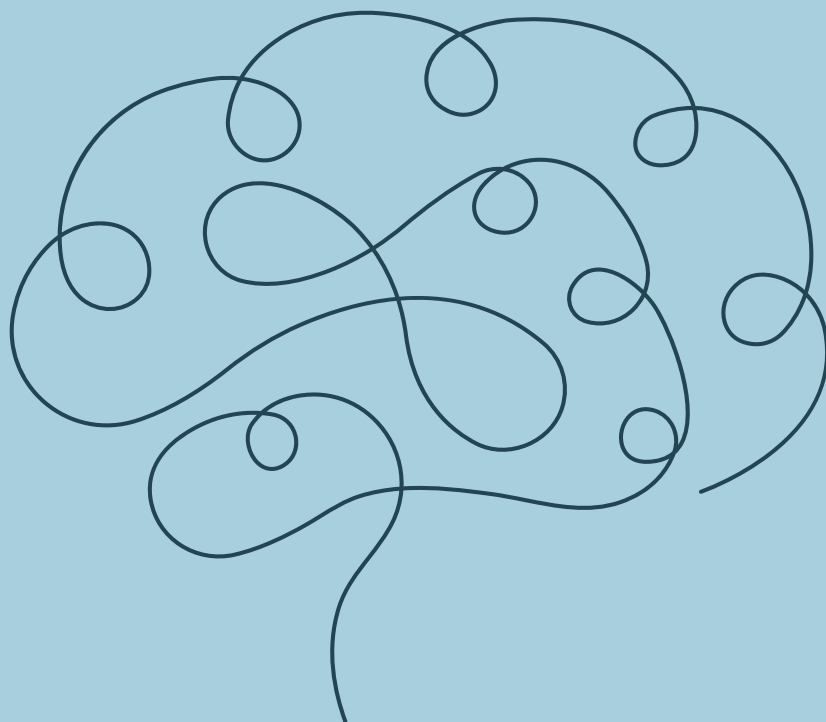


Hilde Synnøve Blix, Kristina Junttila Valkoinen og  
Bengt Arve Haugseth (red.)

# Vurderingsformer og tilbakemelding i kunstfaglig utdanning



# Vurderingsformer og tilbakemelding i kunstfaglig utdanning



Hilde Synnøve Blix, Kristina Junittila Valkoinen og  
Bengt Arve Haugseth (red.)

# **Vurderingsformer og tilbakemelding i kunstfaglig utdanning**

CAPPELEN DAMM FORSKNING

Utgivelsesår: 2025

1. utgave

© 2025 Hilde Synnøve Blix, Anne Eriksen, Tarje Gaustad, Bengt Arve Haugseth, Lillian Grethe Jensen, Esa Lilja, Bodil Kvernenes Nørsett, Maria Medby Tollefsen og Kristina Junttila Valkoinen.

Bokens design og sats: © 2025 Cappelen Damm AS.

Dette verket omfattes av bestemmelsene i *Lov om opphavsretten til åndsverk m.v.* av 1961. Verket utgis Open Access under betingelsene i Creative Commons-lisensen CC-BY 4.0. Denne tillater tredjepart å kopiere, distribuere og spre verket i hvilket som helst medium eller format, og å remixe, endre, og bygge videre på materialet til et hvilket som helst formål, inkludert kommersielle, under betingelse av at korrekt kreditering og en lenke til lisensen er oppgitt, og at man indikerer om endringer er blitt gjort. Tredjepart kan gjøre dette på enhver rimelig måte, men uten at det kan forstås slik at lisensgiver bifaller tredjepart eller tredjeparts bruk av verket. Enhver bruk av hele eller deler av verket som input eller som treningskorpus i generative modeller som kan skape tekst, bilder, film, lyd eller annet innhold og uttrykk, er ikke tillatt uten særskilt avtale med rettighetshaverne.

Lisensvilkår: <http://creativecommons.org/licenses/by/4.0/>

ISBN heftet utgave: 9788202895358

ISBN PDF: 9788202882952

ISBN EPUB: 9788202899950

ISBN HTML: 9788202899974

ISBN XML: 9788202899967

DOI: <https://doi.org/10.23865/cdf.269>



Dette er en fagfelleverdert antologi, med unntak av kapittel 7.

Sitering: Blix, H. S., Valkoinen, K. J. & Haugseth, B. A. (2025). (Red.) *Vurderingsformer og tilbakemelding i kunstfaglig utdanning*. Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.269>.

Cappelen Damm Forskning er redaksjonen for åpen forskningspublisering i Cappelen Damm Akademisk.

Omslagsdesign: Carine Fløystad, Cappelen Damm AS

Omslagsfoto: Getty Images

Cappelen Damm Forskning

[forskning@cappelendamm.no](mailto:forskning@cappelendamm.no)

# Innhold

<b>Forord</b> .....	<b>7</b>
<b>Kapittel 1 Vurderingsformer og tilbakemelding i kunstfaglig utdanning - en introduksjon</b> .....	<b>9</b>
<i>Hilde Synnøve Blix, Kristina Junttila Valkoinen og Bengt Arve Haugseth</i>	
<b>Kapittel 2 Vurdering av gehør- og musikkteorikompetanse i opptaket til høyere musikkutdanning</b> .....	<b>29</b>
<i>Maria Medby Tollefsen, Esa Lilja og Tarje Gaustad</i>	
<b>Kapittel 3 Berekraft i vurderingspraksis ved vokaleksamen i høgare populærmusikkutdanning</b> .....	<b>67</b>
<i>Bodil Kvernenes Nørsett</i>	
<b>Kapittel 4 Elevmedvirkning i praksisveiledning. Et dialogisk perspektiv på tilbakemeldingsprosesser - en pilotstudie</b> .....	<b>101</b>
<i>Hilde Synnøve Blix</i>	
<b>Kapittel 5 «Å få en rund ball til å passe inn i et firkantet hull» - sensorveiledningers rolle i vurdering av eksamenskonsserter</b> .....	<b>129</b>
<i>Lillian Grethe Jensen</i>	
<b>Kapittel 6 Læringsfremmende vurderingsformer i musikk lærerutdanning</b> .....	<b>157</b>
<i>Bengt Arve Haugseth &amp; Hilde Synnøve Blix</i>	
<b>Kapittel 7 Studenters egenvurdering som læring</b> .....	<b>185</b>
<i>Anne Eriksen</i>	



# Forord

Vi i redaksjonen vil med dette takke alle bidragsytere for godt samarbeid. Takk til bokas forfattere for solid arbeid over lang tid. Det er en lang prosess å utføre og framskrive forskning- og utviklingsarbeider, og vi har lært mye sammen i denne prosessen. Vi vil også takke våre særdeles dyktige fagfeller som har bidratt med sine *vurderinger og tilbakemeldinger*. Kvaliteten på artiklene er positivt påvirket av fagfellens innspill, og vi har brukt mye tid på å drøfte selve tilbakemeldingsprosessene i tillegg til å arbeide med å forbedre tekstene. I tillegg er det viktig for oss å få takket forlaget, med Maria Braadland i førersetet, for profesjonell og smidig oppfølging. Tusen takk.

Som leseren av de følgende artiklene helt sikkert vil legge merke til, er det slående lite snakk om såkalt *kunstig intelligens* (KI) i artiklene i antologien. Det er det i hovedsak to grunner til, mener vi. For det første er studienes datainnsamling gjort litt i forkant av KI-utviklingens knytning til plagiat- og juks-problematikk. Noe som gjør at ingen av informantene eller annet datamateriale tar opp dette eksplisitt. Men kanskje aller viktigst er det at nettopp de kunstfaglige vurderingsformene som er omtalt i artiklene, har kvaliteter hvor KI foreløpig ikke er så relevant, verken som verktøy eller mulig snarvei til et godt resultat. Her vil det nok skje endringer i nær framtid; endringer vi vanskelig kan forutsette idet denne antologien blir publisert og et område med stort behov for forskning og utviklingsarbeider. Det er grunn til å tro at kunstfagenes vurderingsformer i denne sammenheng har potensial til å være modeller for endringene som vil måtte komme i andre fagfelt.

Prosjektet er støttet av UiT Norges arktiske universitet.

Hilsen redaksjonen,

Hilde Synnøve Blix, Bengt Arve Haugseth og Kristina Junttila

August 2025



## KAPITTEL 1

# Vurderingsformer og tilbakemelding i kunstfaglig utdanning – en introduksjon

*Hilde Synnøve Blix, Kristina Junttila Valkoinen og  
Bengt Arve Haugseth*

UiT Norges arktiske universitet, Musikkonservatoriet

## Innledning

*Når vi starter å skrive denne introduksjonen til den ellefte antologien i serien MusPed:Research, er det eksamenstid ved UiT Musikkonservatoriet, og vi blir daglig minnet på hvor mange hensyn og behov som skal ivaretas i vurderingssammenheng. Og hvor fryktelig vanskelig, og til tider meningsløst, det kan oppleves å dytte kreative, levende uttrykk inn i forhåndsdefinerte, kjølige kategorier av bokstavkarakterer eller bestått eller ei.*

I denne antologien vil du møte en samling tekster om vurdering og tilbakemelding i kunstneriske og kunstpedagogiske fag. Antologien består av sju tekster som undersøker ulike sider ved vurderingsformer og tilbakemelding i utdanning av musikere, scenekunstnere og kunstfagpedagoger. Antologiens forfattere er ansatte ved utøvende og pedagogiske musikk- og dramafag ved UiT Musikkonservatoriet i Tromsø og Universitetet i Agder/Ansgar høyskole. Forfatterne er også en del av forskningsgruppen *InOvation*, som de siste årene har jobbet med å utvikle og undersøke

Sitering: Blix, H. S., Valkoinen, K. J. & Haugseth, B. A. (2025). Vurderingsformer og tilbakemelding i kunstfaglig utdanning – en introduksjon. I H. S. Blix, K. J. Valkoinen & B. A. Haugseth (Red.), *MusPed:Research: Vol. 11. Vurderingsformer og tilbakemelding i kunstfaglig utdanning* (s. 9–28). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.269.ch1>  
Lisens: CC BY 4.0

vurderings- og tilbakemeldingsformer gjennom egne praksiser og gjennom systematiske undersøkelser og forskning i kunstutdanningsfeltet.

Vi har utforsket og forsket på hva vurderingsformer og -situasjoner innebærer for studentene, lærerne, enkeltfagene, elevene, utdanningene og kunsten selv. Hva vurderes, og hvordan? Hva er læringspotensialet i forskjellige vurderingsformer? Hvordan henger vurderings- og tilbakemeldingsformene i et fag sammen med den øvrige utdanningen og med de profesjonene studentene utdannes til? Antologien har som mål å gi forskningsbasert kunnskap om hva slags landskap vurdering og tilbakemelding i kunstneriske fag kan være i dag, og gir i tillegg et godt utgangspunkt for større bevissthet og kritisk refleksjon om kunst og kvalitet på et mer generelt grunnlag. Antologien vil også ha relevans for andre fagområder hvor utøvende, kreative og praktiske eksamener og vurderingsformer er i bruk. Artikkelen tar opp tematikker som opptaksprøver som vurderingsarena, elevmedvirkning i praksisveiledning, sensorveiledningens rolle i vurderingssituasjoner, egenvurdering og tilbakemelding fra medstudenter, bærekraftig og læringsfremmende vurdering og måter det jobbes med nyskapende vurderingspraksiser i kunstfaglig utdanning på. Antologien adresserer et økende behov for å drøfte og ivareta kunstens særpreg og uforutsigbarhet i møte med et utdanningsfelt hvor målbarhet og klare vurderingskriterier synes å ha forrang, ikke minst i en tid hvor såkalt kunstig intelligens gjør sitt inntok i den akademiske hverdagen.

Vurdering av kvalitet på kunstneriske uttrykk har alltid vekket debatt. Det henger sammen med at det å bli vurdert er en viktig del av et kunstnervirke, samtidig som det er svært krevende å sette ord på hva som gjør kunst god eller dårlig. Kanskje går vi glipp av noe banebrytende nytt fordi vi ikke har opplevd det før og derfor ikke kan forklare dets kvaliteter?

For mange er vurderinger av egne kunstneriske arbeider noe som starter allerede lenge før de søker på en høyere utdanning i et kunstfag, eksempelvis i kulturskolen gjennom konkurranser eller i utvelgelsesprosesser når noen vil starte et band. Mange har også et nært og bevisst forhold til egen kunst uavhengig av nevnte konkurranser og utvelgelsesprosesser, og utvikler selv-vurdering og refleksjon gjennom personlige arbeider og prosjekter. Under utdanningene trenes studenter i å vurdere hva som er bra med egne og andres utøvelse og kunstuttrykk, og blir samtidig selv

vurdert systematisk gjennom utdanningsløpet. Å uttrykke noe kreativt og dypt personlig, og deretter stille seg til disposisjon for tilbakemeldinger, endringsforslag og bedømmelse, kan være skremmende for noen – kanskje til og med slutten på ønsket om å bli utøver. På den andre siden kan det å ta imot og gi tilbakemelding midt i en kunstnerisk prosess være en måte å sikre at verdifull ny innsikt og eksperimentering ikke går tapt ved at man kun «finsliper» uttrykket i *slutten* av en kunstnerisk prosess.

Tatt i betraktning at vi stadig opplever uenigheter mellom fagfolk, må vi regne med at vurderingene også hviler på mer subjektive faktorer som bedømmernes personlige preferanser, i tillegg til kontekst, dagsform og «innpakning». Bedømmere ser og hører etter forskjellige ting, og de overskrider kanskje avtalte vurderingskriterier fordi «det er noe spesielt her», eller fordi «dette vil aldri kunne bli noe mer». Relevante spørsmål blir derfor også: Er det faglig og etisk forsvarlig med kriteriebasert vurdering av kunstneriske uttrykk? Hvilke konsekvenser kan det få for den eksperimenterende og skapende kunsten? Og på hvilke måter kan vurderinger og tilbakemeldinger føre til læring i fag hvor kreativitet og nyskaping er en del av målsettingene?

I det følgende skal vi kort si noe om hva vurdering og tilbakemelding i kunstoffag kan være, og hvilke former dette kan ta i høyere utdanning. Vi ser på hva internasjonal forskning sier om tematikken, drar opp noen sentrale problemstillinger som er særegne for de utøvende kunstoffagene og gir en kort presentasjon av kapitlene i antologien. Til slutt i dette introduksjonskapittelet åpner vi opp for måter vi alle kan øve oss på vurdering og tilbakemelding i samhandling med framtidens kunst og kunstutøvere.

## Summativ og formativ vurdering

I utdanningsammenheng skiller vi i dag mellom formativ og summativ vurdering, hvor summativ vurdering «summerer opp» kunnskaper og ferdigheter en student har tilegnet seg. Summative vurderingsformer har tidligere vært den primære måten å vurdere studenters læring og prestasjoner på i høyere utdanning, gjerne gjennom avsluttende eksamen en gang i året eller semesteret (Lauvås, 2024). I dag er slike oppsummerende prøver og eksamener ofte mindre i omfang og større i antall, blant

annet på grunn av erfaringer med frafall og studenters ønsker om å vite «hvor de står» underveis i et studieforløp. Med summative vurderinger følger ofte et karakteruttrykk. I kunstutdanninger i Norge brukes karakterskalaen A–F og bestått / ikke bestått i slike summative vurderingssammenhenger. Tendensen er at vurdering med bestått / ikke bestått på utøvende eksamener blir stadig mer vanlig i høyere musikkutdanninger.<sup>1</sup> I Universitets- og høyskolerådets *retningslinjer for karakterskalaen bestått/ikke bestått* (UHR, 2015) står det at karakterskalaen A–F i 2015 fremdeles var den mest brukte i summative vurderinger i høyere utdanning i Norge. Bestått / ikke bestått ble brukt på 16,3 prosent av besvarelser, og mest i fag som kunst- og design-, scenekunst- og utøvende musikkutdanninger (UHR, 2015). Argumentene for å bruke graderte karakterer (A–F) på eksamen er vanligvis at det gir mer detaljert informasjon om prestasjoner enn bestått / ikke bestått, og at studentene motiveres til ekstra innsats (Raaheim, 2019). Det vil også kunne legge grunnlag for videre studier eller for hvem som får en jobb. Argumentene *mot* utstrakt bruk av graderte karakterer er blant annet at det skaper stress for enkelte, at det gjør studentene mer opptatt av kriterier for å få gode karakterer enn den indre motivasjonen for utforskning av faget, at det er veldig reduserende og ofte treffer på siden av kjernen i estetisk kompetanse (Olsson, 2014), og at karakterer viser seg å være av mindre interesse for kommende arbeidsgivere (Lauvås & Raaen, 2017).

I utøvende kunstfag er det svært vanlig at graderte karakterer på kunstneriske prestasjoner etterfølges av en samtale med studentene sammen med sensor og faglærer (se Jensen i denne antologien), blant annet med tips til videre arbeid som kunstner/utøver etter endt studium. Dette gjør eksamensformen både summativ og formativ, og har en spesiell status i en utøvers utdanningsløp og videre kunstnerskap (Rønningen, 2018).

Den *formative* vurderingen gjøres av lærere og elever/studenter underveis i studieforløpet, og fokuset er på studentens langsiktige læring og utvikling. Uttrykket «vurdering for læring» blir ofte brukt synonymt

---

1 For eksempel har bachelorstudiet i musikkpedagogikk på Norges musikkhøgskole karakterene bestått / ikke bestått på sine utøvende eksamener, emnene musikkpedagogikk I og musikkdidaktikk I samt musikerhelse, motivasjon og øving (se Norges musikkhøgskole, 2023). UiA har bestått / ikke bestått på alle hovedinstrumenteksamener på bachelornivå (2025).

med formativ vurdering, og denne typen vurdering blir gjerne ansett som en sentral del av såkalt livslang læring (NOU 2019: 12). Boud (2000) bruker begrepet *bærekraftig vurdering* (*sustainable assessment*) om vurdering som har langsiktige mål for studentenes læring og som forbereder dem til framtidige læring, og å selv bli gode til å vurdere egen og andres produkter og prosesser (se Nørsett i denne antologien). Lauvås (2024) argumenterer for mer bruk av det han kaller *ren formativ vurdering*, og skiller mellom formative vurderingsformer som er knyttet til bedømmelser av resultat i tillegg til tilbakemeldinger, og det å få «kun tilbakemeldinger [...] uten antydning av summativ vurdering» (s. 129).

I flere av antologiens artikler drøftes fordeler og ulemper med summative og formative vurderingsformer i kunstpedagogutdanninger, både gjennom opptaksprøver, arbeidskrav, eksamener og fra sensorenes ståsted.

## Tilbakemelding

Økt fokus på studentaktiv læring i academia fører til økt behov for tilbakemeldinger fordi studentene produserer mer og forventer å lære gjennom tilbakemeldinger på produktene/aktivitetene. Tilbakemelding, eller feedback, forbindes med formative vurderingstyper hvor studenten gjennom regelmessige tilbakemeldinger blir mer bevisst på egne læreprosesser. Gamlem (2019, s. 14) definerer tilbakemelding slik: «Informasjon gitt eller søkt etter av en lærer, elev eller en medelev angående kvalitetsaspekt av egen (eller andres) prestasjon (for eksempel framføring, respons og produkt) med det formålet å peke på en retning for videre arbeid.»

Tilbakemeldinger kan komme fra lærer til student, men også fra student til lærer, og studenter imellom. Jevnlige tilbakemeldinger har som mål å involvere studenten i egen læreprosess, og gi mulighet for kritisk refleksjon underveis i utviklingen av kompetanse. Hattie og Timperley (2007) framsetter tre hovedspørsmål som både lærere og studenter bør søke å følge opp gjennom læreprosessen: «Hvor går jeg?», «Hvordan går jeg?» og «Hvor skal jeg videre?». De tre spørsmålene må henge sammen, og det er vesentlig å stadig søke tydelige mål i læreprosessen. Samtidig er det også viktig å lete etter nye og større utfordringer når man har nådd målene (Hattie & Timperley, 2007).

## Vurdering og læringsutbytte

Vurdering inngår som en obligatorisk del av studieprogramstrukturen i høyere utdanning, og det står sentralt å vurdere om studentene har nådd læringsutbyttet som beskrevet i emneplanen (NOKUT, u.å.). Vurderingen foregår vanligvis gjennom skriftlige og/eller muntlige tilbakemeldinger på eksamen og arbeidskrav. Det varierer fra utdanning til utdanning hvor mange arbeidskrav som kreves. I UH-sektoren diskuteres det stadig om det vil være fruktbart å redusere antall arbeidskrav til fordel for mer bruk av hyppige tilbakemeldinger og medstudentvurdering (Haugan & Lysebo, 2018; Lauvås, 2024). Arbeidskrav kan være både disiplinerende og gi godt læringsutbytte, men er også forbundet med noe lite motiverende for studenter og dissonans mellom lærerens forventninger og studentenes innsats (Kristiansen, 2020). Eksamen som vurdering kan ta mange former og er i større grad enn arbeidskrav oppsummerende og skal i hovedsak måle i hvilken grad målene for faget er oppnådd. Eksamensresultater måler også til en viss grad utdanningsinstitusjonenes faglige nivå.

I kunstfagene kan det være utfordrende å ha tydelige læringsutbyttebeskrivelser som igjen kan ha konsekvenser for vurderingskriterier og eksamen. Typiske formuleringer kan være: Kandidaten «har instrumentale ferdigheter til å realisere musikalske intensjoner» og «viser kunstnerisk uttrykksevne og kreativitet» (UiT Norges arktiske universitet, 2024). Hvilke eller hvem sine musikalske intensjoner det kan være snakk om, og hva som kan være kunstnerisk kreativitet, kan vise seg problematisk å tolke både for studenter, lærere og sensorer. Samtidig kan det være vanskelig å forestille seg hvordan man skulle ha formulert helt klare og utvetydige læringsutbytter og vurderingskriterier i denne typen kreative fag. I en rapport om kriterier og skjønn i evaluering i musikk uttrykker Gynnild (2010) utfordringene slik:

Læringsmålene er i mange tilfeller enten fraværende, utydelige, utilstrekkelige eller simpelthen lite egnet for eksaminasjon. Vurderingen forutsetter ofte et tilsvarende diffust skjønn som neppe kan forklares eller forstås fullt ut av studentene, og som i enkelte tilfeller bidrar til mange ubegrunnede klager på sensurvedtak. (s. 15)

Vurdering i kunstfag viser seg altså å være et komplekst problemområde. Dette illustreres blant annet av Waagen (2021), som i likhet med Vinge (2014) skisserer to hovedretninger i vurdering av musikalske prestasjoner. På den ene siden kan man vurdere (eksamens)prestasjonen holistisk, ved at man på bakgrunn av sin kompetanse i faget avgjør hvilket nivå studenten ligger på, basert på en helhetlig vurdering av prestasjonen og produktet. Motstykket til dette er en kriteriebasert, analytisk vurdering, hvor en rekke kriterier er listet opp på forhånd og som gjør at sensor må starte med å «besvare» disse før en endelig karakter kan settes (Waagen, 2021, s. 48). Rønningen (2018) viser til hvordan man i en holistisk vurdering av en praktisk eksamen i drama- og teaterfaget vil vurdere eksamensforestillingen ut ifra et helhetsinntrykk, mens man i en analytisk vurdering vil dele opp prestasjonen i ulike kategorier og vurdere disse enkeltvis (Rønningen, 2018). Studentene i Rønningens studie uttrykte at de foretrakk en blanding av holistisk og analytisk vurdering – altså tilbakemeldinger på både helhet og sentrale detaljer. Det kan være fordeler og ulemper med begge tilnærminger, og Waagen (2021) sier at i holistiske praksiser bør sensorer og lærere øve seg på å ha et tilbakemeldingsspråk, slik at studentene kan forstå hvordan de kan utvikle seg videre. På den andre siden kan en analytisk praksis som fokuserer ensidig på vurderingskriterier, miste nyanser og kvaliteter som er vanskelig/umulig å verbalisere. Bengt Olsson (2014) spør: «Hur fångar man språkligt centrala icke-verbala aspekter av musikaliska framföranden, kompositioner, dans, bilder etc, dvs hur kan skilda representationsformer transformeras till språkliga yttranden?» (s. 137).

## Læringsfremmende vurdering og studentaktive læringsformer

I Meld. St. 16 (2020–2021) *Utdanning for omstilling. Økt arbeidslivsrelevans i høyere utdanning* argumenteres det for en større vektlegging av studentaktive læringsformer. Det skal hjelpe studentene med å forstå og kunne mestre overgangen til en framtidig yrkespraksis, noe som krever at arbeidsformene bør ligne på det studentene vil møte i arbeidslivet (Meld. St. 16 (2020–2021)). Dette skal gjennomføres ved å tone ned

enveiskommunikasjonen som ofte oppstår på tradisjonelle forelesninger og i stedet vektlegge andre arbeidsformer: «For eksempel kan praksis, simulering, caseoppgaver, diskusjoner og studentdeltakelse i forskning og utviklingsarbeid være egnet for å gi innsikt i arbeidsformer som brukes i arbeidslivet» (s. 47). *Læringsfremmende vurdering* har med samme argumentasjon hatt et økende fokus i høyere utdanning. Læringsfremmende vurdering og studentaktive læringsformer henger sammen fordi begge har et spesielt søkelys på studentens utvikling. Gjennom læringsfremmende vurdering, eller *vurdering for læring* (Dysthe, 2008; Wiliam, 2011), skal studenten bli mer bevisst på egen læring gjennom selvvrdering/refleksjon, tydelige læringsmål og kontinuerlige tilbakemeldinger på varierte arbeidsformer (Hattie, 2013; Wiliam, 2011). Du kan lese mer om læringsfremmende vurdering i Blix og Haugseths artikkel i denne antologien.

## Tidligere forskning og litteratur om vurdering

Internasjonal forskning på vurdering i høyere utdanning viser at tilbakemelding er sentralt for læring og utvikling (Hattie & Timperly, 2007; Prøitz et al., 2015), og at vurderingsformer som blir integrert i undervisning som en del av dialogen mellom elev og lærer, har bedre læringseffekt (Hattie & Yates, 2014; Karlsen & Johansen, 2019; Wiliam, 2011): «There is now a strong body of theoretical and empirical work that suggests that integrating assessment with instruction may well have unprecedented power to increase student engagement and to improve learning outcomes» (Wiliam, 2011, s. 13).

I sin omfattende gjennomgang av forskning på effekter av formativ vurdering, viser Wiliam (2011) til internasjonal enighet om at engasjement i egne læreprosesser, og tilbakemeldingsformer som har til hensikt å hjelpe den lærende å se veien videre, gir best resultater. Studenten bør gis informasjon om riktig resultat/svar, noen forklaringer på hva som mangler, og deretter spesifikke instruksjoner om hva som kan gjøres for å forbedre seg. Forskning viser også at på tross av innføring av prinsipper om *vurdering for læring*, tolkes prinsippene relativt smalt, noe som fører til mindre endringer av vurderingspraksiser enn forventet (Torrance, 2007).

Sentralt i nyere forskning om læringsfremmende vurderingsformer er at de reelle praksisene studentene skal ut og virke i, endres i stadig økende tempo. Det betyr at utdanningene må forberede studentene på at selve kunnskapsgrunnlaget kan endre seg raskt, og at den viktigste kompetansen består i å kunne sette seg i stand til å stadig lære noe nytt, og å være kritisk reflekterende og kreativ, altså «lære for en ukjent fremtid» (Raaheim, 2019, s. 84). Dette medfører behov for en større debatt om hva dette har å si for valg av vurderingsformer i utdanningene.

I Norge har man forsket på tilbakemelding i høyere utdanning i for eksempel lærerstudenters grammatikklæring i engelsk (Cabot, 2020). Her fremheves verdien av muntlig tilbakemelding (s. 122), samt at hyppig tilbakemelding gir lærerstudenter en «dobbel gevinst» i form av både økt læring om fag og et viktig verktøy i framtidig læreryrke (Cabot, 2020, s. 126). Karlsen (2016) har gjennom sitt ph.d.-arbeid utviklet en modell for tilbakemelding for studenter i lærerutdanningen. Modellen viser fire aspekter i tilbakemeldingsprosessen, hvor hvert aspekt har sine egenskaper: *form* (to av egenskapene er korreksjon og forklaring), *fokus* (to av egenskapene er produkt og prosess), *formål* (kognisjon, affeksjon) og *tidsretning* (bakovermelding, framovermelding) (Karlsen, K., 2016, s. 76–77).

Forskningsstudier som undersøker vurdering og tilbakemelding i kunstfeltet, viser i hovedsak de samme tendensene som vurderingsforskning generelt (Elliot et al., 2019). På kunstfeltet i høyere utdanning er kritiske, selvstendige og kreative studenter vanligvis et særlig mål for læringsprosesser, og vurderings- og tilbakemeldingsformer er således viktige å studere og diskutere i denne konteksten. Murphy og Espeland (2007) beskriver det de regner som de viktigste dilemmaene når det gjelder vurdering og evaluering i kunstutdanningene. Først og fremst poengterer de at vurdering i kunstfagene er preget av uklare, lite eksplisitte kriterier, og i mange tilfeller heller egenutarbeidede eller innforståtte standarder i feltet (se også Blix et al., 2019; Oltedal, 2018; Vinge, 2014). Selv om de ikke primært ser på undervisning i høyere utdanning, har dilemmaene Murphy og Espeland (2007) stiller opp generell relevans i vår sammenheng:

Often, arts educators operate without explicit criteria while in other cases they create their own criteria or they rely on a common ethos or standard accepted in the specific field, basing their judgment on what they see, and on what they know about how the artwork came to be. (Murphy & Espeland, 2007, s. 339)

*Kriterier* for vurdering er et tilbakevendende tema i forskning om vurderingspraksiser. Forskere som har undersøkt bruk av kriterier og vurderingsskjemaer i forskjellige fag, påpeker at selv om forhåndsbestemte kriterier er nyttige, er det viktig med individuelle tilpassinger (Sadler, 2010; Torrance & Pryor, 2001). Språkforskeren Sadler (2010) opererer med begrepet *latente kriterier*, som dreier seg om kvaliteter ved en elevtekst som først og fremst blir tydelige når den ikke er spesielt god, men som ikke kommer tydelig fram når teksten fungerer godt. Det er grunn til å tro at dette i stor grad også gjelder kunstneriske framførelser/produkter (Gynnild, 2010).

I 2019 samlet Elliot, Silverman og McPherson nyere internasjonal musikkpedagogisk forskning i utgivelsen *The Oxford Handbook of Philosophical and Qualitative Assessment in Music Education*. Til sammen utgjør artiklene en god innsikt i musikkfeltets forskning på vurdering. Redaktørene understreker behovet for kvalitative studier i feltet som har som mål om å forstå «the fundamental conceptual nature of and assumptions about assessment and evaluation across global context» (Elliot et al., 2019, s. 15). Et hovedbudskap som oppsummerer forskningen på feltet, er at det er viktig å stadig utfordre vurdering som fenomen i kunsthøgskolen (s. 4). Studiene viser blant annet at tilbakemeldinger bør støtte opp om studentenes arbeid med personlige, kreative uttrykk (Hill, 2019, s. 293). Bell har undersøkt egenvurdering og hverandrevurdering i faget musikkteknologi og konkluderer med at dette, i kombinasjon med en blanding av formative og summative vurderingsformer, er ideelt for kunstneriske fag, blant annet på grunn av fagenes situerte natur (Bell, 2019, s. 315). Abrahams (2019, s. 257) viser i sin artikkel til at bare når vurderingsformene fører til musikalsk agentskap og transformativ læring, kan de være et verdifullt bidrag til læring i musikkfaget. Kertz-Welzel konkluderer i siste del av håndboken med at utfordringene med vurdering i kunsthøgskolen er mange: «Music's nature as art or social action, the subjectivity of musical experiences, the diversity of musical cultures and competencies, and issues of talent» (Kertz-Welzel, 2019, s. 517).

I undersøkelsen av vurdering i fagene kunst og håndverk og musikk i ungdomstrinnet, dokumenteres det i rapporten *Bedre vurdering i kunstfagene* at presis fagterminologi er et savn i vurdering av fagene fordi det handler om kunstneriske prosesser og personlige uttrykk, noe som er utfordrende i vurderingssammenheng (Limstrand & Abrahamsen, 2009). Rapporten konkluderer med å oppfordre til forskning som ser nærmere på hvordan kvaliteten på personlig uttrykk og kreative prosesser bør vurderes i skolen, og hvordan vurderingskompetanse undervises i kunstfagene i lærerutdanningene. Ifølge Shuler et al. (2016) er problemet med vurdering i kunstfagene at autentisitet og «capacity to create new artwork» (s. 187) blir sidestilt med dokumenterbare, «gyldige» vurderingskriterier. De viser til et behov for at kunstfagfeltet øker kompetansen på vurdering, både gjennom faglig dialog og refleksjon, og gjennom forskning (Shuler et al., s. 2016). Artiklene i den foreliggende antologien har som ambisjon å være bidrag til nettopp en slik dialog.

## Om de enkelte artiklene

Den første artikkelen i denne antologien er en studie av opptaksprøvene i gehør og teori som en søker må bestå for å få en studieplass i høyere utøvende musikkutdanning i Norge. Studien er skrevet av tre forfattere, Maria Medby Tollefsen, Esa Lilja og Tarje Gaustad, hvorav to av dem er gehør- og musikkteorilærere på to forskjellige utøvende musikkutdanninger. Artikkelen gir en analyse av en slik opptaksprøve som vurderingsform og er basert på en kvalitativ undersøkelse, og sammenligning, av resultatene på to forskjellige prøveformater; en tidligere skriftlig prøve, og en ny, digital prøve innført nasjonalt i 2017. Forskerne har også analysert opptaksprøvenes design og oppgavetyper i lys av teori om kategorisering av læringsmål som rammeverk. I artikkelen argumenteres det for nødvendigheten av at prøvene *vurderer* startkompetansen til studentene på en relevant måte, både av institusjonelle og individuelle årsaker.

Den neste artikkelen er skrevet av Bodil Kvernenes Nørsett og handler om *bærekraftig vurderingspraksis*. Nørsett har undersøkt vurderingspraksiser i vokaleksamen i høyere populærmusikkutdanning, og sett på hvordan studentmedvirkning og egenvurdering reflekteres i

studie- og emneplaner i utdanningen og hos vurderingsaktørene (studenter og vurderere). Gjennom dokumentanalyser og intervjuer med vokallærere og -studenter viser hun hvordan målsettinger om livslang læring og studenters egenvurderingskompetanse er vanskelige å formulere i læreplanverk, og at tydelige kriterier og tilbakemeldingsformater savnes. Nørsett konkluderer med at bærekraftige vurderingspraksiser i forbindelse med vokaleksamen bør være mer studentsentrerte gjennom dialogiske tilbakemeldinger og gjennom egenvurderinger som en del av eksamensprosessen.

Artikkelen *Elevmedvirkning i praksisveiledning*, skrevet av Hilde Synnøve Blix, baserer seg på en pilotstudie som undersøker hva som skjer hvis praksiselever inkluderes i tilbakemeldingsprosesser i praksisveiledning. Eleven i studien fikk være med på å gi tilbakemeldinger til instrumentallærerstudentens undervisning. Spilletimen ble filmet, og deretter så elev, veileder og student innspillingen sammen og diskuterte og ga tilbakemeldinger til studenten. I artikkelen diskuteres dialogene som oppstod med utgangspunkt i teorier om dialogisitet, resonans og demokrati. Studien viser hvordan tilbakemeldingsprosessen endrer seg gjennom å tilføre andre aktører og prosesser til veiledningskonteksten.

Lillian Grethe Jensens artikkel *Å få en rund ball til å passe inn i et firkantet hull* ser på hvilken rolle sensorveiledningen spiller i vurdering av eksamenskonserten, og i hvilken grad sensor anvender skjønn og ekspertkompetanse i sensureringen. Studien baserer seg på fem kvalitative forskningsintervju med fagpersoner som har hatt sensoroppdrag i høyere musikkutdanning. I artikkelen viser Jensen eksempelvis at det at man blir berørt og begeistret – at konserten treffer – spiller en større rolle for vurderingen enn de kriteriene som er oppgitt i sensorveiledningen.

Den neste artikkelen har *læringsfremmende vurderingsformer* som omdreiningspunkt. Bengt Arve Haugseth og Hilde Synnøve Blix gjennomførte i 2023 gruppeintervjuer med seks nyutdannede musikkklærere fra to utdanningsinstitusjoner om deres erfaringer med eksamener i musikkfagene. Studien viste at opplevelsen av *eierskap* til vurderingsprosessene var en av hovedkildene til læring i de forskjellige vurderingsformene studentene beskrev å ha vært gjennom. Det viste seg også at det i musikkklærerutdanningene ser ut til å være uklare grenser mellom det

som oppfattes som undervisnings- og vurderingssituasjoner, og studentene mente at de lærte mer når tilbakemeldinger på prestasjoner var vevd inn i undervisningskontekster med læring som tydelig mål.

Antologien avrundes med en fagartikkel skrevet av dosent i drama og teater, Anne Eriksen. Artikkelen gir oss innblikk i hvordan studenter kan jobbe med egenvurdering som læring. Eriksen tar utgangspunkt i et pedagogisk utviklingsarbeid, der studenter i et kunstfaglig studieemne har brukt metoden «gjort – lært – lurt» i forbindelse med samlingsstunder i musikk og drama. Studentenes vurderingslogger utgjør materialet som diskuteres i relasjon til refleksjonsteori og vurdering som læring. Erfaringene fra prosjektet viser hvordan egenvurdering kan bidra til læring gjennom systematisert refleksjonsarbeid, og viser samtidig at refleksjon krever øvelse, språk om kunstnerisk kvalitet og tilpasset vurderingskompetanse knyttet til kunstneriske fag.

Fagfellevurderingene av artiklene i antologien representerer også en type vurderingsform. Vi i redaksjonen ønsket å også ha søkelyset på denne typen tilbakemeldinger og prosessene rundt vurderingene og tilbakemeldingene. Spesielt diskuterte vi den litt «hemmelige» prosessen med *blind review*, og på bakgrunn av erfaringer og drøftingene våre, valgte vi å tilby fagfellene å kunne ta kontakt med redaksjonen og diskutere tilbakemeldingene de ga til de enkelte artiklene. Det var fire av fagfellene som valgte å gjøre dette, og de uttrykte at det gjorde jobben med å vurdere tryggere og mindre «ensom», samtidig som de i stor grad fikk bekreftet sine vurderinger, og i noen tilfeller formulert disse enda tydeligere. Dette gjorde også at jobben redaksjonen hadde, som et slags mellomledd mellom forfatterne og fagfellene, var basert på mer informasjon og bedret kommunikasjonen mellom leddene i vurderingsprosessene.

## Om å øve seg på vurdering

Om vi skal øke vår kompetanse på vurdering og finne nye måter å vurdere hverandre, må vi øve oss på det. Gjennomgående for artiklene i denne antologien er at de går i dybden på konkrete eksempler fra vurderingssituasjoner i høyere kunstpedagogisk utdanning. Om vi løfter blikket fra de konkrete situasjonene og spør hvorfor vi skal vurderes, så er det et spørsmål

om hva slags kunstutøvere og kunstpedagoger vi ønsker oss i framtiden. Dette er det ikke noe entydig svar på. Vi ønsker et mangfold av kunstutøvere og -pedagoger, som ikke utøver ut ifra en enkel norm eller et verdenssyn. Dette mangfoldet får vi ikke ved å følge en enkelt oppskrift på hvordan vi vurderer og hvordan vi opplever å bli vurdert. Ved å rette blikket mot at vi utdanner *kunstutøvere* og -pedagoger må fokuset være på *kunst*, og ikke på instrumentalisering som forsøker å oppnå et bestemt resultat. Når selve resultatet ikke nødvendigvis er det som skal vurderes og det ikke er én autoritet (les: læreren) som har entydige svar når det gjelder hva som fungerer, så handler vurdering og tilbakemelding mye om å øve oss på, og skape verktøy for, å jobbe i prosess og å være i det ukjente (Valkoinen, 2023). Denne kunnskapen er et av de viktigste verktøyene vi kan gi en framtidig utøver som skal virke som kunstner eller pedagog i et stadig skiftende felt.

*Kritikken* av en stadig økende satsing på vurdering i både skole og høyere utdanning, har vært at utdanningen kan bli ensidig og reduksjonistisk og at det ikke tas høyde for hvor kompleks læring er og hvor mange variabler som spiller inn i en lærings- og undervisningssituasjon (Andersen, 2010; Waagen, 2021; Æsøy, 2022).

Vurdering kan absolutt misbrukes. I en søken etter håndgripelige, fattbare og målbare kriterier for kunstfaglige prestasjoner, kan det oppstå et slags vurderingsregimets tyranni, hvor både lærere og studenter kontinuerlig er på jakt etter å måle læring og kvalitet. Kan det av og til være like greit å være fri til å lære uten å måtte vurdere seg selv eller bli vurdert av andre hele tiden? I hvilke situasjoner vil det være problematisk, eller sågar bli helt feil å vurdere? Våre erfaringer som pedagoger er at hyppige vurderinger *kan* bli bremseklosser for skaping og kreativitet fordi man hele tiden stopper opp og vurderer seg selv og andre istedenfor å være i «flyt» (Csikszentmihalyi, 1990). I tillegg kan vi i vår søken i å gi alle studenter individuell oppfølging og fortløpende vurdering, miste det viktige *felleskapet* i kunstfagene. Vinge (2014) synliggjør dette i sin doktoravhandling om vurdering i musikkfaget i ungdomsskolen: Stadig økende oppmerksomhet på individuelle tilbakemeldinger, vurderingskriterier og retten til å klage på karakterer, gjør at man beveger seg bort fra opplevelser og produksjoner som dyrker og styrker fellesskapet (tverrfaglige arbeider, forestillinger og konserter) (Vinge, 2014, s. 349).

Det er mye som taler for at kunstfagene trenger en annen type vurdering enn de generiske metodene man opplever i noen av de mer tradisjonelle universitets- og høyskolefag. Gjennom arbeid med kunstfag bør man la seg henføre, leke, skape, imitere, øve og repetere, og dermed er det behov for å etablere tilbakemeldinger, vurderingskriterier, læringsutbyttebeskrivelser og karaktersettinger på disse fagenes premisser. Artiklene i antologien drøfter på forskjellige måter hvordan vurdering kan gjennomføres på måter å som ikke går ut over kreativitet og skapertrangen hos unge kunststuderende. Det kan virke skremmende å måtte få tilbakemelding på noe som er i prosess og som er uferdig. Man kan lett havne i en forsvarssituasjon, heller enn å gå i dialog, og studentene kan lære seg metoder for å «svare riktig» eller svare det læreren vil de skal svare. For å unngå dette bør også måtene vi *organiserer* vurdering og tilbakemelding på reflekteres over. Hvordan kan vi legge til rette for at det å bli vurdert og få tilbakemelding på, er en måte å bli satt faglig pris på og at det er en kontinuerlig dialog? En dialog som setter alle og alt i mellomposisjoner, en dialog mellom lærer og elev, intuisjon og fakta, språk og kunst, institusjonen og hjemmet, fortid, nåtid og framtid? Undervisningsinstitusjoners strenge krav og lange tradisjoner gjør at det kan være utfordrende å endre vurderingspraksiser. Til slutt presenterer vi derfor noen forslag til måter vi kan øve oss på vurderingssituasjonen, inspirert av Raaheim (2019). Disse øvelsene vil nok ikke fungere i alle akademiske eller praktiske kontekster, men kanskje det gir perspektiv på hva som er mulig og umulig for å endre praksis i vurdering? At vi først og fremst øver oss på å bli framtidens kunstutøvere og -pedagoger (Valkoinen, 2023)?

- Gjør hele vurderingssituasjonen i blinde.
- Snakk om hva studenten drømte om for fem år siden i forhold til hvor hen er i dag. Snakk om hvor studenten drømmer om å være om fem år.
- Ta med deg et barn og en pensjonist til vurderingssituasjonen og la de være ærlig på hva de synes.
- Ha tilbakemelding hjemme hos studenten.
- Bli enig om å kun snakke om følelser i vurderingssituasjonen: Hva følte læreren når hen mottok, og hva følte studenten når hen utførte?

- Prøv å lage en liste over alt som påvirker vurderingssituasjonen.
- Skap noe sammen mellom utøver og den som skal komme med tilbakemelding.
- Kommuniser din vurdering gjennom et kunstnerisk uttrykk.
- Eksperimenter sammen med studentene med kunstig intelligens og hvordan KI ville vurdert hvordan studentene uttrykker seg musikkalsk og hvordan de formidler følelser.
- Få studenten til å forstå at hen er først og fremst ansvarlig for egen læring.
- Å sammen finne ut hva det er man vil ha tilbakemelding på
- At lærer kommer til studenten sin arena (prøvesal, konsert, prosessvisning, hjem osv.) og observerer og reflekterer sammen der.
- At studenten får ta del i lærer sin profesjonelle praksis og de vurderer sammen hvordan det gikk.

## Forfatterbiografier

**Hilde Synnøve Blix** er professor i hørelære og musikkdidaktikk ved UiT Norges arktiske universitet, Musikkonservatoriet. Hennes forskningsinteresser er blant annet gehørutvikling, musikkliteracy og læremidler, i tillegg til kjønnsforskning og likestillingsforskning. Hun har ledet flere forskningsprosjekter, og har publisert en rekke forskningsartikler og bøker nasjonalt og internasjonalt med tematikk som spenner fra gehørundervisning og notedidaktikk til likestilling og kunstnerisk utviklingsarbeid. Hun er leder for forskningsgruppen InOvation ved UiT og FoU-leder ved Musikkonservatoriet.

**Kristina Junttila Valkoinen** er kunstner, underviser og forsker. I dag jobber hun freelance som kunstner, og da spesielt med scenekunstkompaniet Ferske Scener og det kunstpedagogiske prosjektet *Snakk for deg sjøl*. Hun har undervist i scenekunst på alle nivå, for barn, eldre, profesjonelle og kunststudenter, og har vært førsteamanuensis i scenekunst ved UiT Norges arktiske universitet. Hun er utdannet teaterinstruktør og performancekunstner fra Kunstakademiet i Turku og Teaterhøgskolen i Helsinki, samt at hun har en PhD i kunstpedagogikk fra kunstakademiet

i Tromsø, UiT Norges arktiske universitet. Hun sitter i flere styrer og verv, blant annet i scenekunstutvalget i Kulturrådet, styret i Kvääniteatteri, Danseinformasjonen og Tromsø Kunstforening.

**Bengt Haugseth** er førstelektor i musikk ved UiT Norges arktiske universitet, Musikkonservatoriet og Institutt for lærerutdanning og pedagogikk. Hans forskningsinteresser er gehørutvikling, musikk og bevegelse, tverrfaglig arbeid og vurdering i musikkfaget. Haugseth har vært i styret for nasjonalt nettverk for musikk i lærerutdanningene og medleder i det internasjonale nettverket Teacher Education Network. Han har samarbeidet med flere andre fagfelt, de siste årene blant annet i ReKomp (Regional ordning for kompetanseutvikling i barnehage) og DeKomp (Desentralisert ordning for kompetanseutvikling i grunnskole og videregående skole) med temaet musikk og språk.

## Referanser

- Abrahams, F. (2019). A case for integrative assessment from a Freirian perspective. I D. J. Elliot, M. Silverman & G. McPherson (Red.), *The Oxford handbook of philosophical and qualitative assessment in music education* (s. 257–274). Oxford University Press.
- Andersen, H. L. (2010). «Constructive alignment» og risikoen for en forsimplende universitetspedagogikk. *Dansk universitetspedagogisk tidsskrift*, (9), 30–35.
- Bell, A. P. (2019). Enter the feedback loop: Assessing music technology in music education with personal bests. I D. J. Elliot, M. Silverman & G. McPherson (Red.), *The Oxford handbook of philosophical and qualitative assessment in music education* (s. 299–317). Oxford University Press.
- Blix, H. S., Mittner, L. & Gjærum, R. G. (2019). I kjønne orden? Kvalitet og betydningen av kjønn i vurdering av professor- og dosentkompetanse i kunstfagene. *Nordic Journal of Art & Research*, 8(1). <https://doi.org/10.7577/information.v8i1.2853>
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151–167. <https://doi.org/10.1080/713695728>
- Cabot, M. A. (2020). Grammatikk og feedback i lærerutdanningens skriveopplæring i engelsk. *Viden om literacy*, (27), 122–129.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper and Row.

- Dysthe, O. (2008). Klasseromsvurdering og læring. *Bedre skole*, (4).
- Elliot, D. J., Silverman, M. & McPherson, G. E. (Red.). (2019). *The Oxford handbook of philosophical and qualitative assessment in music education*. Oxford University Press.
- Gamlem, S. M. (2019). *Tilbakemelding for læring og utvikling*. Gyldendal Norsk Forlag.
- Gynnild, V. (2010). *Kriterier og skjønn i evaluering. En kasusstudie i utøvende musikkutdanning* (NOKUT-rapport nr. 2010-4). NOKUT.
- Hattie, J. (2013). *Synlig læring: Et sammendrag av mer enn 800 metaanalyser av skoleprestasjoner*. Cappelen Damm Akademisk.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 88–118.
- Hattie, J. A. C. & Yates, G. C. R. (2014). Using feedback to promote learning. I V. A. Benassi, C. E. Overson & C. M. Hakala (Red.), *Applying science of learning in education: Infusing psychological science into the curriculum* (s. 45–58). Society for the Teaching of Psychology.
- Haugan, J. & Lysebo, M. (2018). Hvorfor antall arbeidskrav bør reduseres. Om formativ vurdering i ingeniørutdanning. *UniPed*, 41(3), 347–360. <https://doi.org/10.18261/issn.1893-8981-2018-03-13>
- Hill, J. (2019). Cultural imperialism and the assessment of creative work. I D. J. Elliot, M. Silverman & G. McPherson (Red.), *The Oxford handbook of philosophical and qualitative assessment in music education* (s. 277–297). Oxford University Press.
- Karlsen, K. H. (2016). *Feedback under lupen. Utvikling og utprøving av en modell til bruk i forskning og utdanning* [Doktorgradsavhandling]. Universitetet i Oslo.
- Karlsen, S. & Johansen, G. (2019). Assessment and the dilemmas of a multi-ideological curriculum: The case of Norway. I D. J. Elliot, M. Silverman & G. McPherson (Red.), *The Oxford handbook of philosophical and qualitative assessment in music education* (s. 447–463). Oxford University Press.
- Kertz-Welzel, A. (2019). International perspectives on assessment in music education. I D. J. Elliot, M. Silverman & G. McPherson (Red.), *The Oxford handbook of philosophical and qualitative assessment in music education* (s. 513–527). Oxford University Press.
- Kristiansen, C. W. (2020). *Arbeidskrav i lærerutdanninger. En studie av muntlige arbeidskrav som didaktisk vurderingspraksis* [Doktorgradsavhandling]. Universitetet i Stavanger.
- Lauvås, P. (2024). Utviklingen av vurderingsformer i høyere utdanning. I D. Husebø, L. Ferguson, O. R. Stalheim, I. C. Eriksen, R. Isaksen, A. Mavroudi & P. Wallin (Red.), *Det universitets- og høyskolepedagogiske vitenskapsområdet i Norge – fremvekst, grunnlagstenkning og «state of the art»*. Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.225>

- Lauvås, P. & Raaen, K. (2017). *Passion, cooperation and JavaScript: This is what the industry is looking for in a recently graduated computer programmer* [Konferansepaper]. NIK.
- Limstrand, K. & Abrahamsen, G. (2009). *Bedre vurdering i kunstfagene*. Nasjonalt senter for kunst og kultur i opplæringen.
- Meld. St. 16 (2020–2021). *Utdanning for omstilling. Økt arbeidslivsrelevans i høyere utdanning*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-16-20202021/id2838171/>
- Murphy, R. & Espeland, M. (2007). Prelude: Making connections in assessment and evaluation in arts education. I L. Bresler (Red.), *International handbook of research in arts education* (s. 337–340). Springer.
- NOKUT. (u.å.). *Vurdering og eksamen*. <https://www.nokut.no/publikasjoner-og-kunnskap/kvalitetsoppsummeringer--hoyere-utdanning/vurdering/>
- Norges musikkhøgskole. (2023, 7. august). *Bachelorstudiet i musikkpedagogikk*. <https://student.nmh.no/studiehandboker/startkull-2023/studier/bachelorstudiet-i-musikkpedagogikk>
- NOU 2019: 12. (2019). *Lærekraftig utvikling. Livslang læring for omstilling og konkurranseevne*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/nou-2019-12/id2653116/>
- Olsson, B. (2014). Vad är grunden för musikaliskt lärande och bedömningar? I T. K. Häikiö, M. Lindgren & M. Johansson (Red.), *Texter om konstarter och lärande* (s. 132–161). Göteborgs universitet.
- Oltedal, E. (2018). Challenges of assessing music performance: Teachers' perceptions. I E. Georgii-Hemming, Ø. Varkøy, A. Kallio & F. Pio (Red.), *Nordisk musikkpedagogisk forskning. Årbok 18* (s. 241–270). Norges musikkhøgskole.
- Proitz, T. S., Hovedhaugen, E. & Aamodt, P. O. (2015). Tilbakemeldinger i høyere utdanning: hva sier studenter og lærere? I N. Frølich (Red.), *Hva skjer i universiteter og høyskoler? Perspektiver fra vitenskapelig ansatte og studenter* (s. 37–49). Universitetsforlaget.
- Rønningen, L. H. (2018). Vurdering av praktiske eksamener i drama og teater. I V. Aune & C. Haagensen (Red.), *Teaterproduksjon. Ti produksjonsetetiske innganger* (s. 243–261). Cappelen Damm Akademisk. <https://doi.org/10.23865/noasp.43>
- Raaheim, A. (2019). *Eksamensrevolusjonen: Råd og tips om eksamen og alternative undervisningsformer* (2. utg.). Gyldendal.
- Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5), 535–550.
- Shuler, S. C., Brophy, T. S. & Sabol, F. R. (2016). Arts assessment in an age of accountability: Challenges and opportunities in implementation, design, and measurement. I H. Braun (Red.), *Meeting the challenges to measurement in an era of accountability* (s. 193–226). Routledge.

- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning, *Assessment in Education: Principles, Policy & Practice*, 14(3), 281–294. <https://doi.org/10.1080/09695940701591867>
- Torrance, H. & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, 27(5), 615–631. <http://dx.doi.org/10.1080/01411920120095780>
- UiT Norges arktiske universitet. (2024). *MUS-3301 Hovedinstrument – 10 stp*. [https://uit.no/utdanning/emner/emne?p\\_document\\_id=841673](https://uit.no/utdanning/emner/emne?p_document_id=841673)
- Universitets- og høyskolerådet. (2015). *Retningslinjer for karakterskalaen bestått / ikke bestått*. UHR.
- Valkoinen, K. J. (2023). *What makes us act? On the potentials of exercises in live art education and performances* [Doktorgradsavhandling, UiT Norges arktiske universitet]. UiT Munin. <https://hdl.handle.net/10037/28763>
- Vinge, J. (2014). *Vurdering i musikkfag. En deskriptiv, analytisk studie av musikk læreres vurderingspraksis i ungdomsskolen* [Doktorgradsavhandling]. Norges musikkhøgskole.
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37, 3–14. <https://doi.org/10.1016/j.stueduc.2011.03.001>
- Waagen, W. (2021). Vurderingskriterier i utøvende musikk – et verktøy for pålitelig vurdering av studentenes bachelorkonsert? I E. Angelo, J. Knigge, M. Sæther & W. Waagen (Red.), *Higher education as context for music pedagogy research* (s. 41–64). Cappelen Damm Akademisk. <https://doi.org/10.23865/noasp.119.ch2>
- Æsøy, K. O. (2022). Kritisk blikk på vurdering for læring. *Bedre skole*, (3), 68–71.

## KAPITTEL 2

# Vurdering av gehør- og musikkteorikompetanse i opptaket til høyere musikkutdanning

Maria Medby Tollefsen<sup>1</sup>, Esa Lilja<sup>2</sup> og Tarje Gaustad<sup>3</sup>

<sup>1</sup>UiT Musikkonservatoriet

<sup>2</sup>Universitetet i Stavanger

<sup>3</sup>Handelshøyskolen BI

## Abstract

This study examines the entrance exams in aural skills and music theory for higher music education in Norway, with a particular focus on the national exam introduced in 2017. The study aims to explore how factors such as task types, scoring systems, and the complexity of the knowledge being tested can affect the exam's ability to assess applicants' prior knowledge. The study includes detailed analyses and comparisons of the task types and knowledge levels tested in the current and previous exam sets, as well as testing on first-year students at two educational institutions. The findings suggest that the current entrance exam does not adequately capture the variation in applicants' knowledge levels, which may result in applicants without sufficient prior knowledge being admitted. This can lead to challenges in meeting the demands of higher music education. The authors argue for a revision of the entrance exam to enhance its ability to effectively identify students who possess the necessary competencies to succeed in their studies. At the same time, it is highlighted that a potential revision may result in fewer students being offered admission, with the consequences this may have for educational institutions.

**Keywords:** entrance exam, prior knowledge, aural skills, music theory, assessment, higher music education

Sitering: Tollefsen, M. M., Lilja, E. & Gaustad, T. (2025). Vurdering av gehør- og musikkteorikompetanse i opptaket til høyere musikkutdanning. I H. S. Blix, K. J. Valkoinen & B. A. Haugseth (Red.), *MusPed:Research: Vol. 11. Vurderingsformer og tilbakemelding i kunstfaglig utdanning* (s. 29–65). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.269.ch2>  
Lisens: CC BY 4.0

For de fleste yrkesmusikere i Norge starter musikkutdanningen lenge før de starter i høyere musikkutdanning. Tradisjonelt har opplæringen foregått i etablerte institusjoner som kulturskole, korps, storband og kor, på musikklinjer i videregående skole og i talentprogrammer. I tillegg til å lære å spille et instrument eller å synge, lærer elevene, mer eller mindre systematisk, fagbegreper, noter, lytteteknikker og mer. Når unge musikanter søker seg til musikkutdanning, forventes det at de ikke bare har et høyt nivå på hovedinstrumentet, men at de også har tilegnet seg grunnleggende kompetanse i gehør og musikkteori (GT).

I opptaket til høyere utdanning testes søkeren gjennom prøvespill der hovedinstrumentferdighetene vurderes, samt gjennom en gehør- og teoriprøve (GT-prøve). Et av GT-prøvens viktigste mål er å sikre at søkeren har tilegnet seg et tilstrekkelig fundament av kunnskap og ferdigheter til å kunne mestre musikkutdanningens ulike emner.

Forfatterne av denne artikkelen er gehør- og musikkteoripedagoger ved to norske offentlige studiesteder for utøvende musikk. Vi har inntrykk av at nivåspredningen i førsteårsstudentenes GT-ferdigheter har økt etter innføringen av en ny, nasjonal opptaksprøve i GT i 2017, her omtalt som *2017-prøven* (Lilja & Tollefsen, 2023). Vi møter naturligvis mange studenter som har solid forkunnskap i gehør og musikkteori, men også studenter som, selv om de har bestått GT-prøven, ikke kan fagstoff vi anser som grunnleggende. Disse studentene kan ha en lang og vanskelig vei å gå for å kunne gjennomføre undervisning og eksamen på universitetsnivå i emner som arrangering, komponering, ensembleledelse og gehør.

Høsten 2022 startet vi et tverrinstitusjonelt samarbeid der målet er å skaffe kunnskap om GT-prøven og dens kapasitet til å vurdere søkeres forkunnskap. Gjennom den foreliggende studien belyser vi følgende spørsmål:

*På hvilke måter kan faktorer ved opptaksprøven som oppgavetyper, poenggivings-system og kompleksitet i kunnskapen som testes, påvirke dens kapasitet til vurdering av søkeres forkunnskap?*

## Kunnskapsgrunnlag

Kunnskapsgrunnlaget for studien består av vitenskapelige publikasjoner og dokumentasjon fra arbeidet med å utvikle 2017-prøven. Forskningen det

refereres til i følgende avsnitt, kaster på ulikt vis lys over opptaksprøvene i GT. Den brukes videre til å diskutere slike prøvers formål og funksjon.

I denne studien benyttes publikasjoner av Aliferis og Stecklein (1953), Arenson (1983), Baysal et al. (2018), Bergby (2003, 2013), Ruokonen et al. (2018) og Sloboda (1985/2005) som kunnskapsgrunnlag. Noen av disse publikasjonene dokumenterer prosessen med å utvikle opptaksprøver (Aliferis & Stecklein, 1953; Baysal et al., 2018), og drøfter blant annet fordeler og ulemper med ulike typer vurderingsmåter. Enkelte studier sammenligner forskjellige studiesteders prøver (Ruokonen et al., 2018), mens andre studerer sammenhengen mellom studentenes resultater på opptaksprøvene, på eksamener i emner som gehør og musikkteori og på utøvende eksamen (Arenson, 1983; Bergby, 2003, 2013).

På bakgrunn av publikasjonene kan vi anta at det å inkludere prøver i GT i opptaket til høyere musikkutdanning er en utbredt praksis i den vestlige delen av verden. Det kan også virke som om ulike lands prøver tester lignende kunnskap og ferdigheter, på tross av ulikheter i musikktradisjon, utdanningssystem og prøveform (Aliferis & Stecklein, 1953; Arenson, 1983; Baysal et al., 2018).

De offentlige utøvende musikkutdanningsinstitusjonene i Norge har lang tradisjon for bruk av felles nasjonale opptaksprøver i GT. Et uttalt mål med prøvene, både her til lands og i andre land, er å sikre at nye studenter har nødvendig forkunnskap for å mestre studiets mange emner<sup>1</sup> (Aliferis & Stecklein, 1953; Baysal et al., 2018; Bergby, 2016). Bergby (2016) omtaler denne forkunnskapen som *grunnleggende kompetanse*. Forventningen til grunnleggende kompetanse gjenspeiles i emneplanene til ulike førsteårsemner gjennom læringsmålformuleringer som tar utgangspunkt i, og bygger på, denne kompetansen. Eksempelvis kan det tolkes som at det ligger en implisitt forventning til forkunnskap om notesystemet og annet når det i emneplanen til «Satslære 1» ved bachelorstudiet i utøving ved Norges musikkhøgskole står at studentene skal lære å «analysere og beskrive harmonikk, stemmeføring og form i repertoar for kor, piano og mindre ensembler» (Norges musikkhøgskole, 2023).

---

1 Begrepet *emne* brukes her for å beskrive studiepoenggivende kurs innen ulike fagområder. Begrepet er vanlig innen høyere utdanning i Norge.

Opptaksprøvene tjener også funksjoner ut over å sikre at søkerne har nødvendig forkunnskap. Prøvene gir studiestedets lærere et bilde av hver enkelt søkers kunnskaps- og ferdighetsnivå. Informasjonen kan brukes til kartlegging og inndeling av studentene i nivåhomogene grupper i emner der det er hensiktsmessig, og er et viktig verktøy for å kunne gi differensiert undervisning (Aliferis & Stecklein, 1953; Bergby, 2013). Prøven gir i tillegg signaler til videregående skoler, søkere og deres lærere om hva nye studenter forventes å kunne ved studiestart (Bergby, 2013, 2016).

Opptaksprøvene har røtter tilbake til tester som *Seashores measures of musical talent* (1919) og andre musikalitetstester fra første halvdel av 1900-tallet (Bergby, 2003). Hva de nevnte prøvene egner seg til å måle, er omdiskutert. Både Aliferis og Stecklein (1953), Bergby (2003) og Sloboda (1985/2005) understreker at GT-prøver ikke bør sees på isolert fordi de kan gi ensidige og upålitelige bilder av søkerne, og at de må sees i sammenheng med prøvespill på hovedinstrument og eventuelt andre tester. Bergby finner i en studie fra 2003 at det er liten eller ingen sammenheng mellom en students resultat på opptaksprøven i GT og resultatet på utøvende eksamen ved slutten av studiet, og at GT-prøven ikke gir noen indikasjon om søkerens fremtidige suksess som musiker. Studien «Relationships between entrance tests and exams in music performance and aural skills at the Norwegian Academy of Music» (Bergby, 2013) viser at det er en sammenheng mellom resultatet på opptaksprøven og gehøreksamen, men at den samme sammenhengen ikke kan sees når det gjelder opptaksprøven, gehøreksamen og nivå på hovedinstrumentutøvingen. Hvorvidt en students muligheter i yrkeslivet kun beror på det utøvende nivået, problematiseres ikke i studiene. Gode ferdigheter i å arrangere, instrumentere, transkribere, transponere, med mer, kan være vel så avgjørende for studentens suksess i arbeidslivet som hovedinstrumentferdighetene alene. Resultater fra emner der slike ferdigheter læres, vil kanskje kunne vise seg å samsvare bedre med resultatene på GT-prøven enn hva resultatene på utøvende eksamen gjør.

Bergby (2013) konkluderer i sin studie med nødvendigheten av en nasjonal debatt om hensikten med GT-opptaksprøven, der det også drøftes hvilke kunnskaper og ferdigheter den skal måle og hvordan dette skal skje. Dokumenter fra arbeidet med å utvikle 2017-prøven viser at en slik

debatt fant sted i det norske gehør- og musikkteoretiske fagfellesskapet i 2014 (Bergby, 2014b). Gjennom diskusjonen ble det uttrykt ulike syn på opptaksprøvens formål og funksjon. Enkelte mente at prøven kan ekskludere gode instrumentalister fra å studere musikk, og at den kun bør ha en kartleggende funksjon. Andre mente at det er viktig med en opptaksprøve som sikrer at studentene har grunnleggende kompetanse i GT for å kunne mestre studiet (Bergby, 2014b), og at prøveresultatene bør vektlegges i opptaket.

Referater, høringsnotater og annen dokumentasjon fra arbeidet med å utvikle en ny nasjonal prøve i årene frem mot 2017 gir innsikt i prosessene som lå til grunn for utarbeidelsen av prøven, og sier noe om hvilke diskusjoner og avveininger som fant sted. Vi gir i det følgende et innblikk i prosessen med å utvikle prøven, samt en redegjørelse for likheter og ulikheter mellom 2017-prøven og det foregående prøvesettet, her omtalt som *de gamle opptaksprøvene*.

## Utarbeidelse av en ny nasjonal opptaksprøve i gehør og musikkteori

Paraplyorganisasjonen *Fagorgan for skapende og utøvende musikkutdanning* (FUM) setter rammer for de norske GT-opptaksprøvene. Eksempelvis legger organet føringer for hvor lang prøven skal være og hvorvidt den skal være mulig å ta flere ganger, og det er også de som avgjør når en ny prøve skal lages. 2017-prøven ble laget på bakgrunn av et mandat fra FUM, som i 2014 oppnevnte en arbeidsgruppe bestående av gehør- og musikkteorilærere fra flere av de offentlige musikkutdanningsinstitusjonene.<sup>2</sup> FUM ønsket en prøve som skulle skille seg fra det foregående prøvesettet på flere områder: Gehør og musikkteorikunnskap skulle bli testet i en og samme prøve, ikke i to separate. Prøven skulle være egnet for søkere med et spekter av sjangersspesialiseringer, den skulle være digital og kunne gjennomføres og rettes automatisk. Prøvens oppgaver skulle ha entydige svar, slik at skjønnsmessige vurderinger ville være unødvendige. En særlig beskrivelse av prøvens faglige nivå ble ikke

<sup>2</sup> Norges musikkhøgskole, UiT, UiB og NTNU.

formidlet gjennom mandatet. Det ble allikevel spesifisert at den må måle at søkerne har den grunnleggende kompetansen som anses som nødvendig ved studiestart, både med hensyn til utøvende emner og basisemner<sup>3</sup> i en musikkutdanning (Råd for utøvende musikkutdanning [RUM], 2014).

Begrepet *grunnleggende kompetanse* benyttes ved flere anledninger i dokumentene fra arbeidet med 2017-prøven (Bergby, 2016; RUM, 2014). Det spesifiseres i liten grad hvilken kunnskap denne kompetansen består av. Fordi opptaksprøven skal benyttes på søkere fra flere ulike sjangerfordypninger, må det kunne antas at det menes kunnskap og ferdigheter som kan sees på som grunnleggende uavhengig av musikalsk sjanger. Debatten i det nasjonale GT-fagmiljøet i 2014 viste at kollegiet har ulike synspunkter rundt hva grunnleggende kompetanse innebærer. Det ble eksempelvis påpekt at musikalsk terminologi oftest er sjangerspesifikk, og at en prøve som forutsetter kjennskap til terminologi, eller som benytter oppgaver typiske for en musikktradisjon, vil være ekskluderende for søkere utenfor den spesifikke musikktradisjonen (Bergby, 2014a).

Enkelte av studiene i kunnskapsgrunlaget definerer hvilken kompetanse opptaksprøver til høyere musikkutdanning bør måle. Evne til auditiv-visuell diskriminering av rytmiske, melodiske og harmoniske elementer, samt ferdigheter i å gjenkjenne intervaller, melodiske og treklangsvise melodiske bevegelser, kadenser, akkorders relative forbindelser og alminnelige rytmefigurer, fremheves både i Aliferis og Steckleins studie fra 1953 og Baysal et als studie fra 2018. Begge studiene fokuserer på opptaksprøver til musikkutdanning innen klassisk musikk, og tar utgangspunkt i et sett med kunnskaper og ferdigheter som denne musikktradisjonen tradisjonelt krever. Det er derfor usikkert hvor dekkende kompetansedefinisjonene er for den norske opptaksprøven, som skal være relevant for søkere fra et større sjangermangfold.

Arbeidsgruppen som utarbeidet 2017-prøven formulerte spesifikke kompetanseområder som ble liggende til grunn for utformingen av prøvens oppgavemateriale. Valgene som ble gjort, er ikke redegjort for i

---

3 Med basisemner menes emner som gehør, musikkteori, arrangering/komponering, musikalsk analyse, musikkhistorie og lignende. Slike emner inngår i studieplanene for utøvende musikk og musikkpedagogikk hos samtlige studiesteder som bruker den nasjonale opptaksprøven, men emnene benevnes gjerne med ulike navn ved forskjellige studiesteder.

dokumentasjonen som er tilgjengelig. Det er nærliggende å tro at debatten i fagmiljøet (Bergby, 2014a), emneplaner fra utdanningsinstitusjonene, tradisjon, samt tilgjengelig forskning rundt opptaksprøver, kan ha hatt betydning for valg av kompetanseområder. For å legge til rette for at prøven dokumenterer og sikrer kunnskap uten å ekskludere søkere, ble det bestemt å lage en demonstrasjonsprøve som alle søkere har tilgang til. Demonstrasjonsprøven gjør det mulig for søkere å forberede seg målrettet og grundig, uavhengig av bakgrunn og sjangertilhørighet (Bergby, 2016). Demoprøven ble senere publisert på YouTube, og er allment tilgjengelig.

En gjennomgang av kunnskapen som testes i det gamle prøvesettet og 2017-prøven, viser at prøvene tester mange av de samme kunnskapsområdene, men at det også er enkelte forskjeller. Forskjellene gjenspeiler en etterstreben etter å gjøre prøven mindre sjangerspesifikk, og viser at synet på hva grunnleggende kompetanse er, er i endring. Tabell 1 viser hvilket fagstoff som testes i de to prøvesettene.

**Tabell 1.** Fagstoff som testes i de to prøvesettene

Fagstoff og hvordan det testes	Gammelt prøvesett	2017-prøve
Auditiv gjenkjenning av dur og moll i musikksekvenser	Ja	Ja
Auditiv gjenkjenning av melodiske sluttintervaller	Ja	Ja
Auditiv gjenkjenning av harmoniske intervaller	Ja	Nei
Bestemme intervaller i notebilde (g-nøkkel og f-nøkkel)	Ja	Ja
Bestemme skalaer og tonearter	Ja: dur, ulike typer moll, modale skalaer og heltoneskala.	Ja, men kun dur og moll.
Bestemme parallelltonearter	Nei	Ja
Bestemme treklang-/akkordtyper i notebilde	Ja	Ja
Bestemme akkordomvendinger auditivt	Ja	Ja
Bestemme treklanger auditivt (dur/moll/forminsket/forstørret)	Ja	Ja
Harmonisk analyse av firstemt sats (notert i system med g-nøkkel og f-nøkkel)	Ja	Ja
Koralsatsharmonisering	Ja	Nei
Lage motstemme (kontrapunkt)	Ja	Nei
Auditiv harmonisk analyse	Nei	Ja

(Forts.)

**Tabell 1.** (Forts.)

Fagstoff og hvordan det testes	Gammelt prøvesett	2017-prøve
Melodidiktat	Ja	Nei
Melodikorreksjon	Ja	Ja
Samsvar lyd og melodi (enkelte melodier notert i g-nøkkel, enkelte i f-nøkkel)	Nei	Ja
Enstemt melodiforandring	Ja	Nei
Tostemt melodiforandring	Ja	Nei
Grunntonebestemmelse	Ja	Nei
Ta stilling til intonasjon av sure toner	Ja	Nei
Transponering og lesing/notasjon i andre nøkler enn g- og f-nøkkel	Ja	Nei
Taktartsbestemmelse	Ja	Ja
Rytmekorreksjon	Ja	Ja
Rytmediktat	Ja	Ja (i ny prøve gjøres rytmediktaten gjennom seleksjon av rytmefigurer).
Bestemme taktart i noteeksempel	Nei	Ja
Samsvar mellom lyd og rytme	Nei	Ja
Plassering av taktstreker i notebilde	Nei	Ja

Mens det gamle prøvesettet ble løst med blyant og papir og ved hjelp av innspilt lyd på CD, er 2017-prøven digital. 2017-prøven ble utviklet i samarbeid med programvareleverandøren Inespera Assessment (Bergby, 2016). Inespera var allerede i 2015 en erfaren programvareleverandør når det gjaldt eksamensløsninger til utdanningsinstitusjoner, og kunne blant annet tilby gode løsninger for sikker nettleser. De hadde ikke erfaring med å lage digitale løsninger for musikkutdanninger, og kunne eksempelvis ikke levere oppgaveformer som krevde at søkerne skrev inn noter selv. Dette førte til at enkelte ønskede oppgaveformer, som diktat og utfyllingsoppgaver, ikke kunne inkluderes i prøven. Det ble i stedet innført andre oppgaveformer som vurderer tilsvarende kunnskap, innenfor rammen av datidens tilgjengelige teknologi.

Tiden søkerne har hatt til disposisjon, er ulik i de to prøvesettene. I det gamle prøvesettet varer gehørprøven og teoriprøven i 1 time hver, mens

2017-prøven varer i 1,5 time. I begge prøvesett gjennomføres prøven i et rom sammen andre søkere. I den gamle gehørprøven ble det avspilt en CD som inneholdt det auditive oppgavematerialet og instruksjoner for oppgaveløsingen, og tempoet i prøven ble styrt av dette. I 2017-prøven bruker søkeren pc og hodetelefoner og lytter til lydfilene i eget tempo. Mens de fleste lydeksempelene i den gamle prøven ble spilt av en eller to ganger, har søkeren mulighet til å spille av hvert eksempel tre ganger i 2017-prøven.

2017-prøven er altså noe forkortet i forhold til det gamle prøvesettet. Fordelingen mellom gehør- og teorioppgaver er dessuten ulik i de to prøvesettene, både med tanke på antall oppgaver og hvordan de vektet med poeng. I det gamle prøvesettet måtte gehør- og teoriprøvene bestås hver for seg, noe som innebar at studenten måtte oppnå 60 prosent av maksimal poengsum på gehøroppgavene og 43 prosent av maksimal poengsum på teorioppgavene. I 2017-prøven må studenten oppnå 60 prosent av maksimal poengsum, men det stilles ingen krav til antall poeng på henholdsvis gehør- og teorioppgaver.

Sammenligningen av 2017-prøven og gammelt prøvesett indikerer at prøvene skiller seg fra hverandre også når det gjelder hvilke oppgavetyper som benyttes og hvor kompleks kunnskap som testes. Vi ønsket å undersøke hvilken betydning disse aspektene har for opptaksprøvens kapasitet til pålitelig vurdering av søkerne, og valgte derfor å benytte en modell for klassifisering av læringsmål og kunnskapsnivåer, samt litteratur omkring ulike prøveformers betydning for pålitelig testing, som teoretisk rammeverk for studien. Rammeverket ga oss et godt utgangspunkt for analyse av opptaksprøvene, samt for å belyse og forklare funn vi har gjort.

## **Teoretisk utgangspunkt**

### **Klassifisering av læringsmål og kunnskapsnivåer**

Det fins ulike modeller for klassifisering av læringsmål, deriblant Blooms taksonomi (Bloom et al., 1956; Simkin & Kuechler, 2004). Taksonomien ble først beskrevet i boken *The Taxonomy of Educational Objectives, The Classification of Educational Goals* (Bloom et al., 1956), og har hatt stor betydning for det internasjonale utdanningsfeltet (Anderson & Krathwohl, 2001). Taksonomien er i denne sammenheng valgt fordi den

systematiske og hierarkiske måten den beskriver kunnskap på, egner seg godt som rammeverk for analysene av opptaksprøvene. Flere andre studier som tar for seg problematikk rundt testing i undervisnings- og utdanningssammenheng, bygger på samme taksonomi (eksempelvis Kastberg, 2003; Simkin & Kuechler, 2004). I disse studiene har den vist seg som et nyttig rammeverk for å belyse og forstå sammenhengen mellom læringsmål, kunnskapsnivåer og oppgaveformer.

Blooms taksonomi beskriver seks nivåer av læringsmål, hvor kompleksiteten i den tilegnede kunnskapen øker for hvert nivå. Læringsmålene i det første nivået handler om å kunne gjenkjenne og gjengi kunnskap. På det øverste nivået er målet at kunnskapen skal kunne brukes til å konstruere noe nytt (Anderson & Krathwohl, 2001). Når det skal vurderes hvorvidt en elev eller student har oppnådd et bestemt læringsmål, er det studentens *kunnskapsnivå* som vurderes. Det er et mål at kunnskapsnivået skal være i tråd med beskrivelsen av læringsmålet når vurderingen finner sted. På den måten kan Blooms taksonomi også sies å beskrive ulike kunnskapsnivå. Fordi den i denne studien handler om vurdering av kunnskap, vil klassifiseringen her først og fremst brukes for å beskrive kunnskapsnivå. I den videre teksten brukes derfor begrepet kunnskapsnivå istedenfor læringsmål i beskrivelsen av taksonomiens kategorier.

I det følgende redegjøres det for taksonomiens nivåer satt i en gehør- og musikkteori-kontekst. Klassifiseringen henter inspirasjon fra Anderson og Krathwohls (2001) beskrivelser av, og eksempler på, de ulike nivåene i taksonomien. Eksemplene er hentet fra grunnskoleundervisning i fag som naturfag og historie, og lar seg ikke problemfritt overføre til gehør og musikkteori. Mye av kunnskapen som ansees som elementær i gehør- og musikkteorifeltet forutsetter kognitive prosesser på et annet nivå enn kunnskapen som kreves for å huske årstallet for en historisk hendelse, eller annen faktakunnskap. Allikevel er det ulike grader av kompleksitet i læringsmålene også i gehør- og musikkteorifagene. En hierarkisk klassifisering av disse muliggjør analyser av opptaksprøvene som kan gi oss innsikt i hvor avansert kunnskap som måles. Klassifiseringen av læringsmål i gehør- og musikkteorifeltet innebærer en viss grad av subjektiv vurdering. Oss bekjent fins det ikke tilsvarende klassifiseringer i fagfeltene som vi kan støtte oss til. Vurderingene har derfor skjedd gjennom systematisk diskusjon i forskningsgruppen.

**Nivå 1: Huske.** Det første nivået i taksonomien beskriver to delnivåer av elementær kunnskap: det første (nivå 1a) handler om å kunne gjenkjenne og identifisere et læringsinnhold. På neste delnivå (nivå 1b) kan kunnskapen gjenkalles fra langtidsminnet og gjengis (Anderson & Krathwohl, 2001). I gehør- og musikkteorifeltet kan slik kunnskap være å huske hvilke fortegn ulike tonearter har, å kunne gjenkjenne og gjengi intervaller, å kunne gjenkjenne og gjengi grunnleggende rytmefigurer, eller å vite hvilke harmoniske funksjoner en bestemt harmonisk progresjon er bygget opp av.

**Nivå 2: Forstå.** Nivå to beskriver et kunnskapsnivå der mening kan konstrueres fra muntlige, skriftlige eller grafiske beskjeder. Kunnskapsnivået innebærer evne til å tolke, parafrasere, eksemplifisere, klassifisere, relatere, oppsummere, generalisere, konkludere, sammenligne og modellere. På dette kunnskapsnivået kan kunnskapsinnehaveren endre et budskaps representasjonsform (Anderson & Krathwohl, 2001). I gehør- og musikkteorifeltet kan et slikt kunnskapsnivå komme til uttrykk gjennom evnen til å skape en auditiv representasjon av et lyd- eller notebilde, å gjengi et lydbilde med musikalsk notasjon, å beskrive et modulerende musikkstykkets tonale forløp eller å tolke et harmonisk forløp og forstå hvordan akkordene relaterer seg til hverandre.

**Nivå 3: Anvende.** Tredje nivå av Blooms taksonomi beskriver evnen til å anvende kunnskap i situasjoner som strukturelt ligner situasjonene eller materialet læringen tok utgangspunkt i (Simkin & Kuechler, 2004). I gehør- og musikkteorifeltet kan dette innebære å kunne bruke generell kunnskap om modulasjon til å tolke, forstå og uttrykke et musikalsk forløp som modulerer til andre tonearter enn det studenten har møtt på tidligere. Et eksempel kan være å bruke prinsipper som ligger til grunn for metrisk modulasjon til å forstå og utføre andre metriske modulasjoner enn det som har blitt gjennomgått i undervisningen.

**Nivå 4: Analysere.** Det fjerde nivået i taksonomien beskriver et kunnskapsnivå der kunnskap kan brukes til kompleks analyse av et materiale. Analyse av et verks strukturelle oppbygning krever kunnskap på dette nivået. Slik kunnskap kreves også for å analysere et komplekst lydbilde, eksempelvis for å avdekke hva som kan gjøres for å korrigere lydbildet.

**Nivå 5: Evaluere.** På nivå fem har kunnskapsinnehaveren evne til å evaluere et materiale eller en prosess. Å kunne vurdere hvordan fremføringen av et stykke musikk fremhever musikkens strukturelle oppbygning, krever et slikt nivå av kunnskap. Det krever også et slikt nivå av kunnskap å evaluere et notemanuskript for å finne stilistiske avvik eller andre typer feil i manuskriptet.

**Nivå 6: Skape.** Taksonomiens sjettede nivå beskriver evnen til å bruke spesifikk kunnskap til å skape noe nytt (Anderson & Krathwohl, 2001). Blant annet kreves et slikt nivå av kunnskap når en musiker improviserer over et akkordskjema, når en komponist komponerer en symfoni, eller når en forsker syntetiserer empiri og teori og utleder ny kunnskap av dette.

Nivå fire til seks i taksonomien beskriver læringsmål som er aktuelle i løpet av et musikkstudium. Med utgangspunkt i hvilken kunnskap som vurderes i ny og gammel opptaksprøve, kan det virke som om det ikke forventes at studenter som skal til å *starte* i høyere utdanning har kunnskap på disse nivåene.

## Oppgavetyper

I skriftlige prøver skiller det vanligvis mellom to hovedtyper av oppgaver: såkalte *selected response items* (SR) og *constructed response items* (CR) (Utdanningsdirektoratet, 2021).

**Selected response Items (SR).** SR-oppgaver er oppgavetyper som besvares gjennom å velge ett av flere svaralternativer. Det fins mange varianter av denne typen oppgaver. En av de mest brukte er *flervalgsoppgaver*, eller *multiple choice*, som består av et spørsmål og to eller flere svaralternativer. Et svaralternativ, *nøkkelen*, er korrekt, mens de andre, *distraktorer*, er ukorrekte (Utdanningsdirektoratet 2021). En annen variant av SR-oppgaver er *matching*, hvor det skal trekkes forbindelser mellom to ord, begreper, symboler, eller lignende som hører sammen. *Riktig/galt-oppgaver* og *alternate choice*, hvor det skal tas stilling til to alternative svar, er andre velbrukte varianter av SR (Seifert & Sutton, 2009).

Den første publiserte prøven basert på flervalgsoppgaver som vi kjenner til, ble laget i USA i 1914 (Sirnes, 2015). Vurderingsformen er i dag

veletablert også i Norge, og brukes blant annet i nasjonale prøver i lesing, regning og engelsk (Olsen, 2021). Det er flere fordeler med SR-oppgaver. Oppgavene er som regel betydelig raskere å besvare enn CR-oppgaver. En prøve kan derfor inneholde mange ulike oppgaver, noe som kan bidra til at store deler av et pensum kan dekkes i én test (Kastner & Stangl, 2010). Velutformede SR-oppgaver gir entydige, objektive svar, og gjør bruk av skjønn i vurderingen unødvendig, noe som kan bidra til å sikre en rettferdig sensur (Kastner & Stangl, 2011; Ventouras et al., 2009). SR-prøver kan dessuten rettes automatisk. I situasjoner der mange kandidater avlegger samme prøve, slik tilfellet er i opptaksprøvene til høyere musikkutdanning, er dette en faktor av betydning.

Det er problematiske aspekter ved SR-baserte prøver, først og fremst knyttet til prøvenes reliabilitet. Fordi det ikke er uendelig med plausible distraktorer i en oppgave, vil den som tar en SR-prøve kunne oppnå kunstig gode poengsummer gjennom å gjette på svar. Burton (2001) gjorde en studie med mål om å kvantifisere andelen flaks i *multiple choice*- og *true/false*-tester. Studien konkluderte med at diskrepansen mellom kunnskapsnivå og oppnådd poengsum i slike tester er så stor at den ikke kan sees bort fra i vurderingen av resultatene. Særlig i prøver med kritisk kunnskapsnivå på mellom 50 og 60 prosent av oppgavens innhold, eksempelvis ved en beståttgrense på 55 poeng av 100, kan flaks være avgjørende for om prøven består (Burton, 2001). I en norsk undersøkelse fra 2010, der 275 åttendeklasseelever deltok gjennom å avlegge naturfagsprøver utformet i tråd med studiens formål, oppnådde elevene signifikant bedre resultater i en SR-basert prøve enn i en CR-basert prøve der de fikk tilsvarende spørsmål. Avviket mellom de to prøvenes resultat var størst hos de svake elevene og minst hos de sterke elevene, men tydelig hos alle grupper av elever (Ravlo, 2010).

I Burtons studie fra 2001 var prøvene bygget opp av spørsmål som studentene med stor sannsynlighet visste svaret på og spørsmål som det var svært usannsynlig at de visste svaret på. Når studentene valgte riktig svaralternativ på spørsmål som var langt utenfor deres kompetanseområde, skyldtes dette høyst sannsynlig ren flaks. En SR-basert prøve i en reell utdanningssituasjon vil som regel kun være bygget opp av spørsmål innen et pensum eller kunnskapsområde studenten eller eleven er eller

bør være kjent med. Det vil være muligheter for at den som tar prøven, kjenner igjen det riktige svaret blant svaralternativene selv om hen ikke ville kommet på det uten hjelp (Gjersvik, 2020; Kastner & Stangl, 2010), eller at gale alternativ kan elimineres på bakgrunn av *delvis kunnskap* om det som det spørres om (Burton, 2001; Simkin & Kuechler, 2004). Burton kaller dette *assistert besvarelse*, og argumenterer for at dette svekker SR-prøvenes reliabilitet.

Denne problematikken adresseres også i studier som ser på opptaksprøver i gehør og musikkteori. Baysal et al. (2018) mener SR-baserte GT-prøver kan brukes til å filtrere ut de kandidatene som absolutt ikke er kvalifiserte til høyere musikkutdanning, og at bruk av SR-oppgaver på den måten kan gjøre opptaksprosedyrene mindre ressurskrevende. De mener videre at en annen type prøve vil være nødvendig for å kartlegge om søkerne som passerer SR-testen har det nivået som forventes i høyere musikkutdanning.

En svakhet med SR-oppgaver er at de kun viser om et svar er riktig eller galt. Resultatene sier lite om hva den som har tatt prøven har tenkt, hva som gjorde at vedkommende ikke fikk til oppgaven og hvilke strategier som lå til grunn for besvarelsen. Dette kan gi et ufullstendig bilde av kunnskapsnivået til den som har avlagt prøven, og dessuten avgrense prøvens bruksområde.

Det fins ulike teknikker for å styrke SR-baserte prøvers reliabilitet. Dersom en oppgave har mange alternativer til gale svar (distraktorer), kan sannsynligheten for å gjette riktig reduseres. Effekten oppstår kun dersom svaralternativene er plausible. Det å finne mange gode svaralternativer oppleves ofte som arbeidskrevende og vanskelig (Burton, 2001).

Såkalt *negativ poengsetting* kan øke prøvenes reliabilitet. Teknikken går ut på at det gis minuspoeng for gale svar, og har til hensikt å gjøre det mindre attraktivt å gjette på svarene (Burton, 2001). Metoden er omdiskutert fordi den kan føre til at studenter svarer blankt i frykt for å få minuspoeng, og dermed underpresterer (Harvei & Allern, 2023).

**Constructed response items (CR).** CR-oppgaver er oppgaver uten svaralternativer, der svaret konstrueres av den som løser oppgaven. Svaret kan gis skriftlig eller muntlig, i form av ord, nummer, noter eller andre symboler (Kastner & Stangl, 2010). Det skilles mellom *kortsvarsoppgaver*,

*brief constructed response* (BCR) og mer omfattende oppgaver, *extended constructed response* (ECR) (Seifert & Sutton, 2009). I en BCR-oppgave skal svaret gjerne oppgis i form av et ord, et tall, et symbol eller et begrep. En ECR kan eksempelvis være en essay-oppgave, hvor oppgaven løses ved hjelp av en lengre tekst. Oppgaver som løses ved å sammenfatte informasjon i en graf eller et notebilde, er andre eksempler på ECR.

CR-oppgaver har noen åpenbare fordeler. Mens SR-oppgaver kun viser om studenten har avlagt riktig eller galt svar, kan CR-oppgaver gi et mer nyansert bilde av hva studenten kan. CR-oppgaver kan dessuten formuleres slik at refleksjon, metode og prosess fremkommer i besvarelsen.

I mange fagfelt vil CR-oppgaver ligne mer på utfordringer i yrkeslivet enn det SR-oppgaver gjør (Gjersvik, 2020; Simkin & Kuechler, 2004). Problemene og dilemmaene i yrkeslivet kommer som regel ikke med ledetråder eller forhåndsbestemte svaralternativer, men krever kritisk og kreativ kunnskapsanvendelse. Det kan derfor argumenteres for at arbeid med CR-oppgaver er en nødvendig del av en utdanning.

En svakhet ved CR-prøver er at de er tidkrevende, både å avlegge og å sensurere. Retting og vurdering vil ofte kreve tolkning og skjønn, og sensuren kan ikke nødvendigvis gjøres automatisk (Kastner & Stangl, 2010; Slørdahl, 2017). Undersøkelser har vist hvordan resultatet av sensur gjort av ulike sensorer kan variere, selv med sensorveiledninger som inneholder klare vurderingskriterier (Jensen, 2024; Robson, 2002). Behovet for skjønnsmessig vurdering kan svekke prøvens reliabilitet, og dessuten gjøre prøvesituasjonen mindre rettferdig.

En annen utfordring knyttet til CR-baserte prøver, og særlig dersom de inneholder mange ECR-oppgaver, er at de krever god leseforståelse og gode skriveferdigheter hos den som avlegger prøven (Kastner & Stangl, 2010). Dersom oppgaveteksten feiltolkes og besvarelsen ikke gjenspeiler oppgaven, eller dersom den skriftlige fremstillingen er så svak at det faglige innholdet ikke går tydelig frem av besvarelsen, vil prøven i liten grad avdekke studentens kunnskapsnivå.

I takt med SR-formatets stigende popularitet, har et økende antall studier undersøkt hvilken type kunnskap CR- og SR-baserte prøver egner seg til å vurdere, og om prøveformene kan sees på som likeverdige. Resultatene av studiene er lite ensartede (Simkin & Kuechler, 2004).

Studier gjort av forskere som Bridgeman og Rock (1993), Walstad og Becker (1994), Kennedy og Walstad (1997) og Bacon (2003), konkluderer med at SR- og CR-baserte prøver må ansees som likeverdige med tanke på validitet og reliabilitet, og at testformene har potensial til å vurdere samme type kunnskap. Andre studier, deriblant Birenbaum og Tatsouka (1987), Becker og Johnston (1999) og Dufresne et al. (2002), konkluderer med at SR- og CR-basert testing både måler ulike type kunnskap og ulikt nivå av kunnskap, og at de *ikke* kan sees på som likeverdige.

I de fleste av studiene som ser på prøveformer, behandles SR-baserte prøver som en homogen gruppe tester, uten at det spesifiseres hvilket nivå av kunnskap de måler (Simkin & Kuechler, 2004). I studier der kunnskapsnivå spesifiseres, har det vist seg at SR-baserte tester som måler *forståelse* av det aktuelle kunnskapsfeltet (Blooms nivå 2) samsvarer bedre med CR-oppgaver enn hva SR-oppgaver som kun måler evne til *gjenkjennelse og gjengivelse* av kunnskap (Blooms nivå 1) gjør (Simkin & Kuechler, 2004; Martinez, 1999). I testing som dreier seg om gjenkjennelse og gjengivelse av kunnskap, virker oppgaveformatet å være av stor betydning for resultatene som oppnås. Å diskriminere mellom ulike svaralternativer krever ikke de samme kognitive prosessene som å lete etter og formulere et svar fra eget hode (Birenbaum & Tatsouka, 1987; Martinez, 1999).

## Metodisk innretning

Som nevnt innledningsvis, skiller 2017-prøven seg fra det foregående prøvesettet blant annet gjennom hvilke oppgavetyper som benyttes, hvilket kunnskapsnivå som testes, og på hvilken måte de ulike delene av prøven vektet poengmessig. I denne studien undersøker vi disse faktorene ved opptaksprøven for å se i hvilken grad de kan påvirke dens vurderingskapasitet.

Studien består av en analyse av oppgavetyperne som benyttes i 2017-prøven og de gamle prøvene, og hvilke kunnskapsnivå de ulike oppgavene tester. Teori rundt prøve kategorier og Blooms taksonomi blir brukt som rammeverk for analysene. I tillegg er poengfordelingssystemet studert gjennom å kartlegge antall oppgaver i de to prøvesettene som er

henholdsvis gehør- og teorioppgaver, hvor stor andel av total poengsum som fordeles til gehør- og teorioppgaver, og hvor stor del av henholdsvis gehør- og teorioppgavene som må bestås for å få prøven godkjent.

I studien ble det foretatt en testing av førsteårsstudenter ved Musikonservatoriet i Tromsø (UiT) og Universitetet i Stavanger, Institutt for musikk (UiS) høsten 2022, 2023 og 2024. Testingen foregikk gjennom at studentene avla de gamle opptaksprøvene. Besvarelsene ble brukt til å sammenligne hvordan de samme studentenes kunnskap ble vurdert i den gamle prøven i forhold til i 2017-prøven, som studentene hadde avlagt i opptaket til studiet. Prøveresultatene ble sammenholdt med de innledende sammenligningene av prøvene. Til dette er det brukt multivariate statistiske analyser (ANOVA, t-test og Levene's test).

Det ble sendt ut invitasjon til alle førsteårsstudentene ved UiT og UiS om å delta i studien, samt informasjon om at deltakelse ville innebære å avlegge gamle opptaksprøver og å la oss bruke besvarelser og resultater fra opptaksprøven i GT (2017-prøven) til forskning. Høsten 2022 mottok samtlige studenter ved UiT forespørselen, mens praktiske hensyn gjorde at kun studenter med klassisk sjangerspesialisering ble spurt ved UiS. 18 studenter valgte å delta, noe som utgjorde 42 prosent av alle førsteårsstudentene på de to institusjonene. Høsten 2023 og 2024 mottok samtlige førsteårsstudenter forespørselen, både ved UiT og UiS. 75 av disse deltok, tilsvarende 74 prosent av førsteårsstudentene. Til sammen besto altså utvalget av 96 studenter. Mens det høsten 2022 var en overvekt av studenter med klassisk sjangerspesialisering, var klassiske og rytmiske studenter jevnt representert høsten 2023 og 2024.

Årsaken til lavere deltakelse i studiets første år var trolig måten vi organiserte testingen på. En del av studentene kunne ikke møte på grunn av kolliderende aktivitet i studiet. Da vi gjennomførte testingen i 2023 og 2024, lyktes vi bedre i organiseringen, med det resultat at deltakelsen var større. Vi kan ikke utelukke at studentene som uteble, ville ha påvirket resultatene dersom de deltok i testingen. Allikevel er mener vi at utvalget er tilstrekkelig for å kunne gjøre pålitelige analyser. Studentene ble gjort oppmerksomme på at deltakelsen var frivillig, at de når som helst kunne trekke seg fra studien, og at deltakelse ikke ville få noen konsekvens hverken for undervisningen eller vurderingen av dem i emner i

studiet. Studentene ble også informert om at resultater fra studien ville bli behandlet på en slik måte at individuell anonymitet ble ivarettatt.

For å sikre anonymitet, ble studentenes navn erstattet med numre da resultatene fra undersøkelsen ble lagret. Dokumenter som gjorde det mulig å koble navn og nummer ble umiddelbart slettet etter at dataene var blitt registrert. Fordi studentkullene på begge studiesteder er små, utelot vi å inkludere bakgrunnsinformasjon som kjønn, instrument, sjangertilhørighet og annet som kan gjøre det mulig å gjenkjenne enkelt-individer i studien.

Sammenligning av besvarelser og resultater på de gamle GT-prøvene og 2017-prøven har knyttet enkelte usikkerhetsmomenter til seg. Da studentene avla 2017-prøven i forbindelse med opptaket til studiet, sto studie-plassen på spill. Det er sannsynlig at studentene hadde forberedt seg grundig på fagstoffet som testes i prøven, og at de hadde øvd spesifikt på oppgavetyper som ligner oppgavene i opptaksprøven.

Studentene ble forespurt om å delta i denne studien et par uker forut for testingen, og hadde dermed begrenset tid til rådighet for å forberede seg. Det er derfor sannsynlig at studentene møtte bedre forberedt til 2017-prøven enn til testingen. På den andre siden hadde de av studentene som gikk musikklinje på videregående skole eller folkehøgskole hatt et ekstra semester med undervisning mellom ordinær opptaksprøve og testingen, noe som kan ha påvirket resultatene i motsatt retning.

Vi valgte i denne studien å studere prøvene gjennom deres bruk av oppgaveformer, kompleksitet i oppgavene og poengfordelingssystem. Det fins andre aspekter ved prøvene som også kan være av betydning for studentenes resultater, slik som det at den ene prøven er en papirprøve mens den andre er digital. Studier viser at digitalisering av prøver kan ha en effekt på testresultatene (Pengelley et al., 2025). Gjennom måten vi designet studien på, var det vanskelig å måle i hvor stor grad det er tilfelle her, og vi har derfor valgt å utelate aspektet.

Gjennom analysene og resultatene som følger, vises sider ved opptaksprøvene som kan svekke prøvenes kapasitet til pålitelig vurdering av søkerens kompetanse. Usikkerhetsmomentene som beskrives, gjør at resultatene ikke kan tolkes med sikkerhet. De gir likevel et godt grunnlag for videre arbeid med å forbedre opptaksprøven, og viser måter man,

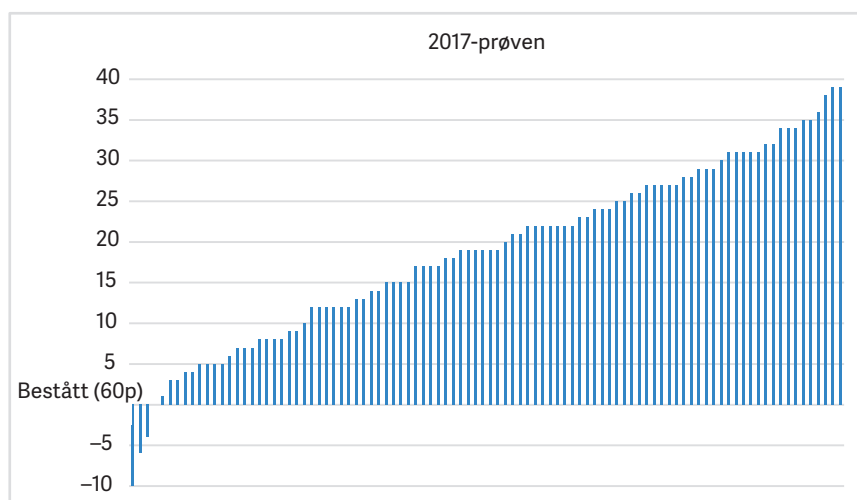
gjennom testing av eventuelle forbedringstiltak før de iverksettes, kan bedre kvaliteten på prøvene.

## Resultater

Presentasjonen av resultater innledes med å vise studentenes totale score på de ulike prøvesettene. Dette første resultatet viser at studentenes kunnskapsnivå i gehør og musikkteori er sprikende. Resultatet gir også et inntrykk av hvordan og i hvor stor grad de to prøvesettene fanger opp dette spriket. Videre presenteres analyser av oppgavetyper, kunnskapsnivå og poengfordelingssystem. Analysene sammenholdes med studentenes resultater på begge prøvesett.

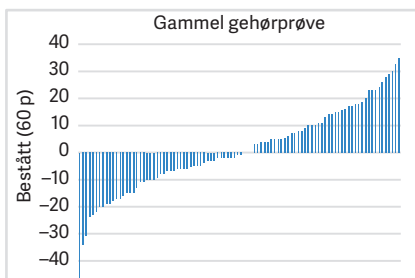
## Totale poengsummer

Nesten alle studentene sto naturligvis på 2017-prøven (kun tre studenter fikk studieplass selv om de ikke besto prøven), mange med god margin i forhold til kravet på 60 poeng. Figur 1 viser resultatene for hver enkelt student relativt til opptakskravet, der 0 på x-aksen er opptakskravet. I gjennomsnitt fikk studentene 78,67 ( $SD = 10,87$ ,  $n = 96$ ) poeng av 100 mulige.

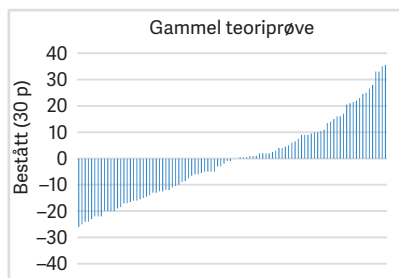


**Figur 1.** Resultater på 2017-prøven, differanse fra opptakskravet (60 poeng).  $N = 96$ .

De samme studentene oppnådde signifikant dårligere resultater da de tok den gamle *gehørprøven* ( $F(1, 189)$ ,  $p < ,001$ ). I gjennomsnitt fikk studentene 61,26 ( $SD = 14,79$ ) av 100 mulige poeng. Opptakskravet var 60 poeng (60 % av maksimal poenggrense). Under halvparten (49 %) av studentene oppnådde opptakskravet. På den gamle *teoriprøven* oppnådde studenten i gjennomsnitt 29,52 ( $SD = 15,52$ ) av 70 mulige poeng. Opptakskravet var 30 poeng (43 % av maksimal poenggrense). Halvparten (50 %) av studentene oppnådde en slik poengsum. Figur 2 og 3 viser resultatene til hver enkelt student på de gamle prøvene relativt til opptakskravene på henholdsvis 60 og 30 poeng.



**Figur 2.** Resultater på gammel gehørprøve, differanse fra opptakskravet (60 poeng).  $N = 96$ .



**Figur 3.** Resultater på gammel teoriprøve, differanse fra opptakskravet (30 poeng).  $N = 96$ .

I gjennomsnitt var differansen fra opptakskravet 18,68 for 2017-prøven, mens gjennomsnittlig differanse fra de respektive opptakskravene var 0,60 for den gamle gehørprøven og  $-0,59$  for den gamle teoriprøven. I den gamle gehørprøven var altså gjennomsnittlig oppnådd poengsum under opptakskravet. En ANOVA-test av gjennomsnittlig differanse fra de respektive opptakskravene, viser at det er signifikante forskjeller ( $M_{NY} = 18,68$ ,  $M_{GGehor} = ,60$ ,  $M_{GTeori} = - ,59$ ;  $F(2, 285) = 54,36$ ,  $p < ,001$ ). Kontrastanalyser viser at differansen fra opptakskravet er signifikant høyere for 2017-prøven, sammenlignet med den gamle gehørprøven ( $p < ,001$ ) og den gamle teoriprøven ( $p < ,001$ ). Relativt til opptakskravene, presterte studentene altså signifikant bedre på 2017-prøven enn på de gamle prøvene.

Det er verdt å merke seg at det er relativt stor forskjell i standardavvikene rundt gjennomsnittsresultatene mellom 2017-prøven ( $SD = 10,87$ )

og de gamle prøvene ( $SD = 14,79$  og  $SD = 15,52$ ). Det indikerer at det kan være større sprik i resultatene i de gamle prøvene, mens resultatene er mindre varierende mellom studentene i 2017-prøven. En test av homogeniteten i variasjonen (Levene's test) viser at det *er* forskjell i variasjonen i resultatene på de ulike prøvene ( $p < ,003$ ). Altså, ikke bare gjorde studentene det bedre på 2017-prøven, men det var også mye mindre variasjon i resultatene. En tolkning av dette resultatet er at 2017-prøven ikke får frem de samme kunnskapsforskjellene mellom studentene som det de gamle prøvene avdekker.

Detaljerte analyser av resultatene viser at noen studenter oppnådde svært lave poengsummer på de gamle prøvene på tross av at de besto 2017-prøven: 2 prosent av studentene oppnådde mindre enn 30 prosent av maksimal poengsum på gammel gehørprøve. 34 prosent av studentene oppnådde tilsvarende lave poengsummer (mindre enn 30 %) på den gamle teoriprøven. Resultatene fra de gamle gehør- og teoriprøvene til studenter som oppnådde 80 prosent eller mer av mulig poengsum på 2017-prøven ( $n = 46$ , 48 % av informantene) var følgende:

- 12 studenter (26 %) oppnådde 80 prosent eller mer av mulig poengsum på gammel *gehørprøve*
- 6 studenter (13 %) oppnådde 80 prosent eller mer av mulig poengsum på gammel *teoriprøve*
- 11 studenter (24 %) oppnådde under 60 prosent (= beståttkrav) på gammel *gehørprøve*
- 21 studenter (47 %) oppnådde under 43 prosent (= beståttkrav) på gammel *teoriprøve*

Tallene viser at minst 47 prosent av studentene som oppnådde 80 prosent eller mer på 2017-prøven ikke ville ha bestått den gamle prøven.

## Studentenes prestasjoner på de ulike oppgavene i gammelt prøvesett

**Gehørprøven.** Tabell 2 viser resultatene på oppgaver i den gamle gehørprøven kategorisert etter om de tester fagstoff som også testes i

2017-prøven eller ikke. Ettersom det er mulig å oppnå ulike poengsummer på ulike spørsmål, viser tabellen prosentandelen av mulig poengsum som er oppnådd i gjennomsnitt.

**Tabell 2.** Gjennomsnittlig resultat på oppgavene i gammel gehørprøve.  $N = 96$

Gehørprøven	Oppnådd andel av mulige poeng	SD
<b>Oppgaver med fagstoff som ikke testet i 2017-prøven</b>	66 %	<b>15,06</b>
Oppgave 3: Auditiv gjenkjenning av harmoniske intervaller	64 %	29,89
Oppgave 7: Enstemt melodiforandring	88 %	19,29
Oppgave 8: Tostemt melodiforandring	67 %	22,84
Oppgave 12: Melodidiktat	53 %	30,75
Oppgave 13: Grunntonebestemmelse	44 %	31,96
Oppgave 14: Ta stilling til intonasjon av sure toner	82 %	21,62
<b>Oppgaver med fagstoff som testes i 2017-prøven</b>	65 %	<b>14,16</b>
Oppgave 1: Auditiv gjenkjenning av dur og moll	84 %	24,99
Oppgave 2: Auditiv gjenkjenning av dur og moll	97 %	11,43
Oppgave 4: Auditiv gjenkjenning av melodiske sluttintervaller	53 %	29,95
Oppgave 5: Rytmekorreksjon	55 %	28,55
Oppgave 6: Rytmediktat	32 %	25,64
Oppgave 9: Melodikorreksjon	29 %	26,41
Oppgave 10: Auditiv gjenkjenning av ulike typer treklanger	75 %	23,82
Oppgave 11: Auditiv gjenkjenning av akkordtyper og omvendning	67 %	21,42
Oppgave 15: Taktartsbestemmelse	90 %	14,14

En t-test av forskjellen i gjennomsnittlig andel av mulige poeng oppnådd på oppgaver med fagstoff som ikke ble testet ( $M_{\text{ikke testet}} = 66\%$ ,  $SD = 15,06$ ) og oppgaver som ble testet i 2017-prøven ( $M_{\text{testet}} = 65\%$ ,  $SD = 14,16$ ) var ikke signifikant ( $t(77) = 1,15$ ,  $p = ,25$ ). Det var altså ingen forskjell på hvordan studentene gjorde det på oppgaver med fagstoff som også ble testet i 2017-prøven sammenlignet med hvordan de gjorde det på oppgaver med fagstoff som ikke testes i 2017-prøven.

**Teoriprøven.** Tabell 3 viser resultatene på oppgaver i den gamle teori-prøven kategorisert etter om de tester fagstoff som også testes i det nye prøvesettet eller ikke.

**Tabell 3.** Gjennomsnittlig resultat på oppgavene i gammel teoriprøve.  $N = 96$ 

Teoriprøven	Oppnådd andel av mulige poeng	SD
<b>Oppgaver med fagstoff som ikke testet i 2017-prøven</b>	28 %	<b>22,92</b>
Oppg. 3: Gjenkjenne skalaer i notebilde, notere skalaer	43 %	34,20
Oppg. 5: 4-stemt korallharmonisering	29 %	32,38
Oppg. 6: Lage motstemme til melodi. Harmonikk og rytmikk er oppgitt	18 %	29,51
Oppg. 7: Transponering og nøkkellesing	23 %	29,07
<b>Oppgaver med fagstoff som testes i 2017- prøven</b>	58 %	<b>23,11</b>
Oppg. 1: Gjenkjenne intervaller i notebilde	68 %	17,39
Oppg. 2: Bestemme akkordtyper og omvendinger i notebilde	55 %	33,64
Oppg. 4: Harmonisk analyse av akkorder i notebilde	51 %	35,03

En t-test av forskjellen mellom gjennomsnittsresultatene på oppgaver med fagstoff som ikke ble testet ( $M_{\text{ikke testet}} = 28 \%$ ,  $SD = 22,92$ ) og oppgaver som ble testet i 2017-prøven ( $M_{\text{testet}} = 58 \%$ ,  $SD = 23,11$ ) viste en signifikant forskjell ( $t(78) = -17,32$ ,  $p < ,001$ ). Studentene gjorde det altså signifikant bedre på teorioppgavene med fagstoff som også ble testet i 2017-prøvene.

## Oppgavetyper i 2017-prøven og gammelt prøvesett

Tabell 4 viser fordelingen mellom SR- og CR-oppgaver i 2017-prøven og det gamle prøvesettet. Enkelte av oppgavene i gammelt prøvesett rommer flere oppgavetyper. Samlet antall oppgaver er derfor ikke samsvarende med tallene i kolonnene for SR og CR.

Noen av CR-oppgavene i det gamle prøvesettet ligner SR-oppgaver. Disse er oppgaver hvor de fleste begreper, tall og lignende som trengs for å svare, oppgis i oppgaveteksten. Til forskjell fra i SR-oppgaver, oppgis ikke alle nødvendige begreper og tall.<sup>4</sup>

4 Eksempel på en slik oppgave er «bestem intervaller i notebildet»-oppgaven i gammel teoriprøve. Her oppgis oppgavetekst: «Bestem følgende intervaller (prim = 1, sekund = 2 etc. — ren = r, liten = l, stor = st, forstørret = fst, forminsket = fm)». Kandidaten vet ut fra oppgaveteksten at det skal oppgis intervallstørrelse i form av tall og forkortelse for liten, stor, ren, forminsket eller forstørret variant. Kandidaten må selv vite hvilke tall som er mulige, samt hvilke intervaller som fins i stor/liten versjon og hvilke intervaller som er rene.

**Tabell 4.** Fordeling SR- og CR-oppgaver i 2017-prøven og gammelt prøvesett

Fordeling CR- og SR-oppgaver i de to prøvesettene	SR	CR
2017-prøve (totalt 19 oppgaver)	100 %	0 %
Gammelt prøvesett (totalt 22 oppgaver)	31 % (8 oppgaver)	69 % (18 oppgaver)

Vi testet hvorvidt studentene oppnådde forskjellige resultater på SR-oppgave sammenlignet med CR-oppgavene på de gamle prøvene. To t-tester viste at studentene i gjennomsnitt oppnådde bedre resultater på SR-oppgavene enn på CR-oppgavene, både i den gamle gehørtesten ( $M_{SR} = 72\%$ ,  $SD = 12,68$ ;  $M_{CR} = 55\%$ ,  $SD = 18,11$ ,  $t(77) = 10,69$ ,  $p < ,001$ ) og den gamle teoriprøven ( $M_{SR} = 55\%$ ,  $SD = 33,64$ ;  $M_{CR} = 38\%$ ,  $SD = 21,02$ ,  $t(78) = 6,28$ ,  $p < ,001$ ). I testen er det ikke tatt hensyn til at noen CR-oppgaver ligner på SR-oppgaver. Testen kan derfor anses som en streng test. Det ville trolig blitt enda større forskjeller om vi tok hensyn til CR som ligner SR.

## Kunnskapsnivå som testes i 2017-prøven og gammelt prøvesett

Tabell 5 viser hvilke oppgaver som tester hvilke kunnskapsnivåer i de to prøvesettene:

**Tabell 5.** Oversikt over hvilke oppgaver som tester ulike kunnskapsnivå

Kunnskapsnivå	Gammel teoriprøve	Gammel gehørprøve	2017-prøve
1A: Gjenkjenne	Oppg. 1: Gjenkjenne intervaller i notebilde. 2: Bestemme akkordtyper og omvendinger i notebilde. 3a: Gjenkjenne skalaer i notebilde.	1 og 2: Auditiv gjenkjenning av tonekjønn. 3: Auditiv gjenkjenning av harmoniske intervaller. 7: Enstemt melodiforandring. 8: Tostemt melodiforandring. 10: Auditiv gjenkjenning av ulike typer treklanger. 11: Auditiv gjenkjenning av akkordtyper og omvendinger. 14: Ta stilling til intonasjon av sure toner. 15: Taktartsbestemmelse.	1a-d <sup>5</sup> : Auditiv gjenkjenning av tonekjønn og intervaller. 1e-f: Gjenkjenning av intervaller i notebilde. 2a-c: Auditiv treklangbestemmelse. 2d-f: Auditiv firklangbestemmelse. 2g-i: Auditiv bestemmelse av akkordomvendinger. 2j-l: Navngi enkeltakkorder. 3e-g: Taktartbestemmelse. 3d: Rytmediktat gjennom seleksjon av rytmefigurer. 3j-k: Gjenkjenning taktart i noteeksempler.

5 1a-d tester samme kunnskap, og regnes som én oppgave. Det samme gjelder 1 e-f, 2a-c osv.

Kunnskapsnivå	Gammel teoriprøve	Gammel gehørprøve	2017-prøve
1B: Gjenkalle/ Gjengi	3b: Notere skalaer. 7d: Nøkkelukunnskap.		1g: Tonearter og fortegn. 2m: Bestemme parallelltonearter.
2: Forstå	4: Harmonisk analyse av akkorder i notebilde.	4: Auditiv gjenkjennelse av melodiske sluttintervaller. <sup>6</sup> 5: Rytmekorreksjon. 6: Rytmediktat. 9: Melodikorreksjon. 12: Melodidiktat.	2n-p: Harmonisk analyse av akkorder i notebilde. 3i: Rytmekorreksjon. 3a-c: Rytmelesing. 1k-l: Melodikorreksjon. 1h-j: Melodilesing. 3h: Samsvar mellom lyd og notasjon. 2q-r: Auditiv bestemmelse av akkordprogresjoner.
3: Anvende	5: 4-stemt koralharmonisering. 6: Lage motstemme til melodi. 7a-c: Transponering og nøkkellesing	13: Grunntonebestemmelse	

Tabell 6 viser fordelingen av oppgaver på ulike kunnskapsnivå i det gamle prøvesettet og 2017-prøven.

**Tabell 6.** Oversikt over fordelingen av oppgavetyper på ulike kunnskapsnivå

Kunnskapsnivå	Gammelt prøvesett	2017-prøve
1A: Gjenkjenne	50 % (12 oppgaver)	48 % (9 oppgaver)
1B: Gjenkalle	8 % (2 oppgaver)	16 % (3 oppgaver)
2: Forstå	25 % (6 oppgaver)	37 % (7 oppgaver)
3: Anvende	17 % (4 oppgaver)	0 % (0 oppgaver)

Analysen viser at det gamle prøvesettet har litt flere oppgaver som tester det aller enkleste nivået av kunnskap, mens det i 2017-prøven er én oppgave mer som tester på nivå 1B og 2. Kunnskapsnivå 3 testes kun i et fåtall oppgaver i gammelt prøvesett. Det viser seg også at oppgaver som tester

6 Denne oppgaven er plassert på nivå 2 fordi den, i tillegg til å kreve at intervallets størrelse gjenkjennes, krever forståelse av intervallets to toners funksjon i tonearten. Eksempelvis har intervallene stor sekst og forminsket septim samme størrelse, de spenner over 4,5 skalatrinn. Dersom intervallet spilles fra durskalaens kvint opp til ters, utgjør intervallet en stor sekst. Spilles det fra mollskalaens høye ledetone opp til sekst, utgjør det en forminsket septim.

samme tematikk i de to prøvene, for det meste, men med enkelte unntak, tester samme kunnskapsnivå.

## Studentenes resultater på oppgaver som tester ulike kunnskapsnivå

Vi utførte en ANOVA for å undersøke om det var forskjeller i resultater avhengig av hvilke kunnskapsnivå som ble testet i de ulike oppgavene i de to prøvesettene. For å sikre et likt sammenligningsgrunnlag ble prosentandel av antall poeng som kan oppnås per oppgave benyttet som grunnlag.

**2017-prøve.** 2017-prøven tester kunnskap på nivå 1a (gjenkjenne), 1b (gjenkalle) og 2 (forstå). ANOVA avdekket signifikante forskjeller i resultatene mellom oppgaver på nivå 1a ( $M_{1A} = 73,29 \%$ ,  $SD = 12,38$ ), nivå 1b ( $M_{1B} = 92,14 \%$ ,  $SD = 13,66$ ) og nivå 2 ( $M_2 = 68,76 \%$ ,  $SD = 14,58$ ) ( $F(2, 68) = 91,79$ ,  $p < ,001$ ). Kontrastanalyser viste at resultatene var signifikant bedre på oppgaver på 1a-nivå, enn oppgaver på nivå 2 ( $F(1, 69) = 8,92$ ,  $p < ,01$ ). Imidlertid var resultatene på 1b-oppgavene bedre enn på begge de to andre nivåene (1b vs. 1a:  $F(1, 69) = 100,85$ ,  $p < ,001$ ; 1b vs. 2:  $F(1, 69) = 129,28$ ,  $p < ,001$ ).

**Gammel gehørprøve.** Den gamle gehørprøven tester kunnskap på nivå 1a (gjenkjenne), 2 (forstå) og 3 (anvende). ANOVA viste at det var signifikante forskjeller i resultatene mellom oppgaver på nivå 1a ( $M_{1A} = 79,37 \%$ ,  $SD = 12,23$ ), nivå 2 ( $M_2 = 51,55 \%$ ,  $SD = 17,40$ ) og nivå 3 ( $M_3 = 44,32 \%$ ,  $SD = 31,96$ ) ( $F(2, 75) = 82,41$ ,  $p < ,001$ ). Kontrastanalyser viste at resultatene var signifikant bedre på oppgaver på nivå 1a, enn på de to andre nivåene (1a vs. 2:  $F(1, 76) = 328,08$ ,  $p < ,001$ ; 1a vs. 3:  $F(1, 76) = 107,05$ ,  $p < ,001$ ). Det var også signifikant bedre resultater på oppgaver på nivå 2 enn på nivå 3 ( $F(1, 76) = 4,71$ ,  $p < ,04$ ).

**Gammel teoriprøve.** Den gamle teoriprøven tester kunnskap på nivå 1a (gjenkjenne), 1b (gjenkalle), 2 (forstå) og 3 (anvende). Oppgave 3 tester både kunnskapsnivå 1a (3a) og 1b (3b), mens oppgave 7 tester både kunnskapsnivå 1b (7d) og 2 (7a–c). Oppgavene 3 og 7 ble følgelig ikke inkludert i analysene av resultater på oppgaver med ulike kunnskapsnivå. Dette

innebærer at kunnskapsnivå 1b ikke er inkludert i de videre analysene, da dette kun er testet i oppgavene 3b og 7d.

ANOVA viste at det var signifikante forskjeller i resultatene mellom oppgaver på nivå 1a ( $M_{1A} = 61,51\%$ ,  $SD = 22,06$ ), nivå 2 ( $M_2 = 50,58\%$ ,  $SD = 35,03$ ) og nivå 3 ( $M_3 = 23,81\%$ ,  $SD = 26,88$ ) ( $F(2, 76) = 82,56$ ,  $p < ,001$ ). Kontrastanalyser viste at resultatene var signifikant bedre på oppgaver på nivå 1a enn på de to andre nivåene (1a vs. 2:  $F(1, 77) = 10,38$ ,  $p < ,01$ ; 1a vs. 3:  $F(1, 77) = 230,32$ ,  $p < ,001$ ). Det var også signifikant bedre resultater på oppgaver på nivå 2 enn på nivå 3 ( $F(1, 77) = 74,17$   $p < ,001$ ).

## Samsvar i resultater på oppgaver som måler samme fagstoff i de to prøvesettene

I mange av oppgavene som tester samme fagstoff i de to prøvesettene, oppnådde studentene samsvarende resultater. Dette gjelder oppgaver som tester auditiv og teoretisk identifikasjon av intervaller, auditiv identifikasjon av tre- og firklanger, identifikasjon av taktart, samt melodi- og rytmekorreksjon. Felles for oppgavene hvor studentenes resultater samsvarte, er at de testet på samme kunnskapsnivå og med samme eller nesten samme<sup>7</sup> oppgavetype.

I teorioppgavene som dreier seg om å bestemme akkorder og omvendinger, samt harmonisk analyse av koralsats, var resultatene i ny prøve vesentlig bedre enn i den gamle. I harmonisk analyse-oppgavene testes samme kunnskapsnivå, men oppgavetypen er ulik (SR i 2017-prøve, CR i gammel prøve). I akkordbestemmelsesoppgaven er oppgavetypen lik, og det testes på samme kunnskapsnivå. Her skiller prøvene seg fra hverandre innholdsmessig, både når det gjelder antall akkorder som skal bestemmes og antall mulige svaralternativer. Dette kan ha påvirket resultatet.

I gehørprøven oppnådde kandidatene bedre resultater i 2017-prøven enn i gammel prøve på oppgaver som tester notasjon og lesing av rytme og melodi, selv om det også i 2017-prøven er stor variasjon i hvor godt studentene har mestret disse oppgavene. Melodioppgavene i de to

<sup>7</sup> Med *nesten samme* menes oppgavetyperne SR og CR som ligner SR.

prøvesettene tester samme nivå av kunnskap, mens rytmeoppgavene tester ulike nivå. Oppgaveformene som benyttes er ulike, i gammel prøve finner vi CR-oppgaver, mot SR-oppgaver i 2017-prøven.

I oppgaver som tester evne til å gjenkjenne dur og moll i musikkutdrag var resultatet dårligst i 2017-prøven, på tross av at samme kunnskapsnivå testes, og med samme oppgavetype. I 2017-prøven er oppgaven en kombinasjonsoppgave der det også skal bestemmes sluttintervall, mens det i den gamle prøven kun skal bestemmes tonekjønn. Dette kan ha påvirket resultatet.

## Poengsystem

I gammelt prøvesettet må studentene oppnå 60 prosent av maksimal poengsum på gehøroppgavene og 43 prosent av maksimal poengsum på teorioppgavene for å bestå. I 2017-prøven må studentene oppnå 60 prosent av maksimal poengsum, men det stilles ikke krav til antall poeng på henholdsvis gehør- og teorioppgaver. Tabell 7 viser antall oppgaver i de to prøvesettene, samt fordelingen mellom gehør- og teorioppgaver.

**Tabell 7.** Antall oppgaver i 2017-prøven og gammelt prøvesett, samt fordelingen mellom gehør- og teorioppgaver

Prøvesett	Totalt antall oppgaver	Maksimal poengsum	Andel gehøroppgaver	Andel teorioppgaver	Maksimal poengsum gehøroppgaver	Maksimal poengsum teorioppgaver
2017-prøve	19	100	63 % (12 oppgaver)	37 % (7 oppgaver)	44 % (44 poeng)	56 % (56 poeng)
Gammelt prøvesett	22	170	68 % (15 oppgaver)	32 % (7 oppgaver)	59 % (100 poeng)	41 % (70 poeng)

Det er verdt å merke seg at 2017-prøven har en større andel gehøroppgaver enn teorioppgaver, men poengmessig gis det større uttelling for teorioppgavene (56 % av poengene gis til 37 % av oppgavene).

Vi undersøkte om skjevheten i poengfordeling var synlig i studentenes resultater på 2017-prøven. Tabell 8 viser gjennomsnittlig resultat på henholdsvis prøvens gehør- og teorioppgaver. Tabellen viser også den studentbesvarelsen der det var størst skjevhet mellom poengsum oppnådd fra gehør- og teorioppgaver.

**Tabell 8.** Poengsummer på gehør- og teorioppgaver i 2017-prøven.  $N = 96$ 

	<b>Totalt</b>	<b>Gehøringsoppgaver</b>	<b>Teorioppgaver</b>
Gjennomsnittlige poengsummer (2017-prøve)	78,5 poeng	30 poeng (38 % av poengene)	48,5 poeng (62 % av poengene)
Besvarelse med størst skjevhet i oppnådd poengsum	75 poeng	17 poeng (27 % av poengene)	46 poeng (73 % av poengene)
Besvarelse med minst skjevhet i oppnådd poengsum	85 poeng	41 poeng (48 % av poengene)	44 poeng (52 % av poengene)

Tabellen viser at studentene i gjennomsnitt oppnådde bedre resultat på teoridelen av prøven enn på gehørdelen, også når vi tar høyde for at poengene i utgangspunktet ikke er likt fordelt mellom gehør- og teorioppgaver. I mange tilfeller var det stor ulikhet i studentenes resultater på gehør- og teorioppgaver. Besvarelsen vist i tabellen er det tydeligste eksemplet på dette. Her har studenten oppnådd 17 av 44 mulige poeng på gehøringsoppgavene, altså langt under halvparten av mulig score, og allikevel bestått testen. Fire besvarelser hadde jevnere poengfordeling enn den opprinnelige fordelingen mellom gehør- og teorioppgaver, mens ytterligere tre oppnådde en poengsum tilsvarende hvordan gehør- og teorioppgavene er fordelt. I samtlige andre besvarelser er resultatet på teorioppgavene vesentlig bedre enn resultatet på gehøringsoppgavene.

## Diskusjon

Testing av førsteårsstudenter tre påfølgende år indikerer at 2017-prøven ikke fanger opp den samme variasjonen i studentenes kunnskapsnivå som det foregående prøvesettet gjorde. Det er viktig både for studentenes mestringsfølelse og evne til å gjennomføre studiet at opptaksprøven fungerer etter hensikten. En prøve som i mindre grad fanger opp variasjon i kunnskapsnivå, kan være dårligere egnet til å sikre nødvendig kompetanse hos søkerne. Det kan føre til at studenter som ikke har gode nok forutsetninger for et musikkstudium (vel å merke slik musikkstudiene tradisjonelt er oppbygd), tilbys studieplass og starter på studiet i den tro at de har den nødvendige kompetansen. Å skulle mestre undervisning og eksamen i emner som forutsetter kunnskap en ikke har, kan innebære en

arbeidsmengde som er uproporsjonalt stor i forhold til emnenes størrelse i studiepoeng, og kan føre til for stor totalbelastning i studiet.

I tillegg til at opptaksprøven skal sikre at søkerne har nok forkunnskap, brukes den på flere studiesteder som verktøy for nivåinndeling i ulike emner. En opptaksprøve som ikke får tydelig frem variasjonen i kunnskapsnivå, kan ha lavere verdi som kartleggingsverktøy.

## Faktorer som påvirker opptaksprøvens vurderingskapasitet

Vi har gjennom studien forsøkt å identifisere faktorer som påvirker prøvens vurderingskapasitet. Analysen av oppgavetyper viser at 2017-prøvesettet utelukkende består av SR-oppgaver, i motsetning til det gamle prøvesettet, hvor en overvekt av oppgavene er CR-oppgaver. Studentene gjorde det signifikant dårligere på CR-oppgaver enn på SR-oppgaver i det gamle prøvesettet. I oppgaver som tester samme kunnskap i de to prøvesettene, men der oppgaven er en CR-oppgave i gammelt prøvesett, oppnådde studentene vesentlig bedre resultater i 2017-prøven. Funnet resonnerer med blant andre Burtons (2001) forskning på SR-basert testing. Burton mener at flaks kan være avgjørende i SR-baserte tester, og særlig dersom den beståttgrensen ligger på 50–60 prosent. Det kan ikke utelukkes at flaks har vært avgjørende for noen av studentene i 2017-prøven, som har en beståttgrense på 60 prosent. Funnet indikerer at 2017-prøven kan gjøres mer pålitelig gjennom å endre oppgavetyper. Å lage digitale BCR-oppgaver (kortsvarsoppgaver) er enkelt med dagens teknologi, og vil kunne være en forholdsvis kjapp og rimelig måte å forbedre opptaksprøven på.

Analysen av hvilke kunnskapsnivåer som testes, viser at oppgavene som tester samme fagkunnskap i de to prøvesettene, stort sett tester tilsvarende kunnskapsnivåer. Prøvene som helhet inneholder omtrent like store andeler oppgaver som tester de ulike kunnskapsnivåene, med unntak av oppgaver som tester kunnskap på nivå tre (anvende). Slike oppgaver fins kun i det gamle prøvesettet.

I testingen var det få studenter som fikk til nivå 3-oppgavene i det gamle oppgavesettet. Det kan skyldes at fagstoffet som testes i nivå

tre-oppgavene er valgt vekk i 2017-prøven. Det er sannsynlig at lærestoffet bortprioriteres i videregående skole og andre før-konservatorieutdanninger når det ikke lenger er en del av opptaksprøven, ikke minst fordi mye av fagstoffet har sterk sjangermessig klassisk tilhørighet og kan oppleves lite relevant for mange.

Fordi det ikke er noen stor forskjell mellom de to prøvesettene når det gjelder fordelingen av nivå 1- og nivå 2-oppgaver, og fordi nivå 3-oppgavene er få, også i det gamle prøvesettet, er det ikke sannsynlig at kunnskapsnivåene som testes i 2017-prøven har påvirket dens kapasitet til å vurdere søkerens kompetanse sammenlignet med den foregående prøven. Forskning på SR-basert testing viser allikevel at resultater fra SR-oppgaver samsvarer bedre med resultater fra CR-oppgaver når forståelseskunnskap (nivå 2) eller sammensatt kunnskap vurderes (Martinez, 1999; Simkin & Kuechler, 2004), sammenlignet med når kunnskap på nivå 1 vurderes. Dersom opptaksprøven i fortsettelsen baseres på SR-oppgaver, kan en mulighet for å styrke prøvens kapasitet til god vurdering være å øke antall oppgaver som tester på nivå 2.

Analysen av poengsystem viser at det er stor ulikhet mellom gammelt prøvesett og 2017-prøven i dette henseendet. I gammelt prøvesett kreves det at studentene oppnår 43 prosent av total poengsum på teoriprøven og 60 prosent av total poengsum på gehørprøven. I 2017-prøven kreves 60 prosent av total poengsum for å bestå, men det er ikke spesifisert hvilke deler av oppgavesettet poengene må hentes fra. En mulig konsekvens av skjev poengfordeling er at prøven kan bestås gjennom å ha kunnskap kun om enkelte deler av temaene som testes, eksempelvis kan den bestås med «full pott» på teorioppgavene og kun 4 poeng på gehørøppgavene.

Analysen av studentenes poengsummer på 2017-prøven viser at de fleste studentene oppnår betydelig svakere resultater på prøvens gehørøppgaver enn på teorioppgavene. Studenten med lavest score på gehørøppgavene oppnådde kun 17 av 44 poeng. Sammenlignet med hvilket prosentmessige resultat som var nødvendig for å bestå gehørdelen i gammel prøve, er resultatet svært lavt. Funnet forteller oss at 2017-prøven sannsynligvis kan gjøres mer pålitelig ved å endre poengsystem. Ved eksempelvis å innføre separate poengkrav for gehør- og teorioppgaver, vil en kunne sørge for at søkere har nødvendig kompetanse både i gehør og musikkteori

når de godkjennes for opptak. Det kan også tenkes at prøven vil fungere bedre som kartleggingsverktøy med en annen form for poenggivning. Å endre poengsystem er en nærmest kostnadsfri og tidseffektiv måte å øke prøvens kapasitet til pålitelig vurdering på. En annen mulighet som vil kunne gi tilsvarende resultat, er å tilbakeføre ordningen fra gammelt prøvesett med en todelt opptaksprøve der delene har separate beståttgrenser.

## Konklusjon og veien videre

Dersom de utøvende musikkutdanningene i fremtiden skal organiseres på samme måte som i dag, med de samme kravene til forkunnskap for å henge med i undervisningen, viser funnene i denne studien at en revisjon av opptaksprøven kan være fornuftig. I en revisjon vil det særlig være hensiktsmessig å se på hvilke oppgavetyper prøven skal bestå av og hvordan poengsystemet og beståttkravet skal være. Fordi det er usikkert hvilke konsekvenser en revisjon av prøven vil gi for de ulike institusjonenes studentopptak, vil et grundig forarbeid være nødvendig. I et slikt forarbeid vil kartlegging av søkermassen og deres vei til høyere utdanning være relevant. Det vil også være viktig å se på hvordan videregående skoler og folkehøgskoler med musikk forbereder elevene på opptaksprøven gjennom sitt fagtilbud, slik at det er samsvar mellom hva en elev kan forventes å ha lært der, og hva som testes i opptaksprøven til høyere utdanning.

FUM hadde gode grunner til å innføre en ny opptaksprøve i 2017. Andelen søkere til høyere utdanning med annen sjangerbakgrunn enn klassisk hadde vært økende i flere år (Tønsberg, 2012), og tiden var i 2015 moden for å utvikle en mer sjangerbred eller sjangernøytral prøve enn den foregående, som ble laget i en tid der høyere musikkutdanning i størst grad var forbeholdt musikere med klassisk sjangerfordypning. En sjangerbred opptaksprøve er viktig både for at den skal oppleves relevant for alle søkere og for at den skal teste så generell kunnskap at den er gyldig for musikere i mange sjangere. En utfordring med å lage en prøve som skal favne studenter med ulik sjangerbakgrunn er at fagtradisjonene fra en sjanger kan kollidere med kravet til startkompetanse i en annen sjanger. Eksempelvis vil musikere som har lært å spille gjennom en gehørbasert musikktradisjon, og der studiet også i stor grad baserer seg på

undervisning i denne tradisjonen, kunne møte på problemer i møte med en opptaksprøve som også må sikre at studenter innen i en notebasert musikktradisjon har nødvendig kompetanse i notesystemet.

Fordi 2017-prøven brukes på studenter fra ulike sjangertilhørigheter, vil det være av interesse å studere hvorvidt og i hvor stor grad sjangertilhørighet har noe å si for studentenes resultat. I denne studien valgte vi å behandle informantene som én gruppe, uten å se på testresultatene i lys av studentenes sjangertilhørighet. Et kjapt overblikk over det empiriske materialet viser ikke tydelig forskjell i resultater fra studenter med klassisk sjangerbakgrunn og studenter med annen sjangertilhørighet. Det vil allikevel kunne være nyttig å gjøre grundigere analyser av materialet for å undersøke om det fins en forskjell i hvordan studenter fra ulike sjangere presterer på prøven. Om forskjellen skulle vise seg å være stor, vil det nødvendiggjøre en debatt om hvorvidt det er hensiktsmessig å ha felles opptaksprøve på tvers av sjangere.

Denne studien inngår i første fase av prosjektet «Opptaksprøvene i gehør og teori til høyere musikkutdanning – et prosjekt i fem faser». Prosjektets *fase to* inneholder kartlegging og analyse av andre nordeuropeiske musikkutdanningsinstitusjoners opptaksprøver. Målet med arbeidet er å få innblikk i hva som testes, og hvordan det testes, i land med sammenlignbare utdanningssystem som oss. Prosjektets *fase tre* adresserer temaet «*Hva trenger en førsteårsstudent å kunne ved studiestart?*». Her gjøres det spørreskjemaundersøkelser og intervjuer av lærere i høyere musikkutdanning for å få innblikk hva lærere i ulike emner, og som representerer ulike sjangerfordypninger, mener musikkstudenter må kunne ved studiestart. I prosjektets *fase fire* planlegges en undersøkelse av hvem dagens søkere til høyere musikkutdanning er, og hvordan deres vei til å søke høyere musikkutdanning er. I *femte og siste fase* av prosjektet planlegges en revisjon av opptaksprøvene i gehør og teori. Revisjonen vil skje på bakgrunn av funn i de fire første prosjektdelene, og vil skje i tett samarbeid med det nasjonale fagmiljøet og med FUM.

Opptaksprøven er et viktig ledd i vurderingen av søkere til høyere musikkutdanning. Ved å sørge for at studentene har et visst fundament av kunnskap og ferdigheter ved studiestart, bidrar prøven til at de har gode forutsetninger for å mestre studiet. Vi håper at denne studien

kan være et nyttig bidrag til arbeidet med å videreutvikle prøven som vurderingsform.

## Forfatterbiografier

**Maria Medby Tollefsen** er førstelektor i hørelære og studieprogramleder for Bachelorstudiet i musikkutøving ved UiT Norges arktiske universitet, Musikkonservatoriet. Hennes forskningsinteresser er blant annet utvikling av gehørmetodikk, bruk av digitale læringsressurser i høyere musikkutdanning, samarbeid mellom forsknings- og praksisfeltet, samt vurdering av gehør- og teorkompetanse i opptaket til høyere musikkutdanning. Hun leder det nasjonale forskningsprosjektet *Opptaksprøven til høyere musikkutdanning*, og har utviklet det hørelæredidaktiske nettstedet *Musificium*. Maria er svært aktivt i et europeisk nettverk av gehørpedagoger, og holder jevnlig hørelæredidaktiske workshops i ulike europeiske land. Hun underviser også på en ettårig mastergrad i gehørdidaktikk ved Conservatorio Arrigo Boito i Parma.

**Esa Lilja** er førsteamanuensis og leder for musikkteoriseksjonen ved Universitetet i Stavanger, samt adjunktprofessor i musikkvitenskap ved Universitetet i Helsingfors, Finland. Han er forsker, musikkteoretiker, pedagog, komponist og rocketrommeslager, og han arbeider i sjangrene euroklassisk musikk og populærmusikk. Siden 1997 har han tilegnet seg variert arbeidserfaring fra alle nivåer av musikkutdanning, fra musikk-skole til høyere utdanning. Hans akademiske arbeid er hovedsakelig fokusert på musikkteori og analyse, heavy metal og musikkpedagogikk.

**Tarje Gaustad** er førsteamanuensis i markedsføring ved Handelshøyskolen BI og Handelshøgskolen ved UiT Norges arktiske universitet. Hans forskningsinteresser er forbrukerpsykologi og merkevarer, særlig hvordan forbrukere reagerer på bedrifters markeditiltak og bærekraftsinitiativ. Gaustad har publisert i flere ledende, internasjonale forskningstidsskrifter, slik som *International Journal of Research in Marketing*, *Journal of Business Research*, *Psychology & Marketing* og *Journal of Cleaner Production* m.fl.

## Referanser

- Aliferis, J. & Stecklein, J. E. (1953). The development of a college entrance test in music achievement. *Journal of Research in Music Education*, 1(2), 83–96. <https://doi.org/10.2307/3344543>
- Anderson, L. W. & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: Complete edition*. Addison Wesley Longman.
- Arenson, M. A. (1983). The validity of certain entrance tests as predictors of grades in music theory and ear training. *Bulletin of the Council for Research in Music Education*, 75, 33–39.
- Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice and short-answer questions in a marketing context. *Journal of Marketing Education*, 25(1), 31–36.
- Baysal, O., Bozkurt, B., Sağer, T. & Doğrusöz, N. (2018). *Towards a hybrid assessment model for music conservatory entrance exams* [Paperpresentasjon]. Computer Science.
- Becker, W. E. & Johnston, C. (1999). The relationship between multiple choice and essay response questions in assessing economics understanding. *Economic Record*, 75(4), 348–357.
- Bergby, A. K. (2003). *Forsøksprosjekt med tilrettelagt undervisning i gehørtrening for studenter som har fått underkjent opptaksprøven i gehør*. NMH-publikasjoner.
- Bergby, A. K. (2013). Relationships between entrance tests and exams in music performance and aural skills at the Norwegian Academy of Music. I I. E. Reitan, S. K. Bergby, V. C. Jakhelln, G. Shetelig & I. Fanavoll (Red.), *Aural perspectives. On musical learning and practice in higher music education* (s. 7–23). Norges musikkhøgskole.
- Bergby, A. K. (2014a). *Hvorfor har vi opptaksprøver i gehør og teori?* Referat fra debatt på Nasjonal gehørkonferanse, Trondheim. Upublisert.
- Bergby, A. K. (2014b). *Høringsnotat*. Upublisert.
- Bergby, A. K. (2016). *Digital opptaksprøve i gehør og musikkteori, rapport*. Upublisert.
- Birenbaum, M. & Tatsuoaka, K. K. (1987). Open-ended versus multiple-choice response formats – it does make a difference for diagnostic purposes. *Applied Psychological Measurement*, 11, 385–395.
- Bloom, B., Englehart, M., Furst, E., Hill, W. & Krathwohl, D. (1956). *A taxonomy of educational objectives: Handbook I. Cognitive domain*. David McKay Company.
- Bridgeman, B. & Rock, D. A. (1993). Relationships among multiple-choice and open-ended analytical questions. *Journal of Educational Measurement*, 30(4), 313–329.
- Burton, R. F. (2001). Quantifying the effects of chance in multiple choice and true/false tests: Question selection and guessing of answers. *Assessment & Evaluation in Higher Education*, 26(1), 41–50. <https://doi.org/10.1080/02602930020022273>

- Dufresne, R. J., Leonard, W. J. & Gerace, W. J. (2002). Making sense of students' answers to multiple-choice questions. *Physics Teacher*, 40(3), 174–180.
- Gjersvik, P. (2020). Eksamenslære for dummies. *Tidsskriftet Den norske legeforening*, 140(9). <https://doi.org/10.4045/tidsskr.20.0142>
- Harvei, K. N. & Allern, E. (2023). *Multiple choice tasks in teaching and assessment*. Universitetet i Oslo: EILIN network for learning and teaching.
- Jensen, L. (2024). Sensorveiledninger i utøvende musikkutdanning. I H. S. Blix, B. A. Haugseth & K. Junttila (Red.), *MusPed:Research: Bd. 10. Vurderingsformer og tilbakemelding i kunstfaglig utdanning*. Cappelen Damm Akademisk.
- Kastberg, S. (2003). Using Bloom's taxonomy as a framework for classroom assessment. *Mathematics Teacher*, 96(6), 402–405.
- Kastner, M. & Stangl, B. (2010). *Multiple choice and constructed response tests: Do test format and scoring matter?* International Conference on Education and Educational Psychology.
- Kennedy, P. & Walstad, W. B. (1997). Combining multiple-choice and constructed-response test scores: An economist's view. *Applied Measurement in Education*, 10(4), 359–375.
- Lilja, E. & Tollefsen, M. (2023). *Opptaksprøven i gehør og teori, innlegg og plenumsdebatt* [Konferansepresentasjon]. Skandinavisk gehørpedagogisk konferanse, Stavanger.
- Martinez, M. E. (1999). Cognition and the question of test item format. *Educational Psychologist*, 34(4), 207–218.
- Norges musikkhøgskole. (2023, 18. desember). *Satslære 1*. <https://student.nmh.no/studiehandboker/startkull-2024/emner/satslaere-i-2>
- Olsen, V. (2021, 29. november). *Flervalgsoppgaver – en papegøye fra Amerika?* Utdanningsnytt. <https://www.utdanningsnytt.no/flervalgsoppgaver-laering-rolf-vear-olsen/flervalgsoppgaver-en-papegoye-fra-amerika/158202>
- Pengelly, J., Whipp, P. R. & Malpique, A. (2025). A testing load: A review of cognitive load in computer and paper-based learning and assessment. *Technology, Pedagogy and Education*, 34(1), 1–17.
- Ravlo, G. (2010). *Hvilken sammenheng er det mellom elevers prestasjoner på en test i naturfag og formatet til oppgavene? En sammenligning av resultater oppnådd på lukkede oppgaver (SR) og tilsvarende oppgaver i åpent format (CR)* [Masteroppgave]. Norges teknisk-naturvitenskapelige universitet.
- Robson, C. (2002). *Real world research. A resource for social scientists and practitioner-researchers*. Blackwell Publishers.
- Ruokonen, I., Kiilu, K., Sepp, A. & Ruismäki, H. (2018). A comparison of the entrance examinations for post-graduate music education at selected Estonian and Finnish universities. *Future Academy, EjSBS* 2, 89–102. <https://doi.org/10.15405/ejsbs.233>

- Råd for utøvende musikkutdanning. (2014). *Mandat til arbeidsgruppe om utarbeidelse av nye opptaksprøver i gehør og musikkteori*. Upublisert.
- Seifert, K. & Sutton, R. (2009). *Educational psychology* (2. utg.). The Saylor Foundation
- Simkin, M. G. & Kuechler, W. L. (2004). Multiple-choice tests and student understanding: What is the connection? *Decision Sciences Journal of Innovative Education*, 3(1), 73–98.
- Sirnes, S. (2015, 23. oktober). *Omstridt prøveform fyller 100*. Utdanningsnytt. <https://www.utdanningsnytt.no/omstridt-proveform-fyller-100/145833>
- Sloboda, J. A. (1985/2005). *The musical mind. The cognitive psychology of music*. Oxford University Press.
- Slørdahl, T. S. (2017). *Kortsvarsoppgaver, en undervisningsveileder*. NTNU, universitetssykehuset.
- Tønsberg, K. (2012). Rekrutteringen til de utøvende musikkutdanningene i Norge. *Norsk tidsskrift for musikkforskning*, 1, 1–36.
- Utdanningsdirektoratet. (2021). *Rammeverk for eksamen LK 20 og LK 20S*. <https://www.udir.no/eksamen-og-prover/eksamen/rammeverk-skriftlig-eksamen-i-lk20-og-lk20s/?depth=0&print=1>
- Ventouras, E., Triantis, D., Tsiakas, P. & Stergiopoulos, C. (2009). Comparison of examination methods based on multiple-choice questions and constructed-response questions using personal computers. *Computers & Education*, 54(2), 455–461.
- Walstad, W. B. & Becker, W. E. (1994). Achievement differences on multiple-choice and essay tests in economics. *The American Economic Review*, 84(2), 193–196.



## KAPITTEL 3

# Berekraft i vurderingspraksis ved vokaleksamen i høgare populærmusikkutdanning

*Bodil Kvernenes Nørsett*

Ansgar høyskole

**Abstract**

This article investigates sustainable assessment practices in the context of vocal performance examinations within higher popular music education in Norway. Focusing on three institutions, it explores how sustainability principles – emphasizing student participation, self-assessment, and dialogic feedback – are reflected in syllabi, examiner guidelines and assessment practices. The study draws on qualitative analyses of institutional documents and semi-structured interviews with students and assessors to examine the alignment between assessment frameworks, learning outcomes and the lived experiences of participants. Findings highlight a fragmented approach to sustainable assessment. While some elements, such as student reflection and assessment competence, are present, explicit integration of student participation into summative assessment processes remains limited. The article concludes that sustainable assessment in higher popular music education necessitates a systematic strategy integrating collaborative feedback, self-and peer-assessment, and instrument-specific examiner guidelines. Such reforms are essential to align assessment practices with the pedagogical goals of fostering lifelong learning and preparing students for dynamic careers in the music industry.

**Keywords:** sustainable assessment, higher popular music education, constructive alignment, lifelong learning, popular vocal pedagogy.

Sitering: Nørsett, B. K. (2025). Berekraft i vurderingspraksis ved vokaleksamen i høgare populærmusikkutdanning. I H. S. Blix, K. J. Valkoinen & B. A. Haugseth (Red.), *MusPed:Research: Vol. 11. Vurderingsformer og tilbakemelding i kunstfaglig utdanning* (s. 67–100). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.269.ch3>  
Lisens: CC BY 4.0

## Innleiing

Vurdering i høgare populærmusikkutdanning spelar ei avgjerande rolle i å forme læringsprosessane til musikkstudentar. Denne artikkelen undersøker korleis berekraftig vurdering kjem til uttrykk i summativ vurdering (eksamenskonserten) i høgare populærmusikkutdanning, og kva oppfatning vurderingsaktørane (studentar og vurderarar) har av vurdering som læring.

Denne artikkelen, som er basert på datamateriale frå eit doktorgradsprosjekt, tek for seg vurderingspraksis ved eksamenskonserten slik denne praksisen vart opplevd av utøvande vokalistar ved tre høgare populærmusikkutdanningsinstitusjonar i Noreg i tidsrommet 2020–2022<sup>1</sup>. Høgare populærmusikkutdanning er ein paraplyterm som inkluderer utdanning i populærmusikkbransjen, musikkvitskap, pedagogikk, utøving, produksjon, teknologi, teoretisk forskning og låtskriving (Reinhert, 2018, s. 4). Definisjonen av populærmusikk er, ifølgje Powell (2011, s. 4), open for «ei subjektiv tolking av kva som er populært», noko som speglar dei dynamiske og stadig skiftande normene innan feltet. I norsk kontekst vert omgrepa *rytmisk* og *populærmusikalsk* nytta synonymt for å beskrive groove-baserte stilartar som omfattar alle sjangrar som stammar frå jazz, rock, country, blues, pop, soul og funk, til dei nyaste stilartane innanfor hiphop, elektronika og laptop-produisert musikk (Danielsen, 2002, s. 131).

Berekraftig vurdering (sustainable assessment) er eit omgrep David Boud introduserte for å knytte vurdering nærare til FN-måla (Boud, 2000). Boud (2022) legg vekt på at vurdering i høgare utdanning kan forståast som tre typar aktivitetar: sertifisering gjennom karakterar, tilbakemelding for å støtte læring og bygging av studentane sin kapasitet til eigenvurdering. Den siste dimensjonen, som handlar om å utvikle evna til å vurdere eige arbeid over tid, vert også kalla berekraftig vurdering (Boud, 2022, 0:15–1:13). Perspektivet legg vekt på at vurdering ikkje berre

---

<sup>1</sup> Alt datamateriale vart samla inn under koronapandemien. Nedstenginga fekk store konsekvensar for eksamensgjennomføringane i denne perioden, der nokre av informantane fekk endra eksamensforma til ein sjølvproduisert video i staden for konsert. Intervjumaterialet er likevel basert på den levande eksamenskonserten som utgangspunktet for samtalen.

skal måle kva studentane har lært, men støtte læringsprosessane deira også etter avslutta utdanning (Boud, 2000; Falchikov & Boud, 2007).

Sjølvs om forskning har peika på behovet for meir variasjon i vurderingspraksisar, syner studiar frå norsk høgare utdanning at summative, instruktørstyrte tilnærmingar framleis dominerer, noko som tyder på at praksis ikkje har følgt med teoriutviklinga (Hansen, 2022; Lauvas et al., 2000; Raaheim, 2019). Kristiansen (2012) og Nikolaisen (2020) kommenterer at vurdering som læring, ein kjernekomponent i berekraftig vurdering, er lite implementert i kunstfag som drama og musikk. Raaheim (2019) peikar på at dette kan avgrense den aktive deltakinga og sjølvstendet til studentane i læringsprosessane. Eigne erfaringar som rytmisk vokalpedagog støttar desse funna. Samstundes spelar den summative eksamenskonserten ei viktig rolle for studentane som ei motiverande avslutning på studiet.

Under tematikken berekraftig vurdering i norsk høgare populærmusikkutdanning svarar studien på desse to forskingsspørsmåla:

- 1 I kva grad legg vurderingsgrunnlaget for eksamenskonserten til rette for studentmedverknad og eigenvurdering?
- 2 Korleis opplever vurderarar og vokalstudentar vurderingspraksisane, og korleis kan berekraftige vurderingspraksisar implementerast for å støtte livslang læring?

For å svare på desse spørsmåla, undersøker studien korleis berekraftig vurdering kjem til uttrykk i emneplanar, sensorrettleiingar og gjennom erfaringane til vurderingsaktørane ved tre norske institusjonar for høgare populærmusikkutdanning. Gjennom kvalitativ analyse av dokument og intervju vert det vurdert i kva grad vurderingsrammene støttar eller hindrar berekraft, og korleis aktørane opplever praksisane som støttande for livslang læring. Artikkelen er forankra i teoriar om berekraftig vurdering (Boud, 2000; Boud & Soler, 2016; Raaheim, 2019) og diskuterer korleis praksisar som dialogbasert tilbakemelding, studentrefleksjon og transparente vurderingskriterium kan styrkje læringsutbyttet og vurderingskompetansen i norsk høgare populærmusikkutdanning. Målet med artikkelen er å utvide kunnskapen om korleis studentar og lærarar oppfattar berekraft og studentorienterte

vurderingspraksisar, og dessutan å identifisere høva for vidare utvikling av vurderingspraksisar som fremjar livslang læring.

## Konteksten for studien: tidlegare forskning på vurdering i kunstfag

Forskingslitteraturen om vurdering av praktiske kunstfag framhevar gjennomgåande at vurdering er ei kompleks og utfordrande oppgåve (Elliott et al., 2019; McPherson & Schubert, 2004; Rønningen, 2018; Smart & Dixon, 2002; Vinge, 2011). Kjernen i utfordringa ligg i at vurdering i utdanning krev klare kvalitetsstandardar (Sadler, 2009), medan kunstnarlege kvalitetar ofte er dynamiske og kontekstuelle og unndreg seg standardisering (Arlov, 2016). Dette gjer vurdering særleg sårbart for utøvarar som vert vurderte på grunnlag av kreative og populærmusikalske framføringar. For rytmiske vokalistar er utfordringa ytterlegare skjerpa, ettersom stemma ikkje berre er eit teknisk verktøy, men også ei personleg og kroppsleg forankra uttrykksform knytt til individet sin identitet (Hughes, 2013; Juntunen et al., 2023; Lemon-McMahon, 2022; Schei, 2007). Powell med kollegaer (2019) presiserer denne utfordringa ved å påpeike at populærmusikkutdanning handlar om å «støtte og utvikle den musikalske verda som allereie bur i studentane» (2019, s. 24, mi omsetjing). I 2021 vart 200 delegatar frå 26 ulike nasjonar samla i ein digital verdsallianse for kunstutdanning (World Summit on Arts Education). Dette resulterte i *The guiding principles for the assessment of arts learning*. Desse retningslinjene understrekar at berekraftig vurdering er viktig for læring i kunstfag, og at flytande og dynamiske standardar må tilpassast kreative og individuelle uttrykk (Brophy et al., 2021, s. 7).

Rytmask vokalpedagogikk har fått auka merksemd i forskingslitteraturen dei siste tiåra (Benson, 2020; Keskinen et al., 2024; Mesiä, 2019), men i den nordiske konteksten meiner Mesiä (2019) at det er lite forskning på vurdering i rytmisk vokalundervisning. Det er framleis utfordringar knytt til vurdering av rytmisk vokalfremføring, særskilt i samband med sjanger og smak, noko til dømes von Germeten har utforska i musikalundervisning (2022), men også ved at både det personlege uttrykket og

opplevinga til publikum spelar inn på vurderinga av ei vokalframføring (Juntunen et al., 2023, s. 2). I motsetnad til klassisk song, der vurdering tradisjonelt er knytt til tradisjonar og komponisten sin intensjon (Coimbra et al., 2001; Gynnild, 2010), er rytmisk vokal meir kontekstuell og basert på ekspressivitet, individualitet, kreativitet og stilistisk integritet (Hughes, 2014, 2015, s. 179). Samstundes skapar nettopp den kreative fridomen og den ofte tause kunnskapen unike utfordringar for å etablere konsistente vurderingsprinsipp (Isbell, 2022; Powell & Smith, 2019; Waagen, 2021, s. 49).

Berekraftig vurdering har som mål å styrkje studentane si evne til å reflektere over og evaluere eigen praksis, ikkje berre i utdanninga, men gjennom heile yrkeslivet (Bartleet et al., 2019; Fastré, 2023; Raaheim, 2019). Perspektivet er særleg relevant i rytmisk vokal, der kunstnarleg utvikling og kreativ fridom er sentrale komponentar. Hughes (2020) og Hallam (2019) peikar på at summative vurderingar som ved eksamenskonserten kan fungere som informasjonskjelde for framtidig læring dersom vurderingspraksisane er tilrettelagde med dette føremålet. Boud og Soler (2016, s. 407) understrekar at summative eksamensformer ofte vert oversette som læringsarenaer, og etterlyser vurderingspraksisar som støttar langsiktige læringsmål (s. 401). To døme som peikar på moglege berekraftige løysingar for vurderingspraksis ved musikkframføring i høgare populærmusikkutdanning er forhandlingsvurdering (negotiating assessment), utvikla i Storbritannia (Kleiman, 2007, 2009; Powell & Smith, 2019), eller deltakande vurdering (collaborative assessment), utvikla i Finland og Australia (Partti et al., 2015). I Noreg peikar Haakstad (2011, s.79) på at vurderingspraksisane i siste instans handlar om faglege spørsmål, og at det kan vere behov for å integrere berekraftige vurderingspraksisar meir systematisk, kunnskapsbasert og dokumentert i utdanningane for å sikre at vurdering støttar læring og transparens.

Studien adresserer altså fleire forskingshol innan vurdering av rytmisk vokal i høgare musikkutdanning. For det første undersøker han korleis berekraftige vurderingspraksisar kjem til uttrykk i emneplanar og sensorrettleiingar ved tre norske populærmusikkinstusjonar, eit tema som tidlegare har vore lite belyst i forskingslitteraturen. For det

andre inkluderer han perspektivet til både studentar og lærarar og undersøker korleis desse opplever summative vurderingar som ein del av læringsprosessar. Gjennom ei triangulering av funn frå dokumentanalyse og kvalitative intervju bidreg studien til eit meir heilskapleg bilete av vurderingspraksis i norsk høgare populærmusikkutdanning. Målet er å vise korleis vurderingspraksis for eksamenskonserten kan utviklast for å støtte livslang læring, samstundes som han tek omsyn til dei særskilte utfordringane og høva til utvikling på fagområdet for rytmisk vokalpedagogikk.

## **Teoretisk rammeverk: teori om berekraftig vurdering**

All vurderingspraksis, inkludert vurdering av eksamenskonserten, har forankring i verdiar og ideologiar. Boud (2000) definerer berekraftig vurdering som vurdering som «møter samtidige behov og samstundes førebur studentane for framtidige læringsbehov» (s. 151, mi omsetjing). I norsk kontekst vert berekraftig vurdering omtalt som vurdering *som* læring, eller studentmedverkande vurdering, ei tilnærming som legg vekt på at studentane er aktivt med i vurderingsaktivitetar (Raaheim, 2019, s. 29). Dette inkluderer både vurdering *av* læring (summativ vurdering) og vurdering *for* læring (formativ vurdering). Berekraftig vurdering fokuserer på å styrkje studentane si evne til refleksjon og eigenvurdering, noko som støttar deira profesjonelle utvikling og livslange læring. Denne studien nyttar berekraftig vurdering synonymt med studentaktiv vurdering for å understreke tilknyttinga til livslang læring. Omgrepet er forankra i ei forståing av vurdering som ein prosess som fremjar sjølvstende, kritisk refleksjon og evne til å ta ansvar for eiga læring. Hughes (2015) demonstrerer korleis berekraftig vurdering kan gjennomførast i rytmisk vokalundervisning, ved å peike på betydninga av tilbakemeldingsstrukturen i emneplanen. Emneplanen tilrettelegg for både formell og uformell tilbakemelding, som saman med refleksjon og refleksivitet er avgjerande for læring og utvikling av rytmiske vokalistar (Hughes, 2010, s. 256). Teoriane om

meningsskapande samsvar (constructive alignment), utvikla av Biggs og Tang (2015) og sett i norsk kontekst av Raaheim (2019), legg vekt på samanhengen mellom læringsutbytte, undervisningsaktivitetar og vurderingspraksis. Dette rammeverket er avgjerande for å forstå korleis vurderingsrammer kan styrkje eller hemme læring. Berekraftig vurdering gjev ytterlegare mening til dette samsvaret: Om studenten sjølv er delaktig i vurderingsarbeidet, kan vurderinga i større grad tilpassast læringsmåla og gjere studenten ansvarleg for læringsprosessen på ein meir demokratisk måte enn når studenten berre er objektet for vurdering (Raaheim, 2019, s. 29). I vurdering av praktisk-estetiske fag er omgrepsparet holistisk og analytisk (kriteriebasert) vurdering ein sentral teori (Vinge, 2011; Waagen, 2021). Waagen (2021, s. 60) diskuterer bruken av sensorretteiingar i samband med eksamenskonserten i høgare musikkutdanning og argumenterer for behovet for å styrkje fagmiljøa ved å knyte vurderarane sin intuitive og holistiske kvalitetsoppfatning til språkleg artikulert kunnskap.

Boud og Soler (2016, s. 409) minner om at sjølv om utforminga på vurderingsoppgåva er viktig i berekraftig vurdering, må student-medverknad òvast på og integrerast i undervisning og læring. Dei viser også til Falchikov og Boud (2007), som foreslår fem nøkkelement som kan utvikle ei informert vurderingsevne hjå studentar: 1) å identifisere seg som ein aktiv lærande; 2) å identifisere sitt eige kunnskapsnivå og hola i dette; 3) å øve på å vurdere og bedømme; 4) å utvikle desse evnene over tid, og 5) inkorporere refleksivitet og eit forpliktande engasjement til dette (Boud & Soler, 2016, s. 402). Informert vurderingsevne inngår i omgrepet vurderingskompetanse (Fjørtoft & Sandvik, 2020). Boud presiserer at vurderingskompetanse er eit element som ikkje studenten kan utvikle åleine, det er ein ferdigheit som må utviklast gjennom læreplanar og pedagogikk i tilknytning til dei ulike læringsutbytta innanfor faget (Boud & Soler, 2016; O'Neill & Boud, 2023). Gjennom å knyte dei teoretiske perspektiva til empiriske funn undersøker denne studien korleis berekraftige tilnærmingar kjem til uttrykk i vurderingsrammer, og korleis slike praksisar kan utviklast for å støtte livslang læring.

## Metode

For å undersøkje i kva grad vurderingsgrunnlaget for eksamenskonsertern legg til rette for studentmedverknad, og korleis vurderarar og vokalstudentar opplever vurderingspraksis i norsk høgare populærmusikkutdanning, kombinerer denne studien dokumentanalyse og intervjuanalyse. Spesifikt er det analysert læreplanar og sensorrettleiingar, samt intervju med vurderingsaktørane, det vil seie vurderarar og studentar. Vurderarane i studien er vokallærarar ved dei undersøkte institusjonane som i tillegg har roller som interne sensorar eller eksterne sensorar. Ei triangulering av læreplandokument, lærarar sine tolkingar og studentar sine erfaringar kan gje innsikt i korleis overordna verdiar vert reflekterte på tvers av nivåa. Dette perspektivet gjev eit breiare grunnlag for å forstå vurderingspraksis i rytmisk vokalutdanning.

Dokumentanalysen omfattar læreplanar og sensorrettleiingar frå tre norske institusjonar for høgare populærmusikkutdanning, valde basert på storleik (tre av dei største rytmisk utøvande utdanningane i Noreg) og breidde i sjangertilbod:

- 1) Universitetet i Agder (UiA): utøvande rytmisk musikkutdanning
- 2) Høgskolen Kristiania (HK): populærmusikkstudiet
- 3) NLA Høgskolen (NLA): utøvande rytmisk musikkutdanning

Dokumentanalysen fokuserer på læreplanar og sensorrettleiingar som var i bruk i perioden 2020–2022. Læreplanane vart henta frå institusjonane sine nettsider, medan sensorrettleiingane vart innhenta via direkte kontakt med administrasjonen. Analysen er avgrensa til dei delane av læreplanane som omhandlar eksamenskonsertern for hovudinstrument tredje studieår. Etersom læreplandokument er dynamiske og kontinuerleg under revisjon, har studien fokusert på dokument som var relevante for dei studentinformantane som vart vurderte i det same tidsrommet.

Både dokument- og intervjuanalyse byggjer på kvalitative metodologiske prinsipp, der tekstar vert undersøkte og tolka for å skape innsikt og gje forståing og kunnskap (Bowen, 2009, s. 28). Som forskar har eg ei hermeneutisk-fenomenografisk tilnærming innanfor eit

sosialkonstruktivistisk perspektiv. Denne tilnærminga legg vekt på mi rolle som deltakande forskar i å skape empirisk kunnskap gjennom dialog med informantane (Braun & Clarke, 2022, s. 183). Ei fenomenografisk tilnærming vert nytta for å utforske variasjonar i gruppeforståinga av berekraftig vurderingspraksis ved vokaleksamen (Marton, 1981; Marton & Pong, 2005), medan det sosialkonstruktivistiske perspektivet tillèt ei djupare tolking av strukturar og språk. Med bakgrunn som vokalstudent, vokalpedagog og læreplanutviklar i høgare populærmusikkutdanning tolkar eg materialet både frå innsida og frå utsida av feltet. Braun & Clarke (2022) meiner at eit slikt innsideperspektiv i kvalitativ analyse kan vere ein fordel, då det gjev forskaren fleire tolkingstilgangar. Det er viktig for meg i eit forskingsetisk perspektiv å formidle historier som vert opplevde som sannferdige, både for aktørane som vert intervjuva, og når det gjeld institusjonsdokumenta som vert analyserte (Braun & Clarke, 2022, s. 214).

Tabell 1 gjev ei oversikt over det empiriske materialet som forskings-spørsmåla er forankra i.

**Tabell 1.** Oversikt over empiriske analysekjelder

<b>Datasett</b>	<b>Datakjelder</b>
Institusjonsdokument <b>UiA, MUR312 Utøvende hovedområde 4 (høst 2020)</b> <b>NLA 4MHI300 Hovedinstrument III (2020-21)</b> <b>HK MUT300 Hovedinstrument 3 (2020 høst)</b>	Emneplanar for utøvande hovudinstrument Sensorrettleiing
<b>Intervju med informantar</b>	<b>Intervjutraskript</b>
Vurderarar (lærarar/sensorar): UiA, NLA og HK	<b>Lærer 1-9</b>
Vokalstudentar (bachelor/master): UiA, NLA og HK	<b>Student 1-15</b>

For å triangulere dokumentanalysen vart semistrukturerte intervju gjennomførte med ni vurderarar og femten bachelor- eller masterstudentar med utøvande rytmisk vokal som hovudinstrument ved dei tre institusjonane. Informantgruppa bestod av fem menn og nitten kvinner, og intervjuva fokuserte på informantane sine forventingar, erfaringar og opplevingar med vurderingspraksis. Utvalet av informantar vart gjort gjennom eit strategisk utval (Thagaard, 2018, s. 54) for å sikre variasjon i erfaringar og perspektiv, men også kjønnsbalanse og

balanse med omsyn til institusjonsrepresentasjon. Informantane vart rekrutterte gjennom kontakt med institusjonsadministrasjonen. Alle intervjuar vart gjennomførte med videoprogramvaren Zoom med ein semistrukturert intervjuguide. Både lyd og bilete vart teke opp, men berre lydopptaket var gjenstand for transkripsjon og analyse. Intervjumaterialet vart analysert og tematisert ved hjelp av refleksiv tematisk analyse (Braun & Clarke, 2022) og avdekka fire store tematikkar kring vurderingspraksis ved eksamenskonserten: vurdering, tilbakemelding, læring og kvalitetar (vokale verdiar). Tabell 2 syner korleis undertematikkar om berekraft er kopla til tre av dei fire hovudtematikane. Kvalitet og vokale verdiar er knytte opp til vurderingskriterier og vart ikkje direkte påverka med tanke på studentmedverknad og berekraft, men sju undertematikkar kunne koplant til vurdering, tilbakemelding og læring. To av undertematikkane er tydeleg delte ut frå vurderarane og studentane sine ulike perspektiver, medan dei øvrige er felles.

**Tabell 2.** Empiriske hovudtema og berekraft i vurdering som undertema hjå vurderarar og studentar

Hovudtema intervju	Undertema lærarintervju	Undertema studentintervju
<b>Vurdering</b>	Formativ vurdering er studentorientert	Summativ vurdering er sensororientert
	Vurdering har berekraftrelevans for populærmusikkbransjen	
	Oppfatning av eksamenskonserten som autentisk (relevant) vurderingsform	
	Graden av studentmedverkande vurdering i rytmisk vokalundervisning	
<b>Tilbakemelding</b>	Oppfatning av tilbakemelding og karakter som (ir)relevante vurderingsuttrykk for livslang læring	
<b>Læring</b>	Oppfatning av læringsutbytte av eksamenskonserten. Læringserfaringar med livslangt perspektiv	
	Berekraftig vurdering på oppfatta og gjennomført læreplannivå	Berekraftig vurderingskunnskap på erfart læreplannivå

Forskinga vart godkjend av Norsk senter for forskingsdata AS (NSD, no Sikt) i 2018, og alle informantar har gjeve informert samtykke. Institusjonane er namngjevne, då dokumenta som er analyserte, er offentlege, i tråd med forskningsetiske retningslinjer for openheit og transparens. Namngjevinga er likevel vurdert opp mot potensielle ulemper for institusjonane.

Føremålet er ikkje å peike på kritikkverdige forhold, men å identifisere strukturelle mønster som kan bidra til utvikling av berekraftige vurderingspraksisar i kunstfagutdanningar. Ingen personlege opplysningar som kan identifisere informantane, er inkluderte, og anonymitetsprinsippet avgrensar analytiske koplingar mellom informantgrupper og dei namngjevne institusjonane.

## **Kva finst av berekraft i læreplandokumenta for utøvande hovudinstrument?**

Dei kommande avsnitta fram til intervjuanalysane tek utgangspunkt i spørsmålet «I kva grad er berekraft synleg i det formelle vurderingsgrunnlaget for høgare rytmisk vokalutdanning i Noreg?». Berekraft i denne studien er forstått som synonymt med studentmedverkande vurdering og vurdering som læring (Raaheim, 2019). For å gjere dokumentanalysen transparent for lesaren har eg samanstillt emneplanane for hovudinstrument ved dei tre institusjonane i tabell 3 og 4. Emneplanar i norsk høgare utdanning følgjer ein nasjonal mal, som strukturerer læringsutbytta i kategoriane kunnskap, ferdigheiter og generell kompetanse (NOKUT, 2023). Samstundes er det faglege innhaldet utforma av faglærarane ved kvar enkelt institusjon, noko som har ført til variasjonar i korleis vurderingspraksis vert uttrykt og operasjonalisert, til dømes ved musikkutdanningane (RUM, 2018). For oversikt har eg delt emneplanane i to hovuddelar: Tabell 3 skildrar emna og vurderingspraksis, og tabell 4 summerer opp læringsutbytta. Analysen baserer seg på emneplanar frå Universitetet i Agder (Universitetet i Agder, 2023), NLA Høgskolen (NLA Studietilbud, 2021) og Høgskolen Kristiania (Kristiania, 2022) slik dei var formulerte i tidsrommet 2020–2022.

## **Berekraft i emneplanane**

Fokuset for berekraftig vurdering i denne studien er den avsluttande eksamenskonserten, så sjølv om emneplanane også peikar på formative vurderingssituasjonar, særleg gjennom samspelgrupper og ulike

bandsmanssettjningar, avgrensar eg kommentarane hovudsakleg til eksamenskonserten. Høgskolen Kristiania (HK) vektlegg eit livslangt læringsperspektiv og understrekar at «Emnet skal bidra til å utvikle både håndverksmessige kvalifikasjoner og kunstnerisk særpreg, samt sette studenten i stand til å kunne fortsette egen kompetanseutvikling og spesialisering på en stadig mer selvstendig basis» (HK MUT300). Variasjonen mellom emneplanane speglar ulike prioriteringar, men det er uklart korleis berekraft vert operasjonalisert i vurderinga av eksamenskonserten.

Når det gjeld den summative vurderinga, nemner emneplanane lengda på konserten (45 minutt ved UiA, 30 minutt ved NLA og 20 minutt ved HK), men berre NLA Høgskolen (NLA) konkretiserer korleis vurderinga vert gjennomført, når det står: «Studenten vurderes i forhold til emnets læringsmål. Vurderingen gjøres av minst én ekstern og to interne sensorer, hvorav normalt én er studentens hovedinstrumentlærer. Eksamenkommisjonen skal gi en muntlig uttalelse om eksamenskonserten til studenten» (NLA 4MHI300). Universitetet i Agder (UiA) omtalar ingen strukturelle element som indikerer studentmedverkande vurdering. NLA skriv at studenten får ei munnleg tilbakemelding, men formuleringa «Eksamenkommisjonen skal gi en muntlig uttalelse om eksamenskonserten til studenten» (NLA 4MHI300) legg ikkje opp til dialog og speglar snarare einvegskommunikasjon. HK kombinerer eit skriftleg refleksjonsnotat med konserten, men ut frå emneplanen er det uklart om tilbakemeldinga etter konserten fungerer som ein del av vurderingsgrunnlaget eller som ein separat sekvens. Desse variasjonane tyder på at praksisen for studentmedverkande vurdering har ulike prioriteringar og praksisar mellom institusjonane.

Tabell 3 nedanfor syner korleis hovudinstrumentundervisninga er lagt opp. Den viser korleis skilnadene i studiepoeng vert reflekterte i arbeidsomfanget, og den viser skilnadar i arbeidskrav/vilkår som skal oppfyllest for å gå opp til eksamen.

**Tabell 3.** Samanstilling av emneplanar for hovudinstrument (2020–2022) ved utøvande rytmiske studium: Universitetet i Agder, NLA Høgskolen og Høgskolen Kristiania

Kategori	Universitetet i Agder (UIA), MUR312 Utøvende hovedområde 4 (høst 2020)	NLA Høgskolen (NLA) 4MHI300 Hovedinstrument III (2020–21)	Høgskolen Kristiania (HK) MUT300 Hovedinstrument 3 (2020 høst)
Studiepoeng	15	20	15
Arbeids- og undervisnings-former	<p>Emnet inneholder hovedinstrumentundervisning, samspill, seminarer og prosjekt.</p> <p>Hver student gis individuell undervisning på sitt hovedinstrument. Undervisningen kan organiseres i grupper om det er hensiktsmessig. Samspillundervisningen foregår i faste grupper og i ulike prosjekter.</p> <p>I løpet av emnet deltar studenten på interpretasjon, samt ulike obligatoriske prosjekter og/eller seminarer.</p> <p>For studentene på programmet i utøvende musikk – rytmisk foregår samspillundervisningen i faste samspillband og i ulike prosjekter.</p> <p>På slutten av høstsemesteret har samspillbandet en studiosesjon der man spiller inn noe av materialet man har jobbet med dette semesteret.</p> <p>For studentene på programmet i elektronisk musikk er samspillundervisningen organisert slik at studenten kan velge ulike områder som for eksempel DJ, co-writing, studioband, ensemble eller sampling og Live Remix ensemble. Valgbarhet i samspillundervisningen kan variere på grunnlag av studentantall.</p> <p>Deler av undervisningen er obligatorisk. Disse delene fremlegges i Canvas ved semesterstart.</p>	<p>Instrumentalundervisning: 10 timer à 45 minutter pr. semester over to semestre med instrumentallærer.</p> <p>Det blir gjennomført 40 undervisningstimer i samspill fordelt på hele skoleåret.</p> <p>Repertoaret bestemmes av studenter i samarbeid med faglærer.</p> <p>Skolen setter sammen band/ensemble.</p>	<p>Emnet skal bidra til å gi studenten større musikalske og tekniske ferdigheter på sitt hovedinstrument, og øke studentens forståelse for instrumentets egenart og uttrykksmuligheter.</p> <p>Emnet skal bidra til å utvikle både håndverksmessige kvalifikasjoner og kunstnerisk særpreg, samt sette studenten i stand til å kunne fortsette egen kompetanseutvikling og spesialisering på en stadig mer selvstendig basis.</p> <p>Studenter velger i dette emnet mellom hovedinstrumentene vokal, gitar, tangent, bass, trommer, blåseinstrumenter og live elektronikk (strykerinstrumenter vil også kunne vurderes).</p> <p>Emnet er praktisk utøvende.</p> <p>En vesentlig del av emnets arbeidsmengde utgjøres av egenøving på instrumentet. Det gis følgende former for undervisning:</p> <p>individuell undervisning</p> <p>individuell undervisning med ekstern lærer</p> <p>fellestimer / masterclass</p> <p>For studentene på tredje år er det egne fellestimer, på tvers av instrumentgrupper, og noen fellestimer med første og andre år. Dette semesteret gjennomføres også undervisningstimer med en ekstern lærer etter eget ønske/avtale.</p>

(Forts.)

Tabell 3. (Forts.)

Kategori	NLA Høgskolen (NLA) 4MHI300 Hovedinstrument III (2020–21)	Høgskolen Kristiania (HK) MUT300 Hovedinstrument 3 (2020 høst)
<b>Arbeidsomfang</b>	Forventet arbeidsomfang er 27 timer pr. studiepoeng. (405 t)	600 timer Deltakelse i undervisning og veiledning: 48–72 timer Egenøving: 300 timer Forberedelse til fellestimer: 10 timer Gjennomføring av og forberedelse til eksamen: 18–42 timer Anbefalt tidsbruk totalt: 400 timer
<b>Arbeidskrav/ vilkår</b>	Godkjent fremmøte til obligatorisk undervisning. Innlevering av repertoarliste godkjent av hovedinstrumentlærer innen fastsatt frist. 1–3 godkjente samspillkonserter/ prosjekt i løpet av studieåret. 1–3 godkjente obligatoriske oppgaver Godkjent deltakelse på minst 1 interpretasjon i løpet av studieåret. En godkjent ferdighetsprøve på sitt hovedinstrument. Ferdighetsprøven gjennomføres i løpet av emnet. Nærmere informasjon gis i Canvas ved semesterstart.	Følgende arbeidskrav må være godkjent før en kan gå opp til eksamen: Fremmøtekrav: Minimum 80 %. Fremmøte i henhold til retningslinjer for NLA Høgskolen. Arbeidskrav: Tilstrekkelig egenøving til å tilegne seg de ferdigheter som er beskrevet i læringsutbytte. Arbeidskrav: Tilstedeværelse på samspillsundervisning. Arbeidskrav: Være delaktig på en konsert hvert semester med samspillband (varighet 20 min.). Hovedinstrumentrapport.

<b>Kategori</b>	<b>NLA Høgskolen (NLA) Høyskolen Kristiania (HK) MUT300 4MH300 Hovedinstrument Hovedinstrument 3 (2020 høst) III (2020-21)</b>	
<b>Vurderingsuttrykk for arbeidskrav</b>	Karakteruttrykk: Bestått / Ikke bestått	Godkjent / Ikke godkjent
<b>Eksamensform</b>	Konsert på ca. 45 minutter. Studentene vurderes individuelt. Repertoar settes i samråd med hovedinstrumentlærer.	Eksamenskonsert Konsertens innhold bestemmes av studenten selv i samråd med læreren. Konsertens lengde skal være 30 minutter. Repertoarliste for alle tre årene skal medbringes på eksamen. Studenten vurderes i forhold til emnets læringsmål. Vurderingen gjøres av minst én ekstern og to interne sensorer, hvorav normalt én er studentens hovedinstrumentlærer. Eksamenskommisjonen skal gi en muntlig uttalelse om eksamenskonserten til studenten.
<b>Vurderingsuttrykk for konsert</b>	Karakteruttrykk: Bestått / Ikke bestått	Gradering: Nasjonal karakterkala Bestått / Ikke bestått (B/IB)

## Tilbakemelding som element i berekraftig vurdering

Eit interessant funn er at tilbakemelding som eksplisitt omgrep er fråverande i alle emneplanane, trass i at forskning framhevar dette som ein nøkkelfaktor for at summativ vurdering skal fungere som berekraftig (McPherson et al., 2022). Intervjuinformantane bekrefta at vurderings- og tilbakemeldingspraksis var meir omfattande enn det emneplanane skildra. Studentane ved HK fekk både karakter og munnleg tilbakemelding, medan dei to andre institusjonane nytta bestått / ikkje bestått i tillegg til munnleg tilbakemelding. Winstone & Boud (2022) viser korleis distinksjonen mellom vurdering og tilbakemelding har implikasjonar for korleis studentane oppfattar vurderinga som ein del av læringsprosessen. Desse perspektiva vil drøftast nærare i diskusjonsdelen.

Den neste tabellen, tabell 4, er ei samanstilling av læringsutbytteskildringane i emneplanane. Den viser læringsmåla dei ulike institusjonane har sett opp i hovudinstrument, og avsnittet i forkant av tabellen syner eit utval av det eg tolkar som berekraftige element i læringsutbytteskildringane.

## Berekraft i læringsutbytteskildringane

Læringsutbytteskildringane ved dei tre institusjonane inneheld fleire formuleringar som kan tolkast i berekraftig retning. UiA sitt læringsutbytte «innsikt i ulike arbeidssituasjonar innanfor det profesjonelle musikkliv» under «Generell kompetanse» peikar på eit livslangt læringsperspektiv som strekkjer seg utover utdanninga. NLA gjer berekraftig vurderingskompetanse meir eksplisitt ved å nemne vurdering som eit konkret læringsutbytte, både under «Kunnskap» og «Generell kompetanse». Likevel er det uklart om desse formuleringane refererer til formative vurderingshandlingar i prosessen fram mot eksamen, eller om dei også omfattar summative vurderingar. HK sitt læringsutbytte, der studenten skal «Lære å kunne ivareta egen kompetanseutvikling og spesialisering», representerer eit tydeleg berekraftig perspektiv med fokus på livslang læring og heilskapleg tilnærming. Desse nyansane i formuleringane reflekterer institusjonane sine ulike prioriteringar, men viser også at berekraftig vurdering er ei delvis synleg, men ulikt konkretisert, retning i læringsutbytta.

Nyansar i ordbruk og prioriteringane i læringsutbytteskildringane tyder på at dei tre institusjonane vektlegg ulike sider ved musikalsk

kompetanse. NLA legg særleg vekt på samspel og stilforståing, medan UiA prioriterer stilkunnskap og «innsikt i arbeidssituasjonar». HK framhevar instrumentkunnskap og -ferdigheiter og løftar fram refleksjon som ein viktig del av vurderingsgrunnlaget. Der UiA skildrar utfordringar som «håndtere kunstneriske utfordringar» (UiA MUR312), vektlegg NLA det meir personlege ved å framheve at studenten skal ha «utviklet en personlig musikalsk uttrykksform», «dypere identitet» og «større bevissthet» (NLA 4MHI300). HK er den einaste institusjonen som eksplisitt nemner «refleksjon omkring egne valg og arbeidsprosess» (HK MUT300) og integrerer både skriftleg og munnleg som ein del av vurderingsgrunnlaget. Dei varierende formuleringane av læringsutbytta synleggjer korleis dei ulike institusjonane vektlegg ulike aspekt ved berekraft i si vurdering av studentane. Samstundes reflekterer læringsutbytta institusjonsspesifikke tilnærmingar til rytmisk musikkutdanning og tilpassingar av vurderingspraksis i tråd med eigne pedagogiske og kunstnarlege målsetjingar.

**Tabell 4.** Samanstilling av læringsutbytte for hovudinstrument (2020–2022) ved utøvande rytmiske studium: Universitetet i Agder, NLA Høgskolen og Høgskolen Kristiania<sup>2</sup>

Kategori	UiA MUR312 Utøvande hovedområde 4 (høst 2020)	NLA 4MHI300 Hovedinstrument III (2020–21)	HK MUT300 Hovedinstrument 3 (2020 høst)
<b>Kunnskap</b>	<p>Studenten skal</p> <ul style="list-style-type: none"> <li>- Ha bred kjennskap til ulike stiltrekk innenfor rytmiske musikkjanger i forhold til sitt hovedområde</li> </ul>	<p>Studenten (har):</p> <ul style="list-style-type: none"> <li>- kunnskap om planlegging, innstudering og gjennomføring av ulike konserter</li> <li>- bredere kunnskap om fremføring, vurdering og lytting til musikk</li> <li>- har kunnskap om improvisasjons- strategier og teori</li> <li>- bred ensemblekunnskap</li> <li>- har kunnskap om hvordan man kan lage nye versjoner av eksisterende komposisjoner</li> <li>- har kunnskap om hvordan skape et eget uttrykk i et samspill</li> </ul>	<p>Studenten:</p> <ul style="list-style-type: none"> <li>- har bred kunnskap om sitt instruments funksjoner i ulike musikalske sammenhenger</li> <li>- ha god kunnskap om sitt instrument både musikkhistorisk og i samtid</li> <li>- kan ivareta egen kompetanseutvikling og spesialisering</li> </ul>

(Forts.)

2 Læringsutbytta er direkte sitat frå dei respektive læreplanane, og er ikkje bearbeida språkleg frå mi side.

Tabell 4. (Forts.)

Kategori	UiA MUR312 Utøvende hovedområde 4 (høst 2020)	NLA 4MHI300 Hovedinstrument III (2020-21)	HK MUT300 Hovedinstrument 3 (2020 høst)
<b>Ferdigheter</b>	<p>Studenten skal:</p> <ul style="list-style-type: none"> <li>- Ha opparbeidet seg teknisk fleksibilitet innenfor sitt hovedinstrument.</li> <li>- Tilegnet seg et personlig uttrykk i sin musisering og i sitt kreative arbeid</li> <li>- Kunne anvende stilkunnskap knyttet til sitt hovedområde i en ensemblesituasjon innenfor den rytmiske musikktradisjonen</li> </ul>	<p>Studenten:</p> <ul style="list-style-type: none"> <li>- kan et bredt og allsidig repertoar</li> <li>- behersker prima vista-spill</li> <li>- kan improvisere innenfor ulike stilarter</li> <li>- kan være med å prege ulike ensembler</li> <li>- behersker et repertoar bestående av flere sjangere</li> <li>- kan improvisere variasjoner og skape egne løsninger i et samspill</li> </ul>	<p>Studenten:</p> <ul style="list-style-type: none"> <li>- viser god bruk av sitt instrument i ulike stilretninger innenfor populærmusikk</li> <li>- behersker en god og fleksibel instrumentalteknikk</li> <li>- kan anvende en hensiktsmessig øvingsmetodikk, og kunne begrunne og videreutvikle denne</li> <li>- holder seg orientert innen musikalsk utvikling omkring eget instrument</li> <li>- viser personlig uttrykk og egenart</li> <li>- jobber mot et helhetlig sceneuttrykk i en formidlingssituasjon</li> </ul>
<b>Generell kompetanse</b>	<p>Studenten skal:</p> <ul style="list-style-type: none"> <li>- Ha innsikt i ulike arbeidssituasjoner innenfor det profesjonelle musikkliv</li> <li>- Kunne håndtere kunstneriske utfordringer på en kreativ, undersøkende og analytisk måte</li> </ul>	<p>Studenten:</p> <ul style="list-style-type: none"> <li>- har utviklet en dypere identitet som musiker</li> <li>- har utviklet en personlig musikalsk uttryksform</li> <li>- har en større bevissthet omkring fremføring, vurdering og lytting til musikk</li> </ul>	<p>Studenten:</p> <ul style="list-style-type: none"> <li>- kan vise refleksjon omkring egne valg og arbeidsprosess, kommunisere om dette og gjennom dette bidra til utvikling av god praksis</li> <li>- kan kommunisere om estetiske kvaliteter i musikk</li> <li>- kan vise god kreativitet og ferdigheter i improvisasjon</li> <li>- kan holde seg oppdatert i musikkproduksjon med eget instrument</li> <li>- kan vise kjennskap til nytenkning og innovasjonsprosesser basert på kunstnerisk utviklingsarbeid og tradisjon</li> </ul>

## Berekraft i sensorrettleiinga

Sensorrettleiingane for hovudinstrument inneheld detaljerte instruksjonar og informasjon som er retta mot vurderarane. Då kravet om sensorrettleiingar for alle eksamenar i UH-sektoren vart innført i 2017, bygde desse i stor grad på læringsutbytta i emneplanane (Hermansen, 2019). Sjølv om sensorrettleiingar ikkje er eksplisitt nemnde i emneplanane, viser institusjonane sine nettsider ofte til meir utfyllande informasjon om eksamen på studentane sine læringsplattformer, som Canvas. Både læringsutbytte og sensorrettleiingar er generelle når det kjem til sjanger og instrument, med eit unntak for UiA, som har egne retningslinjer for bruk av laptop som instrument. Sensorrettleiinga for UiA presiserer også at eksterne sensorar bør ha same instrumentbakgrunn som instrumentlæraren, noko som peikar på ein instrumentspesifikk praksis. UiA inkluderer dessutan ein teknisk prøve for hovudinstrument, der studentane vert vurderte på sentrale tekniskar og refleksjon kring eigen utøving, noko som kan forståast som eit element av berekraftig vurdering. NLA legg vekt på vurderingsevne som eit kriterium i sensorrettleiinga, men spesifiserer ikkje korleis dette skal operasjonaliserast. HK framhevar derimot den skriftlege refleksjonen som eit grunnlag for studenten si forståing av musikkestetiske kvalitetar.

Tabell 5 oppsummerer dei holistiske og analytiske vurderingskategoriane som vert nytta i sensorrettleiingane ved dei tre institusjonane. Denne inndelinga reflekterer ei todelt tilnærming til vurdering (Vinge, 2011). Dei holistiske kriteria prioriterer det heilskaplege og kunstnarlege uttrykket, medan dei analytiske kriteria vurderer spesifikke utøvar tekniske eigenskapar som til dømes intonasjon, dynamikk og teknikk. Sjølv om denne inndelinga gjev eit nyansert bilete av vurderingsgrunnlaget, manglar det eksplisitte referansar til berekraftige prinsipp som studentmedverknad, transparens og langsiktig læring. Sadler (2015, s. 12) nemner transparens som ein avgjerande faktor for vurderingskompetanse hjå studentar og set omgrepet i samband med føreseieleg vurdering, og byggjer på at vurderinga er reliabel/påliteleg (vurderer det som skal vurderast) og valid/rettferdig (like vurderingsrammer). Her kunne sensorrettleiingane inkludert element som oppmuntra til refleksjon over eigen praksis, eller dialog med sensorpanelet.

Slik sett kan sensorrettleiingane seiast å vere meir orienterte mot å dokumentere prestasjon enn å støtte studentane si framtidige utvikling som skapande musikarar.

**Tabell 5.** Samanstilling av vurderingskategoriar i sensorrettleiingar ved UiA, NLA og HK

Kategori	Sensorrettleiing UiA MUR312 Utøvende hovedområde 4 (høst 2020)	Sensorrettleiing NLA 4MHI300 Hovedinstrument III (2020-21)	Sensorrettleiing HK MUT300 Hovedinstrument 3 (2020 høst)
<b>Holistisk vurderingskategori (kriterium)</b>	Utøve og formidle musikk på høgt nivå Formidling og eigenart	Musisering: Integritet/formidling Vurderingsevne Sjølvtende	Forståing for musikken sine estetiske kvalitetar (med tilvising til kandidatens skriftlege del) Personleg uttrykk, kreativitet
<b>Analytisk vurderingskategori (kriterium)</b>	Stilforståing på eige hovudinstrument Teknikkar og utøvar kontroll på instrumentet Musikalske parameter som time, intonasjon og dynamikk Samspel og produksjon (laptop)	Basisferdigheiter: Tone/klang, teknikk, samspel, time Musisering: Stilforståing, foredrag, frasering, dynamikk, originalitet	Meistring av instrumentet i vald stilretning/sjanger Instrumentalteknikk Instrumentet sin funksjon i sampelsamheng Ferdigheiter i improvisasjon

Tekstane i emneplanar og sensorrettleiingar gjev innsikt i kva institusjonane ynskjer og forventar av studentane, og kva som ligg til grunn for vurdering i konteksten for vokaleksamen. Likevel kan ikkje dokumenta åleine gje eit fullt bilete av praksis, og det er naudsynt å setje tekstinnhaldet i læreplanane saman med praksiserfaring for å forstå «kva som faktisk skjer» (Prior, 2010, s. 96, mi omsetjing). Ein annan viktig faktor er sjølve «gjenstanden» som skal vurderast - den levande musikalske framføringa av studenten. Denne framføringa inngår i studien berre gjennom studentane sine egne opplevingar av henne. Til sist er vurderarane sine oppfatningar og tolkingar av framføringa avgjerande for vurderingspraksisen. Studien kombinerer difor analyser av dokumenta med studentane sine opplevingar av framføringa og vurderarane sine erfaringar, for å oppnå eit heilskapleg bilete av vurderingspraksis for rytmisk vokal. Dei fylgjande avsnitta utforskar korleis studentane og vurderarar oppfattar berekraft i vokal vurderingspraksis i høgare populærmusikkutdanning.

## Funn om berekraft hjå vokalstudentinformantane

Informantane fekk intervjuguidane tilsendt i forkant av intervjuet, og dei fleste hadde førebudd seg, mellom anna på spørsmålet om kjennskap til emneplan og læringsmål for studiet. Under sjølve intervjuet opplyste dei fleste at dei hadde lese emneplanen på førehand, men berre to studentar hadde gjort det før studiestart. Dette kan tyde på at læreplanen i stor grad blir erfart indirekte gjennom undervisninga, og at studentane i praksis orienterer seg etter eigne læringsagendaer, meir eller mindre medvite.

## Vurderinga si rolle i musikkbransjen

Musikkbransjen som kontekst spelar ei viktig rolle for studentane sine vurderingar av vurderingsformene. Fleire vokalstudentar uttrykte frustrasjon over manglande samsvar mellom eigne forventingar og den faktiske vurderingspraksisen, og retta kritiske spørsmål mot relevansen av vurderingsmåten slik den vart gjennomført for rytmiske vokalistar. Kritikken retta seg ikkje berre mot innhaldet i sjølve vurderinga, men også mot opplevinga av at vurderingsforma var forankra i ein klassisk tradisjon som ikkje spegla dei populærmusikalske arbeidsmåtane dei vart utdanna i. Student 8 reflekterte:

Det diskuterte vi faktisk etter eksamen med en del andre i klassen, at vi føler: Okei, de sier vi går på et popstudie, men vi har jo en eksamensform som om vi skal bli vurdert som klassisk skolerte sangere, som føles veldig rart da. Så vi, eller jeg, tenker at den vurderingsformen, er dette det beste for det jeg jobber med på dette studiet? Nemlig å være en popartist? (S8)

Denne typen refleksjon peikar mot eit behov for betre samsvar mellom vurderingsform, innhald i utdanninga og studentane sine framtidige yrkesroller. Det at dei diskuterte dette i klassen, tyder på at denne frustrasjonen ikkje berre var individuell, men ei kollektiv fagleg forståing og frustrasjon. Kritikken ber òg med seg spørsmål om fagleg identitet og kva slags musikalsk uttrykk og profesjonalitet som vert anerkjend gjennom vurderinga. Som Waagen (2021, s. 8) peikar på, har vurderingspraksis

ved eksamenskonsertar i norsk musikkutdanning djupe røter i klassiske musikktradisjonar, noko som kan skape spenningar når vurderinga skal tilpassast det rytmiske feltet.

## Karakterar som vurderingsuttrykk

Studentane hadde ulike haldningar til karakterar som vurderingsuttrykk. Fleire uttrykte at karakterar er ubrukelege i ein musikkbransjekontekst der det er den kunstnarlege prestasjonen som vert vurdert: «Ein A&R-person<sup>3</sup> hadde vore meir læringsrikt og meir relevant i forhold til studiet, fordi det er ingen som skal sjå desse karakterane i etterkant. [...] Viss du skal gå til eit plateselskap, driv du ikkje og slår karakterar i bordet!» (S9). Samstundes viste analysen grupperingar blant studentane: Ei gruppe var ambivalente til karakterar, men takksame for kvalitativ tilbakemelding saman med vurderingsuttrykket bestått / ikkje bestått. Desse hadde høge krav til seg sjølv og understreka at dei ikkje ville ha godteke noko anna enn toppkarakter, og for desse reduserte vurderingsuttrykket Bestått/Ikkje bestått karakterpresset. Den andre gruppa studentar var meir pragmatiske og var nøgde med karakteren; nokre fordi dei fekk toppkarakter, andre fordi dei tenkte at nest beste karakter var ei betre kjelde til motivasjon og vidare utvikling enn toppkarakteren.

## Tilbakemelding som berekraftig praksis

Tilbakemelding vart framheva som ein nøkkelfaktor for berekraftig vurdering. Fleire studentar meinte at medstudentane sine tilbakemeldingar var minst like verdfulle som dei frå sensor: «Eg bryr meg jo mest om kva medelevarane mine seier, eigentleg. Sei, opp mot eksamen, så hadde eg brydd meg nesten meir om kva fem av mine medelevar sa, enn kva den eine sensoren sa» (S8). I tillegg vart lærarane sine tilbakemeldingar, som kunne spegle utvikling over tid, opplevde som meir meningsfulle enn tilbakemeldingane frå sensor, som baserer seg

---

3 Talentspeidar i musikkbransjen. A&R står for Artist & Repertoire.

på ei enkeltoppleving: «Sensor har jo ikkje følgt meg gjennom tre år, men eg fekk ei tilbakemelding frå læraren min på førre interpret, og det synest eg var skikkeleg fint, for hen sa: 'Du har utvikla deg mykje sidan første klasse!', og det synest eg kanskje er meir nyttig» (S13). Slike tilbakemeldingar som inkluderer både fortid og framtid, gjev studenten ei heilskapleg forståing av eiga utvikling og støttar livslang læring. Dette understrekar betydninga av å prioritere kvalitativ tilbakemelding som ein integrert del av vurderingspraksisen framfor å fokusere på karakterar åleine.

## Funn om berekraft hjå vurderarinformantane

Vurderarane framhevar korleis vurderingspraksis i høgare populærmusikkutdanning ikkje alltid samsvarar med realitetane i musikkbransjen. Som studentane kommenterer dei at karakterar er lite relevante som argument i musikkbransjen: «I og med at me har opptaksprøve, så er utøveren vurdert på nivå i utgangspunktet. At ein i dei faga får ein A eller ein D, eller ein B eller ein C på ein eksamenskonsert, så det er ikkje det som gjer om du får jobb i denne bransjen.» (L1). Lærarane ser derimot eksamenskonserten som eit høve til å gje studentane tilbakemeldingar frå ein meir nøytral og bransjenær posisjon, noko som kan hjelpe dei å utvikle seg vidare i sin yrkeskarriere: «Vurdering er fantastisk å få av ein sensor eller ein lærar som ikkje er ute etter å ta deg, men ute etter å hjelpe deg til å utvikle deg og vere klar til å møte bransjen, for der vert ein ikkje teken på med silkehanskar.» (L4). På grunnlag av bransjen vokalstudentane skal ut i, gjev vurderarane som gruppe tydeleg uttrykk for at dei finn karakterar ueigna som vurderingsuttrykk og også problematiske i høve til berekraftig vurdering. Samstundes såg nokre lærarar karakterar som eit mogleg pedagogisk verkemiddel for motivasjon, men dei understreka at karakterbasert vurdering i seg sjølv ikkje støtta studentane si læring på same måte som ei kvalitativ tilbakemelding gjer. Dette kan tyde på eit behov for å vurdere om karakterar faktisk tilfører noko pedagogisk verdi i denne konteksten.

## Eigenvurdering og dialogbasert tilbakemelding

Det kom fram av intervjua at alle dei tre institusjonane praktiserte å gjennomføre munnlege tilbakemeldingar rett etter eksamenskonserten, sjølv om dette ikkje var synleg i emneplanar og sensorrett-leiingar. Fleire av vurderarane fortalde om ulike praksisar på korleis studentane kunne ta med tilbakemeldinga: om dei kunne ta henne opp, notere stikkord, eller om dei fekk henne skriftleg notert i etterkant. Vurderarane framheva at det er viktig å styrkje studentane si evne til eigenvurdering og refleksjon. Sjølv om det ikkje var presisert skriftleg at tilbakemeldinga er dialogbasert, la nokre lærarar vekt på å skape rom for at studentane kunne argumentere for sine kunstnarlege val, til dømes gjennom refleksjonsnotat knytte til eksamenskonserten. Eit døme er ein student som forsvarte sitt val om «dårleg engelskuttale» som eit relevant argument, fordi marknaden han sikta seg inn mot, ville ha det:

Han kunne jo vise til spillejobber og anmeldelser, dette kan ikke være feil når markedet vil ha det! Det er jo kvalitet i det, man kan ikke si at det er uten kvalitet å synge på dårlig engelsk. Det sier noe om personligheten, at du er mer opptatt av å kommunisere enn å kommunisere riktig, istedenfor at man skal være opptatt av å gjøre ting korrekt. (L6)

Slike vurderingssituasjonar kan samanliknast med det Kleiman (2009) kallar forhandlingsvurdering, der studentane får høve til å forklare og forsvare sine kunstnarlege val i dialog med sensorane. Dette dømet tyder på eit potensial for å gjere dialogbaserte vurderingspraksisar meir systematiske. Vurderarane sine perspektiv syner at det finst element av berekraftige vurderingspraksisar i norsk høgare populærmusikkutdanning, men også at desse manglar ein systematisk integrasjon. Vurderarane verdset dialogbasert tilbakemelding og refleksjon, men ser utfordringar i korleis vurderingsrammene, særleg karakterar som vurderingsuttrykk, kan hemme både studentane sin motivasjon og deira evne til å lære av vurderinga. Dette understrekar behovet for å utvikle ein meir heilskapleg strategi for berekraftig vurdering som inkluderer både eigenvurdering, kvarandrevurdering og profesjonelle refleksjonsprosessar.

## Diskusjon: I kva grad er vurderingspraksis ved eksamenskonserten berekraftig?

### Emneplanar og sensorretteiingar: eit fragmentert bilete

Analysen av emneplanar og sensorretteiingar i norsk høgare populærmusikkutdanning avdekkjer variasjonar i korleis berekraftige prinsipp, som eigenvurdering, kvarandrevurdering og refleksjon, er integrerte. Høgskolen Kristiania (HK) vektlegg refleksjon og sjølvstendig kompetanseutvikling i læringsutbytta sine, medan Universitetet i Agder (UiA) prioriterer eit profesjonelt læringsperspektiv. NLA Høgskolen (NLA) fokuserer på vurdering som ein kompetanse i seg sjølv. Likevel manglar alle tre institusjonane ein heilskapleg strategi for å operasjonalisere studentmedverknad i summative vurderingar.

Sensorretteiingane reflekterer liknande utfordringar. Dei nyttar overordna omgrep som fangar heilskapen i framføringa, men manglar spesifikke retningslinjer for korleis instrumentspesifikke element skal vurderast. Dette kan skape uvisse både for sensorar og studentar med tanke på tilbakemelding, særleg dersom studentane ikkje har lese sensorretteiinga på førehand. Forsking syner at transparente vurderingsrammer og tydelege kriterium fremjar læring og tillit (Boud, 2000; Higgins et al., 2001; Sadler, 2009), men effekta av dette vert redusert dersom studentane ikkje har tilstrekkeleg vurderingskunnskap.

Når det gjeld vurdering av musikkframføring, foreslår fleire forskarar instrumentspesifikke sensorretteiingar som eit tiltak for å styrkje transparens (Gynnild, 2010; Walton, 2020; Waagen, 2021). I eit intervju med Geraldine O'Neill, presiserer Boud at eigenvurdering og kvarandrevurdering ikkje oppstår automatisk hjå studentar, men må utviklast gjennom systematisk trening i læringsprosessen (O'Neill & Boud, 2023, 15:49–16:10). Mangelen på rammeverk for systematisk læring av vurdering kan indikere at studentmedverknad i vurderingsprosessen framleis er lite utvikla i norsk høgare populærmusikkutdanning. Sjølv om NLA inkluderer vurdering i læringsutbytta og HK nyttar refleksjonsnotat som ein del av vurderinga, er det behov for ei meir gjennomgripande tilnærming til studentmedverknad i summative vurderingar.

## Studentane si rolle: frå passive mottakarar til aktive deltakarar

Vokalstudentane si oppleving av vurdering viser eit klart potensial for å innføre berekraftig vurdering i samband med eksamenskonserten. Studentane framheva verdien av å delta aktivt i vurderingsprosessen, særleg gjennom tilbakemeldingar frå medstudentar og lærarar som kjente utviklingsprosessen deira. Intervjua avdekte at studentane hadde erfaring med eigenvurdering og kvarandrevurdering frå undervisninga, men desse vurderingsformene syntest ikkje å vere ein del av praksisen etter eksamenskonserten. Falchikov og Bouds (2007) teoriar understrekar at studentar treng både medvit om vurdering og systematisk øving for å utvikle vurderingskompetanse, og i tråd med dette peikar denne studien på ein skilnad mellom undervisningspraksis og summative vurderingar. Nokre studentar framheva munnleg tilbakemelding etter eksamenskonserten som den mest verdifulle kommunikasjonsforma for tilbakemelding, fordi den opna for dialog og høve til å stille spørsmål. Dette kan indikere eit behov for større studentmedverknad i eksamenskonserten som summativ vurderingspraksis, slik at studentane i større grad får bidra til å definere kva som skal vurderast, og korleis, i staden for å vere passive mottakarar av informasjon.

Vidare stilte studentane spørsmål ved relevansen av den summative vurderingsforma for rytmisk vokal. Nokre meinte at eksamenskonserten var prega av tradisjonar frå klassisk musikkutdanning, noko dei oppfatta som lite tilpassa deira sjangertilnærming og yrkesrelevans. Karakterar på vokaleksamenen vart heller ikkje oppfatta som autentisk for studentane. Denne typen refleksjon viser at studentane etterspør ei vurderingsform som er meir tilpassa deira behov og framtidige karrierar, noko som kan knytast til diskusjon kring autentisk vurdering (Reid & Ballantyne, 2019). Dette peikar på eit behov for å utvikle vurderingsformer som ikkje berre speglar studentane sine faglege mål, men også musikkbransjen sine krav og forventingar.

## Vurderarane si rolle: nøkkelen til berekraftig vurdering?

På grunnlag av intervjua med vurderarane kom det fram at alle institusjonane praktiserte ein tilbakemeldingssamtale med studentane etter

eksamenskonserten. Vurderarane understreka at karakterar som vurderingsuttrykk ikkje er berekraftige i musikkbransjen, men kan fungere som ein pedagogisk motivasjon eller vere eit administrativt krav. Eit større medvit hjå både vurderarar og studentar kring distinksjonen mellom vurdering (tilbakeskodande) og tilbakemelding (framoverskodande) (Winstone & Boud, 2022) kunne støtta utviklinga av ei meir dialogbasert vurderingssamtale, der fokus vert flytta frå ei einseitig tilbakemelding til ei gjensidig utveksling. Diskusjonen om karakter versus bestått / ikkje bestått engasjerte vurderarane sterkt fordi dei meinte at tilbakemeldinga hadde størst verdi for studentane, medan mange studentar sette sin verdi i å få toppkarakter. Frustrasjonen over karakterar som vurderingsuttrykk kan også vere eit uttrykk for at det mangla ein tillit mellom sensor og studentar i tilbakemeldingssituasjonen. Samsvaret mellom undervisning, læreplanar og sensorrettlegg er med på å fremje tilliten til vurderinga, men dette samsvaret er noko som må førebuast, først og fremst av lærarane.

Fleire av vurderarane peika på refleksjon som ein avgjerande komponent for berekraftig vurdering. Dette er i tråd med Hughes' (2015, s. 255) teoriar om å implementere systematisk refleksjon og eigenvurdering for rytmisk vokal læring, også i samband med summative vurderingar. Ein overgang frå einvegs til dialogbasert tilbakemelding krev ikkje berre ressursar, men endring i vurderingskultur. Med sin praksis om å ha med eit refleksjonsskriv har Høyskolen Kristiania (HK) eit høve til å utvikle og styrkje involveringa av studentane på ein refleksiv måte, slik til dømes Partti med kollegaer (2015) ser deltakande vurdering som eit høve til å diskutere framføringa i etterkant og grundig diskutere kvifor alle kunstnarlege val er gjorde. Her vert det heller ikkje gjeve ein karakter, men både studentar, medstudentar og sensor får høve til ein fagleg diskusjon på grunnlag av eksamenskonserten (Partti et al., 2015, s. 483–484).

Det kom fram kritiske kommentarar knytta til vurderingsuttrykket ved eksamenskonserten. Medan vurderarar ved institusjonen som nytta karakter uttrykte ønske om endring, vart det ytra takksemd frå dei som hadde gått over til karakteruttrykket Bestått/Ikkje bestått. Det var ingen av aktørane som eksplisitt ytra ynske om endringar i vurderingsprosedyrane ved den summative eksamenskonserten, noko som kan tyde at både

vurderarar og studentar opplevde vurderingspraksisen som tilfredsstillende. Det er lite i dei formelle emneplanane eller sensorretteiingane som foreslår at studentane skal medverke i vurderingspraksis ved den avsluttande vokaleksamenen. Skal berekraftig vurdering verte ein realitet i norsk høgare populærmusikkutdanning, impliserer dette at ein berekraftig vurderingspraksis og tilbakemelding må inn som uttalte punkt i læreplanane. Først då kan berekraftig vurdering effektuerast. Å integrere studentmedverknad i summative vurderingar vil ikkje berre støtte læring i utdanninga, men også gjere studentane betre rusta til å møte krava frå ein dynamisk musikkbransje.

## Oppsummering og konklusjon

Denne artikkelen har som mål å belyse to sentrale spørsmål knytte til berekraftig vurdering ved eksamenskonserten i høgare populærmusikkutdanning. Det første spørsmålet handlar om korleis berekraft er integrert i vurderingsgrunnlaget for tre norske rytmiske vokalutdanningar. Analysen viser at berekraft, i tydinga studentaktiv vurdering, er relativt svakt representert i læreplanane. Nokre læreplanar inkluderer element som sjølvrefleksjon og vurderingskompetanse, men det manglar tydelege krav til at eksterne sensorar vurderer desse aspekta. Dette avgrensar kor synleg og operativ studentmedverknad er i praksis. Intervjufunna framhevar at både studentar og lærarar verdset dialogbaserte og framoverskodande vurderingsformer, men at desse formene er tilfeldig integrerte i den summative eksamenspraksisen. Samla sett syner analysen at vurderingspraksisar i samanheng med eksamenskonserten inneheld berekraftige element, men at det manglar ein heilskapleg strategi for å operasjonalisere prinsippa for studentmedverknad i summativ vurdering. Trass i at vurderarar og studentar hadde ulike perspektiv på vurderingspraksis, og at fleire informantar uttrykte kritiske haldningar til mangelen på berekraftige vurderingselement, framstod den summative eksamenskonserten likevel som ein katalysator for læring og ein sterk pedagogisk motivasjonsfaktor for både studentar og lærarar. Dette understrekar behovet for å styrkje og integrere berekraftige og studentmedverkande vurderingspraksisar, slik at eksamenskonserten ikkje berre støttar læring

i høgare populærmusikkutdanning, men også bidreg til å utruste studentane for livslang læring og framtidig yrkespraksis.<sup>4</sup>

## Forfatterbiografi

**Bodil Kvernenes Nørsett** er høgskulelektor i musikkfag ved Ansgar høyskole og PhD-stipendiat ved avdeling for Populærmusikkforskning, Institutt for Kunstfag ved Universitetet i Agder i Kristiansand. Hennar forskningsinteresser er mellom anna vurdering og vurderingskompetanse i musikkutdanning, stemmeforskning, rytmisk vokalpedagogikk og kor- og ensembleleiing. I tillegg er ho aktiv utøvar som kordirigent og songar.

## Referanser

- Arlov, T. B. (2016). *Kvalitet i kunstnerisk virksomhet* (Rapport). NTNU 2016. [https://www.ntnu.no/documents/1262755726/1268954456/Kvalitet\\_i\\_kunstnerisk+virksomhet\\_endelig\\_april-2016.pdf/495635d1-91e2-4601-8cfb-0ecea1692f4](https://www.ntnu.no/documents/1262755726/1268954456/Kvalitet_i_kunstnerisk+virksomhet_endelig_april-2016.pdf/495635d1-91e2-4601-8cfb-0ecea1692f4)
- Bartleet, B.-L., Ballico, C., Bennett, D., Bridgstock, R., Draper, P., Tomlinson, V. & Harrison, S. (2019). Building sustainable portfolio careers in music: Insights and implications for higher education. *Music Education Research*, 21(3), 282–294. <https://doi.org/10.1080/14613808.2019.1598348>
- Benson, E. A. (2020). *Training contemporary commercial singers*. Compton Publishing Ltd.
- Biggs, J. & Tang, C. (2015). Constructive alignment: An outcomes-based approach to teaching anatomy. I L. K. Chan & W. Pawlina (Red.), *Teaching anatomy: A practical guide* (s. 31–38). Springer International Publishing. [https://doi.org/10.1007/978-3-319-08930-0\\_4](https://doi.org/10.1007/978-3-319-08930-0_4)
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151–167. <https://doi.org/10.1080/713695728>
- Boud, D. (Regissør). (2022, 24. januar). *David Boud: Purposes of assessment* [Videoopptak]. YouTube. <https://www.youtube.com/watch?v=r-RmxgmaxEM>

---

4 I arbeidet med denne artikkelen, vart ChatGPT, eit KI-basert verktøy, nytta som språkleg støttepelar for å sikre klar, korrekt og transparent tekst (ChatGPT, 2024).

- Boud, D. & Soler, R. (2016). Sustainable assessment revisited. *Assessment & evaluation in higher education*, 41(3), 400–413. <https://doi.org/10.1080/02602938.2015.1018133>
- Bowen, G. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9, 27–40. <https://doi.org/10.3316/QRJ0902027>
- Braun, V. & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.
- Brophy, T. S., Krieger, C., Leite, M. C., McCaffrey, M., Palmer Wolf, D. & Poulin, J. M. (2021, 17. november). The guiding principles for the assessment of arts learning. *CREATIVE GENERATION*. World Alliance for Arts Education's Executive Forum, 2021 Virtual Summit. <https://creative-generation.org/publications/the-guiding-principles-for-the-assessment-of-arts-learning>
- ChatGPT. (2024). *OpenAI* (Versjon 4.0) [Programvare].
- Coimbra, D., Davidson, J. & Kokotsaki, D. (2001). Investigating the assessment of singers in a music college setting: The students' perspective. *Research Studies in Music Education*, 16(1), 15–32. <https://doi.org/10.1177/1321103X010160010301>
- Danielsen, A. (2002). Estetiske perspektiver på populærmusikk. I J. Gripsrud, *Populærmusikken i kulturpolitikken*: Vol. nr 30 (s. 129–155). Norsk kulturråd.
- Elliott, D. J., Silverman, M. & McPherson, G. E. (Red.). (2019). *The Oxford handbook of philosophical and qualitative assessment in music education*. Oxford University Press.
- Falchikov, N. & Boud, D. (2007). *The role of assessment in preparing for lifelong learning: Problems and challenges*. Deakin University. <https://doi.org/10.4324/9780203942185>
- Fastré, G. M. J. (2011). *Improving sustainable assessment skills in vocational education*. [Doktoravhandling, Open Universiteit].
- Fjortoft, H. & Sandvik, L. V. (2020). *Vurderingskompetanse i skolen: Praksis, læring og utvikling* (1. utg.). Universitetsforlaget. [https://urn.nb.no/URN:NBN:no-nb\\_pliktmonografi\\_000020870](https://urn.nb.no/URN:NBN:no-nb_pliktmonografi_000020870)
- Gynnild, V. (2010). *Kriterier og skjønn i evaluering: En kasusstudie i utøvende musikkutdanning*. NOKUT – Nasjonalt organ for kvalitet i utdanningen.
- Hallam, S. (2019). The influence of assessment on learning and teaching: Using assessment to enhance learning. I D. J. Elliott, M. Silverman & G. E. McPherson (Red.), *The oxford handbook of philosophical and qualitative assessment in music education* (s.167-185) Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190265182.013.5>
- Hansen, G. (2022). *Use of feedback in a higher education learning context. "It's not like it's too late to learn. And this is an important message, you can learn even if you have misunderstood"* [Doktoravhandling, NTNU]. <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3005448>
- Hermansen, H. (2019). *Utforming av og vurderingskriterier og sensorveiledninger – sentrale spørsmål og avveininger*. (F-LINK-serien), Universitetet i Oslo.

- Higgins, R., Hartley, P. & Skelton, A. (2001). Getting the message across: The problem of communicating assessment feedback. *Teaching in Higher Education*, 6(2), 269–274. <https://doi.org/10.1080/13562510120045230>
- Hughes, D. (2010). Developing vocal artistry in popular culture musics. I S. D. Harrison (Red.), *Perspectives on teaching singing* (s. 244–258). Australian Academic Press.
- Hughes, D. (2013). An encultured identity: Individuality, expressivity and the singing-self. *Australian Voice*, 15, 13–19. <https://search.informit.org/doi/10.3316/informit.237047104572417>
- Hughes, D. (2014). contemporary vocal artistry in popular culture musics: Perceptions, observations and lived experiences. I S. D. Harrison & J. O'Bryan (Red.), *Teaching singing in the 21st century* (Bd. 14, s. 287–301). Springer Netherlands. [https://doi.org/10.1007/978-94-017-8851-9\\_17](https://doi.org/10.1007/978-94-017-8851-9_17)
- Hughes, D. (2015). Assessment and feedback in curricula design for contemporary vocal studies. I D. Lebler, G. Carey & S. D. Harrison (Red.), *Assessment in music education: From policy to practice* (s. 251–268). Springer International Publishing. [https://doi.org/10.1007/978-3-319-10274-0\\_17](https://doi.org/10.1007/978-3-319-10274-0_17)
- Hughes, D. (2020). Transformational insights and the singing-self: investigating reflection and reflexivity in vocal and musical group learning. I J. Encarnacao, & D. Blom (Eds.), *Teaching and evaluating music performance at university: beyond the conservatory model* (s. 52–64). Routledge, Taylor and Francis Group.
- Haakstad, J. (2011). Læringsutbytte: Begrepets anvendelighet i kvalitetsvurdering av høyere utdanning. *Uniped*, 34(4), 72–81. <https://doi.org/10.18261/ISSN1893-8981-2011-04-06>
- Isbell, D. S. (2022). Curriculum and assessment in popular music education. *Journal of Popular Music Education*, 6, (s.349–366). [https://doi.org/10.1386/jpme\\_00096\\_1](https://doi.org/10.1386/jpme_00096_1)
- Juntunen, M.-L., Arlin, E. P. & Liira, K. (2023). Expression in popular music singing as embodied and interpersonal. *Frontiers in Education*, 8. <https://doi.org/10.3389/educ.2023.1092736>
- Keskinen, K. A., Juntunen, M.-L., & Nerland, M. (2024). Expanding professionalism in popular music voice teaching: A framework synthesis. *Research Studies in Music Education*, 46(3), 420–455. <https://doi.org/10.1177/1321103X231223414> (Original work published 2024)
- Kleiman, P. (2007). *Beyond the tingle factor: Creativity and assessment in higher education*. [https://www.academia.edu/64204472/Beyond\\_the\\_Tingle\\_Factor\\_creativity\\_and\\_assessment\\_in\\_higher\\_education](https://www.academia.edu/64204472/Beyond_the_Tingle_Factor_creativity_and_assessment_in_higher_education)
- Kleiman, P. (2009). *Negotiating Assessment, an approach to assessing practical work, including assessment criteria*. [https://www.academia.edu/1508971/Negotiating\\_Assessment\\_2009\\_](https://www.academia.edu/1508971/Negotiating_Assessment_2009_)

- Kristiania (2022). *Bachelor i populærmusikk* [Studiemodell]. <https://www.kristiania.no/studier/bachelor/populærmusikk/>
- Kristiansen, C. W. (2012). Vurderingens kunst: En diskusjon om vurdering som læring i dramafaget [Masteroppgave, Universitetet i Agder]. <https://uia.brage.unit.no/uia-xmlui/handle/11250/139186>
- Lauvas, P., Havnes, A. & Raaheim, A. (2000). Why this inertia in the development of better assessment methods? *Quality in Higher Education*, 6(1), 91–100. <https://doi.org/10.1080/13538320050001090>
- Lemon-McMahon, B. (2022). *The singing voice and vocal identity: Investigating the perceptions and beliefs of singers and teachers* [Doktorgradsavhandling, Macquarie University]. <https://doi.org/10.25949/19438148.v1>
- Marton, F. (1981). Phenomenography- describing conceptions of the world around us. *Intructional Science* 10, (s.177–200). <https://doi.org/10.1007/BF00132516>
- Marton, F. & Pong, W. Y. (2005). On the unit of description in phenomenography. *Higher Education Research and Development*, 24(4), 335–348. <https://doi.org/10.1080/07294360500284706>
- McPherson, G., Blackwell, J. & Hattie, J. (2022). Feedback in music performance teaching. *Frontiers in Psychology*, 13. <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.891025>
- McPherson, G. & Schubert, E. (2004). Measuring performance enhancement in music. I A. Williamson (Red.), *Musical excellence. Strategies and techniques to enhance performance* (s. 61–82). Oxford University Press. [https://www.academia.edu/30238476/McPherson\\_and\\_Schubert\\_2004\\_Performance\\_Enhancement\\_pdf](https://www.academia.edu/30238476/McPherson_and_Schubert_2004_Performance_Enhancement_pdf)
- Mesiä, S. (2019). Developing expertise of popular music and jazz vocal pedagogy through professional conversations: A collaborative project among teachers in higher music education in the Nordic countries. [Doktoravhandling, Sibelius Academy of the University of the Arts Helsinki]. <https://taju.uniarts.fi/handle/10024/6748>
- Nikolaisen, T. (2020). *Hvordan vurderer man egentlig i musikkfaget? En kvalitativ studie om ungdomsskolelæreres vurderingspraksiser i musikkfaget*. [Masteroppgave, UiT Norges arktiske universitet]. <https://munin.uit.no/handle/10037/18841>
- NLA Høgskolen Studietilbud, O. (2021, 31. desember). *4MHI200 Hovedinstrument II 2020–2021*. <https://www.nla.no/studietilbud/emner/2020/4mhi200/>
- NOKUT. (2023). *Læringsutbyte og Nasjonalt kvalifikasjonsrammeverk for livslang læring*. NOKUT. <https://www.nokut.no/fagskule--hogare-yrkesfagleg-utdanning/regelverk-for-fagskoler/laringsutbyte-og-nasjonalt-kvalifikasjonsrammeverk-for-livslang-laring/>
- O'Neill, G. & Boud, D. (Regissørar). (15. Mars 2023). *Some critical issues in enabling impactful assessment* [Video]. YouTube. <https://www.youtube.com/watch?v=dZoDokRGlaU>

- Partti, H., Westerlund, H. & Lebler, D. (2015). Participatory assessment and the construction of professional identity in folk and popular music programs in Finnish and Australian music universities. *International Journal of Music Education*, 33(4), 476–490. <https://doi.org/10.1177/0255761415584299>
- Powell, B. J. (2011). *Popular music ensembles in post-secondary contexts: A case study of two college music ensembles* [Doktoravhandling, Boston University]. <https://search.proquest.com/openview/e872bff6d8752cdbb9d97c5f1e69aod/1?pq-origsite=gscholar&cbl=18750>
- Powell, B. & Smith, G. D. (2019). Philosophy of assessment in popular music education. I D. J. Elliott, M. Silverman, & G. E. McPherson (Red.), *The Oxford handbook of philosophical and qualitative assessment in music education* (s. 343-363). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190265182.013.32>
- Powell, B., Smith, G. D., West, C. & Kratus, J. (2019). Popular music education: A call to action. *Music Educators Journal*, 106(1), 21–24. <https://doi.org/10.1177/0027432119861528>
- Prior, L. (2010). Using documents in social research. I D. Silverman (Red.), *Qualitative Research: Issues of Theory, Method and Practice* (3.utg.). SAGE Publications Ltd.
- Reid, A. & Ballantyne, J. (2019). Building a Culture of Ethical, Comparable, Authentic Assessment: Music education in Queensland. I D. J. Elliott, M. Silverman & G. E. McPherson (Red.), *The Oxford handbook of philosophical and qualitative assessment in music education* (s. 464-481). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190265182.013.35>
- Reinhert, K. (2018). Developing Popular Music Programs in Higher Education: Exploring Possibilities. I *ProQuest Dissertations and Theses*. [Doktorgradsavhandling, University of Miami]. <https://www.proquest.com/docview/2051302719/abstract/404AF7571DE4A5CPQ/1>
- RUM. (2018). *Rapport om sensorveiledninger i høyere utdanning. Rapport fra en arbeidsgruppe oppnevnt av Rådet for utøvende musikkutdanning, desember 2018*. <https://nmh-nettsted.s3.amazonaws.com/files/Studier/Sensorveiledninger-i-h%C3%B8yere-musikkutdanning.pdf> (25.06.2025)
- Rønningen, L. (2018). Vurdering av praktiske eksamener i drama og teater. I V. Aune og C. Haagenen (Red.), *Teaterproduksjon: Ti produksjonsetetiske innganger* (s. 243–262). Cappelen Damm Forskning. <https://doi.org/10.23865/noasp.43.ch10>
- Raaheim, A. (2019). *Eksamensrevolusjonen: Råd og tips om eksamen og alternative vurderingsformer* (2. utg.). Gyldendal.
- Sadler, R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159–179. <https://doi.org/10.1080/02602930801956059>

- Sadler, R. (2015). Backwards assessment explanations: Implications for teaching and assessment practice. I D. Lebler, G. Carey & S. D. Harrison (Red.), *Assessment in music education: From policy to practice* (s. 9–19). Springer International Publishing. [https://doi.org/10.1007/978-3-319-10274-0\\_2](https://doi.org/10.1007/978-3-319-10274-0_2)
- Schei, T. B. (2007). *Vokal identitet: En diskursteoretisk analyse av profesjonelle sangeres identitetsdannelse* [Doktorgradsavhandling, Universitetet i Bergen]. <http://hdl.handle.net/1956/2549>
- Smart, J. & Dixon, S. (2002). The discourse of assessment: Language and value in the assessment of group practice in the performing arts. *Arts and Humanities in Higher Education*, 1(2), 185–204. <https://doi.org/10.1177/1474022202001002005>
- Thagaard, T. (2018). *Systematikk og innlevelse: En innføring i kvalitative metoder* (5. utg.). Fagbokforlaget.
- Universitetet i Agder. (2023). *Våre studier*. Studieplaner for utøvende musikk, rytmisk, bachelor, Universitetet i Agder (25.06.2025) <https://www.uia.no/studier/program/utovende-musikk-rytmisk-bachelor/studieplaner/>
- Vinge, J. (2011). Analytiske og holistiske tilganger til vurdering – en teoretisk drøfting av vurderingsstrategier og vurderingshandlinger i musikkfaget. I S.-E. Holgersen, S. Graabræk Nielsen og L. Väkevä (Red.) *Nordic Research in Music Education. Yearbook: Vol. 13* (s. 199–220). Norges musikkhøgskole.
- von Germeten, G. (2022). We are also music lovers: Testing vocal tastes in higher musical theater education. *Research Studies in Music Education*, 44(3), 554–569. <https://doi.org/10.1177/1321103X221081787>
- Walton, J. (2020). *Making the grade: Theorising musical performance assessment* [Doktoravhandling, Griffith University]. [https://www.academia.edu/43255189/Making\\_the\\_Grade\\_Theorising\\_Musical\\_Performance\\_Assessment](https://www.academia.edu/43255189/Making_the_Grade_Theorising_Musical_Performance_Assessment)
- Winstone, N. E. & Boud, D. (2022). The need to disentangle assessment and feedback in higher education. *Studies in Higher Education*, 47(3), 656–667. <https://doi.org/10.1080/03075079.2020.1779687>
- Waagen, W. (2021). Vurderingskriterier i utøvende musikk – et verktøy for pålitelig vurdering av studentenes bachelorkonsert? I E. Angelo, J. Knigge, M. Sæther & W. Waagen (Red.), *Higher education as context for music pedagogy research* (s. 41–64). Cappelen Damm Akademisk. <https://doi.org/10.23865/noasp.119.ch2>

## KAPITTEL 4

# Elevmedvirkning i praksisveiledning. Et dialogisk perspektiv på tilbakemeldingsprosesser – en pilotstudie

*Hilde Synnøve Blix*

UiT Norges arktiske universitet

**Abstract**

An essential part of music teacher education is supervised teaching practice. In instrumental teaching, this practice usually comprises a student teaching one or more pupils on their instrument, a schoolteacher, and/or a practice supervisor that gives feedback to the student. The rationale for this format is that supervised practice is part of a socialization process where future teachers must acquire knowledge, norms, and values as a basis for their work as teachers. In this article, pupil participation in feedback in the supervision of instrumental teacher students is explored, based on theories of dialogic pedagogy and philosophy. The starting point for the discussion is a pilot study where we explored the learning potential of including the pupil's voice in the supervised practice in a music school. In the academic year 2021/2022, the study was carried out where an instrumental teacher student received feedback on his teaching from both the supervisor and the pupil being taught. The supervised practice comprised a session where all three participants watched a recording of the teaching situation, and the pupil participated in the discussion about, and the analysis of, the content and execution of the lesson. The results are discussed in light of theories about dialogicity in art educational practices understood through, among others, Mikhail Bakhtin, Hartmut Rosa and Anne Bogart. The study shows that there are new possibilities for dialogic processes in the development of professional competence as an instrumental teacher by involving the pupils in the supervised teaching practice.

**Keywords:** student participation, feedback, practicum supervision, dialogism, resonance, instrumental teaching

Sitering: Blix, H. S. (2025). Elevmedvirkning i praksisveiledning. Et dialogisk perspektiv på tilbakemeldingsprosesser – en pilotstudie. I H. S. Blix, K. J. Valkoinen & B. A. Haugseth (Red.), *MusPed:Research: Vol. 11. Vurderingsformer og tilbakemelding i kunstfaglig utdanning* (s. 101–127). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.269.ch4>  
Lisens: CC BY 4.0

## Introduksjon

En sentral arbeidsform i utdanningen av kommende instrumentaledagoger er veiledet undervisningspraksis på kulturskoler, musikklinjer og tilsvarende undervisningsarenaer. Praksisveiledningen foregår vanligvis ved at studenten underviser en eller flere elever på sitt instrument, og en faglærer<sup>1</sup> (øvingslærer) fra skolen, og/eller en praksisveileder fra studentens utdanningsinstitusjon, kommer med tilbakemeldinger på studentens undervisning (Aglen, 2021; Bjørndal, 2008; Skagen, 2013). Ved mange lærerutdanningsinstitusjoner samarbeider faglærer og veileder fra studiestedet om praksissamtalene med studentene, ofte kalt treparts-samtaler (Gloppen, 2013). Denne arbeidsformen har lange tradisjoner i lærerutdanningene i Norge og er bygget på en forståelse av at veiledet praksis er en viktig del av en sosialiseringssprosess hvor kommende lærere skal utvikle kunnskaper, normer og verdier som grunnlag for sitt lærervirke (Lejonberg, 2016, s. 6). Tilbakemeldinger på undervisningspraksis involverer sjeldnere *elev(e)* som er blitt undervist (Pettersen & Lejonberg, 2022). I det følgende er det tilbakemelding i praksissamtalen som står i sentrum; en tilbakemeldingsform hvor den underviste eleven medvirker.

Det er mange former for tilbakemelding knyttet til i denne typen læringsarenaer, både eksplisitte faglige vurderinger og mer indirekte vurderinger, og også praktiske demonstrasjoner. Vi kan snakke om flere forskjelligartede *dialoger* som konstituerer veiledningspraksisene. I denne artikkelen skal vi se på tilbakemelding som *faglig dialogisk tilbakemeldingsinteraksjon* hvor målet med tilbakemeldingene er å ta utgangspunkt i de lærendes ståsted og samarbeidet om å vurdere veien videre for læringsarbeidet (Gamlem, 2019, s. 114). Jeg er her spesielt opptatt av *elev(e)* som en relevant og interessant samtalepartner og ressurs i praksissamtaler (Pettersen & Lejonberg, 2022).

I utdanningen av musikkpedagoger ønsker vi at våre studenter skal være gode på elevmedvirkning og elevaktive læringsformer i sin undervisning.

---

<sup>1</sup> Begrepsbruken knyttet til veiledere fra praksisstedet og veiledere fra utdanningssted er relativt uklart i forskningen på yrkespraksis for kommende lærere, og jeg har her valgt å kalle veilederen fra praksisstedet for faglærer og veilederen fra studiestedet for praksisveileder.

Det krever at vi som driver lærerutdanning også må praktisere en slik pedagogikk. Flere internasjonale studier viser at musikkfagets instrumentalundervisning fremdeles er preget av lærerstyrt undervisning med en form for mesterlæremodell som ideal (Blix, 2012; Harrison & Grant, 2015; West & Rostwall, 2003). Elevens medvirkning i undervisningskonteksten, særlig i en-til-en-undervisning, er ofte begrenset til å ha innspill på musikk de ønsker å spille. I den norske opplæringsloven (2023) framholdes at elevene «skal ha medansvar og rett til medverknad» (§ 1-3), noe som også følges opp av planverket for kulturskolene (Norsk kulturskoleråd, 2016).

Med dette som bakgrunn gjennomførte jeg i samarbeid med en instrumentallærerstudent og hans messingelev i 2022 et forprosjekt hvor vi prøvde ut en modell for elevmedvirkning i veiledning med musikk lærerstudenter. Utgangspunktet for prosjektet var et ønske om å inkludere elevens opplevelser og unike kunnskaper i tilbakemeldinger til lærerstudenter som mottar praksisveiledning i instrumentalundervisning. Jeg tar i denne artikkelen utgangspunkt i følgende problemstilling: *Hva karakteriserer de dialogiske prosessene som oppstår når eleven medvirker i praksissamtalen som tilbakemelding på undervisningspraksis, og hva slags læringspotensial for lærerstudenten kan elevens medvirkning bidra til?* I drøftingen av dette problemområdet vil jeg bruke eksempler fra pilotstudien gjennomført ved UiT Musikkonservatoriet. Studien vil fungere som illustrasjon og omdreiningspunkt i denne undersøkelsen av dialog, resonans og demokrati i tilbakemelding og veiledning.

## Elevmedvirkning

I plan for grunnopplæringen som gjelder for grunn- og videregående skole slås det fast i *Overordnet del* at:

Elevmedvirkning må prege skolens praksis. Elevene skal både medvirke og ta medansvar i læringsfellesskapet som de skaper sammen med lærerne hver dag. Elever tenker, erfarer og lærer i samspill med andre gjennom læringsprosesser, kommunikasjon og samarbeid. (Kunnskapsdepartementet, 2017, s. 15)

Demokrati og medvirkning fremheves som sentrale verdier i læreplanverket Kunnskapsløftet (LK20), og har som mål å utvikle elevenes

kompetanse som ansvarlige samfunnsborgere. Med dette vektlegges elevens rett til å bli hørt og respektert. De siste års elevundersøkelser fra Utdanningsdirektoratet viser at elever i norsk skole mener at de i liten grad får være med og foreslå hvordan de skal jobbe med fagene, og i enda mindre grad får delta i vurderingsprosesser (Utdanningsdirektoratet, 2021). På tross av at føringene i planverket er nokså konkrete, både når det gjelder hva elevene skal kunne medvirke i og på hvilke måter, ser det ut til å være utfordrende for skolen å utøve reell medvirkning på alle skolens arenaer (Mikalsen et al., 2013). Det å utøve musikk bringer med seg demokratiske kompetanser som samhandling, lytting og det å uttrykke tanker, meninger og følelser. Dette er noe framtidens musikk lærere må ha et bevisst forhold til.

Kulturskolen representerer en viktig praksisarena og fremtidig arbeidsplass for musikk lærerstudenter. I *Rammeplan for kulturskolen. Mangfold og fordypning* (Norsk kulturskoleråd, 2016) understrekes det at elevmedvirkning skal stå sentralt både i undervisning, vurdering og skolens organisering. I rammeplanen kan vi lese om elevmedvirkning flere steder, eksempelvis:

Elevmedvirkning og motivasjon henger sammen. Gjennom dialog mellom eleven og læreren, der elevene får være med å legge planer, utvikle evne til egen-vurdering, si sin mening om miljøet på kulturskolen, undervisninga og hvordan de lærer, får eleven innflytelse på egen læringsprosess. [...] Vurdering for læring kan ha form som utviklingssamtaler med dialog og refleksjon som sentrale elementer og der ansvarliggjøring og elevmedvirkning er viktige aspekter. (Norsk kulturskoleråd, 2016, s. 15, 28)

Medvirkning innebærer altså både rettigheter og muligheter, og at elevene både skal kunne få si sin mening og ha innflytelse på det som angår dem på skolen. Dette kan handle om selve undervisningen, hvordan eleven(e) lærer, miljøet eleven(e) er en del av, mangfolds- og inkluderingsproblematikk og utvikling av skolens regler og systemer.

I henhold til planverket skal altså dialogen mellom elev og lærer være preget av respekt for hverandre, og elevene skal oppleve at synspunktene deres blir tatt hensyn til. Instrumentalundervisning i Norge foregår vanligvis i en-til-en-settninger, og man kan anta at eleven i en slik

sammenheng opplever seg sett og hørt, i alle fall på noen områder. I hvilken grad dette også dreier seg om å påvirke innholdet i undervisningen, eller formen på denne, har vi foreløpig lite kunnskap om (Blix, 2012; West & Rostwall, 2003). Det er rimelig å anta et elever i kulturskolen som har spilletimer alene med én lærer, relativt sjelden opplever at de i denne konteksten er med på å forme utviklingen av skolen som helhet, noe de sannsynlig har større muligheter til i kontekster som samspill, kor, band og orkester i dette skoleslaget. Medvirkning inkluderer i denne studien også det å kunne ha innflytelse på den didaktiske innretningen på spille-timen som undervisningspraksis. Elevmedvirkning i undervisning er en kompetanse som må læres, og studiens design har dermed også funksjon som eksempel på organisering hvor studenten får øvd seg på en refleksiv, kreativ dialog med eleven om selve undervisningen.

## Tidligere studier av praksisveiledning og elevmedvirkning

Forskning på praksisveiledning i musikk har de senere år økt i omfang, og vi ser eksempelvis studier av utvikling av profesjonskompetanse (Bröske & Johansen, 2012; Johansen, 2009; Thorgersen & Johansen, 2012), profesjonsidentitet (Angelo, 2012; Georgii-Hemming, 2013) og sammenhenger mellom praksisfelt og utdanningsinstitusjonen (Ballantyne & Retell, 2020; Bandlien, 2023). Noen enkeltstående studier ser også på veiledningskompetanse og innhold og organisering av veiledning (Aglen, 2021; Thorgersen, 2010).

I sin studie av *praksissamtalen* i kulturskolelærerutdanninger undersøkte Aglen (2021) hvordan praksissamtalen fungerer som arena for musikkstudenters utvikling av profesjonsforståelse, og hva slags forståelser av kulturskolelærerprofesjonen som var i spill mellom aktørene i slike veiledningssituasjoner. Hun observerte praksissamtaler hvor student, øvingslærer (faglærer) og praksisveileder diskuterte undervisningen studenten hadde både observert og gjort selv ved praksisstedet. Undersøkelsen viste blant annet at ekspertene (lærerne) ledet an diskusjonene om undervisningspraksis, og at studentene lett kom i bakgrunnen. Asymmetrisk maktbalanse og forskjeller i profesjonsforståelse blant

ekspertene som veileder studenter, er interessante dimensjoner ved disse dialogene. Dette har også Bjørndal (2008) undersøkt i sin avhandling om symmetri og asymmetri i veiledningssamtalen, hvor han viste at manglende refleksjon over, og bevissthet om, at makten man som veileder har, kan stå i veien for studentenes læring.

Thorgersen og Johansen (2012) intervjuet musikk lærerstudenter, veiledere og faglærere ved musikkutdanningsinstitusjoner i fire nordiske land om kvaliteten på undervisning og læring i musikk lærerutdanning. Deres resultater viste tre hovedområder som hadde betydning for studentenes læring: (a) *organisatoriske faktorer*, hvor samarbeidet og sammenhengen mellom utdanningsinstitusjon og praksissted hadde varierende grad av felles fokus og målsetting for studentene, (b) *veiledere og faglærere*, som i varierende grad hadde felles oppfatninger om det didaktiske, og (c) *musikkstudentene selv*, hvor flere studenter rapporterte at de opplevde seg som overlatt til seg selv når det gjaldt å skape sammenhengen mellom forskjellige læringsbaner. Lignende problemstillinger tas også opp av Eik et al. (2021) i deres undersøkelse av veiledning i grupper, og studentenes erfaringer fra veiledningssamtalene. Her kom det fram at studentene var opptatte av hvor viktig det var å lytte, spesielt for å unngå å overkjøre hverandre som veiledere, og betydningen av *metasamtalen* som utgangspunkt for refleksjon over egen og andres læringsprosess.

I forskningslitteraturen finner vi fremdeles relativt få studier som ser på inkludering av de underviste elevene som en del av disse praksissamtalene, altså et *firepartsamarbeid*. I en nyere norsk studie av praksisveiledning undersøkte Pettersen og Lejonberg (2022) hvordan elevrespons kan bidra til refleksjon i veiledningssamtaler med lærerstudenter. De testet ut bruken av et spørreskjema som elevene som ble underviste skulle svare på, hvorpå elevenes anonymiserte svar ble brukt i veiledningssamtaler mellom studenter og praksisveiledere. Selv om undersøkelsen ikke var av stort omfang, antyder resultatene at elevrespons i veiledning kan «bidra til å støtte, utfordre og forankre refleksjon hos veisøkere» (s. 12). Engh (2007) drøfter også elevmedvirkning i  *vurderingsarbeid*, og vektlegger i begrunnelsene for medvirkningen at det bidrar til bærekraftige læringsbaner og bygger opp under demokratiske skoler og samfunn generelt.

Sentralt i forskningen på feltet står også veilederens rolle som en som bidrar til å belyse forholdet mellom praksis og pedagogisk teori. Skandinavisk forskning på praksisveiledning er opptatt av kritisk bevissthet hos profesjonsutøvere med utgangspunkt i at kunnskap og kompetanse skapes gjennom en dialogisk fundert veiledningspedagogikk (Bjørndal, 2008; Dysthe, 2001; Eik et al., 2021; Føinum, 2019). På tross av dette ser vi av Skagens (2013) gjennomgang av modeller for praksisveiledning, at modellene for veiledning i svært liten grad tar opp eleven som aktør og ressurs i praksisveiledning (Skagen, 2013).

## Dialog, resonans og demokrati

Denne artikkelen har en overordnet kritisk-pedagogisk inngang til praksisveiledning, med vekt på dialogiske prosesser som essensielle i utvikling av yrkeskompetanse som instrumentallærer. Av særlig interesse er de demokratiske aspektene som ligger i det å involvere de underviste elevene i studentenes bevisstgjøringsprosesser på vei mot lærerprofesjonen. Dette kommuniserer også indirekte til, og bevisstgjør, studentene om rollen de kan og bør ha i framtidig lærerpraksis som noen som lytter, spør, deler, skaper og investerer seg selv i dialogene med sine elever. En rekke tidligere skandinaviske studier av praksisveiledning har anvendt teoretikere som Bakhtin (1981), Rommetveit (1992) og Linell (2000) i analyser av dialogiske aspekter ved praksisveiledning (se for eksempel Bjørndal, 2008; Eik et al., 2021; Rinholm et al., 2023; Skagen, 2011), spesielt med vekt på meningsskaping i mellommenneskelige interaksjon og læring. Det teoretiske ståstedet i det foreliggende arbeidet er, i likhet med disse, i grunnholdning opptatt av demokrati, kritisk våkenhet, kreativitet, resonans og flerstemmighet i dialog, og av samtalens grenseløse potensial i lærings situasjoner.

Inspirert av teoretikere som Mikhail Bakhtin (1986), Anne Bogart (2021), Hartmut Rosa (2019), Paulo Freire (2011), Antonia Darder (2015), og bell hooks (2003) har jeg vært interessert i forskjellige måter å åpne opp for mer radikal dialogisitet<sup>2</sup> som setter *ting på/i spill* i veiledningssamtaler.

2 Linell (2017) skiller mellom dialog, dialogisme og dialogisitet, hvor *dialog* defineres som den faktiske samtalen mellom individer i sosiale praksiser, mens *dialogisme* refererer til et epistemologisk rammeverk, og *dialogisitet* til dialogiske «properties [...] of language, discourse and cognition» (s. 5)

Disse teoretikerne representerer alle kritikk av pedagogiske metoder som bidrar til *schoolification* (Bakhtin), eller det Freire kaller *banking* (putte kunnskap i elevene som om de var sparegriser), i likhet med dagens kritikere av nyliberale, målstyrte undervisningsregimer (Darder, 2015; Giroux, 2011). bell hooks (2003) argumenterer for *radikal åpenhet* i sine refleksjoner om hva som skal til for å skape meningsfull læring og undervisning gjennom dialog. Hun sier at det er i det marginale, det risikable, at endring kan skje i individet og i samfunnet. I arbeidet med å tolke datamaterialet i denne pilotstudien, har det derfor stått sentralt å drøfte om, og hvordan, en dialogisk pedagogikk kan bidra til å legge til rette for deling av makt, og å utfordre «det bestående» gjennom å se elever/studenter/lærere/veiledere som likeverdige meningshavere.

Vi kan betrakte dialogisme som epistemologiske og teoretiske tanke sett som utgjør et sammenhengende rammeverk, men det kan ikke sees på som én retning rent teoretisk. Linell (2000) viser til tre kjennetegn som karakteriserer dialogisme: antakelsen om at tenkning og læring innebærer interaksjon med andre (*interaksjonisme*), antakelsen om at diskurser er gjensidig avhengig av andre diskurser (*kontekstualisme*), og antakelsen om at mening og kunnskap er skapt i kommunikative prosesser. Linell (2000) kaller dette *kommunikativ konstruksjonisme*. I analysen av samtaler og ytringer må vi således ta i betraktning betydningen kontekstene har, de historiske og institusjonelle bakgrunnene til dialogpartene, samtalen som sjangre og lytterens medansvar i dialogen.

Et fruktbart samspill mellom deltakere i en veilednings- og tilbakemeldingsprosess krever at deltakerne opplever en form for kontakt eller forståelse hos de andre; at noe resonnerer. Begrepet *resonans* brukes av blant andre den amerikanske teater- og operaregissøren Anne Bogart (2021) for å forklare forholdet mellom aktører i kommunikative prosesser. Hun definerer resonans som det som «ripples and radiates; one energetic being influences the vibrations of another» (Bogart, 2021, s. 5). Sentralt i hennes argumentasjon står at det dialogiske aspektet ved all språklig og ikke-språklig kommunikasjon har behov for en eller annen form for resonans hos aktørene for å kunne samskape og lære sammen. I analyser av samtaler mellom mennesker viser eksempelvis språkforskeren Anna Joaquin (2013) til at resonans i form av repetisjoner og syntaktisk

stringens kan være fundert på et menneskelig instinkt for interaksjon: «The innate desire to attune to, immitate, and communicate with others» (Joaquin, 2013, s. 82). Joaquins definisjon av resonans er i hovedsak knyttet til språklige ytringer: «Resonance is the tendency of the qualities of one person's speech to match the qualities of another person's speech – phonologically, lexically, syntactically, and semantically» (Joaquin, 2013, s. 58). I tillegg til verbal resonans, understreker Joaquin at også kroppslig (*behavioral*) resonans spiller en stor rolle i dialog mellom mennesker. Hun er altså på linje med Bogart (2021), som sier at alle store kunstneriske opplevelser skaper og skapes av resonans (s. 6). En mer overgripende resonansteori finner vi hos sosiologen Hartmut Rosa (2019), som mener at søken etter resonans er «løsningen» på krisene skapt av en stadig akselererende verden. Ved å forholde oss responsivt og kreativt lyttende til omgivelsene, hverandre og oss selv, kan det skapes rom for en gjenklang/resonans som legger grunnlag for transformasjon (Rosa, 2019). I en analyse av dialoger er den delen av dynamikken som handler om hvordan ytringer *mottas og brukes* også vesentlig. I boken *Resonanspædagogik* viser Rosa og Endres (2016) hvordan konsekvensene for pedagogen dreier seg om å ivareta individet i bærekraftige læringsprosesser gjennom å omfavne det uventede, det relasjonelle og det de kaller læringsgnist.

Alle deltakerne i en dialog har et medansvar for den meningsskapingen som foregår, og forutsetningene for meningsskapingen vil blant annet dreie seg om interaksjonen mellom kontekst og deltakere. For at dialogen skal utvikles og skape noe nytt, må alle bidra med sine ulike perspektiver (Freire, 2011). Dette forutsetter at deltakerne/aktørene omfavner hverandre som likeverdige, med kjærlighet, ydmykhet og respekt (Darder, 2015). På linje med disse har Bakhtins forståelse av dialog som selve essensen i menneskelig væren (Sidorkin, 1996, s. 7) stått sentralt gjennom hele dette prosjektet: «I realize myself initially through others: from them I receive words, forms, and tonalities for the formation of the initial idea of myself [...] a person's consciousness awakens wrapped in another's consciousness» (Bakhtin, 1986, s. 138).

Bakhtin (1986) introduserte begrepet *heteroglossia*, som viser til kompleksiteten i språket, og mener at det er i spenningen og konfrontasjonen mellom ulike stemmer at ny innsikt og forståelse oppstår. Dialogen kan

være både ekstern (mellom to mennesker), og intern (mellom versjoner av selvet), og også den interne dialogen er spenningsfull og kontekst-avhengig (heteroglot) (Stensæth, 2008). Olga Dysthe snakker i bakhtinsk ånd om *det flerstemte*, hvor ulike, ikke-homogene perspektiver kan være utgangspunkt for skaping av nye forståelser der deltakerne kan bruke hverandres ytringer som tankeredsaker for meningssskaping (Dysthe, 2001, s. 126). Sammen med resonans og felles agenda, er flerstemte meningsforhandlinger og «uenigheter» sentrale i dialogen som potensial for læring.

*Ytringer* står sentralt i Bakhtins dialogfilosofi, og når vi går i dybden på hva en *ytring* er og betyr, ser vi at det som kjennetegner ytringer, er at de rettes mot noe eller noen – altså at de har en *adressivitet* (Bakhtin, 1981). Det innebærer at når vi studerer veiledningssamtaler og tilbakemelding må vi ha fokus på både samtalen, det som uttrykkes og hvem det er adressert til, fordi dialogers komplekse flerstemthet inkluderer ytringenes *rettethet* mot noe og/eller noen. *Den andre* kan være tenkt eller reell, resonerende eller ikke-resonerende, umiddelbart til stede eller i et møte langt frem i tid. Ordene vi bruker vil alltid bli uttrykt med en forventning om et «svar» fra en annen.

The word in living conversation is directly, blatantly, oriented toward a future answer-word: it provokes an answer, anticipates it and structures itself in the answer's direction. Forming itself in an atmosphere of the already spoken, the word is at the same time determined by that which has not yet been said but which is needed and in fact anticipated by the answering word. (Bakhtin, 1981, s. 280)

Ytringer er relasjonelle, og vil alltid vil være «minst tostemt» (Slaattelid, 1998, s. 54). De er formet gjennom vårt forhold til andre, og inngår alltid i en kjede av tidligere og kommende kulturelle og historiske øyeblikk. Meningsforhandlinger basert på kreativ forståelse og kunnskapsutvikling er grunnmuren i all dialog. I sammenheng med instrumentalundervisning vil dette også gjelde kunstneriske og kroppslige ytringer (Østern & Engvik, 2016).

Et sentralt element i dialoger handler om å forsøke å lytte radikalt åpent og uten ferdigproduserte svar liggende klare. Dialogisme i veiledning og

tilbakemelding dreier seg etter mitt syn i stor grad om å behandle det å *reflektere sammen* som en øvelse i å vente med å konkludere. White (2016) sier at «the key idea of dialogism is suspended certainty» med referanse til Bakhtin (s. 27). Dette resonerer med Deweys definisjon av refleksiv tenkning som «judgment suspended during further inquiry; and suspense is likely to be somewhat painful» (Dewey, 1933, s. 13). Dette kan være en utfordring for mange erfarne pedagoger som har mye å formidle (Gloppen, 2013), og begrepet *tilbakemelding* kan i seg selv oppfordre til en ulikeverdige dialog.

Gamlem og Smith (2013) beskriver fire ulike tilbakemeldingstyper basert på innholdet i tilbakemeldinger. De viser forskjellige grader av læringsaktivitet og opplevelse av deltakelse. *Type A: karakterer/poeng* vises til som den typen tilbakemelding hvor eleven er mest passiv og med lite interaksjon. *Type B: kontrollering* og *C: rapportering* er mer preget av interaksjon, mens *Type D: Dialogisk tilbakemeldingsinteraksjon* beskrives som den typen tilbakemelding som elevene sier de lærer mest av fordi det er en integrert del av en arbeidsprosess som både skuer tilbake og peker framover. I denne artikkelen er det spesielt denne siste typen tilbakemelding jeg ser på.

I den foreliggende studien står altså dialogisk tilbakemelding, som gir grobunn for kritisk tenkning og reflektert praksis i studentens virke som lærer og medmenneske, i sentrum. Vi ønsker å utdanne lærere som i samarbeid med sine elever og kolleger skaper og utvikler det pedagogiske feltet, og gjennom en reflekterende praksis kan studentene få anledning til å erfare og øve på dette dialogiske samarbeidet.

## Forskningsdesign

Pilotstudien ble gjennomført våren 2022 i form av en veiledningsøkt hvor en instrumentallærerstudent på et messinginstrument fikk tilbakemelding på sin undervisningspraksis fra både praksisveileder (i dette tilfellet meg) og fra messingeleven som var blitt undervist. Lærerstudenten gjennomførte en planlagt spilletime på 25 minutter med praksisveileder til stede, og deretter foregikk tilbakemeldingen på timen ved at lærerstudenten, eleven og veilederen snakket sammen om undervisningstimen.

Det ble gjort videoopptak av undervisningen, og dette opptaket ble brukt i veiledningssamtalen, som det også ble gjort videoopptak av. Alle tre diskuterte det vi så på innspillingen, ledsaget av spørsmål og kommentarer fra meg som veileder. Videoopptakene av alle leddene i prosessen, og transkripsjonene av disse, utgjør datamaterialet i pilotstudien. Transkripsjonene av observasjonene ivaretok både det som ble sagt, samt mine verbale beskrivelser av kroppslige gester og musikalske uttrykk. I transkripsjonsprosessen ble i hovedsak informasjon som ville ha betydning for analyse- og tolkningsarbeidet vektlagt.

Den valgte forskningsmetoden kalles i forskningslitteraturen for *video stimulated recall*; altså at forskningsdeltakerne får uttale seg om det som er observert gjennom å «stimuleres» av en videoinnspilling av den aktuelle situasjonen (Lyle, 2003; Rowe, 2009). Da blir deltakerne påminnet hva som skjedde, og forskeren kan stille spørsmål om «hva tenkte/følte du da?» med direkte henvisning til hendelser fra opptaket. Deltakerne blir på denne måten mer klar over egne reaksjoner, vaner og kroppslige uttrykk.

Spørsmålene jeg hadde med meg, hadde søkelys på tematikker som *kommunikasjon, innhold i timen, musikkvalg, målsettinger med timen og enkeltoppgaver*. I tillegg diskuterte vi utvalgte hendelser vi så på videoinnspillingen. For meg var det vesentlig at dialogene var forankret i tematikk som ga mening for deltakerne (særlig eleven), både for å være sikker på at vi hadde en felles interesse og forståelse av det dialogen handler om og for å legge grunnlaget for resonans mellom deltakerne. I dette tilfellet var hovedtemaet lærerstudentens undervisning, og gjennom å se på innspilling av timen, ble eleven og lærerstudenten presentert for et konkret og meningsfullt utgangspunkt for samtalen. Målet med å inkludere eleven i dialogen, var å åpne for nysgjerrighet, nye dialoger og kritisk engasjement i et praksisfellesskap som også ivaretok elevens innspill og meninger om undervisning. Eleven fikk rolle som en form for medveileder, og jeg ba ham om å hjelpe meg å gi lærerstudenten råd og tips for å bli en enda bedre messingpedagog. Disse to hadde allerede et etablert lærer/elev-forhold.

Eleven i studien gikk første året på musikklinje og hadde spilt i både storband, korps og symfoniorkester. Han ønsket å ta timer hos

lærerstudenten for å lære mer om improvisasjon, og disse timene har utgjort lærerens praksis i tilknytning til masterstudiet.

Lærerstudent og elev er anonymisert i framstillingen av datamaterialet, og de har begge samtykket til å delta. Forskningsetiske retningslinjer er fulgt i henhold til de nasjonale forskningsetiske retningslinjer for samfunnsfag og humaniora, og studien er vurdert av personverntjenesten Sikt. Min egen rolle som aktiv deltaker i praksisveiledningen har hatt betydning for måten datamaterialet er blitt håndtert på. Det kan være problematisk å få en kritisk avstand til materialet, og både datainnsamlingsprosessen og analysen vil til dels være preget av mine forforståelser og min nærhet til problemområdet. Den teoretiske innretningen har derfor stått sentralt i utøvelsen av innsamling og analyse, og i tillegg gir det å lese datamaterialet i form av transkripsjoner en viss avstand til materialet. Jeg har også lest transkripsjonene flere ganger med god avstand i tid, og med nærlesing av teori mellom disse lesingene. Jeg har loggført mine tanker underveis og har vært spesielt opptatt av å stille spørsmål til mine egne konklusjoner underveis. Christophersen beskriver dilemmaet som utspiller seg mellom fordelene med å ha nærhet til kulturen som undersøkes (i form av forståelse), og ulempene som ligger i involveringen som medfølger nærheten og *interessen* forskeren har til prosjektet (Christophersen, 2010). Hun diskuterer også maktmekanismene som ligger latent i handlingen mellom forskningsdeltakerne, noe som er spesielt aktuelt i den foreliggende studien på grunn av kompleksiteten i aktørenes forskjellige, og til dels flytende, roller. Min rolle, som både veileder og forsker, og den asymmetrien som det medbringer, kombinert med lærerstudentens rolle som både lærer og student, har gjort tolkningen av funnene utfordrende. Dialogene med meg selv har i denne sammenheng vært en måte å innta forskjellige posisjoner på som bidrar til en refleksiv balanse mellom distanse og nærhet. Det har altså vært innebygd i analysen at jeg har reflektert over mine antakelser og begrunnelser, noe som vil være godt synlig for leseren i framskrivningen av resultater og drøfting. På samme måte er forholdet mellom lærerstudenten og eleven asymmetrisk, og dette har til en viss grad konsekvenser for studiens pålitelighet, eksempelvis ved at eleven legger bånd på seg for å ikke såre læreren. Transparens i framskrivningen av resultatene har derfor vært en prioritet.

Datamaterialet er analysert gjennom en kvalitativ, induktiv prosess bestående av en tematisk analyse i fem faser med utgangspunkt i problemstillingens vekt på dialogisitet: nærlesing av datamaterialet, generering av begynnende koder, søking etter tema, gjennomgang av temaene på nytt og navngiving av temaer (Braun & Clarke, 2006). I mitt arbeid var de dialogiske prosessene i fokus for tematiseringen, og en pendelprosess mellom de teoretiske anslagene og empiriske data genererte følgende fire tematikker som jeg valgte å gå videre med i resultatdelen av studien: (a) *Det flerstemte*, (b) *Ytringer, rettethet og å bli hørt*, (c) *Resonans* og (d) *Demokrati gjennom medvirkning*.

## Mellom flerstemthet og resonans i tilbakemelding og læring

### Det flerstemte

I den observerte spilletimen startet lærerstudenten med å be eleven vise hvordan han vanligvis varmet opp på instrumentet og hvilke daglige øvelser han brukte. Eleven viste og snakket seg gjennom sin typiske oppvarming, hvorpå læreren påpekte at rekkefølgen på øvelsene kanskje burde justeres, og at dette «egentlig ikke er så valgfritt». Som veileder oppfattet jeg dette som en lite åpen og elevorientert måte å uttrykke seg på. Da vi så på videoinnspillingen av timen sammen etterpå, spurte jeg derfor lærerstudenten hvorfor han insisterte på denne rekkefølgen på *basic*-øvelsene (som de kalte det). Eleven tok umiddelbart ordet og sa at han synes det var kjempebra at lærerstudenten var så tydelig, for han var egentlig litt i tvil om hva som var lurt å tenke når man varmet opp. Han sa at det ofte ble litt tilfeldig når han øvde selv: «Når han sier noe, så går han ofte litt mer i dybden. Han gir råd, og så går han i dybden på hvorfor jeg bør gjøre det.» Lærerstudenten henvendte seg deretter til eleven og sa at de egentlig hadde snakket om dette, men at han i starten av timen skjønnte at han ikke hadde begrunnet dette med rekkefølgen godt nok på et tidligere tidspunkt, og at det derfor var viktig å si tydelig fra at det var viktig, og hvorfor.

Dette er et av flere eksempler i datamaterialet hvor eleven ikke umiddelbart var enig med meg eller lærerstudenten om tolkningen av det vi så i videoen av undervisningen. Videoen så ut til å skape rom for det

flerstemmige i denne dialogen, og vi innså alle tre noe om hvordan de andre skapte mening i det som hadde foregått i timen. Ved å bringe inn elevens stemme i tilbakemeldingene som en måte å utsette slutningene på, skaptesspenninger og nye forståelser oss imellom (Dewey, 1933; Dysthe, 2001). Hva slags forståelser den enkelte satt igjen med, var selvsagt verken gjennomsliktig eller stabilt, men ved å våge å stille spørsmål ved det som foregikk, og å utfordre hverandres forståelser og meningsskaping, ble det gitt grobunn til videre dialog og utforskning.

Et annet eksempel var da vi snakket om målsettinger med undervisningen, og målet med den aktuelle timen. Vi diskuterte om elevens og lærerstudentens målsettinger stemte overens. Det førte til en samtale om hvorvidt det alltid er bra for eleven å vite hvor lærerstudenten vil med sitt opplegg, og her meldte eleven at han, i motsetning til meg som veileder, mente det ikke alltid var av det gode, fordi det kunne gjøre ham litt nervøs å alltid skulle vite hva lærerstudenten hadde som mål med timen:<sup>3</sup>

- V [til eleven]: Syns du det er viktig for en elev å forstå hva læreren har som mål med en time?
- E: Ja, det kommer jo litt an på hva målene er. Men jeg tror de fleste elevene, i alle fall når man er litt viderekommende, skjønner litt hva det er man vil med timen.
- V: Hvis læreren i dag hadde sagt i begynnelsen av timen at «i dag skal vi prøve å strekke dette med interpretasjon litt, og tolkning» som han nevnte tidligere, hadde det hjulpet noe for deg?
- E: Det hadde kanskje gitt meg litt forventninger, og så hadde jeg blitt litt mer nervøs, så det kan ha negative og positive virkninger, at du skjønner hva han vil gjøre denne timen, og du vet hva han vil, så er du klar for det. Men igjen, så vet du hva han vil, og så tenker du: «Å nei, jeg klarer ikke så langt», og så blir du nervøs, og så blir det bare krøll.

Vi ser at eleven har andre meninger enn veileder i tilbakemeldingene til lærerstudenten når det gjelder hvorvidt det kan ha god effekt for læringen

---

3 V = veileder; E = elev; LS = lærerstudent

at eleven vet hva som er målet med timen. Denne «ustemtheten» blir noe lærerstudenten må stå i, og gjøre til en av vurderingene i planleggingen av fremtidige spilletimer. Uten elevens stemme involvert, ville vi kanskje gått glipp av denne nyanseringen.

## Ytringer, rettethet og å bli hørt

Gjennom å gi eleven en tydeligere rolle som medveileder i samtalen om den videofilmede praksisundervisningen, endret dynamikken mellom deltakerne seg litt, og det kunne merkes at elevens ytringer var rettet mer mot lærerstudenten i denne samtalen enn i spilletimen. Han var i mer opptatt av lærerstudentens perspektiv og utviste en ansvarsfull omsorg for situasjonen og læreren. Eksempelvis spilte de i siste del av timen en låt sammen, noe de ikke hadde gjort så mye tidligere. Akkompagnert av et lydspor spilte lærerstudenten melodien unisont med eleven. Lærerstudenten spilte med ekstra tydelig interpretasjon (dynamikk, musikalsk retning, aksentuering osv.), og i samtalen etterpå sa eleven at «det hjalp meg når han spilte med, fordi da skjønner jeg hva han mener. Og det hjelper meg å få til det han har sagt at jeg skal gjøre».

Lærerstudentens spill kan betraktes som musikalske ytringer rettet mot elevens lesing av musikkens mening. Dette så umiddelbart ut til å resonere med elevens indre forestilling av musikken, og han var opptatt av å respondere musikalsk på lærerens interpretasjon. Ved å sette ord på dette i samtalen etterpå, forstod også eleven at å spille sammen, lytte seg inn på en annen persons interpretasjon var en viktig meningsskapende prosess for ham. Dette utsagnet fra eleven var i utgangspunkt et svar på et spørsmål fra meg, men ytringen var tydelig rettet mot studenten; dette var noe han ønsker å gjøre mer. Studenten sa at effekten av at han spilte sammen med eleven var større enn han hadde tenkt over; en refleksjon som var et resultat av både videoinnspillingen vi så på, og av elevens kommentar.

En annen type dialog oppstod når vi diskuterte det å *syng*e musikken man skal spille, og hvorvidt læreren vanligvis oppfordret til dette:

- V: Synger du ofte sammen med opptak av musikken du spiller?  
 E: Nei, jo, av og til. Det kommer an på [...] men når jeg er alene nynner jeg litt til, eller når jeg sykler til skolen.

- V: Er det noe som læreren har foreslått for deg?
- E: Nei, det er noe som bare skjer automatisk. Fordi det er en fin måte å finne ut av det med frasering; hvordan gjør de på opptaket det?
- LS: Eh, jeg er nok ikke helt enig i at vi ikke har snakket om dette. Vi har snakket en del om å synge det du gjør, spille litt på munnstykket, og så få det på trombone. Når du jobber med pitch så har vi snakket om dette med at du må kunne synge det du spiller [eleven nikker bekræftende].
- V [til studenten]: Det er antakelig lurt å integrere det å synge i timene hvis du vil at elevene skal gjøre det, slik at de husker at du har sagt det.

Som vi ser opplevde lærerstudenten at dette var noe de hadde jobbet med, mens eleven ikke hadde oppfattet dette som en arbeidsmåte som lærerstudenten hadde foreslått, kanskje fordi det ikke var blitt demonstrert og brukt i timene. Slik førte de forskjellige stemmene i dialogen med seg klargjøring som endret tidligere ytringers status. Lærerstudenten innså at det å formidle noe og det å bli hørt og forstått ikke nødvendigvis henger sammen. Jeg tolker elevens utsagn som en fortelling om hans musikalske liv, mitt spørsmål var i hovedsak didaktisk, og lærerstudentens «protest» handlet kanskje både om å fortelle meg som veileder at dette er ivaretatt i timene, og samtidig få understreket til eleven at dette er viktig. Vi ser også at mine spørsmål og råd var preget av en utenfor-rolle som igangsetter og analytisk betrakter. Når jeg leser transkripsjonen, ser jeg tydelig andre veier denne delen av samtalen kunne tatt hvis jeg hadde gitt større rom for felles kreativ undring.

En del av veiledningen i forbindelse med spilletimer vil vanligvis bestå av fysiske og/eller musikalske ytringer. En faglærer eller veileder vil ofte demonstrere både spill og undervisningsmetoder underveis i en time. Likeledes er det lettere for en *elev* å synge, vise og spille for å forklare noe enn å uttrykke det med ord. Dialogen vil dermed inneholde andre ytringer enn språk. Ved å demonstrere på instrumentet, er kommunikasjonen på en måte mer direkte, men vil samtidig mangle nivåer av refleksjon som vanligvis ønskes i pedagogisk veiledning. I musikkfaget er likevel denne kunnskap i form av musikalske ytringer sentral i meningsskapingen mellom deltakerne, og bør tillegges vekt i analyser av praksisveiledning (Østern & Engvik, 2016).

## Resonans

- E: Jeg har funnet ganske mange kule duetter for trombone på internett.
- V: Bruker dere å «lete litt rundt» på nett sammen på timene?
- LS: Ja, vi har kanskje gjort det litt i starten. Vi laget i alle fall en liten Spotify-liste sammen, hvor jeg hadde plukket ut noen musikere som jeg mente han burde sjekke ut. [...] Vi har nerdet litte grann, men vi har ikke sosialisert så mye sånn sett. Det er kanskje litt uvant for meg at den muligheten fins. Det har jeg ikke tenkt på.

Denne typen resonans, hvor en av deltakerne plukker opp noe interessant i samtalen og spinner videre på det, tydelig engasjert, skjedde relativt ofte i denne veiledningssamtalen. Årsaken til at dette skjedde, så ut til å handle om at en kommentar, eller et innspill, «traff» en av deltakerne i noe denne var interessert i fra før, og som satte i gang kreative prosesser; «Wenn es im Klassenzimmer knistert» («Når det gnistrer i klasserommet», min oversettelse) som Rosa og Endres (2016) kaller det. Vi kan ane av lærerstudentens kommentar at han kommer til å se nærmere på hvordan de kan «nerde» mer sammen i timene. En slik form for resonans opplevde vi når alle deltakerne var nysgjerrige og spørrende i dialogen. Det handler kanskje om å møtes på «samme bølgelengde» for at nye tanker om undervisningen skulle oppstå. Andre ganger ble innspill eller ytringer bare hengende i luften uten respons. Sagt med Bakhtin er det gjennom *den andres* evne til å møte og lytte til oss, at vi blir til (Børtnes, 2001). Av interesse for en dialogisk pedagogikk, er derfor spørsmålet om hva som skal til for å sette i gang tanker og ideer hos *den andre*.

Samtalene som utspilte seg mellom oss i denne veiledningen, handlet også om hvordan eleven opplevde kommunikasjonen mellom seg og lærerstudenten, om det å velge musikk sammen, og om eleven syns det var meningsfullt å være med på veiledningstimen. Alle tre deltakerne i samtalen hadde innspill til det å ha kunnskap om hverandres forståelser og om hvordan vi kan oppfatte situasjoner og målsettinger forskjellig (uten å vite det). Lærerstudenten sa det slik: «Jeg begynner å forstå at det

er veldig vanskelig å kommunisere og være sikker på at budskapet er forstått. Det går mer og mer opp for meg hvor forskjellig man kan oppfatte et budskap.»

## Demokrati gjennom medvirkning

Det var spesielt spennende med de delene av samtalen som åpnet for tematikker som ikke så ofte kommer fram i en vanlig undervisningstime, og hvor både elev, lærerstudent og veileder så nye muligheter slik at dialogen ble produktiv for alles læring. I dette eksemplet kom vi inn på temaet motivasjon:

- E: Ja, jeg syns den beste motivasjonen for meg er konserter. Jeg har ikke egentlig spilt mye konserter. Jeg tror jeg har spilt én da jeg var litt yngre, og det gikk dårlig, og så har ikke jeg spilt solokonserter siden.
- LS: Å ja [litt overrasket]
- E: Og så hadde jeg en solokonsert nå hvor jeg spilte med akkompagnement tidligere i år, med «Autum Leaves», og så gikk den litt bra. Jeg følte at jeg kunne vært litt bedre forberedt. Og så har jeg konsert nå på lørdag med den samme. Dirigenten ville ha meg til å spille den, så da må jeg gjøre det.
- V: Ja, så du lar deg motivere av at du skal vise det fram?
- E: Ja.

Her opplevde vi at eleven opptrådte med stor selvstendighet og godt grep om egen framdrift. Måten han framstilte seg selv på, ga inntrykk av en elev som hadde eierskap til egne læringsprosesser og tillit til at han var berettiget til dette. Demokrati, tolket som individers muligheter til aktiv deltakelse, bli hørt og frihet til å utveksle meninger, er betinget av åpne og kreative dialoger. Jeg oppfattet også denne delen av dialogen som et sted der flerstemtheten førte til at vi alle tre innså noe som var nytt for oss. I en slik kontekst kan kanskje selve *medvirknings*begrepet bidra til å sette eleven *utenfor* dialogen, som en stemme som ikke er likeverdig de «voksne» i samtalen, noe jeg mener er sentralt å diskutere videre i en bredere kontekst.

Vi kom også inn på hva som skal til for at elever skal oppleve at de har eierskap til egne læreprosesser, etter å ha snakket om i hvilken grad det er lærerstudentens jobb å ha planlagt hele timen i detalj.

- V: Hvis læreren hadde spurt deg før dere startet på låta i dag: «Hva vil du ha hjelp til?», hva hadde du svart da?
- E: Puh, godt spørsmål. Få den videre til neste nivå. Hva er det jeg gjør verst med stykket, kanskje. Finne ut hva det er jeg har gjort dårligst med dette, og hvordan kan jeg gjøre det bedre.
- V: Nettopp. [Til studenten] Akkurat sånne ting kan være vanskelig for en elev å formulere: «Hva vil jeg ha hjelp til?» Hva tenker du om at man som pedagog kan hjelpe elevene sine til å bli gode på å bestille hjelp, gi dem et språk til akkurat dette?
- LS: Jeg tror nok at jeg har tenkt at vi er på tur dit, for vi kommer nok til å bruke også dette stykket du spilte til å etablere nettopp et slikt språk.
- V: Ja, og noe som «tvinger» ham til å utvikle et språk om hva er det er som fører framover, og som gjør ham gradvis mer selvstendig som musiker.
- E: Helt enig.

Vi ser at selv om vi rettet samtalen mer direkte mot én av deltakerne, opplevde alle seg berettiget til en mening om det vi snakket om. Konteksten og selve temaet vi diskuterte så ut til å spille inn på elevens ønske om å ta ansvar for egne innspill, og at han opplevde seg som relevant selv når vi snakket «om ham». Jeg legger i ettertid også merke til at jeg som veileder tok en rolle som var mer «belærende» (i form av både kommentarer og spørsmål) enn de andre deltakerne, og også at de forventet dette av meg i samtalen.

Vi snakket videre om at man som elev og lærerstudent trenger den energien det ligger i å «ville noe», og hvordan man som lærerstudent kan skape en setting som skaper denne dynamikken. I denne delen av samtalen var eleven spesielt engasjert og viste gjennom begrepene han brukte om lærerstudentens rolle som hans lærer, at han var vant til å være en

relevant meningshaver. Han omtalte eksempelvis lærerstudenten som en han kunne «få råd hos», og som skulle være «til hjelp», en han kunne få «tilbakemelding og tips» fra.

## Avsluttende refleksjon

Vi har i det foregående sett hvordan det oppstår ulike lag av dialoger i praksissamtalene, og hva som karakteriserer disse når eleven medvirker i tilbakemeldingene. Materialet viser at både flerstemthet, motstemmer og resonans er ingredienser i dialogene som har som mål å bevisstgjøre og kreativt endre en praksis, og hvordan disse ingrediensene gir rom for å oppdage noe som deltakerne ikke kunne erfart alene. Som læringskontekst så veiledningssamtalen ut til å være et egnet redskap for deltakernes meningsskapende prosesser, og elevens medvirkning bidro til lærerstudentens kompetanseutvikling på flere måter. Ved at samtalene handlet om en spesifikk undervisningssituasjon, ledsaget av et videoopptak av undervisningen, ble det åpnet for konkrete og kreative læringsprosesser og potensielt nye dialoger for alle deltakerne. Det var i det sosiale rommet mellom deltakerne at dialogisk meningsskaping foregikk (Bakhtin, 1981). Ved å avgrense rommet den dialogiske meningsskapingen foregikk i til å omhandle ett gitt tema med gitte deltakere i en gitt kontekst, fikk deltakerne mulighet og frihet til å fordype seg i det mangfoldet av stemmer som var involvert i dialogene. Meningsskaping er komplekse interpretasjoner av «levde sannheter» (White, 2016, s. 19), og er ikke stabile eller enhetlige. Nettopp i disse komplekse, lett kaotiske tolkningsprosessene erfarte vi som deltakere i studien å oppdage noe nytt, både ved oss selv og «den andre».

Det ble tydelig at *ytringer* i en typisk spilletime også kommer i form av handlinger, korte instruksjoner og musikalsk og kroppslig resonans. De meningsdannende prosessene blir spesielt komplekse i en slik praksis, og samtidig kan det å eksempelvis spille noe på instrumentet gi større innsikt, resonans og mening i dialogen enn å forsøke å forklare med ord. På samme måte var videoen vi brukte som utgangspunkt for samtalen, en del av disse intertekstuelle refleksjonene. Muligheten til å få avstand til egne tanker og reaksjoner, og å kunne betrakte samspillet som helhet, ga ytterligere stemmer til dialogen.

Dynamikken mellom deltakerne i pilotstudien var selvsagt preget av rollene vi hadde i utgangspunktet. Som messingblåsere alle tre, resonerte instrumentaltekniske ytringer lett, mens didaktiske tematikker fikk fram forskjelligartede stemmer hvor arbeidet med å forstå og skape mening ble mer krevende. Gjennom å låne ord fra hverandre og å forhandle om betydningene av hverandres ytringer, brukte vi hverandres erfaringer til å skape våre egne nye forståelser. Det dialogiske resonansrommet (Rosa, 2019) fungerte som en felles arbeidsarena som hadde utvikling av lærerstudentens undervisningskompetanse som mål.

Det å legge grunnlaget for en dialogbasert elevmedvirkning krever en type kompetanse som lærestedene må veilede kommende pedagoger i. Uten å vite hva elevene vet, vil og kan, er det vanskelig å utøve en pedagogikk hvor eleven reelt er i sentrum. Flertallet av undervisningsinstitusjoner i Vesten i dag erklærer at de baserer sin undervisning på et læringssyn som innebærer at den lærende konstruerer sin egen kunnskap gjennom aktivitet og sosiokognitive prosesser som resulterer i læring. En sentral del av lærerprofesjonen bør da bestå av å kunne lytte til elevene og å sette seg inn i hvilke faktorer i deres bakgrunn som konstituerer en slik aktivitetsbasert kunnskapskonstruksjon. Dette krever en jevnbyrdig, reell og åpen dialog hvor det felles målet er å endre seg selv og samfunnet gjennom *handling*. Hos reformpedagoger som Paulo Freire (2011) og Antonia Darder (2015) betraktes ord som handling, og uten handling er ordet dødt. I en musikkpedagogisk praksis vil dynamikken mellom måten man benevner fenomener, og handlinger, påvirke hverandre ved at endring av tankesett fører til endring av praksis. Det ligger samtidig et stort dilemma i vår begrensede mulighet til å vite hva andre tenker og mener. Bakhtin sier at ord/ytringer er bærere av stemmer som har tilhørt andre som har brukt ordene, og dette klinger med hver gang ordet tas i bruk (Bakhtin, 1986). Vi kan ikke få reell innsikt i andres erfaringer og kunnskap, og «hvor forskjellig man kan oppfatte et budskap» (lærerstudenten). Likevel kan vi, gjennom å tenke med Bakhtin, ta tak i erfaringene, betrakte dem fra forskjellige perspektiver og anerkjenne de ulike stemmene bak ytringene.

I studien så jeg i hovedsak på veiledning, tilbakemelding og dialog som inkluderte eleven, og hvor lærerstudentens utvikling av kompetanse

var målet. Uten et dialogisk blikk på denne praksisen, kunne elevens medvirkning forblitt en litt eksotisk ekstragevinst i veiledningssamtalen. I didaktisk litteratur som beskriver elevmedvirkning, brukes ofte fraser som «å få lov til å stille spørsmål», «være med i planlegging», «få anledning til å delta» og «å bli spurt hva du har lyst til» (jf. Mikalsen et al., 2013). Dette kan gi inntrykk av invitasjoner til å være med, men fra et utenfor-ståsted. Elevmedvirkning tolket som demokratisk medborgerskap bør innebære at elevene lyttes til, at de får støtte til å uttrykke sine meninger, at deres meninger vektlegges, at de er involverte i (eller styrer) beslutningsprosesser og at de deler makt og ansvar for beslutninger (Redd Barna, 2011, s. 21). En viktig diskusjon blir da hvordan vi får til dette i praksis. Her lå også en vesentlig del av læringspotensialet for lærerstudenten i denne studien: en konkret erfaring med å ta eleven med i samtaler om elevens egne læringsprosesser, og i dialogen om lærerstudentens undervisning.

Elevmedvirkning som vektlegger rom for innspill, meninger og *makt* over innhold, progresjon og så videre, både i spilletimene og i kulturskolehverdagen, fordrer at eleven får anledning til å utvikle et mer bevisst overblikk og myndiggjøres når det gjelder egne lærings- og tilbakemeldingsprosesser. Dette kan sees på som en demokratisering av spilletimene, noe som innebærer en mer spørrende og mindre «belærende» praksis (Aglen, 2021). En sentral del av lærerprofesjonen handler om å øve på å lytte til elevene og å sette seg inn i deres forkunnskaper, aspirasjoner og måter å lære på, gjennom dialog og å bruke flerstemmighet/heteroglossia som ressurs. Dette vil vi at våre lærerstudenter skal trenes i og ha kunnskap om. Da kan elevmedvirkning, gjerne gjennom et *fire-partsamarbeid*, i praksisveiledning kunne være et nyttig tilskudd til de mest brukte veiledningsmodellene beskrevet i Skagen (2011).

Denne pilotstudien viser hvordan dialogene i tilbakemeldingene ble tilført et resonansrom og en flerstemmighet som ga stort potensial for læring for alle deltakerne. Lærerstudenten uttrykte at kombinasjonen av elevens stemme og den videoinnspilte timen gjorde veiledningssamtalen til en meningsfull læringsprosess og demonstrerte i praksis hvordan eleven kunne medvirke også i en vanlig spilletime. Basert på resultatene av pilotstudien vil disse elementene derfor stå sentralt i det planlagte

aksjonsprosjektet om elevmedvirkning i praksisveiledning, og er også relevant i generelt arbeid med veiledning og tilbakemelding, inkludert å jobbe fram en større vifte av tilbakemeldingspraksiser i veiledning av musikk lærere.

## Forfatterbiografi

**Hilde Synnøve Blix** er professor i hørelære og musikkdidaktikk ved UiT Norges arktiske universitet, Musikkonservatoriet. Hennes forskningsinteresser er blant annet gehørutvikling, musikk literacy og læremidler, i tillegg til kjønnsforskning og likestillingsforskning. Hun har ledet flere forskningsprosjekter, og har publisert en rekke forskningsartikler og bøker nasjonalt og internasjonalt med tematikk som spenner fra gehørundervisning og notedidaktikk til likestilling og kunstnerisk utviklingsarbeid. Hun er leder for forskningsgruppen InOvation ved UiT og FoU-leder ved Musikkonservatoriet.

## Referanser

- Aglen, G. S. (2021). Praksissamtalen i kulturskolelærerutdanninger. En arena for utvikling av studenters profesjonsforståelse. *Nordic Research in Music Education*, 2(1), 134–160. <https://doi.org/10.23865/nrme.v2.3020>
- Angelo, E. (2012). Profesjonsforståelse hos en instrumentalpedagog i en samordnet musikk lærerstilling. *Nordisk kulturpolitisk tidsskrift*, 15(1), 27–45. <https://doi.org/10.18261/ISSN2000-8325-2012-01-03>
- Bakhtin, M. (1981). *The dialogical imagination. Four essays by M. M. Bakhtin* (C. Holquist, Red.). University of Texas.
- Bakhtin, M. (1986). *Speech genres and other late essays* (V. W. Maghee, Overs., C. Emerson & M. Holquist, Red.). University of Texas Press.
- Ballantyne, J. & Retell, J. (2020). Teaching careers: Exploring links between well-being, burnout, self-efficacy and praxis shock. *Frontiers in Psychology*, 10, 1–13. <https://doi.org/10.3389/fpsyg.2019.02255>
- Bandlien, B.-T. (2023). Kulturskole–grunnskole-samarbeid som diskursivt felt: En undersøkelse av skoleslagenes plandokumenter med fokus på musikk. I O. B. Øien, S. S. Kolaas, M. F. Duch & E. Anglo (Red.), *MusPed:Research: Vol. 6. Explorative perspectives in music and perspectives in music and education* (s. 225–254). Cappelen Damm Akademisk. <https://doi.org/10.23865/noasp.200.ch10>

- Bjørndal, C. R. P. (2008). *Bak veiledningens dør: Symmetri og asymmetri i veiledningssamtaler* [Doktorgradsavhandling]. Universitetet i Tromsø.
- Blix, H. S. (2012). *Gryende musikk-literacy. Unge instrumentalelevers tilegnelse av musikk-literacy i et sosiokognitivt perspektiv* [Doktorgradsavhandling]. Norges musikkhøgskole.
- Bogart, A. (2021). *The art of resonance*. Methuen Drama.
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology* 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Brøse, B. Å. & Johansen, G. (2012). *Educating music teachers in the new millennium. Multiculturalism, professionalism and music teacher education in the contemporary society*. Norges musikkhøgskole.
- Børtnes, J. (2001). Bakhtin, dialogen og den andre. I O. Dysthe (Red.), *Dialog, samspel og læring* (s. 91–105). Abstrakt forlag.
- Christophersen, C. (2010). Interesser, nærhet og brudd. Om å forske på egen musikkpedagogiske kultur. I S.-E. Helgesen & S. G. Nielsen (Red.), *Norsk musikkpedagogisk forskning. Årbok 12* (s. 31–42). Norges musikkhøgskole.
- Darder, A. (2015). *Freire and education*. Routledge.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. D. C. Heath.
- Dysthe, O. (Red.). (2001). *Dialog, samspel og læring*. Abstrakt forlag.
- Eik, L. T., Moxnes, A. R. & Vala, H. (2021). Spenning mellom ulike stemmer: Læringsmuligheter i veiledning. *Nordisk tidsskrift i veiledningspedagogikk*, 6(1), Artikkel e3131. <https://doi.org/10.15845/ntvp.v6i1.3131>
- Engh, K. R. (2007). Elevmedvirkning i vurderingsarbeid. *Norsk pedagogisk tidsskrift*, 91, 107–119.
- Freire, P. (2011). *De undertryktes pedagogikk* (2. utg.). Gyldendal Akademisk.
- Føinum, M. (2019). Veiledningssamtaler: Hvordan kan mentor bidra til å utvikle en ny lærers forståelse for undervisningsarbeidet? *Nordisk tidsskrift i veiledningspedagogikk*, 4(1), 17–39. <https://doi.org/10.15845/ntvp.v4i1.1385>
- Gamlem, S. M. (2019). *Tilbakemelding for læring og utvikling*. Gyldendal Norsk Forlag.
- Gamlem, S. M. & Smith, K. (2013). Student perceptions of classroom feedback. *Assessment in Education: Principles, Policy & Practice*, 20(2), 150–169. <https://doi.org/10.1080/0969594X.2012.749212>
- Georgii-Hemming, E. (2013). Music as knowledge in an educational context. I E. Georgii-Hemming, P. Burnard, & S.-E. Hølgersen (Red.). *Professional knowledge in music teacher education* (s. 19–38). Ashgate.
- Giroux, H. (2011). *On critical pedagogy*. Continuum International.
- Gloppen, B. H. (2013). Trepartssamtalen. *Uniped*, 36(1), 88–101. <https://doi.org/10.3402/uniped.v36i1.20949>

- Harrison, S. & Grant, C. (2015). Exploring of new models of research pedagogy: Time to let go of master–apprentice style supervision? *Teaching in Higher Education*, 20(5), 556–566. <https://doi.org/10.1080/13562517.2015.1036732>
- hooks, b. (2003). *Teaching community: A pedagogy of hope*. Routledge.
- Joaquin, A. D. L. (2013). Resonance in dialogic interaction. I A. D. L. Joaquin & J. H. Schumann (Red.), *Exploring the interactional instinct* (s. 57–87). Foundations of Human Interaction. <https://doi.org/10.1093/acprof:oso/9780199927005.003.0004>
- Johansen, G. (2009, 5.–9. juli). Relations between identity and educational quality within pre-service music teacher training [Paperpresentasjon]. The 6th International Symposium on the Sociology of Music Education, Limerick, Ireland.
- Kunnskapsdepartementet. (2017). *Overordnet del – verdier og prinsipper for grunnopplæringen*. Fastsatt som forskrift ved kongelig resolusjon. Læreplanverket for Kunnskapsløftet 2020. <https://www.udir.no/lk20/overordnet-del/>
- Lejonberg, E. (2016). *Hva kan bidra til god veiledning? En problematisering basert på veilederes og veisøkeres perspektiver på veiledning av begynnende lærere* [Doktorgradsavhandling]. Universitetet i Oslo.
- Linell, P. (2000). *What is dialogism? Aspects and elements of a dialogical approach to language, communication and cognition* [Forelesning]. Department of Communication Studies, Växjö University.
- Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal*, 29, 861–878. <https://doi.org/10.1080/0141192032000137349>
- Mikalsen, H., Nes, K. & Dobson, S. (2013). Om «dronninger og prinser, resten er bare i veien» – elevmedvirkning i teori og praksis. *PAIDEIA*, 6, 57–71.
- Norsk kulturskoleråd. (2016). *Rammeplan for kulturskolen. Mangfold og fordypning*. Norsk kulturskoleråd.
- Opplæringslova. (2024). *Lov om grunnskoleopplæringa og den vidaregåande opplæringa* (LOV-2023-06-90-30). Lovdata. <https://lovdata.no/lov/2023-06-09-30>
- Pettersen, M. & Lejonberg, E. (2022). Elevrespons i praksisveiledning: Hvordan kan et verktøy bidra til støtte, utfordring og teoretisk forankret refleksjon?. *Nordisk tidsskrift i veiledningspedagogikk*, 7(1), 1–17. <https://doi.org/10.15845/ntvp.v7i1.3336>
- Redd Barna. (2011). *Si din mening – og bli hørt* [Hefte utviklet av det norske Redd Barna i forbindelse med pilotprosjektet *Si din mening – og bli hørt*]. Redd Barna.
- Rinholm, H., Solem, I. H. & Ulleberg, I. (2023). Dialogues in education: Exploring cases from music and mathematics education. *Dialogic Pedagogy*, 11(1), A60–A84. <https://doi.org/10.5195/dpj.2023.458>
- Rommetveit R. (1992). Outlines of a dialogically based social-cognitive approach to human cognition and communication. I A. H. Wold (Red.), *The dialogical alternative: Towards a theory of language and mind* (s. 19–44). Scandinavian University Press.

- Rosa, H. (2019). *Resonance: A sociology of our relationship to the world*. Polity Press.
- Rosa, H. & Endres, W. (2016). *Resonanspædagogik. Når det knitrer i klasseværelset*. Hans Reitzels Forlag.
- Rowe, V. C. (2009). Using video-stimulated recall as a basis for interviews: Some experiences from the field. *Music Education Research*, 11(4), 425–437. <https://doi.org/10.1080/14613800903390766>
- Sidorkin, A. M. (1996). *An ontological understanding of dialogue in education* [Doktorgradsavhandling]. University of Washington.
- Skagen, K. (2011). Norsk veiledningstradisjon i lærerutdanning – analyse og kritikk. I K. Skagen, *Kunnskap og handling i pedagogisk veiledning* (2. utg.). Fagbokforlaget.
- Skagen, K. (2013). *I veiledningens landskap. Innføring i veiledning og rådgivning*. Cappelen Damm Akademisk.
- Slaattelid, R. T. (1998). Etterord. I M. Bakhtin, *Spørsmålet om talegenrane* (s. 47–89). Ariadne forlag.
- Stensæth, K. (2008). *Musical answerability: A theory on the relationship between music therapy improvisation and the phenomenon of action* [Doktorgradsavhandling]. Norges musikkhøgskole.
- Thorgersen, C. F. (2010). Holistic quality learning in musikkdidaktik from a student perspective: Where, when and how does it occur? *Visions of Research in Music Education*, 15. Hentet 23. april 2024 fra <http://www-usr.rider.edu/~vrme/>
- Thorgersen, C. F. & Johansen, G. (2012). Student music teachers' learning trajectories. A relational perspective. I B. Å. B. Danielsen & G. Johansen (Red.), *Educating music teachers in the new millennium* (s. 187–206). Norges musikkhøgskole.
- Utdanningsdirektoratet. (2021). *Elevundersøkelsen – resultater fra 7. og 10. trinn*. Hentet 22. januar 2025 fra <https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/elevundersokelsen-2021-hovedrapporten/>
- West, T. & Rostvall, A.-L. (2003). A study of interaction and learning in instrumental teaching. *International Journal of Music Education*, 40 (1), 16–27. <https://doi.org/10.1177/025576140304000103>
- White, E. J. (2016). *Introducing dialogic pedagogy. Provocations for the early years*. Routledge.
- Østern, A.-L. & Engvik, G. (Red.). (2016). *Veiledningspraksiser i bevegelse. Skole, utdanning og kulturliv*. Fagbokforlaget.



## KAPITTEL 5

# «Å få en rund ball til å passe inn i et firkantet hull» - sensorveiledningers rolle i vurdering av eksamenskonsserter

Lillian Grethe Jensen

UiT Norges arktiske universitet

## Abstract

In this article, the author seeks to gain insight into the assessment actions that underlie the evaluation of students' examination concerts. What role does the examiner's guidance play in the assessment, and in what ways are professional judgment and expert competence applied? The study is based on five qualitative research interviews with professionals who have served as examiners in higher music education. The research material indicates that being moved and excited – when the concert resonates, and the experience of the examiner or the examiner group as both performing artists, educators, and examiners – play a significant role in how examiners justify their assessments.

**Keywords:** assessment in arts subjects, criteria, holistic and analytic assessment, professional judgment in assessment, assessment guidelines, exam concert

Sitering: Jensen, L. G. (2025). «Å få en rund ball til å passe inn i et firkantet hull» - sensorveiledningers rolle i vurdering av eksamenskonsserter. I H. S. Blix, K. J. Valkoinen & B. A. Haugseth (Red.), *MusPed:Research: Vol. 11. Vurderingsformer og tilbakemelding i kunstfaglig utdanning* (s. 129–156). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.269.ch5>  
Lisens: CC BY 4.0

## Introduksjon

*Salslyset dempes, og praten blant tilhørerne stilner idet kandidaten gjør entre sammen med sine medmusikanter. Kandidaten skal ha sin avsluttende bachelorkonsert, en konsert som skal vise hva som er oppnådd gjennom fire år med øving og konsertdeltakelse. I salen sitter familie, medstudenter og andre publikummere som har funnet veien til konserten. Et sensorpanel bestående av ekstern sensor og to av kantatens faglærere gjennom hele bachelorløpet, skal vurdere prestasjonen og komme frem til en rettferdig karaktersetting.*

I stortingsmeldingen *Kultur for kvalitet i høyere utdanning* (Meld. St. 16 (2016–2017)) legges det under avsnitt om sensur frem at gode ordninger for sensur er nødvendig for å sikre en upartisk og rettferdig karaktersetting. Kvaliteten i utdanningen handler her om sammenhengen mellom de fastsatte læringsutbytteskrivelsene, undervisningsformer, faglig innhold, læringsaktivitet og vurderings- og eksamensordninger (Kunnskapsdepartementet, 2025). Denne tankegangen er inspirert av såkalt *constructive alignment* (Biggs & Tang, 2011) der forhåndsdefinerede læringsmål, undervisningsformer og vurdering ses i sammenheng. *Sensorveiledninger* er et virkemiddel som anbefales som bidrag til «kvalitetssikring av det fastsatte læringsutbyttet og til å redusere ulikhet i sensur mellom sensorer» (Meld. St. 16 (2016–2017), s. 58).

I denne artikkelen søker jeg å få innblikk i hva som ligger til grunn for vurdering av studenters eksamenskonsserter. Gjennom å intervju fagpersoner som har jobbet med sensur innenfor dette fagfeltet undersøkes: Hvilken rolle spiller sensorveiledningen i vurderingen, og på hvilke måter anvender sensorene skjønn og ekspertkompetanse når de begrunner sine vurderinger?

Min motivasjon for å utforske dette temaet er forankret i mangeårig erfaringer som faglig ansatt ved Musikkonservatoriet ved UiT. Jeg har jobbet som faglærer i rytmisk vokal, og har derfor deltatt i en lang rekke vurderingssituasjoner i forbindelse med eksamenskonsserter i høyere utøvende musikkutdanning. Ved mange anledninger har jeg opplevd hvor utfordrende det kan være for sensorpanelet å komme frem til en karakter, og hvordan sensorer har svært ulik tilnærming til hva som legges til grunn for den endelige vurderingen.

Universitets- og høyskoleloven (2024, § 11-6) stiller krav om at «universiteter og høyskoler skal sørge for at studentenes kunnskaper og ferdigheter blir prøvet og vurdert på en upartisk og faglig betryggende måte». Loven sier videre at det skal utarbeides skriftlig sensorveiledning til alle eksamener (Universitets- og høyskoleloven, 2024, § 11-7). På bakgrunn av dette varslede kravet besluttet Rådet for utøvende musikkutdanning (RUM)<sup>1</sup> å opprette en arbeidsgruppe med mandat om å «utrede og utarbeide forslag til sensorveiledninger for sentrale, gjennomgående emner i utøvende og skapende musikkutdanning» (Rådet for utøvende musikkutdanning [RUM], 2018, s. 2). Arbeidet resulterte i maler for sensorveiledninger, med føringer for sensorpanelet arbeid og vurderingskriterier, som skulle være et redskap for den enkelte medlemsinstitusjons arbeid med å ferdigstille sensorveiledninger til sine egne emner. Som deltaker i møter i FUM har jeg erfart at sensorveiledninger, med utgangspunkt i disse malene, nå er i bruk i sensurarbeid ved flere av de høyere musikkutdanningsinstitusjonene i Norge. Noen av disse institusjonene har utarbeidet fagspesifikke karakterbeskrivelser, mens andre benytter generelle beskrivelser av karakterskalaen. Om dette har betydning for informantenes vurderingshandlinger kommer ikke frem i denne studien.

RUMs rapport (2018) anbefaler at den enkelte utdanningsinstitusjon fastsetter vurderingskriterier til sine eksamener. Ifølge Olsson og Nielsen er kriterier «språklige abstraksjoner vilka visar på kunnskapskvaliteter och kompetenser på olika nivåer» (2018, s. 14). Vurderingskriteriene skal utarbeides i samsvar med læringsutbyttebeskrivelsene i emneplanene, slik at det er en sammenheng mellom undervisningsaktivitet, emneplanenes læringsutbytte og vurdering (John Biggs' *constructive alignment*) (Andersen, 2010; Biggs & Tang, 2011). Vurderingskriteriene som er eksemplifisert i RUMs rapport fra 2018 er utformet som tematiserte overskrifter med forslag til konkretisering under hver overskrift. Kriterier som *arbeidet som helhet*, *kunstnerisk integritet*, *musikalsk bevissthet*, *faglig relevans* og *kommunikasjon* gir en ramme som sensor skal forholde seg til, men som

<sup>1</sup> Rådet for utøvende musikkutdanning (RUM) skiftet i 2019 navn til Nasjonalt fagorgan for utøvende og skapende musikk – FUM. Medlemsinstitusjonene er: UiT Musikkonservatoriet, Norges musikkhøgskole, Fakultet for kunsthøgskolen (UiS), Institutt for musikk (NTNU), Fakultet for kunsthøgskolen (UiA), Griegakademiet (UiB), Operahøgskolen og Høgskolen Kristiania.

også er åpen for skjønnsmessig og holistisk vurdering. Olsson og Nielsen (2018) viser til forskning som diskuterer «skarpe» og «uskarpe» kriterier der de skarpe er analytisk og objektive, og de uskarpe er vanskeligere å verbalisere, og har en mer subjektiv og holistisk karakter. Kriteriet som har å gjøre med instrumentbeherskelse som intonasjon, klangbehandling, rytmikk og teknikk blir her omtalt som skarpe kriterier (Olsson & Nielsen, 2018). Dette har sammenheng med at disse kriteriene ofte har et utviklet begrepsapparat (Olsson & Nielsen, 2018). I sensorveiledningen til bachelorkonserten ved UiT legges følgende vurderingskriterier til grunn for karactersettelse og tilbakemelding til kandidaten:

- Konserten som helhet (f.eks. dramaturgi, repertoar)
- Instrumentbeherskelse (f.eks. klangbehandling, dynamikk, rytmikk, teknikk)
- Musikalsk bevissthet (f.eks. kunstnerisk uttrykk, tilstedeværelse, karakter, stilforståelse, tolkning)
- Kommunikasjon (f.eks. samspill, musikalsk interaksjon, publikumskontakt, presentasjon av konserten)

Disse vurderingskriteriene samsvarer i stor grad med de kriteriene som eksemplifisert i RUM-rapporten (2018) og som også er å finne ved andre av institusjonene som inngår i FUM-samarbeidet. I veiledningene går det ikke frem at kriteriene er rangert. Sensorveiledningen som brukes ved eksamen i utøvende musikkutdanning ved UiT fremholder at vurderingen skal skje på bakgrunn av eksamensprestasjonen alene.

## Tidligere forskning

I det følgende har jeg sett på noe av den forskningslitteraturen som omhandler vurdering generelt, og vurdering i kunstfaglig sammenheng, spesielt bruken av sensorveiledning og vurderingskriterier.

Tidlige studier av vurdering og vurderingskriterier diskuterer to hovedinnfallsvinkler til vurdering: normbaserte eller kriteriebaserte (Sadler, 1987). Sadler gir i tillegg en oversikt over en tilnærming kalt standard-referert vurdering, som viser til et bestemt nivå, grad av måloppnåelse

eller kvalitet. Han identifiserer og beskriver fire metoder for å spesifisere og formidle utdanningsstandarder. De fire metodene benytter numeriske grenseverdier, taus kunnskap, eksempler og verbale beskrivelser. Det argumenteres for at de to sistnevnte, når de brukes sammen, gir det mest lovende rammeverket for et standard-referert vurderingssystem der lærere kan gjøre gode kvalitative vurderinger av studentens prestasjoner både for å forbedre læring og for summativ rapportering.

Waagen (2021) ser på hvordan sensorveiledninger med forhåndsdefinerede kriterier og karakterbeskrivelser fungerer i praksis. Waagen tar opp flere interessante problemstillinger knyttet til vurderingspraksiser. Hun hevder at kvalitetsvurderingen må ha et «eksplisitt vurderingsgrunnlag» for å unngå at vurderingen blir basert på esoterisk viten, forstått som kunnskap basert på årelang erfaring, og sier at «kriterier gir potensial for transparent, argumenterende og pålitelig vurderingspraksis» (Waagen, 2021, s. 59). Torrance (2007) mener derimot at man ved å legge for mye vekt på kriterieoverholdelse og måloppnåelse beveger seg bort fra en orientering som fremmer studentautonomi og studentaktiv læring. Han sier: «Such practices are interpreted as techniques to assure award achievement and probably help to produce students who are more dependent on their tutors and assessors rather than less dependent» (Torrance, 2007, s. 282).

Vinge (2011) drøfter vurdering i musikkfaget gjennom begrepsparet *analytisk*, der man har formet eksplisitte vurderingskriterier, og *holistisk*, der kriteriene ikke er formulert på forhånd. Kriteriebegrepet blir også belyst av blant andre Olsson og Nielsen (2018) som diskuterer «skarpe og oskarpe» kriterier (s. 14). Gynnild (2022) sier at pålegget om kriteriebasert vurdering er styrket på bekostning av en helhetlig og skjønnsbasert vurderingspraksis også innen kreative disipliner. På lik linje med Sadler (1987) er han kritisk til kriteriebasert vurdering i *et tolkningsbasert kunnskapsdomene* (Gynnild, 2022). Gynnild påpeker at sensurveiledning innen kreative disipliner kan medføre en risiko for forventningsstyrt sensur. Dette «reducerer tolkningsrommet for kandidatene ved det faktum at besvarelser kan være ulike, men av samme kvalitet» (Gynnild, 2022, s. 49). Andersen (2010) drøfter også hvordan prinsippet om at det skal være en klar sammenheng mellom studieplaner, undervisning og vurdering

(jf. John Biggs' *constructive alignment*) kan medføre en risiko for generalisering og simplificering.

Orr (2010, s. 3) spør: «Can fine art assessment be based in connoisseurship and still be rigorous? » Hun undersøker dette ved å analysere vurderingsmetodene som kunstforelesere benytter, med et spesielt fokus på stilltende praksis og hvordan foreleserne forstår begrepene subjektivitet og objektivitet. Hun ser på den subjektive tilnærmingen til vurdering som foregår i et praksisfellesskap, og sier at flere av respondentene i denne studien kommenterte at tause tilnærminger blir mer tydelige gjennom dialog.

I artikkelen «Low Stakes and High Temperatures» ser Oltedal (2020) på hvordan ulike interesser og faktorer spiller inn på det vurderingsarbeidet som gjøres i en sensorgruppe etter spilleeksamen ved en musikklinje i videregående skole. Hun undersøker hvordan *epistemisk autoritet*, – beskrevet som kjennskap til studenten, spisskompetanse, sjangerforståelse, kjønn og erfaring operasjonaliseres i sensurmøter.

Vurdering av kvalitet i kunstneriske uttrykk blir også belyst i artikkelen «I kjønneste orden», hvor Blix et al. (2019) undersøker hvilket kvalitetssyn komitémedlemmer som skal vurdere professor- og dosentkompetanse i kunstfagene legger til grunn for sin vurdering. Studien viser at komitémedlemmene definerer kunstnerisk kvalitet ulikt, selv om det foreligger kriterier og en forskrift de skal forholde seg til i sin egen vurdering.

Flere av informantene uttrykker at de vet hva god kvalitet er når de ser det. Dette signaliserer at de mener det ikke alltid er nødvendig å beskrive konkrete kunstneriske kvalitetsmarkører, fordi det er tilstrekkelig at en gjennom sin posisjon som professor/dosent på feltet har den autoriteten og kunnskapen som skal til for å avgjøre hva som er av høy kvalitativt. (Blix et al., 2019, s. 14)

Skjønn og bruk av skjønn i profesjonsutøvelse blir belyst av Grimen (2009), som tar for seg ulike teoretikers syn på hvordan skjønn kan forstås. Grimen viser blant annet til den engelske statsviteren og filosofen Robert Goodins syn på at utstrakt bruk av skjønn er uforutsigbart og ikke etterprøvbart, og til filosofen Ronald Dworkins forståelse av skjønn som et rom for frihet innenfor gitte rammer.

Ovennevnte forskningslitteratur viser til ulike vurderingsstrategier, og at man både kan finne støtte for sensorveiledninger med forhåndsdefinerte kriteriebaserte sensorveiledninger innen kunstfagene, men også forskning som viser at det er grunn til å være kritisk til for streng målstyring av undervisning, og til forhåndsdefinerte kriterier i tolkningsbaserte fag. I fortsettelsen vil jeg redegjøre for ulike teorier jeg har anvendt i analysen av datamaterialet i denne studien.

## Vurderingsteoretisk bakteppe

Det teoretiske bakteppe til studien omhandler vurdering og ulike vurderingsstrategier som er relevant for analysen av datamaterialet i denne studien.

I sin bok *Det vurderende øyet* sier Cato Bjørndal (2012) at vurdering er en grunnleggende side ved hvordan mennesker bearbeider inntrykk. Ordet vurdering (*werderen*) har tysk opprinnelse og «betyr å verdsette, gjøre seg opp en mening om, bedømme, tillegge verdi, bestemme verdien av noe» (Bjørndal, 2012, s. 12). Vurdering og evaluering blir ofte brukt om hverandre, men selv om begge begrepene har med verdsetting å gjøre, er det viktig å skille mellom disse to. I dag blir *evaluering* brukt når det er snakk om ulike former for systemevaluering, altså en verdi-setting av en institusjons praksis, rutiner, og systemer. Evaluering er en viktig del av utdanningsinstitusjonenes kvalitetsarbeid. *Vurdering*, det vil si verdsetting av studentarbeid, brukes om all form for prøving av studenter (Raaheim, 2019). Vurdering kan, ifølge Raaheim, (2019) grupperes i tre hovedkategorier: summativ vurdering, formativ vurdering og studentmedvirkende vurdering.

Denne studiens omdreiningspunkt er vurdering av eksamenskonsernten, det vil si vurdering av studentens måloppnåelse etter at undervisningen i emnet er avsluttet, altså summativ vurdering, en vurderingsform der hensikten er å oppsummere og kontrollere den kompetansen studenten har oppnådd. Waagen (2021) beskriver summativ vurdering som den formelle informasjonen til omverdenen om læringsutbyttets kvalitet.

Sadler (2005) sier at vurdering og karaktersetting ikke skjer i et vakuum, og at profesjonelle vurderinger alltid gjøres basert på ulike kontekster. En

faktor er det Sadler (2005) benevner som egenreferert vurdering der man sammenligner studentens tidligere prestasjoner og ferdigheter med den prestasjonen studenten viser på eksamenskonserten, det vil si at det som teller er hvilken forbedring studenten har gjort. Han sier at en slik vurdering ikke er anbefalt i høyere utdanning, men at det i praksis allikevel finner sted (Sadler, 2005). Normbasert vurdering (Sadler, 2005), gjort på bakgrunn av forventning om hva som kan beskrives som en god eller mindre god prestasjon, er en annen faktor. I vurdering av eksamenskonserter kan man se for seg at man i normbasert tilnærming vil sammenligne kandidatens prestasjon med andre kandidaters prestasjon innenfor samme instrument (Raaheim, 2019).

Norm-referenced assessment systems, in which there is more emphasis on relating the achievement of a particular student to the achievements of other students than on the specific nature of the achievements themselves. (Sadler, 1987, s. 192)

## Analytisk og holistisk vurdering

Analytisk og holistisk vurdering blir sett på som to ulike tilnærminger til vurdering. I en analytisk vurderingsstrategi vil man forme eksplisitte vurderingskriterier (Sadler, 2005) «Kriteriene blir gjerne formulert som mestringsutbytte: eleven skal kunne demonstrere at han kan *det* og *det*, vet *det* og *det* osv.» (Vinge, 2011, s. 204). Vurderingskriteriene bør gjøres kjent for studentene i forkant av vurderingssituasjonen, og gjerne ved oppstart av undervisningen. utfordringene ved å lage gode vurderingskriterier løftes frem hos blant andre Waagen (2021) og i RUM-rapporten om sensorveiledning: «Dette er som nevnt et vanskelig punkt, da konkrete kriterier aldri kan beskrive musikkutøvelsen som fenomen fullt ut» (RUM, 2018, s. 4). I rapporten (2018) refereres Newsome (2015), som gjør rede for ulike modeller for vurdering i høyere musikkutdanning. Det som er gjennomgående i den forskningen hun viser til er at ren kriteriebasert vurdering går glipp av viktige estetiske aspekter som det er umulig å verbalisere. «Newsome anbefaler derfor et kriteriesett som først utfordrer sensoren til å ta stilling til konserten som helhet, før det tas stilling til et sett med vurderingskriterier». (s. 3)

Sadler (1987) sier også at det er knyttet bekymring til bruk av kriteriebasert vurdering i fag der kvaliteten av studentens prestasjons best kan vurderes «only by direct qualitative human judgment» (s. 192). Vinge (2011) viser også til argumentasjon mot analytisk vurderingsstrategi i performative, kreative og skapende fag – prestasjoner som ikke «nødvendigvis kan bedømmes som rett eller galt, men snarere som karakteristisk for en eller annen form av kvalitet» (s. 204). I motsetning til en analytisk vurderingsstrategi der man har eksplisitte vurderingskriterier er det i holistisk vurderingsstrategi ikke en enkel del, men helheten i en prestasjon som vurderes (Olsson & Nielsen, 2018). Holistisk tilnærming forklares slik av Vinge (2011, s. 204):

Her legges ingen skriftlige kriterier, forstått som spesifikke distinkte egenskaper ved en kompetanse eller et læringsutbytte, til grunn i forkant av vurderingssituasjonen, men appliseres i etterkant med utgangspunkt i det aktuelle vurderingsobjektet som foreligger.

Kriteriene som ligger til grunn tilpasses i etterkant, og begrunnes ut fra en kombinasjon av den faglige kompetansen og vurderingserfaringen den eller de som vurderer besitter (Vinge, 2011). Det Vinge (2011) sier, er at man i en holistisk strategi begrunner vurderingen «som elementer i kvalitativ *argumentasjon*» (s. 210). Eller, som Olsson og Nielsen (2018) sier: «Den holistiska utgångspunkten bygger på en kvalitativ helhesbedömning där olika faktorer viktas till en samlad uppfattning» (s. 13).

## Skjønn i vurdering

Skjønn har opphav i det gammelnorske ordet *skyn*, som betyr ‘dømmekraft, forstand, vurdering, vett’, og knyttes blant annet til taus kunnskap, praktisk kunnskap, intuisjon, holistisk tenkning og det sanselige. For å belyse ulike syn på bruken av skjønn i vurderingssammenheng støtter jeg meg på Harald Grimen (2009). Han viser blant andre til den engelske statsviteren og filosofen Robert Goodins syn på at utstrakt bruk av skjønn er uforutsigbart og ikke etterprøvbart. Skjønnsutøvelse blir sett på som en motsetning til en vurdering man kan argumentere for (Grimen, 2009). Men skjønnsutøvelse kan også forstås som vurdering gjort innenfor gitte

rammer, og som kan argumenteres for på bakgrunn av kunnskap, innsikt og erkjennelse.

Det er ingen gode grunner til å meine at argument berre verkar om den som argumenterer står under ytre tvang. Det er heller ingen gode grunner til å meine at, når ein person ikkje står under ytre tvang, vil han utelukkande la seg leie av kjensler, ideologiar, makt, eigeninteresse, gruppeinteresser og liknande, og ikkje av argument som han finn overtydande, eventuelt av upartisk argument. (Grimen, 2009, s. 4)

Gynnild (2010) sier at man i vurdering i kunstneriske og utøvende disipliner trenger et tolkende mellomledd, det han også kaller et profesjonelt skjønn. Jeg forstår det slik at profesjonelt skjønn her knyttes til profesjonell kunstnerisk og utøvende erfaring, og at man som ekspert på feltet vurderer prestasjoner på bakgrunn av erfaring og gjenkjennelse av kvalitet.

Mens holistisk og skjønsmessig vurdering kan argumenteres for med utgangspunkt i faglig og kunstnerisk kompetanse og erfaring, kan det se ut som den subjektive vurderingen er mer preget av «en personlig, individuell oppfatning grunnlagt i smak, interesser og erfaring hos den som vurderer og vil følgelig kunne variere fra person til person» (Vinge, 2011, s. 207).

Orr (2010) ser på den subjektive tilnærmingen til vurdering som foregår i et praksisfellesskap. Det vil si at sensorgruppen kan operere i et praksisfellesskap der man til en viss grad deler subjektivitet.

Thus, lecturers, to a certain extent, share a common subjectivity (or inter-subjectivity) that is continually negotiated, contested, agreed and developed in the context of studio-based group marking dialogue. (Orr, 2010, s. 16)

Eller som Raaheim (2019) sier: «Gjennom deltagelse i fellesskapet fikk vi etter hvert en føling med hva som skilte de gode fra de mindre gode prestasjonene, og utviklet på den måten en slags professoral magefølelse» (2019, s. 45).

## Metode

For å undersøke hvilken rolle sensorveiledningen har i vurderingen, og i hvilken grad sensors og sensorgruppens subjektive blikk og faglig

skjønn anvendes i vurderingsarbeidet, har jeg foretatt semistrukturerte kvalitative intervjuer med fem sensorer med utgangspunkt i en tilnærming som handler om å få innblikk i intervjupersonenes opplevelser, ønsker og meninger (Kvale & Brinkmann, 2022). Kvale og Brinkmann (2022) beskriver intervjuet som en instrumentell dialog der samtalen er et middel for å få frem intervjupersonens mening og erfaring om et bestemt tema. Når jeg bruker intervju som metode, ligger det potensielt en svakhet i at det som kommer frem ikke alltid gir et fullstendig bilde av det som faktisk skjer. Det kan være at den enkelte informant ikke er fullt bevisst på hva som faktisk ligger til grunn for sin vurdering, og at det finnes tause kriterier som i større eller mindre grad er sammenfallende med kriteriene i sensorveiledningen. Med de forbehold som naturlig ligger i dette, er det intervjusamtalene som danner empirien i denne studien.

Før intervjusamtalen hadde jeg forberedt noen intervju spørsmål som var designet for å anspore til en samtale rundt studiens problemstilling. Til tross for disse spørsmålene både ville og måtte jeg la det være rom for spontanitet, refleksjoner og utdypninger (Ryen, 2002). Ved å holde en fleksibilitet i samtalen kunne jeg følge opp emner og temaer som kom opp underveis. Intervjusamtalen var ikke utelukkende et middel for å få innblikk i intervjupersonenes erfaringer og meninger, men den ga også rom for refleksjon rundt hendelser og temaer som ble berørt i samtalen, og var en læreprosess for både informanten og intervjueren (Kvale & Brinkmann, 2022). Med bevissthet om at intervjusamtalen er et forskningsverktøy, velger jeg å bruke betegnelsen *samtale* om intervju videre i teksten.

Da jeg skulle finne fagpersoner som kunne bidra til å belyse min problemstilling, var det nødvendig å bringe inn personer med erfaring fra sensur i høyere musikkutdanning, både som eksterne og interne sensorer. For å få ulike perspektiver på problemstillingen ønsket jeg informanter både fra klassisk og rytmisk, og en viss kjønnsbalanse. Prosjektet ble godkjent av Sikt, og jeg sendte forespørsel og informasjonsskriv om prosjektet til ni personer som har hatt sensoroppdrag ved UiT Norges arktiske universitet. Noen av disse hadde både hatt oppgaver som intern sensor ved UiT og vært ekstern sensor ved andre læresteder i Norge. Av

de fem som svarte på min henvendelse, var alle positive til å delta, og de syntes temaet var svært interessant og viktig å belyse.

Det empiriske materialet i artikkelen er basert på samtaler med tre menn og to kvinner, der to har bakgrunn fra rytmiske sjangre og tre har bakgrunn fra vesteuropeisk kunstmusikk. Samtlige var fagpersoner med erfaring som profesjonelle musikere og undervisere i høyere musikkutdanning, og alle hadde erfaring som sensorer ved ulike læresteder nasjonalt. De hadde også erfaring med å være med i sensurering som veileder/faglærer til eksamenskandidater. En av de jeg hadde samtaler med, uttrykte å ha kortere erfaring med eksternt sensur, og tre oppga å også ha erfaring fra sensurarbeid utenfor Norge og Norden. Informantene hadde alle hatt sensoroppdrag eksternt og internt i løpet av de to siste årene, og samtlige hadde fått tilsendt sensorveiledninger før sensoroppdraget etter at sensorveiledninger i høyere musikkutdanning ble innført.

Av praktiske, klimamessige og økonomiske årsaker ble tre av samtalen gjennomført digitalt ved hjelp av video- og konferanseprogramvaren Zoom, og to ved fysiske møter. Tre av informantene var ansatt ved UiT, og vi hadde dermed et kollegialt forhold uten at vi hadde hatt et nært faglig samarbeid. Det kan selvsagt ha gitt en annen kontekst til samtalen enn for de to som var knyttet til andre institusjoner. De to respondentene som ikke var tilknyttet UiT, hadde jeg to møter med, dette for å bli bedre kjent og for å få utdypet noen av de aspektene vi hadde berørt i første samtale. Jeg hadde også en kort oppfølgings samtale med en av informanten fra UiT.

Det ble satt av to timer til hver samtale, der to av disse fylte hele den avsatte tiden. Det ble gjort lydopptak av samtalen slik at jeg kunne være konsentrert om å lytte til den jeg snakket med, og samtidig få en best mulig gjengivelse av innholdet. Samtalene ble transkribert og brukes i artikkelen som grunnlag for analysen. Informantene er anonymisert og benevnes med de fiktive navnene Alf, Bård, Cecilie, David og Ellen.

Med så få informanter er det empiriske materiale relativt begrenset, og ikke egnet til å sammenligne i bredde. Det jeg ønsket, var å gå i dybden hos den enkelte informant for å få innsikt i de erfaringene den enkelte hadde med det å vurdere eksamenskonsserter. Informantene kommer fra et fagfelt jeg selv har jobbet med i mange år, og jeg har god kjennskap

til konsert som vurderingsform, og til de avveiningene sensor kan slite med i karaktersettingen. Min erfaring er at man i sensurering av eksamenskonsert i varierende grad forholder seg til sensorveiledningen, og at man som sensor langt på vei gjør vurderingen ut fra sin erfaring – det vil si at man som profesjonell musiker og pedagog har en oppfatning av hva som kreves for å oppnå en høy karakter. Ekstern sensor er som oftest ikke alene i dette arbeidet, men kommer i samarbeid med intern sensor frem til hvilken vurdering studentens prestasjon skal få. Denne studien er naturlig nok preget av min forforståelse, noe som igjen har påvirket analysen av forskningsmaterialet (Dalen, 2011). Jeg valgte derfor en refleksiv tematisk analyse (TA), som baserer seg på en forskning som alltid vil være preget av subjektivitet, og som aldri er nøytral (Braun & Clark, 2023).

## Analyse

I arbeidet med analysene av materialet fra samtalene fulgte jeg en tema-basert kvalitativ analysestrategi (TA), som er en metode for å identifisere, analysere og undersøke meningsbærende tema i det totale datamaterialet (Braun & Clarke, 2008). Refleksiv TA-tilnærming forholder seg til forskerens subjektive blikk som forskningsressurs, og kodingen vil alltid være et resultat av forskerens tolkning av data (Braun & Clarke, 2023). Det er her snakk om at jeg som forsker frembringer temaer som treffer noe jeg ser som meningsbærende i sammenheng med studiens forskningsspørsmål.

I samtalene var jeg opptatt av å få innblikk i hvordan den enkelte sensor forholdt seg til og brukte sensorveiledningen i sensurarbeidet. Når sensoren ga uttrykk for at man forholdt seg lite eller i varierende grad til sensorveiledningen spurte jeg om hva som da lå til grunn for vurderingen. Etter å ha transkribert hele datamaterialet, lest det grundig og sett etter mønstre og mening, laget jeg en liste med koder. Disse kodene representerte aspekter i datamaterialet som fremsto som interessante å analysere, og som var viktige med tanke på forskningsspørsmålet i studien. Med utgangspunkt i kodene utarbeidet jeg overordnede meningsbærende tema og undertema. Disse temaene er:

- Sensorveiledning som redskap i vurdering av eksamenskonserten
  - o Karakterbeskrivelser
  - o Kriteriene
  - o Usikkerhet og uenighet knyttet til sensuren
- Vurdering basert på skjønn og ekspertkompetanse
  - o Kunstnerisk helhet og formidling
  - o Vurdering i team der faglærer deltar

I den første delen ser jeg på hvordan informantene omtaler sensorveiledningen og sin bruk av den, problematisering rundt kriterier og hvilket verktøy de ser at sensorveiledningen er når det oppstår usikkerhet og uenighet knyttet til sensuren. I den andre delen beskriver jeg hvordan informantene omtaler vurdering på bakgrunn av skjønn og ekspertkompetanse. Som et underpunkt i den andre delen ser jeg på vurdering i sensorteam der faglærer deltar, dette fordi informantene ser at det er svært viktig og avgjørende i arbeid med karaktersetning. Når dette er et underpunkt begrunner jeg det ut fra at vurderingene som gjøres her også er basert på skjønn og ekspertkompetanse. De utvalgte empiriske eksemplene og sitatene har jeg trukket frem fordi de samlet sett er særlig godt egnet til å belyse artikkelens problemstilling.

## Sensorveiledning som redskap i vurdering av eksamenskonserten

### Karakterbeskrivelser

Selv om alle informantene fortalte at de hadde fått tilsendt sensorveiledning før sensoroppdraget, var enkelte uklare på hvordan de hadde forholdt seg til veiledningen. En av informantene gir uttrykk for at hun stort sett *ikke* forholder seg til sensorveiledningene.

Min grunnleggende erfaring er at jeg nok leser mer på emneplanen enn jeg bruker sensorveiledningen. Hva er det egentlig jeg forholder meg til? Jo, stort sett emneplanen, tror jeg. (Ellen)

Denne informanten synes veiledningen er nyttig når det gjelder det praktiske, slik som varighet og hvilket studieprogram studenten har gjennomført.

Sensorveiledningene inneholder vanligvis både informasjon om emnet, varighet, sensorpanelets sammensetning, vurderingskriterier, karakteruttrykk og vurderingsprosedyre. I samtalene fikk jeg til tider opplevelsen av at det først og fremst var snakk om karakterbeskrivelsene når informantene viste til veiledningen.

Når man er blitt mer dreven, dreven *sirkushest*, skummer man gjennom sensorveiledningen for å se om det er noe nytt siden sist. Om det er noen særskilte ting som skiller seg ut fra det vi har på vår institusjon, og det er stort sett det samme. De er jo nasjonale, disse kravene. (Cecilie)

På spørsmål om det er kravene knyttet til kriteriene for hver bokstavkarakter i karakterbeskrivelsene hun vise til, svarer hun bekreftende.

David sier at han alltid leser sensorveiledningen, men opplever at det ikke er alle som er like nøye med det. Han sier samtidig at man bruker sensorveiledningen veldig lite, man gjør vurderingen på rutine. Det som har å gjøre med karakterbeskrivelsen blir derimot brukt i vurderingen. Bård og David opplever at man stort sett bruker karakterene A, B og C.

Alle er så vant til det, før sensorveiledningene kom, alle er bevisste på hvor karaktersystemet ligger, særlig vi som har holdt på i mange år. [...] Vi som sensurerer for hverandre vet hva vi kan forvente. Er det en A så er det spesielt godt. B er ganske vid. Er det litt under forventet så er det en C. Det er stort sett de tre karakterene man bruker. (David)

Jeg synes det er hjelp og ha en eller annen form for veiledning – på en måte. Spørsmålet er om det hjelper å ha den veldig detaljert, og selvfølgelig er det vanskelig å ta karakterskalaen på alvor fordi man vet at man egentlig ikke kan gi en D. Skalaen stopper på en C. Det er det inntrykket jeg har. (Bård)

Bård sier: «Jeg vet ikke om jeg har vært med på å gi en kandidat en D.» Denne sensoren sier at selv om beskrivelsen av karakteren C viser til en god prestasjon er det kun for de dårligste eksamensprestasjonene man i det hele tatt diskuterer en C. Et annet perspektiv kommer fra informant David, som sier:

Kandidaten sto jo med et nødscrik, det var litt uenighet i juryen, hva som var riktig utfall, men vi ble nå enig om det til slutt. Om det hadde blitt slik hvis sensorveiledningen hadde vært fulgt er jeg usikker på. (David)

Det David sikter til er at man ikke alltid er så nøye med å følge prosedyren i sensorveiledningen, og er man ikke det, kan det få betydning for utfallet av vurderingen.

## Kriteriene

Det er jo vanskelig å tenke seg at man ikke skal ha en slik [sensor]veiledning, samtidig som det er et litt sånn klønete landskap å bevege seg i når man skal bruke ord på dette. [...] Det er et forsøk på å få en rund ball til å passe inn i et firkantet hull. (Ellen)

Det Ellen snakker om, er at kriterier som for eksempel *høyt nivå, musikkalsk bevissthet og formidling* er vanskelig å bedømme: «Det er vanskelig å beskrive musikk med ord» (Ellen). Hun mener, på lik linje med noen av de andre informantene, at kriteriene er ambisiøse og lite praktisk anvendbare. Jeg spør David om det er slik at man forholder seg til vurderingskriteriene i sensorveiledningen:

Ja, det må man, man skal jo det, og det opplever jeg at vi gjør. Men vi går ikke inn på alle kriteriene, det blir nesten umulig å vurdere. Man må ta det som en helhet. (David)

En av de andre informantene svarer ikke på hvordan han bruker sensorveiledningen, men problematiserer rundt kriteriene og hvordan man i sensorarbeidet må være enige om hvordan kriteriene skal rangeres og brukes:

Du kan egentlig detaljkjøre et opplegg på en sånn veiledning – sitte på kontoret å få den så korrekt som mulig, men så ryker det på føreste spadetak fordi det er så ulike perspektiver på folk, og da kommer erfaringen inn. [...] Det jeg synes har vært bra med disse veiledningene er, hvis det kommer et fokus på – det har vi snakker om *konserten som helhet, den enkeltes teknikk* og så videre. Så har jeg etterlyst av og til – «skal vi ha prioritert rekkefølge, slik at konserten som helhet er det aller viktigste»? Nei, det var de ikke enig om, og da er vi liksom like langt. (Alf)

Bård snakker også om det å enes om hvilke kriterier som skal brukes, og hvordan disse skal vektas mot hverandre og hvilken betydning det skal få

for karaktersettingen. Måtene kriteriene vektes opp mot hverandre på er også Ellen inne på når hun sier: «Jeg er ikke så opptatt av instrumentferdighet som av musikalsk tilstedeværelse.»

Det kan se ut til at informantene fra rytmisk har en oppfatning av at det lettere å lage vurderingskriterier i klassisk musikk enn i de rytmiske sjangerne. Det kommer til uttrykk hos Alf, som sier: «Jeg tror det er lettere å sette opp parametere for vurdering i klassisk musikk. I rytmisk musikk er sjangerne så ekstremt ulike.» Ellen tror det er vanskeligere å vurderer på rytmisk «fordi det ikke er noen klare preferanser. Man leter etter individualiteter».

## Usikkerhet og uenighet knyttet til sensuren

Jeg synes sensorveiledning er veldig viktig, særlig når det er tvil om kandidaten skal stå eller ikke. (David)

I datamaterialet går det frem at sensorveiledningen med kriteriene og karakterbeskrivelsene brukes som en støtte i sensurarbeidet når sensorpanelet er uenig, det er usikkerhet knyttet til hvilken karakter som skal gis, eller om kandidaten skal stå. «Kriteriene er ikke absolutte – det er ikke matematikk», sier David. Cecilie trekker frem at sensorveiledningen fungerer som en sikkerhet for sensoren i de tilfellene det kommer til *disputt eller uenighet*: «Har man fulgt det som står i veiledningen, har man sitt på det tørre.»

Sensorveiledningen er det eneste håndfaste dokumentet man har hvis det er veldig stort sprik mellom faglærer og sensor. Det er det eneste du har å slå i bordet med, og si at ifølge denne veiledningen så kvalifiserer ikke kandidaten til den og den karakteren. Sånn sett er den et veldig hendig verktøy, men det er synd når man må bruke den sånn, når man ikke blir enig. (Cecilie)

Denne sensoren beskriver også hvordan hun bruker sensorveiledningen i situasjoner der hun selv er usikker: «Er jeg i tvil så bruker jeg jo den, og på en måte argumentere meg selv frem til det jeg tenker kan være et rettferdig og realistisk resultat» (Cecilie). Selv om Cecilie sier at bruken av sensorveiledningen kan bidra til et rettferdig og realistisk resultat, mener

hun at eksamenskonserten i utgangspunktet ikke er en rettferdig situasjon. Det hun sikter til, er at eksamenskonserten ikke nødvendigvis viser kandidatens faktiske nivå, dette fordi kandidaten ikke alltid klarer å yte sitt beste på eksamensdagen.

Slik jeg forstår informantene fungerer sensorveiledningen som en rettesnor sensor kan bruke i diskusjon med de andre i sensorpanelet.

Man ser jo på sensorveiledningen som en hjelp til ikke å begynne å gå for mye i det irrelevante eller det subjektive. Man har en rettesnor, men det er jo noe som står i de sensorveiledningene som kriterier som det er vanskelig å forholde seg til. (Bård)

Informantene har til dels sammenfallende oppfatninger om hvordan de selv som sensorer og andre sensorer forholder seg til de aktuelle sensorveiledningene, og hvilken betydning kriteriene og karakterbeskrivelsene har i vurderingsarbeidet. Informantene forteller hvordan sensorveiledningen kan være et redskap for å komme til enighet om hvilken vurdering man skal gi en studentens prestasjon. Uenigheten, slik jeg oppfatter det, har utspring i at vurdering i kunsthøgskolen sammenheng i stor grad preges av skjønn (Gynnild, 2010) og den enkelte sensors subjektive opplevelse. I tillegg kan det oppstå uenighet fordi faglæreres eller veileders kjennskap til kandidatens faglige utvikling, arbeidsinnsats og ferdighetsnivå som ikke avdekkes i eksamenskonserten, ofte blir et tema i sensurmøte.

## Vurdering basert på skjønn og ekspertkompetanse

### Kunstnerisk helhet og formidling

Informantene gir alle uttrykk for at man ikke kommer bort fra at det til syvende og sist er subjektive og skjønnsmessige vurderinger som ligger til grunn for hvor prestasjonen plasseres på karakterskalaen, eller om det er bestått eller ikke bestått. «Det som er komplisert i kunsthøgskolen er at det ikke er matematikk, det blir mye synsing og personlige preferanser» (Ellen). Selv om kriteriene brukes, tolkes de ulikt. Det subjektive og det skjønnsmessige ligger i at man har ulik erfaring og kompetanse, man lar seg berøre

og begeistre, man har ulik smak, konteksten påvirker på ulikt sett, man hører og legger vekt på ulike ting, og man forholder seg til et sett av normer. Eksamenkonserten blir sjelden vurdert og begrunnet der eksplisitte kriterier får avgjørende betydning, men blir vurdert som en helhetlig prestasjon – en konsertopplevelse. Alf stiller spørsmål ved hvor den enkeltes vurdering av hva som er fremragende ligger, og Bård spør seg om man kan gi en god karakter til noe man ikke liker. Cecilie uttrykker det slik:

Det vanskeligste – du har en kandidat foran deg som egentlig gjør disse tingene som står både i veiledningen og i kriteriene ikke sant, men du liker det bare ikke. Det rører deg ikke, det griper deg ikke, det gir deg ingenting, sant ...  
(Cecilie)

Når jeg spør Cecilie om hun klarer å sette ord på hva som berører eller ikke berører henne, svarer hun: «Det er stort sett formidlingsevnen – det er de små nyansene, de små tingene som ligger i det udefinerte» (Cecilie). Bård trekker også frem at formidlingsevnen er en viktig kvalitet når han sier: «Det må jo være et visst minimum av tekniske ferdigheter, men hvis jeg har følelse av formidlingsevne og vilje – musikalsk engasjerende» (Bård).

Hvis jeg føler at jeg har fått en kunstnerisk opplevelse og en flott konsert, da er du inne på kunstnerisk helhet, formidling – og det trekker jo veldig opp vil jeg si – tenker at jeg må forsøke å stille meg åpen – er dette en konsertopplevelse jeg vil huske? (Alf)

Når Alf her viser til *kunstnerisk helhet* og *formidling* som grunnlag for vurdering kan det forstås som at han forholder seg til de tematiske kriteriene i sensorveiledningen, men at vurderingen gjøres på bakgrunn av skjønn. Alf snakker om en følelse eller en kunstnerisk opplevelse som gir grunnlag for vurdering, og disse aspektene er det vanskelig å romme i et kriteriesett. Ellen er også inne på dette når hun sier at det at konserten «treffer» deg, er avgjørende for karakteren du lander på.

Jeg tenkte, «dette er en A», men hvis jeg plukket det i fra hverandre, hver enkelt ting så var det ikke eksepsjonelt på noen ting, det var bare det at jeg var på konsert og tenkte ikke på noe, det var ingenting som skurret. Og da tenker jeg at det er for meg eksepsjonelt. (Ellen)

Det er helheten som teller – en holistisk vurdering der man ikke vurderer enkeltdeler i prestasjonen, men den helhetlige opplevelsen. Dette er også noe som Bård er inne på når han sier: «Man kan nevne detaljer som kan være mangler, men man kan allikevel få en god karakter. Det er helheten som teller» (Bård).

Flere av sensorene her har lang erfaring både som utøvere, pedagoger og i sensurering av eksamenskonsserter. Cecilie fremhever aspektet ved faglig skjønn på bakgrunn av erfaring, erfaringsbasert kunnskap: «Nå begynner jeg å bli en gammel ringrev i faget, jeg har gjort dette en del år, så mye av dette er erfaringsbasert. Du vet hva du skal høre etter, og du vet hva som er forventet» (Cecilie). David sier at «man gjør det på rutine for man er så vant til det, man bruker sensorveiledningen veldig lite». Både han og Cecilie sikter til er at man, selv om man har lest sensorveiledningen, ikke bruker den med begrunnelse i at man i stor grad har rutiner og erfaringer som gjør at man fort blir enig om sensuren. Man vet hva man skal høre etter, og selv om det ikke er eksplisitte kriterier som ligger til grunn, kan vurdering begrunnes ut fra en kunstfaglig praksis. Problemstillingen ved vurdering på bakgrunn av ekspertkompetanse; skjønnsmessig vurdering der ekstern sensors kompetanse, erfaring og holdning legges til grunn for vurdering, løftes også frem i samtalen med Alf, som sier: «Jeg tror det er viktig å ha sensorer som er veldig åpne fordi musikklivet er i endring. [...] Jeg har valgt å være litt selektiv med hensyn til hvilken ekstern sensor jeg vil ha.»

Denne sensoren gir uttrykk for at man som sensor, hvis man ikke kjenner en sjanger eller ikke har forholdt seg til et spesifikt musikkuttrykk, kan få problem med å høre kvaliteten i den prestasjonen kandidaten leverer. «Da blir det kanskje valg av sensor som blir feil» (Alf). Alf mener her at det ikke hjelper med kriterier fordi man som sensor tolker kriteriene og vurderer kvalitet med utgangspunkt i det man selv har erfaring med. Valg av sensor er også et tema i samtalen med Cecilie, som mener at valg av sensor er viktig for å få et best mulig resultat.

## Vurdering i team der faglærer deltar

Sensorgruppen, som i de fleste tilfeller består av team der faglærer deltar, bruker skjønn og kunstnerisk kompetanse i vurderingen av

eksamenskonserten. Bård snakker om at sensorgruppen bør bli enige på forhånd når det gjelder hva man skal forholde seg til, og at sensorveiledningen da er en rettesnor for både ekstern sensor og for den interne (som ofte er studentens faglærer/instrumentallærer): «Som i et svømmebaseng, man må ta tak i kanten for å stabilisere seg litt. [...] kaste et blikk på kriteriene, og ikke bli for personlig engasjert på feil måte» (Bård).

Det Bård blant annet sikter til er at faglærer som en del av sensorpanelet ofte har en avgjørende stemme i vurderingsarbeidet. I samtale med Cecilie blir også rollene som faglærer og ekstern sensor belyst. Som faglærer vet du at studenten gjorde det kjempebra på generalprøven, men ikke klarer å innfri like godt på eksamen: «Så kommer det inn en ekstern person som setter en karakter som jeg tenker ikke står i forhold til den innsatsen studenten har gjort, og hva jeg mener studenten er god for» (Cecilie). Hun fremholder at det er vanskelig å være helt objektiv eller nullstilt i sensureringen av eksamenskonserten, blant annet fordi du som ekstern sensor enkelte ganger sensurerer noen du har hørt før, og kan høre om det har vært progresjon eller om det har gått tilbake. Hun opplever at det ofte er et stort sprik mellom det hun som ekstern sensor hører og det faglærer hører:

Til syvende og sist blir det en subjektiv vurdering, det er nesten uunngåelig, uansett hvor mye veiledning du har – man kan si at det står i veiledningen – men jeg kan si: «men jeg hørte ikke det». Så kan faglærer si at det var helt tydelig – det gjorde vedkommende, men nei. (Cecilie)

Ellen gir uttrykk for at det både er bra og problematisk å ha så tett dialog med faglærer før og etter konserten:

Den verdien av at det kommer utenfra ... Jeg tror det er ulikt i hvor stor grad ekstern sensor blir instruert. Det kan hende jeg ville vært begeistret, eller mer begeistret av noen, men så har jeg fått høre så mye på forhånd at denne studenten er sånn og sånn, han har gjort det og det, så jeg tar det med meg inn, og ikke tillater meg selv å være nøytral. Så de personlige opplysningene tror jeg man er vel så mye farget av som av sensorveiledningen. De begrepene er liksom tomme. Det er ikke målbart fra null til ti uansett. Opplevelsen blir påvirket av den informasjonen man som sensor får [...] Jeg tror jeg er vel så mye preget av faglærerne som kjenner studenten godt, og at det er både bra, men det kan også være skummelt. (Ellen)

Ellen reflekterer videre over at sensorveiledningen kanskje er vel så viktig som en påminnelse til intern sensor/faglærer fordi man lett blir «blindet» når man har jobbet med en student over lang tid. Ellen tror at ekstern sensor kan se verdier som faglærer ikke oppdager, fordi vedkommende har jobbet så tett med kandidaten over lang tid.

I datamaterialet har jeg funnet at man, selv om man forholder seg til vurderingskriteriene, begrunnes vurdering av eksamenskonserten ut fra skjønn og ekspertkompetanse, og at faglærers rolle i sensorpanelet kan gjøre det utfordrende å kun forholde seg til studentens prestasjon der og da. Som Alf sier: «Du blir jo en slags omsorgsperson for de du skal representere; forsvare.» Det Alf sikter til er den rollen han gjerne har når han deltar i sensorpanelet som faglærer. Det betyr ikke at Alf ikke bruker sin fagkompetanse i sensorgruppen, men som faglærer sitter han inne med et større bilde av hvordan studenten har jobbet i sin utdanning.

## Drøfting

I det foregående har vi sett at samtalene rundt sensorveiledningen som redskap i vurderingen av eksamenskonserten berører temaer knyttet til formålet med sensorveiledningen. I rapport om sensorveiledninger i høyere musikkutdanning (2018) skriver forfatterne at:

Ved å gi et sett vurderingskriterier og definerte karakterbeskrivelser, samt foreslå en mal for hvordan samtalen i sensorpanelet og i møtet med kandidaten i etterkant skal foregå, tror arbeidsgruppen at sensorveiledninger kan bidra til å nå målene som etterspørres i kvalitetsmeldingen. (RUM, 2018, s. 7)

Det kvalitetsmeldingen etterspør, er at man skal sikre en upartisk og rettferdig karaktersetting, og at sensorveiledningen vil kunne bidra til kvalitetssikring av det fastsatte læringsutbyttet, og til å redusere ulikhet i sensur mellom sensorer (Meld. St. 16 (2016–2017)).

I forskningsmaterialet går det frem at skjønn; det at man blir berørt og begeistret, erfaring som utøver, pedagog og sensor ser ut til å spille viktig rolle i vurderingen. Slik vurderingskriteriene i sensorveiledningen ved UiT er formulert, gir de etter mitt syn rom for tolkning. Kriterier som har å gjøre med instrumentbeherskelse kan fremstå som skarpe kriterier

(Olsson & Nielson, 2018), mens vurdering av kriterier som går på konserten som helhet, musikalsk bevissthet og kommunikasjon kan ha et større tolkningsrom og fremstå som mer uskarpe. Informantene i denne studien, som har kompetanse innenfor den rytmiske sjangeren, sier at det kanskje er lettere å lage eksplisitte kriterier for å vurdere en prestasjon innenfor klassisk musikk. Dette begrunnes med at den klassiske tradisjonen har mer uttalte regler for hvordan man skal spille et verk, både med hensyn til teknikk og tolkning. Forskningsmaterialet viser at sensorer fra både klassisk og rytmisk musikk i stor grad ser bort fra detaljer i vurderingsarbeidet. Selv om sensorer kan legge vekt på instrumentferdighet og trekke frem detaljer som kan være mangler, er det en utbredt oppfatning at kandidaten likevel kan få en god karakter, fordi det er helheten som teller. Dette samsvarer med det Vinge (2011) beskriver som holistisk vurdering, og er også et tema som drøftes av Gynnild (2010), som fremholder at det ikke nødvendigvis er sammenheng mellom høy score på det enkelte kriterium og hvordan prestasjonen fremstår som helhet.

Ambisjonene om en mer transparent, argumenterende og pålitelig vurdering av eksamenskonserten i høyere musikkutdanning, som resultat av at man har kriterier som grunnlag i sensurarbeidet (Waagen, 2021), kan fremstå som utfordrende. En vurdering tatt ut fra en subjektiv opplevelse kan være vanskelig å forklare ut over at man, som enkelte informanter uttalte, ble begeistret og berørt. Vurderingen kan selvsagt gjøres transparent og forståelig for studenten gjennom at kriteriene som ligger til grunn tilpasses i etterkant, og begrunnes ut fra en kombinasjon av sensors faglige kompetanse og vurderingserfaring. Hvis tanken er at studenten skal få tilgang til vurderingskriteriene i forkant av vurderingen (Vinge, 2011), vil sannsynligvis ikke en tilpassing av kriterier i etterkant oppleves som uproblematisk, hvis de da ikke kan begrunnes ut fra kriteriene som er skissert i veiledningen.

I mal for sensorveiledningen og i sensorveiledningene som er i bruk ved flere av lærestedene står det at «eksamensprestasjonen alene skal ligge til grunn for karakteren, med mindre noe annet er oppgitt i emneplanen» (RUM, 2018, s. 8). Det er her snakk om en summativ vurdering (Raaheim, 2019), der prestasjonen under eksamenskonserten skal vise hvilket utbytte studenten har i hovedinstrumentemnet, og som igjen skal

gi grunnlag for sensorpanelets karaktersetting. Sensur av eksamenskonsertter skjer gjennom diskusjoner i sensorpanelet, og man kan oppleve at tolkning av kriterier skaper frustrasjon og usikkerhet (Gynnild, 2022). For meg kan det se ut som at det ikke bare er tolkning av kriterier som er problematiske i vurdering av eksamenskonserten. Vurderingen av prestasjonen med hensyn til karakternivå eller standard (Gynnild, 2010) kan mulig skape usikkerhet og uenighet i sensorpanelet. Informantene i denne studien forteller at de stort sett bruker de tre høyeste karakterene når vurderingsuttrykket er gradert karakter, og dette samsvarer med det som kommer frem hos Gynnild (2010). Han sier at det karakternivået ved utøvende musikkutdanninger er kjent for å være høyere enn det som var forutsetningen ved innføringen av den graderte karakterskalaen. (Gynnild, 2010). Sensorveiledningen tar ikke stilling til i hvilken grad kriteriene innfris, og derfor må sensorgruppen i fellesskap diskutere og bli enige om en felles forståelse for hva man ser som godt eller mindre godt (Orr, 2010). Med faglærer som del av sensorgruppen er det i tillegg utfordrende, og kanskje ikke ønskelig, å se bort fra den kunnskapen man har om kandidaten. Dette betyr ikke at faglærer ikke evner å vurdere kvaliteten på det som skjer på scenen der og da, men man kan heller ikke se bort fra at faglærer, som kjenner kandidatens utvikling og tidligere prestasjoner, preges av denne kunnskapen når vurderingen skal skje. Som eksterne sensorer er det vanskelig å ikke forholde seg til historien som fortelles om kandidaten. På den måten spiller faglærer ofte en avgjørende rolle i vurderingen, noe som kan slå begge veier for kandidaten. I tilfeller der faglærer og student har utviklet et problematisk forhold vil faglærers medvirkning kunne få negative konsekvenser for den bedømmelsen kandidaten får. Betegnelsen «epistemic authority» (Oltedal, 2020, s. 179) beskriver hvordan man i et sensorpanel bruker ulike ressurser i forhandlingen rundt karaktersetting, og at kunnskap om kandidatens utvikling og arbeidsinnsats kan brukes til å påvirke vurderingen. Vurderingen kan da skje på bakgrunn av studentens utvikling (Sadler, 2005) og på tidligere prestasjoner, og ikke alene på bakgrunn av prestasjonen der og da. Det at man skal se bort fra utviklingen og det man har fordypet seg i trekkes frem som utfordrende og lite ønskelig, som informanten Cecilie, som ut fra et faglærerperspektiv ser det som utfordrende når vurderingen

ikke står i forhold til den innsatsen studentens har gjort; altså det denne sensoren mener studenten er god for.

Ut fra forskningsmaterialet og tidligere forskning (Gynnild, 2010; Oltedal, 2020; Orr, 2010) ser det ut til at uenighet innad i sensorpanelet ikke er uvanlig, og at rollen man har i panelet kan påvirke hvilken posisjon man tar og hvilken vurderingsform man ønsker å legge vekt på (Oltedal, 2020). Forskningsmaterialet i denne studien viser at sensorveiledningen ses som et nyttig formelt rammeverk som særlig kommer til anvendes i tilfeller der det oppstår konflikt og usikkerhet. I så måte innfris intensjonen om at «en klar regi på sensurarbeidet kan bidra til å forebygge eventuelle konfliktsituasjoner» (RUM, 2018, s. 7). Sensorveiledningen, med beskrivelse av prosedyren, kriteriene og karakterbeskrivelsen, brukes da i argumentasjonen for den vurderingen sensor ønsker å lande på. Ut fra datamaterialet i denne studien ser det allikevel ut til at sensorpanelet i de fleste tilfeller kommer til enighet. At man stort i stor grad kommer til enighet, kan begrunnes med at man som utøvende kunstner og erfaren pedagog og sensor møtes i et praksisfellesskap der man, som Orr (2010) sier, deler en viss grad av subjektivitet. Flere av informantene i denne studien har lang erfaring med vurdering av eksamen i utøvende musikk, både som undervisere og sensorer, og de har en oppfatning av hvor nivået ligger og hva de kan forvente (Raaheim, 2019). Sensurering av eksamenskonsserter har også en sosial dimensjon, og det sosiale aspektet får også betydning for karaktersettingen. Oltedal (2020) løfter frem hvordan de sosiale prinsippene om en *hyggelig tone* i sensurarbeidet kan spille inn på karaktersettingen. Hvis sosiale hensyn kommer i konflikt med formelle rammer og faglige vurderinger kan det tenkes at sensor(ene) sitter igjen med opplevelsen av å ikke ha gjort en god nok jobb. Men, som informant Bård sier: «Jeg trøster meg litt med at alle som går ut fra en musikkutdanning må bevise hva de kan [...] sånn at en litt - ikke helt presist satt karakter - de kan ikke bare flyte på den.»

Norge er et lite land, og de som jobber innenfor samme fagområder kjenner hverandre godt. Det går frem av forskningsmaterialet at faglærere ofte selv velger hvem man vil ha inn som sensor. Man ønsker gjerne en som tenker likt som en selv, underviser ut fra samme metode, eller som har god kjennskap til den musikksjangeren kandidaten kommer til å presentere på konserten. Når en av informantene sier at valg av sensor er

viktig for et godt resultat, kan man spørre seg om hva som ligger i en god vurdering. Er det at man får en saklig og grundig vurdering fordi man kjenner sjangeren, stilen eller repertoaret godt, eller viser det til at man er grei og kjent for «å være grei» med hensyn til karaktergivning.

## Avslutning

I denne studien har jeg spurt hvilken rolle sensorveiledningen spiller i vurdering av eksamenskonserter og i hvilken grad skjønn og ekspertkompetanse anvendes.

Ut fra forskningsmaterialet i denne studien kan det se ut som at informantene opplever at det er opp til dem selv å avgjøre hvordan og om de bruker sensorveiledningen som redskap i sensurarbeidet. Når en av informantene i denne studien sier at man bruker veiledningen litt som man vil, blir jeg hengende ved spørsmålet om hvem som er ansvarlig for at veiledningen blir brukt, og hvorvidt den brukes etter hensikten. Sensorveiledninger med vurderingskriterier er en relativt ny ordning når det gjelder sensur av eksamenskonserter, og det kan tenkes at det kan ta tid å endre en kultur som har vært preget av «'taus' kunnskap og implisitte kriterier» (Gynnild, 2019, s. 1). Studien viser at sensorveiledning på den ene siden ses som et nyttig redskap, men at enkelte opplever at vurderingskriteriene kan være vanskelig å forholde seg til, og at det er utfordrende at kriteriene ikke er rangerte – en forståelse av hva som skal veie tyngst. I forskningsmaterialet kommer det frem at man sjelden bruker hele karakterskalaen. Hva som ligger bak denne praksisen, har jeg ikke gått nærmere inn på, men det kan tenkes at det har å gjøre med avveininger foretatt i sensorgruppen. Selv om det kan være utfordringer knyttet til sensorarbeidet, ser det ut til at sensorgruppen i de fleste tilfeller holder en god tone og blir enige om karakter på bakgrunn av en holistisk vurderingsstrategi der skjønn og ekspertkompetanse er avgjørende.

## Forfatterbiografi

**Lillian Jensen** er universitetslektor i sangdidaktikk og rytmisk vokal ved UiT Norges arktiske universitet, Musikkonservatoriet. I tillegg er hun

visedirektør for utdanning ved Norges arktiske universitetsmuseum og akademi for kunsthøgskolen. Hennes forskningsinteresse er knyttet til stemmen som instrument, sangdidaktikk og vurdering i utdanning. Jensen har lang erfaring som ensemblesanger og som sangpedagog innenfor teater, kulturskole, videregående skole, universitet og privat praksis.

## Referanser

- Andersen, H. L. (2010). «Constructive alignment» og risikoen for en forsimplende universitetspædagogik. *Dansk Universitetspædagogisk Tidsskrift*, 5(9), 30–35. <https://doi.org/10.7146/dut.v5i9.5568>
- Biggs, J. & Tang, C. (2011). *Teaching for Quality Learning at University* (4. utg.). Open University Press.
- Bjørndal, C. R. P. (2012). *Det vurderende øyet*. Gyldendal Norsk Forlag.
- Blix, H. S., Mittner, L. & Gjærum, R. G. (2019). I kjønne orden? Kvalitet og betydningen av kjønn i vurdering av professor- og dosentkompetanse i kunsthøgskolen. *Nordic Journal of Art and Research*, 8(1), 1–18. <https://doi.org/10.7577/information.v8i1.2853>
- Braun, V. & Clarke, V. (2008). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Braun, V. & Clarke, V. (2022). Toward good practice in thematic analysis: Avoiding common problems and becoming a knowing researcher. *International Journal of Transgender Health*, 24(1), 1–6. <https://doi.org/10.1080/26895269.2022.2129597>
- Dalen, M. (2011) *Intervju som forskningsmetode* (2. utg.). Universitetsforlaget.
- Grimen, H. (2009). *Skjønn som resoneringsform*. FORPRO. <https://profesjon.no/wp-content/uploads/2012/09/19124a5241fic7cf7.pdf>
- Gynnild, V. (2010). *Kriterier og skjønn i evaluering: En kasusstudie i utøvende musikkutdanning*. NOKUT. <https://kudos.dfo.no/dokument/6144/kriterier-og-skoen-i-evaluering-en-kasusstudie-i-utoevende-musikkutdanning-oktober-2010>
- Gynnild, V. (2022). «Jeg tror jeg brukte den som taus kunnskap, om du skjønner meg»: En studie av sensurveiledning på bachelornivå. *Uniped*, 45(1), 29–52. <https://doi.org/10.18261/uniped.45.1.5>
- Kunnskapsdepartementet. (2025). *Nasjonalt kvalifikasjonsrammeverk for livslang læring (NKR)*. [https://www.regjeringen.no/no/tema/utdanning/voksnes\\_laering\\_og\\_kompetanse/artikler/nasjonalt-kvalifikasjonsrammeverk/id601327/](https://www.regjeringen.no/no/tema/utdanning/voksnes_laering_og_kompetanse/artikler/nasjonalt-kvalifikasjonsrammeverk/id601327/)
- Kvale, S. & Brinkmann, S. (2022). *Det kvalitative forskningsintervju* (3. utg.). Gyldendal.
- Meld. St. 16 (2016–2017). *Kultur for kvalitet i høyere utdanning*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-16-20162017/id2536007/>

- Olsson, B. & Nielsen, E. (2018). Tolkningspraktiker och bedömningskulturer – ett projekt om bedömning inom högre musikalisk utbildning. I E. Georgii-Hemming, Ø. Varkøy, A. Kallio & F. Pio (Red.), *Nordisk musikkpedagogisk forskning. Årbok 10* (s. 9–23). Norges musikkhøgskole.
- Oltedal, E. (2020). Low stakes and high temperatures: A study of music performance assessment in teacher panels. I K. Smith (Red.), *Validity and value of teacher education research* (s. 175–195). Fagbokforlaget.
- Orr, S. (2010). We kind of try to merge our own experience with the objectivity of the criteria: The role of connoisseurship and tacit practice in undergraduate fine art assessment. *Art Design & Communication in Higher Education*, 9(1), 1–32. [https://doi.org/10.1386/adch.9.1.5\\_1](https://doi.org/10.1386/adch.9.1.5_1)
- Ryen, A. (2002). *Det kvalitative intervjuet. Fra vitenskapsteori til feltarbeid*. Fagbokforlaget.
- Rådet for utøvende musikkutdanning. (2018). *Rapport om sensorveiledninger i høyere musikkutdanning*. <https://s3.eu-north-1.amazonaws.com/nmh-nettsted/files/Studier/Sensorveiledninger-i-h%C3%B8yere-musikkutdanning.pdf>
- Raaheim, A. (2019). *Eksamensrevolusjonen* (2. utg). Gyldendal Norsk Forlag.
- Sadler, D. R. (1987). Specifying and promulgating achievement standards. *Oxford Review of Education*, 13(2), 191–209. <https://doi.org/10.1080/0305498870130207>
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading. *Higher Education Assessment & Evaluation in Higher Education*, 30(2), 175–194. <https://doi.org/10.1080/0260293042000264262>
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education*, 14(3), 281–294. <https://doi.org/10.1080/09695940701591867>
- Universitets- og høyskoleloven. (2024). *Lov om universiteter og høyskoler* (LOV-2024-03-08-9) Lovdata. <https://lovdata.no/lov/2024-03-08-9>
- Vinge, J. (2011). Analytiske og holistiske tilganger til vurdering – en teoretisk drøfting av vurderingsstrategier og vurderingshandlinger i musikkfaget. I S.-E. Holgersen, S. G. Nielsen & L. Väkevä (Red.), *Nordisk musikkpedagogisk forskning. Årbok 13* (s. 199–220). Norges musikkhøgskole. <http://hdl.handle.net/11250/172674>
- Waagen, W. (2021). Vurderingskriterier i utøvende musikk – et verktøy for pålitelig vurdering av studentenes bachelorkonsert? I E. Angelo, J. Knigge, M. Sæther & W. Waagen (Red.), *Higher education as context for music pedagogy research* (s. 41–64). Cappelen Damm Akademisk. <https://doi.org/10.23865/noasp.119.ch2>

## KAPITTEL 6

# Læringsfremmende vurderingsformer i musikk lærerutdanning

*Bengt Arve Haugseth og Hilde Synnøve Blix*

UiT Norges arktiske universitet

**Abstract**

Assessment strategies in higher education in Norway is a subject of growing interest due to national ambitions to develop students' independence, curiosity, critical abilities, and knowledge through student-centered learning and teaching. In this article we investigate assessment in music teacher education in Norway through qualitative group interviews with newly graduated teacher-students. We asked the students to what extent assessment strategies were perceived as support for learning. The study shows that boundaries between assessment and teaching seem to be fuzzy, something which was seen as both positive and negative regarding assessment for learning. We also found that a sense of ownership in relation to the assessment processes seemed to be an important condition for the assessment forms to be considered a support for learning, especially in the music performance subjects. The study suggests that more emphasis on assessment competence in music teacher education is warranted, including a more open dialogue about the different phases and shapes of assessment in the course of a music teacher education.

**Keywords:** assessment strategies, music teacher education, formative assessment, ownership

Sitering: Haugseth, B. A. & Blix, H. S. (2025). Læringsfremmende vurderingsformer i musikk lærerutdanning. I H. S. Blix, K. J. Valkoinen & B. A. Haugseth (Red.), *MusPed:Research: Vol. 11. Vurderingsformer og tilbakemelding i kunstfaglig utdanning* (s. 157–184). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.269.ch6>  
Lisens: CC BY 4.0

Vurdering er en vesentlig del av et utdanningsforløp og kan ha mange former og forskjellige målsettinger. I høyere utdanning skal vurdering fungere både som tilbakemeldinger til studentene om deres styrker, progresjon og forbedringsområder, og som en attest på hva de har lært (Meld. St. 16 (2016–2017); Raaheim, 2019). I større grad enn tidligere vektlegges at vurderingsformene skal lede til læring for studenten, og *læringsfremmende vurderingsformer* er et prioritert område i dag (Meld. St. 16 (2016–2017); NOKUT, 2023).

Det politiske ønsket om at vurderingsformer i utdanningene skal ha læring som mål, kommer til uttrykk blant annet gjennom de seneste revisjoner av vurderingsforskriftene (Opplæringsforskrifta, 2024) hvor vi blant annet ser en tydeliggjøring av forskjellige vurderingsformer og bruken av disse, og i tillegg en økt vektlegging av *vurdering for læring* (i det følgende også kalt formativ vurdering) som et viktig prinsipp i planverkene (Gamlem & Smith, 2013; Meld. St. 16 (2016–2017); Utdanningsdirektoratet, 2019).

## Bakgrunn og problemstilling

Vår interesse for tematikken læringsfremmende vurdering er motivert av mange års arbeid i høyere musikk- og musikkpedagogutdanning, hvor vi har observert holdningsendringer hos både studentene og institusjonene med hensyn til vurdering og læring. Vi har også vært opptatt av hvordan de institusjonelle strukturene rundt vurdering påvirker kunstfaglige prosesser og læring i fag hvor ferdigheter trenger tid til å utvikles; fag som er sårbare for kantete kriterier og likhetstenkning. Vi er begge musikkutdannet og har erfaring fra mange av musikklivets praksiser, og vi ser et stort behov for vurderingsformer som *reelt* virker læringsfremmende, og som er spesielt tilpasset kreative, utøvende fag.

I denne studien undersøker vi derfor hvilke typer vurderingsformer som brukes i musikk i lærerutdanninger i Norge, og på hvilke måter lærerstudentene opplever musikkfagets vurderingsformer som læringsfremmende. Studien er basert på intervjuer med seks nyutdannede lærere i musikk fra to forskjellige norske læresteder, og vi tar utgangspunkt i

problemstillingen: *På hvilke måter opplever lærerstudenter<sup>1</sup> musikkfagets vurderingsformer som læringsfremmende?* Problemstillingen sees i lys av vurderingsteoretiske konsepter som utdanningskvalitet, eierskap, dialog og læring.

## Forskning på læringsfremmende vurdering

Generelt viser forskning på vurdering og læring at *tilbakemelding* fra lærer og/eller medstudenter er det som oppleves som mest læringsfremmende for studenten (Black & Wiliam 2009; Hattie, 2013; Hattie & Timperly, 2007). I norsk sammenheng intervjuet Harildstad (2022) fagskolestudenter i filmproduksjon om deres opplevelse av vurderingsformer som hadde som formål å fremme læring. Hun fant at studentene var opptatte av at tilbakemeldinger på prestasjoner var konkrete og bidro til arbeidet med å forbedre prestasjonene, noe som bekreftes i internasjonal forskning på vurdering (Hattie & Timperly, 2007). Steen-Utheim og Hopfenbeck (2018) viser til at tilbakemeldingene må gi mening for studenten, og at det er av betydning i hvilken grad studentene selv engasjerer seg og tar ansvar i tilbakemeldingsprosessen (s. 81). Når karakterer blir gitt i etterkant av eksamener, og selve prestasjonen dermed ikke kan forbedres, er studentene mindre interesserte i tilbakemeldingene (Taras, 2006). Harildstads (2022) studie viste også at studentene var lite opptatt av karakterer, spesielt var de skeptiske til karakterer i praktiske og kreative fag: «En muntlig vurdering uten karakter kan bidra til at det å gjøre både gode og dårlige erfaringer i læringsprosessen føles trygt» (s. 9). På samme måte understreker Taras (2006) at det er spesielt viktig å ha med gode, formative vurderingsrutiner når målet er å bygge opp «expertise in complex, multi-criterion work such as essays and dissertations» (s. 366). Dette kan også gjelde for kunstfaglige og pedagogiske vurderingsformer, som ofte er praktiske og inneholder komplekse kriteriesett.

Klasseromsforskning på vurdering viser nokså entydig at for at vurderingsformer skal være læringsfremmende, må de foregå kontinuerlig,

---

1 Vi har valgt å kalle informantene for *studenter* gjennom studien selv om tre av de intervjuede nylig var startet å jobbe som lærere. Dette fordi intervjuene i hovedsak dreier seg om studietiden, og de tre øvrige informantene var masterstudenter da intervjuene ble gjennomført.

være bevisstgjorte og systematiske, og de må følges opp (Hattie & Timperley, 2007; Wiliam, 2011). Gibbs og Simpson (2005) undersøkte forutsetninger for at vurdering og tilbakemeldinger skal kunne føre til læring. De konkluderte med at sentrale betingelser er at tilbakemeldingene bør komme kort tid etter arbeidet er gjort og ikke sentrere omkring karakterer eller studentenes individuelle egenskaper. Det må være god kvalitet på selve tilbakemeldingene, og det er viktig at studentene får anledning til å bruke tilbakemeldingene aktivt. All den tid målet er å gjøre studentene i stand til å overvåke og utvikle egen læring, bør vurderingspraksisene legge til rette for dette, både gjennom institusjonelle strukturer og i undervisningsdiskursene. Vinge (2014) undersøkte i sin doktoravhandling 14 musikk læreres vurderingspraksiser og betydningen satsingsområdet «vurdering for læring» hadde for deres musikkfag. Han viser hvordan undervisningens aktiviteter og innhold preges av skolens økende fokus på vurdering, og måten lærerne balanserer formelle krav med sine egne, fagets og elevenes behov og mål. Lignende resultater fant Almqvist et al. (2017) i sin studie av vurdering i musikkfaget i Sverige og Norge, og de drøfter i likhet med Torrance et al. (2007) hvordan konkretisering av læringsmål og vurderingskriterier kan komme til å *være styrende* for læringen.

Boud og Molloy (2012) viser til forskning på vurderingspraksisene til universitetslærere som tyder på at de som ser ut til å lykkes, bruker såkalt bærekraftig tilbakemelding (Boud & Molloy, 2012). Det vil si at de (a) involverer studentene til dialog om læring, (b) hjelper studentene med å observere og vurdere egen læring, (c) øker studentenes mulighet for livslang læring ved at de lærer seg å sette seg mål og planlegge sin læring og (d) lager oppgaver som fremmer studentengasjement over tid og hvor man får tilbakemeldinger fra flere kilder (Boud & Molloy, 2012, s. 704). Harildstad (2022) fant også at både lærere og studenter synes å definere *vurdering* svært bredt, og «de skiller ikke mellom tilbakemeldinger, veiledning, respons og formell vurdering – alt oppleves som vurdering som spiller inn på deres læringsprosess» (s. 10).

I artikkelen «The Power of Feedback» presenterer Hattie og Timperley (2007) en omfattende syntese av forskning om tilbakemelding og dens effekt på læring. Her defineres tilbakemelding som informasjon som gis

til elevene om deres prestasjoner eller forståelse av et emne (s. 81). De ser på ulike former for tilbakemelding og diskuterer hvordan tilbakemelding kan påvirke elevenes prestasjoner. I tillegg til å påpeke viktigheten av å gi tilbakemeldinger som er konkrete, relevante og knyttet til målene for læringen, diskuterer forfatterne betydningen av timing, frekvens og format på tilbakemeldingene. Videre løfter de fram betydningen av at elevene aktivt engasjerer seg i tilbakemeldingsprosessen. De påpeker at elevene må ha en forståelse av læringsmålene og muligheten til å bruke tilbakemeldingen til å regulere, og å oppleve eierskap til sin egen læring.

Brevik og Blikstad-Balas (2014) løfter fram og diskuterer vurdering for læring i klasserommet og viser at en av utfordringene er at elevene kan ha høy kompetanse på noen områder, men at de ikke oppfatter dette som viktig hvis det ikke er en del av lærerens vurderingsområder (Brevik & Blikstad-Balas, 2014, s. 204). Et åpenbart problem som kan oppstå når man fokuserer ensidig på klare mål og vurdering, er at mer umålbare sider ved utdanningen blir underkommuniserte. I musikkfagene er dette spesielt relevant.

## Teoretisk grunnlag

I denne studien ser vi i hovedsak på *undervisning* og *vurdering* som sammenvevde og rettet mot felles mål, og som derigjennom bidrar til å øke studentenes motivasjon, engasjement og kunnskapsnivå. I norsk høyere utdanning svarer såkalt *konstruktivt samsvar* (*constructive alignment*) på Hattie og Timperleys (2007) bud om bevisstgjøring, sammenheng og systematikk i utdanningen. Hovedideen bak konstruktivt samsvar (Biggs, 1996) er at undervisning og vurdering skal være tilpasset læringsmålene. Læringsmål bør formuleres tydelig og være sentrale for planleggingen av undervisningen. Valg av innhold og aktiviteter i undervisningen må understøttes av vurdering av måloppnåelsen. Dette kan inkludere ulike undervisningsteknikker, gruppearbeid og praktiske oppgaver eller prosjekter som engasjerer studentene og hjelper dem med å utvikle de ønskede kompetansene. *Vurdering* i et slikt rammeverk må harmonere med læringsmålene og undervisningsmetodene. Vurderingen skal måle i hvilken grad studentene har oppnådd læringsmålene, og den bør være

autentisk og relevant. I likhet med den foreliggende studien baserer Biggs (1996) sin teori på et konstruktivistisk syn på læring hvor «learners arrive at meaning by actively selecting, and cumulatively constructing, their own knowledge, through both individual and social activity» (s. 348).

Høyere utdanning er blitt kritisert for å ha utilstrekkelige metoder og tilnærminger til tilbakemeldinger til studenter. Black og Wiliam (2009) foreslår fem strategier for effektiv formativ vurderingspraksis: (a) avklare og dele læringsintensjoner og suksesskriterier, (b) konstruere effektive klasseromdiskusjoner og andre læringsoppgaver som synliggjør grader av elevforståelse, (c) gi tilbakemelding som bringer elevene framover, (d) aktivere elever som instruksjonsressurser for hverandre, og (e) aktivere elevene som eiere av egen læring (Black & Wiliam, 2009, s. 8). De understreker også betydningen av kontinuerlig profesjonell utvikling av lærere for å implementere formativ vurdering i praksis. Boud og Molloy (2012) diskuterer to modeller for design av effektive tilbakemeldingsmodeller for læring i høyere utdanning: én hvor læreren er styrende (*the drivers*) for tilbakemeldinger, og en annen som baserer seg på ideen om *bærekraftig vurdering*. I sistnevnte har studentene en nøkkelrolle, hvor de skal både generere og søke tilbakemeldinger. Dette skal legge grunnlag for at studentene skal kunne ha tydeligere eierskap til egne læringsprosesser. I vår studie har vi tatt utgangspunkt i en forståelse av at studenters læring foregår gjennom utøvelse av menings skaping:

A working version of constructivism can be integrated with instructional design at three crucial points: the curriculum or unit objectives are clearly stated in terms of content specific levels of understanding that imply appropriate performances, the teaching methods require students to be placed in contexts that will likely elicit those performances, and the assessment tasks address those same performances. (Biggs, 1996, s. 361)

Vurdering bør i et slikt rammeverk være en tydelig del av læringsprosessen, og ikke bare være et måleinstrument.

Dysthe (2008) framhever også at vurdering ikke bare skal være en måte å måle elevenes prestasjoner på, men en integrert del av læringsprosessen. Hun argumenterer for et bredt vurderingsbegrep uten for store skiller mellom vurdering, undervisning og veiledning. Læringsfremmende

vurdering som foregår kontinuerlig «gir signal til elevane om kva som er viktig. Ho handlar om daglige rutinar, systematikk og variasjon og ikkje minst om å bruke den informasjonen som ein får gjennom vurdering og tilbakemelding» (s. 16). Hun påpeker også at lærere må ha god vurderingskompetanse for å kunne gi effektive tilbakemeldinger og understreker betydningen av elevmedvirkning i vurderingsprosessen.

Kritikere av rigide vurderingsregimer i skolen har gjennom tidene påpekt at overordnede målsettinger med skolen, eksempelvis danning og medborgerskap, vanskelig lar seg kvantifisere og underordnes universelle kriterier. Giroux (2014) har uttrykt bekymring for standardisering og målstyring i utdanningssystemet, og mener at fokuset på standardiserte tester og kvantitative resultater er begrensende og ikke tar hensyn til kompleksiteten og mangfoldet av elevenes læring: «Young people are now openly treated as customers and clients rather than a civic resource» (Giroux, 2014, s. 14). Dette er samfunnsmessige problemstillinger som potensielt skaper dårlige vilkår for kreative fag i skolen. Reformpedagogen Paulo Freire (1970) argumenterte i sin tid for at tradisjonelle vurderingsmetoder, som standardiserte tester og rangeringssystemer, ofte er undertrykkende blant annet fordi de ikke tar hensyn til studentenes erfaringer, kultur og kontekst. Han kritiserer en vurderingspraksis som reduserer læring til kvantifiserbare resultater og som bare vurderer studentene ut fra deres evne til å reprodusere kunnskap. Biesta (2017) mener at målet med utdanning bør være kvalifisering, sosialisering og subjektivering, hvorav den siste dimensjonen handler om måten studentene gjennom erfaring utvikler seg som selvstendige mennesker. Han er kritisk til instrumentelle og resultatorienterte utdanningssystemer, som han mener er fundamentalt udemokratisk. Han mener generelt at vi, blant annet gjennom vurderingssystemene, «turn [education] into a machinery in which what matters educationally – such as the freedom and independence of the student – is ultimately squeezed out» (Biesta, 2017, s. 317).

Som et ledd i dannelsen av samfunnsbevisste borgere forventes studenter i høyere utdanning å ta stort ansvar for egen kunnskaps- og ferdighetsutvikling, og i dag står derfor *studentaktiv læring* i høysetet i universitetspedagogisk sammenheng (NOKUT, 2023). Studentaktiv læring fordrer stor grad av *eierskap* til egen læring. Eierskap til læring

handler om studenters mulighet til å påvirke innhold, organisering og aktivitetene knyttet til læringsmål, noe som krever at det skapes rom for deltakelse og engasjement. I tillegg fordrer det, ifølge Rainer og Matthews (2002), balanse i maktforhold og muligheten for studenter til å være ressurser for hverandre. Med utgangspunkt i et grunnleggende konstruktivistisk syn på læring, betrakter vi her eierskap som en av de fundamentale faktorene i meningsbygging og læring i vurderingsprosesser. I denne forbindelse står individets aktive rolle i kunnskapskonstruksjon sentralt:

When humans become mentally active, they are agents in their own learning. They own that knowledge, and they relate it to their personal experience. Hence, ownership could be described as the linchpin, or a central and cohesive element, of knowledge construction. (Rainer & Matthews, 2002, s. 22)

Opplevelse av reelt eierskap til egne læringsprosesser krever at studenten tar og gis autonomi, makt, ansvar og en tydelig *stemme* (Aaronsohn, 1996; Fredriksen et al., 2023; Rainer & Matthews, 2002). I vurderingssammenheng innebærer dette en stor grad av studentinvolvering mye tidligere i vurderingsprosesser, kanskje spesielt i forbindelse med vurdering av kunstneriske prestasjoner hvor kreative og personlige uttrykk står sentralt.

## Metode

For å få innblikk i hva musikk lærerstudentene opplevde som læringsfremmende ved vurderingsformene de har erfart i sin utdanning, valgte vi å gjennomføre gruppeintervju med seks nyutdannede studenter fra to forskjellige norske læresteder. Målet med å gjøre intervjuene i grupper, var å høre hva slags diskusjoner studentene hadde seg imellom, i tillegg til deres individuelle svar på spørsmålene vi undersøkte. Fordelene med gruppeintervju er blant annet at informantene sammen kan generere utdypinger og ideer ut over det den enkelte alene ville bidratt med (Carey & Asbury, 2012; Thagaard, 2013), og at eventuelle meningsforskjeller kommer tydeligere fram. Ettersom studentene i de respektive gruppene hadde studert sammen, var tanken at gruppeformatet ville kunne stimulere den kollektive hukommelsen og refleksjoner, og at vi med dette kunne

gå mer i dybden på tematikken læringsfremmende vurdering (Amdahl, 2002; Carson, 2007; Dalen, 2004). Vi som forskere valgte å gjennomføre ett intervju hver, og dette ble gjennomført digitalt i mai 2023.

Deltakerne bestod av to grupper med tre studenter i hver gruppe (tre kvinner og tre menn), og rekrutteringen ble gjort ved at vi tok kontakt med undervisere på studiestedene som kontaktet studenter de mente ville kunne være interesserte i å delta. Vi kontaktet de aktuelle studentene på e-post, noe som til sist gav oss tre informanter i hver gruppe. De intervjuede studentene i den første gruppen hadde fullført sin bachelor i faglærerutdanning i musikk i 2022 og gikk alle første året på master i musikk ved samme studiested i tidsrommet for intervjuet. I den andre gruppen hadde alle begynt å jobbe som lærere etter fem års lærerutdanning som inkluderte 30 studiepoeng musikk. Forskjellene på disse utdanningene er blant annet at bacheloren til den første gruppen har opptaksprøver i musikk og er en treårig fordypning i musikkfaget, i motsetning til den andre gruppen med studenter, som hadde andre fordypninger enn musikk i sitt femårige studieforløp. Alle informantene var spesielt opptatt av musikkfaget, og hadde erfaring med å undervise musikk på forskjellige nivå og i forskjellige kontekster. Alle hadde også bakgrunn som musikkutøvere, på forskjellig nivå, og flertallet hadde gått på musikklinje på videregående skole. Informantene i de respektive gruppene kjente hverandre godt fra studietiden.

Datamaterialet i studien består av de transkriberte intervjuene, til sammen omtrent to timer med samtaler. Vi valgte å gjøre en tematisk analyse av intervjuene (Braun & Clarke, 2006), og jobbet i fellesskap med å identifisere, analysere og beskrive mønstre i datamaterialet. Vi fulgte prosedyren for *kollektiv kvalitativ analyse* foreslått av Eggebø (2020), gjennom tett samarbeid om prosessen, som bestod av (1) felles gjennomgang av datamaterialet, (2) temakartlegging, (3) temagrupping og (4) disposisjon og arbeidsplan. Vi identifiserte tematikker på tvers av intervjuene for å belyse ulike synspunkter og opplevelser med vurderingsformer og læring. Sentrale tema vi identifiserte i de to gruppeintervjuene ble deretter sammenlignet med hverandre og sett i lys av studiens problemområde.

I den første fasen av analysen, hvor vi i fellesskap gikk gjennom materialet, kom det fram en rekke tematikker som var relevante å se

nærmere på i relasjon til studiens problemstilling, men også temaer som var på siden av denne. For eksempel valgte vi å *ikke* ta med uttalelser som dreide seg om undervisningsinnhold og oppbygging av studiene. Eksempler på tematikker vi så i denne første fasen, var musikkfagets egenart i vurderingssammenheng, opptatthet av prosess mer enn resultat, karakterenes betydning og viktigheten av eierskap. Etter en første gjennomgang kartla vi mulige temaer som kunne bidra til å kaste lys over spørsmålet om vurderingsformer og erfaringer med disse. Her laget vi en liste med konkrete vurderingsformer studentene fortalte om. Vi gjorde en preliminær koding av disse formene i formative og summative vurderingsformer, og i tillegg hvilke typer fag det ble gitt graderte karakterer i. På samme måte grupperte vi uttalelsene som viste til læringsformenes læringspotensial. Her ble det generert temaer ut fra innholdet i intervjuguiden og tilleggsinformasjon som framkom under intervjuene. I tredje fase av analyseprosessen grupperte vi disse i følgende overordnede temaer: *Vurdering i mange former; Karakterer og «touchy soner»; Vurderingskriterier og kunstfag; Eierskap og læring; Tilbakemelding for læring*. Vi har trukket fram utvalgte eksempler og sitater som samlet sett er særlig egnet til å belyse studiens problemstilling, og samtidig gir mulighet for å kunne drøfte datamaterialet i en videre kontekst.

Studien er meldt inn til Sikt, som ivaretar personvern i forbindelse med forskningsstudier. Ethiske retningslinjer med hensyn til samtykke og anonymisering er fulgt.

## Resultater – læringsfremmende vurdering i musikk

I det følgende gis et innblikk i studentenes betraktninger rundt vurderingsformene i musikkfaget de har hatt erfaring med i studietiden, med spesiell vekt på det læringsfremmende potensialet i forskjellige former for vurdering. Innledningsvis beskrives vurderingsformene studentene rapporterte at de hadde hatt i musikkfaget. Studentene omtales som student 1A, 1B og 1C (intervjugruppe 1) og 2A, 2B og 2C (intervjugruppe 2).

## Vurdering i mange former

Vurderingsformene informantene trakk fram når vi spurte hvilke typer vurdering de har hatt i musikkfagene i studietiden, var mangfoldige og en blanding av sluttvurderinger, arbeidskrav og undervisvurderinger. Følgende former for vurdering ble beskrevet i de to intervjuene:

- Framførelser av musikk (utøvende eksamener med muntlig tilbakemelding)
- Prøver i musikkteoretiske fag (skoleeksamen med karakterer)
- Forberedelse og gjennomføring av undervisningsopplegg i samspill med elever
- Arrangering og komponering av låter/musikk
- Eksamensoppgaver med refleksjonsnotater
- Undervisvurderinger av prestasjoner fra lærere
- Eksamensoppgave hvor de skulle gjøre et opptak av seg selv som leder i et ensemble, inkludert et refleksjonsnotat i etterkant
- Muntlige presentasjoner der de fikk tilbakemeldinger fra medstudenter og lærere

På tross av at dette antakelig ikke gir et fullstendig bilde av alle vurderingsformer informantene erfarte i musikkfagene, gir det et innblikk i de vurderingsformene de husket best. Som vi ser, er det større vekt på utøvende, praktiske vurderingsformer enn i mange andre fag i høyere utdanning.<sup>2</sup> Dette innebærer å skulle prestere i øyeblikket, og også å beherske kunstneriske og praktisk-pedagogiske situasjoner i all sin kompleksitet. Studentenes beskrivelser viser også at flere av vurderingsformene innebærer en viss avhengighet av medstudenter.

Begge gruppene beskrev et skille mellom vurderingsformene i mer teoretiske støttefag som musikkhistorie, analyse, komponering/arrangering og hørelære, og de kreative, utøvende fagene som instrument og samspill. I de utøvende fagene ble vurderingsformene oftere beskrevet

---

<sup>2</sup> De vanligste kategoriseringene av vurderingsformer i retningslinjer for emneplaner i sektoren er: arbeidskrav (inkludert undervisvurderinger), obligatorisk deltakelse, skriftlig skoleeksamen, hjemmeeksamen, oppgave, mappe, muntlig eksamen, praksis og praktisk eksamen.

som prosessorienterte og formative, og dermed også unødvendige å ha bokstavkarakterer i. Vurderingsformene i de musikkteoretiske fagene ble i større grad beskrevet som avgrensede, oppsummerende og med bokstavkarakter: «På bachelor hadde vi gradert karakter i arrangement/komponering, hørelære og musikkhistorie/analyse, og profesjonsfag er gradert. Det som har vært bestått/ikke bestått er samspill og ensembleledelse, hoved- og bi-instrument» (1A).

Selv om det syntes å være en blanding av summative og formative former for vurdering, var det tilsynelatende mange formative og studentmedvirkende vurderingsformer. Eksempelvis refereres det til «såne rene framførelser der vi fikk på en måte umiddelbar tilbakemelding» (2A), og «når vi har hatt presentasjon og fått tilbakemelding, har alle studentene vært med» (1B), og at «undervisvurdering har vært vektlagt, og det er det jeg har fått mest utbytte av også» (1B). Refleksjon over praksis viste seg også å ha relativt stor plass som vurderingsform, enten i grupper eller individuelt: «I samspill har vi en eksamensoppgave der vi skal ta et opptak av oss selv som leder et ensemble, og så skulle det være et refleksjonsnotat etterpå» (1A). Det kom tydelig fram av intervjuene at det var praktisk-estetiske fag det var snakk om, både gjennom at informantene understreket dette i forbindelse med vurdering, og fordi vurderingsformene i hovedsak var av praktisk art.

I likhet med Harildstads (2022) resultater, så vi også i materialet at informantene blandet begrepene vurdering (formativ og summativ), veiledning, tilbakemelding og evaluering. Som en av studentene sa: «Det er egentlig vanskelig å skille vurdering og læring» (1B). Det framkom indirekte at dette kunne skyldes at også utdanningsinstitusjonene bruker forskjellige vurderingspraksiser hvor utøvelsen av vurderingene gled over i tilbakemelding og veiledning, inkludert måten den enkelte lærer oppfattet formålet med vurderingsformene.

I gruppeintervjuene diskuterte vi ikke styrker og svakheter ved de forskjellige vurderingsformene, men det kom likevel fram at studentene foretrakk formative vurderinger hvor det er «rom for å feile» og hvor tilbakemeldinger er en del av vurderingsprosessene. I det følgende ser vi på studentenes betraktninger om vurderingsformene som *læringsfremmende*.

## Karakterer og «touchy soner»

I likhet med studentene i Harildstads (2022) studie, var informantene våre opptatt av at karakterer ikke er noe spesielt godt verktøy for å lære, og spesielt ikke i praktisk-estetiske fag, med noen unntak: «Koralsatsene gir det kanskje litt mer mening at det er gradert karakter, fordi der er det åpenbare feil og riktige ting» (1A).

I begge gruppene vi intervjuet var det en gjennomgående skepsis til bokstavkarakterer i de utøvende fagene. Det ene studiestedet opererte med karakteruttrykket bestått / ikke bestått i fagene som hadde utøvende karakter, og studentene framførte flere argumenter for at dette var bedre enn bokstavkarakterer. Hovedsakelig gikk dette på at karakterer fører til større press og stress som kunne gå ut over prestasjonene, særlig i utøvende kunstfag:

Med en gang det blir snakk om en [gradert] karakter i enden så blir jeg sånn «åååå, jeg må prestere» og så ender det med at jeg underpresterer i stedet. Heller enn at jeg bare kan slappe av og så vite at «ja, men det her det mestrer jeg jo, og det kommer til å bli bestått». Og så *nailer* jeg det fordi jeg har sånn tro på at jeg kan klare det godt når det er snakk om bestått/ikke bestått. Sånn sett er det jo også mer læringsfremmende. (1A)

Hvilken bokstav jeg får har ikke noe å si for min kompetanse som musikk lærer [...], så det er ikke noe jeg har bruk for i det hele tatt, sånn egentlig. (2B)

Studentene trakk likevel fram noen positive sider ved graderte karakterer som stort sett handlet om at det å sette seg mål om en god karakter kan gi en ekstra motivasjon til å jobbe hardt og fokusert fordi «nå gjelder det» (2B). Det handler ifølge studentene om å gå mer i detalj og å legge inn en større innsats enn de ellers ville gjort.

Karakterer gir også en pekepinn på hvordan studenter står i et fag sammenlignet med andre studenter, noe som så ut til å ha betydning for studentenes motivasjon og retning i arbeidet med fagene. Dette forutsatte ifølge studentene at karaktergivingen også inneholdt en tilbakemelding eller begrunnelse. En av studentene mente det burde følge begrunnelser med alle karakterer, også de gode, fordi det ligger viktig læring i å få vite hva som var bra når man får karakteren A også.

På spørsmål om hva som skiller musikkfaget fra andre fag i vurderingssammenheng, var studentene stort sett enige om at de utøvende, kreative fagene er vanskelig å både få og gi karakterer i på grunn av kompleksiteten i det som kan/skal vurderes. Studentene viste både til erfaringer fra egne eksamener og til egen vurderingspraksis som lærere.

En av studentene trakk fram at behovet for graderte karakterer antakelig er mindre i fag som er prosjektorienterte og prosessdrevne fordi man får kontinuerlige tilbakemeldinger fra lærere og hverandre underveis, og at dette foregikk i små, oversiktlige grupper: «Sånn som vi har jobba, og med den oppfølgingen vi har fått, så føler jeg ikke at det har vært et behov [for karakterer]» (2B). Begge utdanningene i studien vektla samspill, og i slike fag er det også en utfordring å gi individuelle karakterer. Begge studentgruppene fortalte at det på studiestedene var en gjennomgående kultur for samarbeid og hjelpe hverandre å prestere godt, også i eksamenssituasjoner. I denne konteksten ble karakterer ansett som en ulempe:

Jeg vil bare påpeke hvor fint det er med et slikt fag som ikke er så veldig karakterrettet [...] Du kan bare utforske og prøve deg fram og har du utvikling i en gruppe, så er det det som er artig. Det å få til noe sammen. (2C)

Studentene i vår studie henviste til praksisen enkelte musikkonservatorier har med bestått / ikke bestått på hovedinstrument, og sa seg enige i at dette er en fordel for både musikken og musikeren. Musikkens personlige preg og utøverens frihet til å kunne uttrykke seg er et tema som gikk igjen i intervjuene.

Det er jo snakk om kunst, og det er så personlig at det kan være vanskelig. Altså ikke kritikk som person, eller smak eller stil, men som går på en måte rent musikkmessig [...] heller enn at det går på kritikk av musikeren på en måte da. Så det kan være noen sånne *touchy soner* som jeg tror nettopp er grunnen til at de velger å ikke gradere akkurat de fagene. (1A)

De «touchy sonene» tenker vi handler om det nokså personlige og utlevende som er involvert i utøvelse av musikk, og at dette er spesielt sårbart i en vurderingssammenheng. Å bare bli belønnet med en bokstavkarakter uten noen påfølgende samtale kan oppleves demotiverende (Chamberlin

et al., 2018). Det ser ut til at de fleste utøvende vurderingsformene studentene har hatt, har vært etterfulgt av samtaler om prestasjonen, og at studentene opplevde seg som trygge og ivaretatt gjennom slike samtaler. De «touchy sonene» kan da sirkles inn litt forsiktig, og man unngår kanskje at karakterer oppleves som «karakter-drap».

En av studentene snakket om det vanskelige ved å skulle gi en karakter med utgangspunkt i kriterier i musikkfaget i sitt virke som lærer: «Og ofte da så for eksempel hvis en elev som har bommet litt på vurderingskriteriene, men som likevel har produsert et produkt som er veldig musikalsk da, bare på en annen måte» (2B). Han konkluderte med at det på tross av kriteriene er viktig å løfte blikket og se hva slags andre kvaliteter som ligger i elevens arbeid, og samtidig at det er problematisk at man dermed også åpner for skjønnsvurderinger som kan skape utydelighet for elevene.

## Vurderingskriterier og kunstfag

På masterstudiet har vi også samspill som fag, og det er gradert [karakter] på, og en eksamensoppgave der går ut på at vi skal ta et opptak av oss selv som leder et ensemble [...] og så skal det være et refleksjonsnotat etterpå, og det merker jeg at jeg synes er vanskelig på en måte: «Okay, hvordan i alle dager skal dette vurderes egentlig da.» (1A)

I Biggs og Tangs rammeverk for konstruktivt samsvar er *kriteriebasert* vurdering, i stedet for *normbasert*, en konsekvens av et skifte fra et undervisningsparadigme til et læringsparadigme i høyere utdanning (Biggs & Tang, 2007, s. 61). Den enkelte students læring er i sentrum, og kriteriene for bedømmelser skal reflektere dette, og ikke primært hvordan denne studenten gjør det sammenlignet med andre studenter. Formelt skal kriteriene komme fram i emneplanenes læringsutbyttebeskrivelser, men dette rapporterte studentene i denne studien er utfordrende fordi emneplanene er vage og vanskelige å tolke. I tillegg var studentene usikre på om emnebeskrivelsene alltid «følges» av de som underviser. Vi tolker dette som at studentene var usikre på om emneplanenes formulerte læringsutbytter var viktige i relasjon til vurderingssituasjonene, og til dels at de tviler på om beskrivelsene er reelle.

Kriteriene blir litt uklare [...] det blir så akademisk og fint skrevet at det tar vekk litt mening for meg i hvert fall. Og det samme gjelder læringsutbytte, det er så fint skrevet at jeg føler ikke vi sitter igjen med noe, annet enn at det står på nettsiden. (1C)

Studentene rapporterte at det i flere av musikkfagene ble gitt en kriterieliste for vurdering i forkant av eksamen, og at dette var noe de satte pris på. Kriteriene som ble nevnt var eksempelvis «hvor godt vi tilpasset musikk og metoder til et ensemble, hvor godt vi kommuniserte musikalsk og verbalt, og hvor godt dine musikalske intensjoner samsvarte med det klingende resultatet» (1A). Andre trakk fram musikalsk bredde som et kriterium på spilleeksamen, og antall rett og galt i koraltsats. Vi spurte også studentene om de hadde vært involvert i å lage kriterier for vurdering i løpet av studietiden, noe de ikke kunne huske å ha vært med på.

I likhet med diskusjonen om gradert karakter, mente studentene at kriterier for vurdering avhenger av hva slags fag det er snakk om. I noen fag opplevde de at det var enklere å forstå hvilke kriterier som gjaldt, for eksempel i de musikkteoretiske fagene. I slike fag mente studentene også at det beste for læringsprosessene var at vurderingene inneholdt direkte og kriteriebaserte tilbakemeldinger som kan bygges videre på. Samtidig beskrev en av studentene hvordan fag som kan måles i antall feil og riktig, også kan ha kunstneriske kvaliteter som blir borte i en slik vurderingsform:

Når vi har hatt arrangering og komponering og skrevet koraltsats så ble det jo målt etter hvor mange feil du hadde i forhold til hvilken karakter du fikk, så med en gang du setter en karakterskala med prosenter så tok det litt den der kunstneriske delen vekk i hvert fall. (1C)

Det var en generell holdning blant studentene at det er viktig å beholde friheten til å gjøre valg i de utøvende, kreative fagene. Vi tolker dette som et uttrykk for at kriterier virker begrensende på kreativitet og kunstuttrykk, og at det samtidig vanskelig kan konkretisere hva som ligger i originalitet og personlig uttrykk: «Det blir veldig rart å gradere kunstneriske valg?» (1B). Her ligger også hoveddilemmaet som trekkes fram av studentene, nemlig at det er ønskelig med forutsigbarhet og klarhet, men at dette samtidig kan være begrensende på kreativitet og personlig uttrykk. Dilemmaet

beskrives som vanskelig å forholde seg til, både som student/musiker og som lærer som skal sette karakterer. Derfor er det også viktig å ha innsikt i «hvor vanskelig det er for studenten å vite hva som gir et bra resultat» (1B).

## Eierskap og læring

I studentenes utsagn kom det tydelig fram at de ser gode muligheter til involvering i de praktiske musikkfagene med tanke på tolkning av emneplaner, utvikling av faglig opplegg og deltakelse i å utforme egen eksamen og vurderingskriterier. Det er glidende overganger mellom lærernes styring av vurderinger og tilbakemeldinger og medstudenters bidrag/hjelp. Her var informantene i begge gruppene stort sett enige om at når de øver sammen, oppfattes det som nyttig og givende: «Det blir jo alltid konstruktivt når vi spiller sammen og finner løsninger sammen [...] og hjelper hverandre helt automatisk» (1C).

Samtidig uttrykkes det at det var tryggest å stole på faglæreren når det gjelder hovedinstrument og tilbakemeldinger på eget spill. En av informantene sa nokså direkte at han er skeptisk til kvaliteten på tilbakemeldinger han fikk av medstudenter: «Tilbake til hovedinstrument, så føler jeg at de eneste tilbakemeldingene jeg har fått, skikkelig ordentlig fast, i hovedinstrumentet er fra hovedinstrumentlæreren» (1B). Eksemplene under illustrerer hendelser informantene trakk fram som viser måter de har tatt eierskap til både prosesser og resultater i vurderingssammenheng:

Fram mot disse eksamenene, var vi jo nesten nødt til å gi hverandre tilbakemeldinger og å øve på dette rollespillet. Vi hjalp jo hverandre. (2B)

Men på den siste så trakk jeg et tema med dans, og det var jo spennende. Og da kom jo hen som var ekstern sensor ... hadde jo en interesse for dans, og for min del så var det på en måte den samtalen med hen – altså den vurderingsformen – som gjorde at jeg fikk litt selvtillit i det å skulle ta med dans i min videre undervisning. (2A)

Vi legger merke til at informanten i det siste sitatet beskrev dette som en *samtale* og ikke en enveis tilbakemelding fra sensor, noe som sier oss at eierskapet til vurderingsformen og egen læring opplevdes som reelt og ble oppfattet som verdifullt for læringen som kan ligge i en slik form for

eksamensavvikling. Dette er etter vår erfaring en vanlig eksamensform i praktiske fag; at man utfører noe praktisk og deretter har en samtale om prestasjonen en til en nokså umiddelbart etter eksamen.

Eksempelene viser også at informantene bruker arbeids- og vurderingsformer fra studietiden i eget arbeid som lærere: «Så har jeg også hatt litt sånn egenvurdering, altså at elevene skal vurdere seg selv, både prosessen og sluttresultatet. Da kan de på en måte reflektere litt rundt det» (2C). I denne sammenheng kom altså eierskapet til egen læring i studietiden til syne gjennom en strukturert bruk og videreutvikling av vurderingsformene de selv hadde opplevd.

Karakteristisk for begge gruppene var også at de mente at de selv hadde tatt mye initiativ til å «hjelpes hverandre» (2B) med produksjoner og samspillopplegg, og at det ble opplevd som uklart hvorvidt det var noe som ble styrt av lærerne, eller om det var noe som de selv organiserte. Vi betrakter dette som positivt i retning av opplevd eierskap, og i stor grad noe som ligger i fagenes «natur», altså at læreren kan berede grunnen, men at studentene må involvere seg aktivt for å gjøre produksjoner og samspill til sitt «eget».

På den andre siden kom det også fram at en del av informantene opplevde lite eierskap når det gjaldt de nedfelte kriteriene for vurderingene, sammenhengen med læringsutbytteformuleringene og de kriteriene som faktisk ble benyttet. Intervjuene gir grunnlag for å anta at tilbakemeldinger og begrunnelser for karakterer sikrer større eierskap til vurderingsprosessen fordi man som student kan være i dialog (og prosess) med egen utvikling (og eventuelt kan påvirke selve prosessen).

## Tilbakemelding for læring

Det som har vært viktigst for meg, som vurderingstype da, er dialogen; det å få umiddelbar tilbakemelding. Da vet du hva du har gjort, og du vet hva du kan gjøre bedre, istedenfor at det går det et par uker, og så detter det langt bak i [bevisstheten]. (1C)

Typer tilbakemeldinger studentene har fått underveis beskrives som prosessorienterte i de pedagogiske fagene og mer produktorienterte i

de utøvende, kreative fagene. De har jobbet med konsertproduksjoner hvor de har fått tilbakemeldinger fra lærere underveis, og spilleprøver med eksamensrettet veiledning i forkant av utøvende eksamen. En av studentene beskrev en kombinasjon av arbeidskrav og eksamen som opplevdes som lærerik: «I musikkhistorie var arbeidskravet på en måte et første utkast til eksamen som vi da fikk tilbake fullt av masse tips og triks til hva vi kunne forbedre» (1A). Dette oppfattet studentene som en svært nyttig form for underveisvurdering; «Ja, det føler jeg var en veldig lærerik greie ... noe jeg definitivt lærte av» (1B). Vi ser også her at «timingene» på tilbakemeldingene er viktig for læringen; at informasjonen gis på et tidspunkt hvor studentene er mottakelige og motiverte (Wisniewski et al., 2020, s. 12). Et annet tilbakemeldingsformat som nevnes, er begrunnelser for karakterer på eksamener, ofte i form av individuelle samtaler med sensor i etterkant av eksamen. Dette var noe studentene vektla som svært nyttig og noe som ga mening til i det hele tatt å ha eksamen i kunstrelaterte fag.

En form for tilbakemelding studentene oppfattet som læringsfremmende, og som er vanlig i fag hvor framføringer av kunstneriske produkter er involvert, var forhåndsgodkjenning i forkant av konserten/forestillingen. Dette er ofte et ledd i kvalitetssikringen av noe som skal vises for et publikum utenfor institusjonen. I den ene intervjugruppen nevntes eksempelvis skolekonsertturneer som en slik praksis. Forhåndsgodkjenningen er en svært yrkesrelevant vurderingsform som gir trygghet for studentene og som må være nokså direkte og konkret.

Studentene hadde, som vi har sett, blandede erfaringer med vurderinger og tilbakemeldinger fra medstudenter. I utøvende, kreative fag, hvor fagkompetansen er veldig spesialisert, slik det ofte er på et hovedinstrument, stolte studentene mest på sine lærere, mens i samspillfag og undervisningspraksis ble vurderinger fra medstudenter beskrevet som vel så relevante som lærernes.

Jeg vet ikke helt hva det har hatt for seg – når du skal gi tvungen tilbakemelding til en medstudent. Det kan bli litt tynt, bare for å få det gjort [...], men jeg vet ikke hvor godt det funket. Det ble veldig sånn ... du må si noe bare for å si noe. (1C)

Noen av disse hverandrevurderingene var lærerstyrte og en del av kravene i studiene, mens andre så ut til å ha form som tilbakemeldinger underveis i arbeidsprosesser som var mer studentinitierte:

Vi leverte en oppgave på Canvas og så fikk vi tildelt en annen [student] som skal lese oppgaven og gi tilbakemelding, så det har vi hatt egentlig ganske mye utover i utdanningen. Men også sånn generelt i samspillsettinger når vi spiller med hverandre, hjelper vi hverandre ved å komme med tilbakemeldinger, så det er jo sånn konstant tilbakemeldinger frem og tilbake. (1C)

En av studentene beskrev en eksamensform hvor det var satt av et antall veiledningstimer i forbindelse med hjemmeeksamen, noe studenten opplevde som svært meningsfullt: «Istedenfor å levere blindt, og så få tilbakemelding etterpå, så får du hjelp med denne oppgaven for å kunne prestere best mulig» (1C). Studenten brukte begrepet «levere blindt» om det å levere en oppgave uten å få tilbakemeldinger underveis i arbeidet med oppgaven. Dette sier noe om den utrykgheten som kan oppstå i forbindelse med eksamensordninger hvor tilbakemeldinger ikke er en del av prosessen.

Studentene etterlyser en mer strukturert hverandrevurdering i studietiden, men sier ikke noe om hva slags form de mener dette bør ha. Igjen viser datamaterialet at studentene opplever undervisningen og undervisvurderinger av forskjellige slag som vanskelig å skille fra hverandre. Studentenes ønske om struktur, kan tolkes som et behov for tydeligere skiller mellom undervisningens innhold og innslag av vurdering/tilbakemelding. Gode rutiner, også i formative vurderingssammenhenger (Taras, 2006), bidrar også til å bygge studentens kompetanse i den typen komplekse arbeidsformer det er snakk om i musikkfaget: «Å vurdere hverandre er også en øvelse i det å gi en tilbakemelding, og egentlig i det å gi en vurdering i musikk. Det trenger vi også øving i selv» (2A).

## Oppsummerende drøfting

Altså i min utopi, eller drømmeverden, så er det en verden hvor det bare er bestått ikke bestått med jevne tilbakemeldinger hele veien da, så man vet hvordan

det går og hva man kan gjøre bedre, hva man mestrer, og at man blir gira på å få til mer, da, gjennom oppmuntring, samtale og oppbacking. (1A)

Da vi startet å jobbe med denne studien, var vi i utgangspunktet opptatt av hva slags nytte vurdering i høyere utdanning kan ha for studenters læring, og om det er noe som skiller vurdering i kunstfag fra vurdering i andre fag. Vi ønsket å høre studentenes syn på disse spørsmålene, og valgte derfor å intervju nylig uteksaminerte studenter om deres erfaringer med vurdering. Som vi har sett, var studentene vi intervjuet i stor grad opptatt av meningsfull og bærekraftig vurdering og overraskende lite opptatt av karakterer. De knyttet vurderingenes læringsfremmende potensial til selve formene for vurdering, og også til måtene vurderingene foregikk på.

På spørsmålene om hva slags former for vurdering som hadde vært tatt i bruk i de to utdanningene vi undersøkte, kom det fram en stor variasjon i vurderingsformer, hvorav mange bestod av tilbakemeldinger på prestasjonene som en del av vurderingsprosessene. Inkludert i disse beskrivelsene fant vi at grensene for hva som var vurderinger og hva som var undervisning ofte var uklare for studentene. Det var også til tider uklart for dem hvorvidt tilbakemeldingsprosesser var lærer- eller studentdrevne. Vi fikk i hovedsak inntrykk av at dette ikke ble opplevd som problematisk fra studentenes side, og kanskje til og med en fordel. I Biggs' rammeverk for konstruktivt samsvar vil disse uklarhetene kunne oppfattes som negative fordi samsvaret mellom innhold i undervisningen og vurderinger blir utydelig og dermed lite effektivt. På den andre siden viser forskere som Dysthe (2008), Hattie og Timperly (2007) og Wiliam (2011) at det å integrere vurdering med undervisning kan gi det Wiliam (2011) beskriver som «unprecedented power to increase student engagement and to improve learning outcomes» (s. 13). I vårt materiale er dette spesielt tydelig i beskrivelsene av kunstneriske, praktiske og kreative sammenhenger, og hvor også medstudenter involveres i vurderingssituasjonene. Hvis man skal tenke nytt om vurdering og tilbakemelding for læring i høyere utdanning, anbefales det å bevege seg fra vurdering som noe læreren gjør alene, til noe studentene gjør sammen med læreren og medstudenter (fra monolog til dialog) (Boud & Molloy, 2012). Hattie og Timperley (2007) understreker viktigheten av at lærere og elever utvikler

en felles forståelse av hva god tilbakemelding innebærer, og hvordan den kan brukes til å støtte læringen. De ser også et behov for at disse prosessene er mer planmessige og går over lengre tid, og ikke bare er en rekke isolerte handlinger. Vårt datamateriale antyder at det er et forbedringspotensial i musikk lærerutdanningene vi undersøkte når det gjelder felles bevissthet om, og planmessighet i, tilbakemelding og vurdering.

Som vi har sett, kom det i samtalen med studentene tydelig fram at de formene for vurdering som oppleves som læringsfremmende, inkluderer en eller annen form for muntlig tilbakemelding som er individuell og fremadrettet. Dette er helt i tråd med rådende kunnskap om vurdering og læring (Hattie & Timperly, 2007; Wisniewski et al., 2020), og våre informanter så ut til å være vant til slike former for vurdering i studiesammenheng, spesielt i musikkfagene. De beskrev en relativt stor variasjon i former for tilbakemeldinger, mer eller mindre formelle, og ikke alle like bevisstgjorte.

Analysen viser at studentene var mer involverte i vurderinger av hverandre enn vi kanskje hadde forventet. De beskrev sine medstudenter som ressurser for egen læring og utvikling gjennom en generell kultur for samarbeid og det å hjelpe hverandre å prestere best mulig. Dette framkommer som en blanding av egeninitierte samarbeider og krav fra lærerne om hverandrevurderinger underveis i studiene. Men, selv om Wisniewski et al. (2020) i sin metastudie om tilbakemelding viser at tilbakemeldinger fra student til student synes å ha stor effekt på studenters læring, viste resultatene våre at det også er en grense for hva studentene mener de kan få ut av sine medstudenter. Dette handlet spesielt om å ha adekvat språk til å formulere kritiske og konstruktive tilbakemeldinger og ikke bare si at «ja, det var bra» (1A). Det handler også om å ha kunnskap og autoritet nok til å bidra, spesielt knyttet til de mer tekniske sidene ved utøving på et instrument.

I vårt materiale fikk vi beskrevet en relativt stor vifte av vurderings- og tilbakemeldingsformer som så ut til å være godt tilrettelagt for å fremme læring. Det er av stor betydning for studentenes læring hva slags *typer* tilbakemelding det er snakk om (forsterkning/straff, korrektiver eller informasjonstette tilbakemeldinger), hva slags medium som brukes (muntlig, skriftlig eller video-/audio-assistert) og hvem som gir og får tilbakemelding (lærer til student, student til lærer eller student til student)

(Wisniewski et al., 2020). Våre informanter rapporterte i stor grad at de fikk tilbakemeldinger i vurderingssammenheng som var informative og fremadrettede. Det var likevel noen fag som skilte seg ut, spesielt musikkteoretiske emner, og individuelle instrumentaltimer, som bestod mer av korrektiver (tilbakemelding på «rett og galt»). Vi mener at det kan være verdt å se på hva slags konsekvenser forskjellene i vurderingsformer har for de forskjellige fagene i lærerutdanningene i musikk, og hvorfor det ser ut til å være slike etablerte forskjeller fra fag til fag.

Et annet funn knytter seg til studentenes karakteristikk av emneplaner og læringsutbyttebeskrivelser. Som vi har sett, var studentene i intervju-gruppe 1 kritisk til måten emnenes mål og innhold artikuleres i plan-dokumentene. I den grad læringsutbyttene skal gi mening for den som skal testes i om disse er oppnådd, må disse gi mening for den lærende, noe studentene uttrykker at de ikke helt gjør. Det kan være flere grunner til dette, blant annet måten planene følges opp av lærerne på, og i hvilken grad fagene i seg selv gjør det mulig å konkretisere læringsutbytter. Spesielt i kunstneriske og kreative fag er dette en utfordring, og åpne formuleringer brukes ofte for å ikke være for styrende på det kunstneriske. Det legger grunnlaget for det tidligere nevnte dilemma, nemlig ønsket om tydelighet og forutsigbarhet, men uten å risikere å begrense det kreative og personlige. Samtidig er det verd å merke seg at noen av studentene uttrykte avstand til emneplanenes «akademiske» språk og at de anser dette nærmest som et «spill for galleriet». Vi vet ikke om dette er en generell problematikk i UH-sektoren, men vi ser for oss at det kan være nyttig å iverksette gode dialoger om mål, innhold og vurdering tettere knyttet til planverket, og å involvere studentene mer direkte i utarbeidelse av emneplaner. Vi har også sett i materialet at utarbeidelse av eksamensoppgaver/vurderingsformer (Bærheim & Meland, 2001) ikke er noe studentene rapporterer å ha deltatt på.

En av studentene på gruppe 2 trakk fram at hen hadde brukt en blanding av kriteriebasert og holistisk vurdering med egne elever etter avsluttet studium. Som studenter nevner, kan noen av elevene bomme på noen av vurderingskriteriene, men likevel vise «et produkt som er veldig musikalsk da, bare på en annen måte» (2B). Ren kriteriebasert vurdering i musikk har åpenbare ulemper. Helheten i studentens prestasjon

kan framstå som av høy kvalitet, men ikke nødvendigvis fanges opp av vurderingskriteriene. I kunstfagene kan det oppleves svært vanskelig å forholde seg til krystallklare kriterier, men hvis man skal ta holistiske vurderingsformer i bruk (Vinge, 2014), er det viktig å forberede studentene på dette gjennom realistiske situasjoner og eksempler.

## Konklusjon og veien videre

I denne artikkelen har vi undersøkt følgende problemstilling: *På hvilke måter opplever lærerstudenter musikkfagets vurderingsformer som læringsfremmende?* Våre funn viser at vurderingsformer oppleves som mer læringsfremmende når studenter føler eierskap til prosessene og er aktivt involvert i utformingen av målsettinger og vurderingskriterier. Studentene i studien uttrykte at de formative tilbakemeldingene fra lærere var spesielt verdifulle. I enkelte fag, spesielt de utøvende fagene, kunne vi også spore en viss skepsis til medstudenters evne til å gi relevante tilbakemeldinger, noe som kan skyldes manglende tillit til medstudenters faglige kompetanse. Dette skiller denne studien fra lignende undersøkelser (eksempelvis Black & William, 2009; Boud & Molloy, 2012).

Det kommer fram at det er behov for å øke vurderingskompetansen hos både studenter og lærere i musikkfaget. Studenter trenger undervisning i hvordan de kan gi konstruktive tilbakemeldinger til hverandre. I tillegg bør institusjonene og lærerne involvere studentene mer i fortolkningen av emneplaner og vurderingskriterier. Dette vil kunne bidra til at studentene får en bedre forståelse av hva som ligger til grunn for vurderingene. Studentene bør føle eierskap til prosessen og være med-drivere, med vurdering som skjer fortløpende gjennom tilbakemeldinger underveis.

Videre er det viktig å anerkjenne kunstfagenes egenart, hvor prestasjon og identitet ofte er tett sammenvevd. Dette kan forklare hvorfor enkelte studenter foretrekker tilbakemeldinger fra lærere framfor medstudenter. I lys av dette kan karakterene bestått / ikke bestått være mer hensiktsmessige i kreative og utøvende fag, noe studentene i studien framhevet som mer relevant og læringsfremmende enn graderte karakterer.

Språket i emneplaner bør også tydeliggjøres for å unngå at det oppfattes som uklart eller «høytsvevende». Institusjonene bør arbeide med

rutiner for å klargjøre læringsmål og arbeidsmetoder for studentene, samt fremme en konstruktiv dialog mellom studenter og lærere. Dette inkluderer utvikling av opplegg som stimulerer til deltakelse og samarbeid, hvor læreren legger til rette for at studentene kan hjelpe hverandre. En slik tilnærming krever at studentene involverer seg aktivt i prosessen, slik at vurdering skjer fortløpende gjennom tilbakemeldinger underveis, med mål om mer bærekraftige vurderingspraksiser i musikk lærerutdanningene.

Studien bekrefter i stor grad tidligere forskning som viser at formative vurderingsformer oppleves som læringsfremmende av studenter, og den viser i tillegg hvordan dette er spesielt relevant i kunstfaglig sammenheng. De utøvende fagenes særpreg kan både påvirke og bli påvirket av vurderingspraksisene, noe som krever at institusjonene gir spesiell oppmerksomhet til vurderingsformene i disse fagene.

## Forfatterbiografier

**Hilde Synnøve Blix** er professor i hørelære og musikkdidaktikk ved UiT Norges arktiske universitet, Musikkonservatoriet. Hennes forskningsinteresser er blant annet gehørutvikling, musikk literacy og læremidler, i tillegg til kjønnsforskning og likestillingsforskning. Hun har ledet flere forskningsprosjekter, og har publisert en rekke forskningsartikler og bøker nasjonalt og internasjonalt med tematikk som spenner fra gehørundervisning og notedidaktikk til likestilling og kunstnerisk utviklingsarbeid. Hun er leder for forskningsgruppen InOvation ved UiT og FoU-leder ved Musikkonservatoriet.

**Bengt Haugseth** er førstelektor i musikk ved UiT Norges arktiske universitet, Musikkonservatoriet og Institutt for lærerutdanning og pedagogikk. Hans forskningsinteresser er gehørutvikling, musikk og bevegelse, tverrfaglig arbeid og vurdering i musikkfaget. Haugseth har vært i styret for nasjonalt nettverk for musikk i lærerutdanningene og medleder i det internasjonale nettverket Teacher Education Network. Han har samarbeidet med flere andre fagfelt, de siste årene blant annet i ReKomp (Regional ordning for kompetanseutvikling i barnehage) og Dekomp

(Desentralisert ordning for kompetanseutvikling i grunnskole og videregående skole) med temaet musikk og språk.

## Referanser

- Aaronsohn, E. (1996). *Going against the grain: Supporting the student-centered teacher*. Corwin Press.
- Almqvist, C. F., Vinge, J., Väkevä, L. & Zandén, O. (2017). Assessment as learning in music education: The risk of «criteria compliance» replacing «learning» in the Scandinavian countries. *Research Studies in Music Education*, 39(1), 3–18. <https://doi.org/10.1177/1321103X16676649>
- Amdahl, E. (2002). *Å gjøre gruppeintervju – å skape grupper* (STS-arbeidsnotat 4/02). STS.
- Biesta, G. (2017). Education, measurement and the professions: Reclaiming a space for democratic professionalism in education. *Educational Philosophy and Theory*, 49(4), 315–330. <https://doi.org/10.1080/00131857.2015.1048665>
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, (32)3, 347–364. <http://dx.doi.org/10.1007/bf00138871>
- Biggs, J. & Tang, C. (2007). *Using constructive alignment in outcomes-based teaching and learning teaching for quality learning at university* (3. utg., s. 50–63). Open University Press.
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Boud, D. & Molloy, E. (2012). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698–712. <https://doi.org/10.1080/02602938.2012.691462>
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qpo630a>
- Brevik, L. & Blikstad-Balas, M. (2014). «Blir dette vurdert, lærer?» Vurdering for læring i klasserommet. I E. Elstad & K. Helstad (Red.), *Profesjonsutvikling i skolen* (s. 191–206). Universitetsforlaget.
- Bærheim, A. & Meland, E. (2001). Hva om medisinstudenter lager oppgaver til egen eksamen? *Tidsskrift for Den norske legeförening*, 121(25), 2971–2973.
- Carey, M. A. & Asbury, J.-E. (2012). *Focus group research*. Left Coast Press.
- Carson, N. (2007). Erfaringer og refleksjoner ved bruk av gruppeintervju i kvalitativ forskning. *Norsk pedagogisk tidsskrift*, 91(3), 220–231. <https://doi.org/10.18261/ISSN1504-2987-2007-03-04>

- Chamberlin, K., Yasué, M. & Chiang, I-C. A. (2018). The impact of grades on student motivation. *Active Learning in Higher Education* 24(2), 109–124. <https://doi.org/10.1177/1469787418819728>
- Dalen, M. (2004). *Intervju som forskningsmetode: En kvalitativ tilnærming*. Universitetsforlaget.
- Dysthe, O. (2008). Klasseromsvurdering og læring. *Bedre skole*, (4).
- Eggebø, H. (2020). Kollektiv kvalitativ analyse. *Norsk sosiologisk tidsskrift*, 4(2), 106–122. <https://doi.org/10.18261/issn.2535-2512-2020-02-03>
- Fredriksen, B., Onsrud, S. V., Rinholm, H. & Lewis, J. (2023). Who takes part in participation? Challenges to empowering student voice in music teacher education, *Music Education Research*, 25(4), 1–16. <https://doi.org/10.1080/14613808.2023.2210159>
- Freire, P. (1970). *Pedagogy of the oppressed*. Seabury Press.
- Gamlem, S. M. & Smith, K. (2013). Student perceptions of classroom feedback. *Assessment in Education: Principles, Policy & Practice*, 20(2), 150–169. <https://doi.org/10.1080/0969594X.2012.749212>
- Gibbs, G. & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3–31.
- Giroux, H. A. (2014). *Neoliberalism's war on higher education*. Haymarket Books.
- Harildstad, S. R. (2022). Vurdering for læring i høyere yrkesfaglig utdanning. *Læring om læring*, 8(1), 1–13.
- Hattie, J. (2013). *Synlig læring: Et sammendrag av mer enn 800 metaanalyser av skoleprestasjoner* (I. C. Goveia, Overs.). Cappelen Damm Akademisk.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Meld. St. 16 (2016–2017). *Kultur for kvalitet i høyere utdanning*. Kunnskapsdepartementet. <https://www.regjeringen.no/no/dokumenter/meld.-st.-16-20162017/id2536007/>
- NOKUT. (u.å.). *Vurdering og eksamen*. Hentet 30. november 2023 fra <https://www.nokut.no/publikasjoner-og-kunnskap/kvalitetsoppsummeringer-hoyere-utdanning/vurdering/>
- Opplæringsforskrafta. (2024). *Forskrift om grunnskoleopplæringa og den vidaregåande opplæringa* (FOR-2024-06-03-900). Lovdata. <https://lovdata.no/forskrift/2024-06-03-900>
- Rainer, J. & Matthews, M. (2002). Ownership of learning in teacher education. *Action in Teacher Education*, 24(1), 22–30. <https://doi.org/10.1080/01626620.2002.10463264>
- Raaheim, A. (2019). *Eksamensrevolusjonen: Råd og tips om eksamen og alternative undervisningsformer* (2. utg.). Gyldendal.

- Steen-Utheim, A. & Hopfenbeck, T. N. (2018). To do or not to do with feedback: A study of undergraduate students' engagement and use of feedback within a portfolio assessment design. *Assessment & Evaluation in Higher Education*, 44(1), 80–96. <https://doi.org/10.1080/02602938.2018.1476669>
- Taras, M. (2006). Do unto others or not: Equity in feedback for undergraduates. *Assessment & Evaluation in Higher Education*, 31(3), 365–377. <https://doi.org/10.1080/02602930500353038>
- Thagaard, T. (2013). *Systematikk og innlevelse: En innføring i kvalitativ metode* (4. utg.). Fagbokforlaget.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14(3), 281–294. <https://doi.org/10.1080/09695940701591867>
- Utdanningsdirektoratet. (2019). *Erfaringer fra nasjonal satsing på vurdering for læring (2010-2018)*. Udir.
- Vinge, J. (2014). *Vurdering i musikkfag. En deskriptiv, analytisk studie av musikk læreres vurderingspraksis i ungdomsskolen* [Doktorgradsavhandling]. Norges musikkhøgskole.
- Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37, 3–14. <https://doi.org/10.1016/j.stueduc.2011.03.001>
- Wisniewski, B., Zierer, K. & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, 10, Artikkel 3087. <https://doi.org/10.3389/fpsyg.2019.03087>

## KAPITTEL 7

# Studenters egenvurdering som læring

*Anne Eriksen*

UiT Norges arktiske universitet

**Abstract**

This article refers to a pedagogical development project centered on students' self-assessment as a tool for learning. Within a drama and music course, students employed the “done-learned-wise”-method (*gjort-lært-lurt*) (Tiller & Gedda, 2017) to evaluate their performance during circle times in kindergarten settings. Drawing from students' assessment logs over three academic years the discussion highlights self-assessment as learning, emphasizing key concepts such as transformative learning and reflection. The experiences from the project suggest that self-assessment can significantly enhance learning, provided that students are adequately prepared and clear assessment criteria are established.

**Keywords:** music and drama in kindergarten, assessment as learning, transformative learning, reflection

Sitering: Eriksen, A. (2025). Studenters egenvurdering som læring. I H. S. Blix, K. J. Valkoinen & B. A. Haugseth (Red.), *MusPed:Research: Vol. 11. Vurderingsformer og tilbakemelding i kunstfaglig utdanning* (s. 185–202). Cappelen Damm Akademisk. <https://doi.org/10.23865/cdf.269.ch7>  
Lisens: CC BY 4.0

## Innledning

*I en gapahuk i nærheten av barnehagen har studentene samlet barna. Studentene er i rolle som ulike karakterer fra boka Grå (Rønne, 2017). De synger noen sanger før Grå inviterer barna med for å lete etter steinen som Gul har tatt fra henne. De vandrer gjennom skogen, balanserer på ei bro og leter mellom trærne. De treffer på Rød og Grønn, og etter hvert finner de Gul, som sitter oppe i et tre. Hun har gjemt seg. Kanskje er hun skamfull fordi hun har stjålet steinen til Grå?*

Utgangspunktet for denne artikkelen er et pedagogisk utviklingsarbeid knyttet til *musikk og drama i barnehagen*, et 30 studiepoengs fordypningsemne i barnehagelærerutdanningens tredje studieår ved UiT Norges arktiske universitet. Målet med utviklingsarbeidet var å utforske og prøve ut en formativ vurderingspraksis med særlig vekt på studenters egenvurdering, som et ledd i å forbedre og videreutvikle min egen undervisning. I dette studieemnet utvikler og gjennomfører studentene samlingsstunder i barnehager i grupper, og fortellingen som innleder artikkelen er et utdrag fra en slik samlingsstund. Studentenes muntlige utsagn de første årene studieemnet ble gjennomført, indikerte at samlingsstundene bidro til positive og lærerike erfaringer. Imidlertid var det uklart hvorvidt studentene reflekterte over endringsmuligheter eller årsaker til at en aktivitet fungerte bra eller mindre bra. Dette dannet bakgrunnen for prosjektet med egenvurdering. Målet for prosjektet var å finne fram til en metodikk som la til rette for at studentene kan beskrive og analysere sine samlingsstunder i barnehagen, og som kunne bidra til utvikling og forbedring. Valget falt på metoden *gjort-lært-lurt* (GLL), utviklet av Tiller og Gedda (2017). Metoden ble første gang utprøvd høstsemesteret 2018. GLL-metoden søker å beskrive gjennomførte aktiviteter, analysere læringsutbyttet og reflektere over mulige forbedringer – både med hensyn til hva som kan endres, styrkes eller fjernes (Tiller & Gedda, 2017, s. 29). Opprinnelig var metoden utviklet for å forenkle arbeidet med å skape lærende organisasjoner og brukes gjerne som en evalueringsmetode i aksjonsforskning og aksjonslæring (Tiller & Gedda, 2017). Metoden brukes også i undervisningssammenheng, for eksempel når sykepleierstudenter har anvendt GLL til å loggføre erfaringer fra praksis og reflektert over hva som kan være *lurt* å gjøre i lignende

situasjoner i framtiden (Mæhre & Vestly, 2012). Ved å innta et forskerblikk på egen undervisning (Postholm, 2007), ser jeg på min praksis som et kvalitetsutviklingsprosjekt der målet er å legge til rette for studentenes læring. Konkret betyr det at jeg gjør analyser og reflekterer over innspill fra studentevalueringer, tilbakemeldinger fra samarbeidspartnere og egen planlegging og gjennomføring av undervisning, og iverksetter ulike aksjoner for å utvikle og forbedre undervisningspraksisen, som igjen blir gjenstand for evaluering og refleksjon. Eksempler på slike tiltak kan være å endre eksamensform for å sikre bedre sammenheng mellom innhold, arbeidsmåter og læringsutbyttebeskrivelser, eller økt vektlegging av studentenes egenvurdering underveis i studieforløpet som en del av vurdering *for* og *som* læring. Det siste er også fokuset i arbeidet som denne teksten bygger på. I denne artikkelen utforsker og diskuterer jeg studentenes egenvurdering av samlingsstunder, uten å avgrense arbeidet til en stramt definert problemstilling. Utdrag fra studentenes vurderingslogger blir analysert og drøftet i lys av teorier om transformativ læring, vurdering som læring og begrepet refleksjon. Avslutningsvis oppsummeres artikkelen med noen refleksjoner rundt framtidig vurderingspraksis med bruk av *gjort-lært-lurt*-metoden.

## Om vurdering i høyere utdanning

Vurdering er en sentral del av undervisningsoppdraget i høyere utdanning og skal bidra til å oppnå intendert læringsutbytte (Sambell et al., 2013). Dette inkluderer både formative og summative praksiser, der formativ vurdering forbindes med utvikling og forbedring, mens summativ vurdering setter søkelys på oppnådde mål og er ofte knyttet til eksamen (Sambell et al., 2013). Vurdering handler med andre ord både om å kontrollere i hvilken grad studenten har oppnådd ønsket læringsutbytte, som for eksempel ved eksamen, og om tilbakemeldinger og egenvurderinger med mål om å stimulere til læring underveis i en læringsprosess. Læring er altså nært forbundet med vurdering (Wiliam, 2011). Vurdering *av* læring forstås som en måling av læringsoppnåelse, mens vurdering *for* læring handler om å identifisere læring gjennom vurderingspraksis underveis i en læringsprosess. Lorna Earl (2003) introduserer en tredje

kategori: vurdering *som* læring. Vurdering som læring fremmer utviklingen av metakognitive og selvregulerte ferdigheter, der studenten selv er aktiv og engasjert i vurderingsprosessen. Ifølge Schellekens et al. (2021) innebærer dette at studentene deltar aktivt i egenvurdering: «This notion includes the active involvement of students in self-assessment and self-directed learning as a distinct function to improve the learning process» (Schellekens et al., 2021, s. 2). Også Christine Watne Kristiansen (2012) er opptatt av vurdering som læring som en integrert del av læringsprosessen. Hun diskuterer i sin masteroppgave egen- og hverandrevurdering og vurdering som læring i dramafaget i videregående opplæring, og konkluderer med at det er et behov for at dramalærere oppdaterer sin kunnskap på dette området.

Selv om man som lærer har en bevissthet om vurdering som læring, kan det være utfordrende å oppnå studentengasjement når det gjelder vurdering. Nancy Falchikov (2005) foreslår egenvurdering for å vurdere egen utvikling, gjerne i kombinasjon med hverandrevurdering, som et svar på spørsmålet om hvordan involvere studenter i vurdering. Hun viser til en metaanalyse av ulike prosjekter med egen- og hverandrevurdering, og diskuterer hvorvidt studenter er i stand til å vurdere eget arbeid. En av hennes konklusjoner er at dersom egenvurdering skal ha verdi, må studentene være godt forberedt og vurderingskriteriene må være klare (Falchikov, 2005). Andre studier viser også hvordan studenter kan engasjeres i egenvurdering (Burner et al., 2011; Duggan & Ofte, 2016). Det er sparsomt med forskning om studenters egenvurdering i drama- og musikkpedagogisk arbeid i barnehagelærerutdanning. Dette utviklingsarbeidet fra studieemnet musikk og drama i barnehagen kan bidra med å gi noen eksempler på egenvurdering som læring.

## Undervisningens teoretiske bakteppe

I min undervisning står Iversens (2014) begrep uenighetsfellesskap sentralt. Han viser til at et uenighetsfellesskap, som for eksempel en skoleklasse eller en studentgruppe, består av mennesker som er havnet tilfeldig sammen og skal løse en felles oppgave. For at uenighetsfellesskapet skal fungere best mulig, må alle kunne komme med sine bidrag (Iversen,

2014). Trygge rammer med et raust menneskesyn er en viktig forutsetning. Et uenighetsfellesskap karakteriseres av at man respekterer ulike holdninger og meninger, og at demokrati ikke nødvendigvis betyr at flertallet skal bestemme (Iversen, 2014). Alle studenters bidrag skal ivaretas, slik at det ikke bare blir de som eksempelvis er kunstfaglig dyktigst eller som er mest utadvendt som bestemmer innhold og hvem som skal ha ulike roller. I undervisningen min legger jeg også vekt på hvordan man kan samarbeide selv om man er uenige, noe som også får betydning for valg av innhold og arbeidsmåter. Eksempelvis er improvisasjon sentralt i teater og drama, og Johnstone (1987) vektlegger evnen til å lytte til medspillere og spinne videre på andres ideer for å få en improvisasjon til å fungere.

En inspirasjon i undervisningen min er Paulo Freires frigjørende pedagogikk, der bevisstgjøringsprosesser står sentralt (Freire, 1999). I denne pedagogiske retningen er det et mål å kunne endre virkeligheten ved å reflektere kritisk over den (Westrheim, 2004). Frigjøringsbegrepet til Freire handlet opprinnelig om frigjøring fra undertrykkelse. I frigjøringsprosessen skaper man endring gjennom handling ved selv å ta ansvar for å stille kritiske spørsmål og reflektere over det man ønsker å endre (Berkaak, 2003; Eriksen, 2009; Freire, 1999). Målet med undervisningen er *empowerment*, eller myndiggjøring, som betyr «å gi kraft til» (Freire, 1999). Mitt håp er å kunne legge til rette for at studentene utvikler seg til å bli trygge og «kraftfulle» profesjonsutøvere, og at deres egenvurderinger kan bidra til dette.

En annen inspirasjon er læring som *transformasjon* (Illeris, 2012; Mezirow, 2000). Mezirow skriver: «Learning is understood as the process of using a prior interpretation to construe a new or revised interpretation of the meaning of one's experience as a guide for future action» (Mezirow, 2000, s. 5). I transformativ læring åpnes og endres referanserammer som ofte er dannet ut fra kulturelle paradigmer (Mezirow, 2000, 2003). For eksempel kan studentenes holdninger og meninger om samlingsstunder og hvordan de kan tilrettelegges, være påvirket av både deres egne barndomsopplevelser i barnehagen og av vanetenkning blant barnehagepersonalet når det gjelder innhold, funksjon og organisering. Dette er eksempler på det Mezirow kaller for *mindlessness* – en ukritisk nedarving

fra andre basert på vaner og rutiner, og ikke valg som resultat av egne mål, følelser, verdier og meninger. Målet er at studentene skal utvikle evnen til å ta reflekterte og gode valg, og få ny innsikt ved å ha en reflekterende og åpen tilnærming til ny erfaring, eller sagt på en annen måte: *mindfulness* (Mezirow, 1991). Transformativ læring kan oppstå når man møter det Mezirow kaller *disorienting dilemma*; en hendelse eller en situasjon som utfordrer ens etablerte referanserammer (Alteren & Eriksen, 2020; Mezirow, 1978).

For å oppnå denne typen læring kreves innsikt i hva *refleksjon* er og evne til å reflektere, men det er ikke entydig hvordan begrepet refleksjon forstås. Ifølge Anna R. Moxnes (2016) glemmer man ofte at det å reflektere skal være et bidrag til utvikling og forbedring. Hun diskuterer i sin artikkel hvordan refleksjonsbegrepet brukes i forskningslitteratur og fagteori som handler om barnehagelærerutdanningen. Hun finner at refleksjon forstås ulikt i ulike fagtradisjoner og således ikke er et entydig begrep. I tillegg setter hun søkelyset på manglende empirisk forskning om refleksjonsarbeid i barnehagelærerutdanningen (Moxnes, 2016, s. 9), og stiller spørsmål om hvorvidt barnehagelærerutdannere har en tilstrekkelig oversikt og forståelse av begrepet, eller om det brukes av mange uten å definere hva som legges i det. Kari Søndena (2004) hevder at refleksjon ofte forklares som tenkning, men påpeker at ordet er avledet fra latin og betyr *tilbakekastelse*. På samme måte som det å spille seg er en refleksjon, kastes tanken tilbake på *noe*. Slik jeg forstår det, betyr det at den tanken man opprinnelig hadde om for eksempel en handling, ikke bare endres til en ny tanke gjennom refleksjonsarbeidet, men at refleksjonen også handler om å forholde seg til den opprinnelige tanken. «Hva tenker du nå om det du tenkte før?» er et spørsmål Søndena stiller. Refleksjonsbegrepet er ikke bare knyttet til handlinger, men til tankene i forbindelse med handlinger, noe hun velger å kalle *kraftfull refleksjon*. Kraftfull refleksjon kan bare utvikles gjennom dialogen med medreflekterende personer; ved at ens synspunkter utvikles i møte med og mot andres synspunkter (Eriksen et al., 2015; Søndena, 2004). Dette har relevans for en profesjonsutdanning som skal utdanne mennesker til å jobbe tett sammen med andre, slik som i barnehagen. Det er mange grunner til at man skal/bør reflektere, og en av dem er den instrumentelle begrunnelsen; for å bli en

bedre profesjonsutøver. Å trene seg i å reflektere kan hindre at man som framtidig barnehagelærer bare glir inn i fastlagte rutiner og tankegods, og at man verdsetter det Søndena beskriver som *pedagogisk uro*, og det å betrakte seg selv som en lærende person (Søndena, 2004). Det kan for eksempel bety å stille spørsmål til sine egne opplegg, å utvikle bevissthet om det man gjør, eller om å utfordre seg selv til alltid å være i utvikling (Eriksen et al., 2022).

Refleksjon og speilmetaforer finner vi også hos teaterpedagogen Augusto Boal: «In this mirror we can see the present but can invent the future of our dreams [...] In the act of changing our image, we are changing ourselves, and by changing ourselves in turn we change the world» (Boal, 2006, s. 62). Speilet som metafor er først og fremst knyttet til det å spille forumteater, der det å endre en handling fra noe uønsket til en ideell situasjon er det sentrale. Men slik jeg ser det, kan den overføres også til annet kunstpedagogisk arbeid. Etter å ha reflektert over det man «ser i speilbildet»; det man har erfart og gjort, endrer man motivet for speilbildet – altså ser hva man kunne gjort annerledes. Slik kan man «endre framtiden» ved å gjøre noe på en annerledes måte. Når dette skjer i en gruppe og ikke bare på et individuelt plan, kan det relateres til Deweys forståelse av refleksjon som en sosial aktivitet der man i fellesskap tenker kritisk over egne erfaringer og bruker dette til å forbedre framtidige handlinger (Dewey, 1910). Mæhre og Vestly (2012) sier i tillegg at refleksjon i seg selv ikke er nok, om man ikke også stimuleres til å prøve ut nye veier i praktisk handling, noe som også oppsummerer min tolkning av begrepet. Jeg lener meg i den forstand primært på min egen fagtradisjon, som her er representert med Augusto Boal og hans teatersyn, der målet er å se muligheter for *endring* på bakgrunn av ens erfaringer.

## Metodisk tilnærming

I dette prosjektet var et av målene å utvikle og eventuelt endre egen praksis for å tilrettelegge for studentenes læring (Postholm, 2007). Ulike retninger innen aksjonsforskning, der kvalitetsutvikling av egen undervisning er hensikten med forskningen (Gjötterud, 2020), har vært sentrale i dette arbeidet. Studentene som har planlagt og gjennomført samlingsstunder

i barnehager, har i grupper skrevet logger der de kortfattet har dokumentert ved bruk av GLL-metoden (Tiller & Gedda, 2017), hva de har gjort, reflektert over hva de har lært og hva de kan gjøre annerledes neste gang. Materialet som analyseres i denne artikkelen er fra tre ulike studieår. Noen av loggene er skrevet for hånd på papir, mens andre er skrevet på den digitale tavlen *Padlet*. Padlet er en dynamisk tavle, så derfor er det laget PDF-filer av tavlene etter hver punktpraksis. Tavlene er nå slettet fra padlet.com. Materialet som trekkes fram i artikkelen, vil ikke kunne identifisere enkeltstudenter eller barnehager. Inspirert av Braun og Clarkes (2022, 2023) refleksive tematiske analyse, samlet jeg de av studentenes stikkord og utsagn som var beslektet innholdsmessig og lagde en tematisk gruppering (Befring, 2015) på bakgrunn av min tolkning. Jeg er bevisst på at valgene av hvilke sitater fra studentvurderingene som trekkes fram her, er farget av min subjektive forståelse. Grupperingen resulterte i fire kategorier: *kommunikasjon med barna*, *barns deltakelse*, *forberedelse* samt *egenutvikling og samarbeid*.

Den tematiske organiseringen fungerte som et godt utgangspunkt for å identifisere konkrete eksempler og utsagn som kunne fortelle noe om studentenes læring. Noe av materialet bestod kun av stikkord, og jeg valgte derfor ut sitater i form av setninger, der meningsinnholdet kom tydelig fram. I analyseprosessen så jeg spesielt på utsagn fra vurderingsprosessens tredje nivå (*lurt*), som tydelig hadde sammenheng med de to foregående (*gjort og lært*), for slik å åpne opp og tolke utsagnene i en helhetlig sammenheng.

## Om «Musikk og drama i barnehagen»

Før drøftingen av studentenes egenvurderinger, vil jeg gi et kort innblikk i studieemnet «Musikk og drama i barnehagen». Det er et fler- og tverrfaglig studieemne, og undervisningen består av fagdelte og tverrfaglige elementer med faglærere både fra drama- og teaterfaget og fra musikk (se UiT Norges arktiske universitet, 2024, for emneplan). Emnet er konsentrert om studentenes utvikling av egenferdigheter i musikk og drama, deres deltakelse i tverrfaglige prosjekter og forestillingsarbeid, og ledelse av musikk- og dramaaktiviteter i samlingsstund.

Studentene gjennomfører tre samlingsstunder for ei barnegruppe i løpet av semesteret. Innhold, form og dramaturgi er sentralt i dette arbeidet (Eriksen, 2020). Samlingsstundene utvikles og gjennomføres av og med 3–5 studenter i hver gruppe. Studentene møter de samme barna hver gang og kan slik legge til rette for en progresjon i samlingsstundene.

Et av målene med emnet «Musikk og drama i barnehagen» er at studentene skal utvikle ferdigheter i å planlegge, gjennomføre og evaluere drama- og musikkaktiviteter for og med barn. Det er også viktig i dette kunstfaglige emnet at studentene utvikler sin kreativitet og uttrykksevne i drama og musikk, noe som skjer gjennom undervisning og egen innsats, for eksempel gjennom øving og egenstudier. Dette innebærer også at studentene skal kunne *vurdere* sitt eget arbeid.

I studieemnet brukes både formativ og summativ vurdering, der summativ vurdering anvendes på eksamen. I gruppeeksamen planlegger og gjennomfører studentene en samlingsstund. Eksamen består også av en individuell, muntlig eksamen, der studentene lager en multimodal presentasjon av et faglig tema. Den formative vurderingen i studieemnet knyttes til øvelser og arbeid med eksempelvis små forestillingsprosesser, der studentene får tilbakemeldinger av faglærer og medstudenter underveis. Studentenes egenvurderinger er også en del av den formative vurderingen. Vurderingen av samlingsstundene underveis i semesteret er studentdrevet, og her trekker de fram egne erfaringer og opplevelser av hva som har fungert godt eller dårlig og hva som bør endres til neste samlingsstund.

Valget om å la studentene vurdere seg selv, både i planleggingen og gjennomføringen av samlingsstundene, var en følge av faglærernes utfordring med hensyn til å være til stede og observere alle gruppene. Samtidig var det et ønske om å gjøre vurderingsprosessen til en læringsprosess – vurdering som læring – og å styrke studentenes vurderingskompetanse. De skal kunne vurdere eget og kollegaers arbeid som framtidige barnehagelærere (Hannevig et al., 2020). Valget om å bruke GLL-metoden ble tatt i samarbeid med de andre faglærerne. Allerede i oppstarten av studieemnet presenterte jeg metoden for studentene, som dermed ble kjent med prinsippene og målsetningene med GLL før de benyttet den til å vurdere sin egen praksis. Metoden *gjort-lært-lurt* består av tre

nivåer: planlegging, gjennomføring, observasjon av handling og refleksjon over handling, som også samsvarer med nivåene i aksjonslæring (Utdanningsdirektoratet, 2022). *Gjort* beskriver for eksempel planer som er laget og aktiviteter som er utført. Det neste nivået, *lært*, er analyse-delen, og det siste nivået, *lurt*, viser til hva man på bakgrunn av de to første nivåene kan gjøre videre (Tiller & Gedda, 2017). Når studentene har gjennomført en samlingsstund, vurderer de gruppens arbeid i etterkant. Gruppen vurderes som en enhet, der det er fokus på gruppeprosess og -produkt og ikke på enkeltstudenters prestasjoner. Av den grunn velger jeg å se på vurderingsarbeidet som egenvurdering, der gruppen forstås som en enhet, og ikke hverandrevurdering, som ville satt sterkere søkelys på enkeltindividene i gruppen. Dette er også i tråd med barnehagens virksomhet, der man i stor grad arbeider i team. Å fokusere på hva man kan få til sammen, mer enn på enkeltprestasjoner, bidrar til å styrke fellesskapet og arbeidsmiljøet (Eriksen et al., 2022).

## Drøfting

Basert på erfaringene fra det gjennomførte utviklingsprosjektet, drøfter jeg i det følgende hvordan studentenes egenvurdering kan forstås som læring, her strukturert ut fra de nevnte fire kategoriene jeg fant i studentenes stikkord i vurderingsloggene: *kommunikasjon med barna*, *barns deltakelse*, *forberedelse* og *egenutvikling og samarbeid*. Avslutningsvis trekker jeg fram utsagn som viser sammenheng mellom nivåene i GLL-metoden.

## Kommunikasjon med barna

Når det gjelder kommunikasjon, trekker studentene fram praktiske ting som at barna skal se det som foregår og se studentene som forteller, og at de skal sitte komfortabelt for å holde fokus. Et poeng som kommer fram av loggene, er det å unngå forstyrrende elementer, som for eksempel: «I uterommet kan det lettere oppstå forstyrrende elementer som gjør at barna fort kan miste fokus, her må vi voksne være enda mer 'på' for å holde på oppmerksomheten deres.» Et annet sitat understreker også at

studentene har lært noe om viktigheten av kommunikasjon med barna: «Passe på at solen ikke treffer oss studenter, slik at vi ser barna godt.» Det er vanskelig å fastslå om studentenes valg når det gjelder organisering av samlingsstunden utendørs var et resultat av vanetenkning, eller hvorvidt studentene var forberedt på det som skjedde, og jeg tolker det som at de ble overrasket. De visste at de skulle ha samlingsstund utendørs på forhånd, men hadde ikke fullt ut vurdert at uterommet krever noe annet enn en innendørs samling. Vi ser at i første del av setningen refererer studentene til hva som har skjedd – at det har vært forstyrrende elementer, mens de i andre del av setningen reflekterer over hva de skal gjøre neste gang. Dette er en type kritisk refleksjon som også inkluderer tanker om framtidige handlinger, i tråd med Dewey (1910) og Mæhre og Vestly (2012).

## Barns deltakelse

Det å bruke barnas innspill trekkes fram flere ganger i loggene. Studentene mener det er viktig at barna aktiviseres, men sier også at barns innspill kan avspore og gjøre at noe tar for lang tid. Et utsagn som illustrerer dette, er: «Improvise rundt barnas interesser / endre fokus framfor å holde oss til bestemt tema.» Å utvikle improvisasjonskompetanse handler både om evnen til samspill med medstudenter og å kunne lytte til barns innspill og bygge videre på dem (Johnstone, 1987). Dette sitatet kan tyde på at vurderingsarbeidet har bidratt til en brobygging mellom undervisning på universitetet og praksis i barnehagen, og slik bidratt til læring. Utsagnet kan også forstås som endring av *referanseramme* (Mezirow, 2000, 2003). Studentene er vant til å forholde seg til planlagte opplegg og fastlagte tema. Denne erfaringen har gitt dem en ny innsikt, også kalt *mindfulness*: man kan i større grad frigjøre seg fra planene og bygge mer på barnas interesser (Mezirow, 1991).

## Forberedelse

Under temaet *forberedelse* er det svært mange utsagn i loggene, som det å planlegge, ha en god dialog med barnehagen på forhånd, innsisere stedet der samlingsstunden skal foregå, og ha god struktur på innholdet.

Sitatene som omhandler forberedelse, kan tolkes mer som en bekreftelse på kunnskaper studentene allerede har, snarere enn ny innsikt. Kanskje viser det at studentene har brakt med seg kunnskaper fra undervisningen, reflektert over disse i relasjon til sine erfaringer, og utviklet bevisstheten om handlinger som lærende personer: «Undersøk uteområdet man skal ha samlingsstund på, tenk gjennom hvilke forstyrrelser som kan oppstå: lyder, leker, bråk fra anleggsmaskiner, trafikk m.m.» «Ha en planlagt struktur på opplegget (med rom for improvisasjon).» Noen utsagn kan også tolkes som transformativ læring, der studentenes referanserammer har blitt endret og at de har møtt utfordrende situasjoner de ikke helt har vært forberedt på å møte: «Å la barna delta i samlingsstunda kan føre til avsporing av tema, gi rom for prat, men ikke så langt at man glemmer av hva man skal snakke om.»

## Egenutvikling og samarbeid

Når det gjelder *egenutvikling og samarbeid* framhever studentene viktigheten av å utfordre seg selv og å tørre å gjøre feil. En viktig forutsetning i dette gruppearbeidet er at alles stemmer blir hørt, og at man lar seg inspirere av de som har andre ideer enn en selv. Mezirow (2000) framhever at det er viktigere å ha en kultur for å få til konsensus, enn en kultur der målet er å vinne en argumentasjon. Dette krever at man lytter til hverandre for å finne løsninger når man er uenige. Samarbeid i gruppen er et moment som går igjen i loggene, samt det å ta seg god tid: «Det går heilt fint å gjøre feil, vi henter oss fort inn igjen.» Et annet eksempel: «Barns deltakelse kan føre til avsporing.» Kanskje kan denne refleksjonen bety at studentene har endret sine tanker om at på tross av gode intensjoner om barnas deltakelse i samlingsstunden (Eriksen, 2020), kan ting få et annet utfall enn planlagt (Isaksen, 2020). Når det gjelder egenutvikling, fant jeg flere utsagn som peker på at studentene har lært noe nytt og utviklet seg selv, som for eksempel: «Lært oss å utfordre oss på nye områder.»

## Gjort, lært og lurt

Hva mer i materialet kan forstås som vurdering som læring? Som jeg allerede har trukket fram, er mange av utsagnene fra studentene i stikkordsform, og det er ikke alltid lett å spore sammenhengen mellom innholdet i de tre nivåene *gjort-lært-lurt*. I noen eksempler er denne sammenhengen klar. «Vi fikk en helt ny barnegruppe med treåringer istedenfor fireåringer som vi hadde møtt før. Det gikk greit og alle var flinke til å gjøre litt om på planen.» Under *lært* skriver de: «Man må være tilpasningsdyktig da det kan skje uforutsette ting og ikke være redd for å improvisere.» Dette utsagnet er knyttet til at studentene måtte endre på sitt opplegg fordi de møtte en annen barnegruppe enn forventet, noe som hindret den planlagte progresjonen mellom de tre samlingsstundene. Her kan en tolkning være at de har anvendt kunnskaper og ferdigheter i improvisasjon som de har utviklet i løpet av studiet. Når studentene skriver at det er «lurt å ha god kommunikasjon med barnehagen», viser det at det har foregått en bevisstgjøringsprosess, noe som kan leses som å «se et spillbilde» av et alternativt valg for framtiden (Boal, 2006). Forhåpentligvis bidrar dette også til det Freire (1999) kaller myndiggjøring, ved at studentene kjenner seg tryggere på sine framtidige valg.

Andre eksempler der studentene trekker fram det de har lært, finnes i en annen logg. Her har en av gruppene reflektert over det de har gjort i samlingsstunden og trukket fram: «Vi skulle også hatt tydeligere kommunikasjon. Vi hadde en tydelig start, men skulle også hatt en tydelig avslutning.» De har stilt kritiske spørsmål og reflektert over det de ønsker å endre neste gang. I en av samlingsstundene inkluderte studentene en sang som barna skulle være med å synge og gjøre noen forhåndsbestemte bevegelser til. Dette ikke gikk helt som planlagt, og studentene rapporterte at de har lært at «barna var litt usikre på hva de skulle gjøre under myggsangen, og her skulle vi heller funnet noen andre bevegelser/leker de kunne gjort under sangen». Denne erfaringen førte til en klargjøring av hva som må endres til en annen gang: «Til fremtiden er det lurt å planlegge bevegelsessanger på en måte som gjør at alle barna kan delta og føle mestring uten at det er fastlagte bevegelser.»

En av gruppene framførte en skyggeteatersekvens for barna. I etterkant reflekterte de over at det er lurt å snakke høyt og tydelig og at ikke

for mange personer beveger figurer samtidig når man spiller skyggespill. Selv om disse forutsetningene for vellykket skyggespill var vektlagt i undervisningen og veiledningen før samlingsstunden, kan det virke som om erkjennelsen rundt disse poengene kom først etter å ha erfart hva som skjer når man *ikke* snakker høyt og tydelig og at for mange figurer beveges samtidig: «Ikke så mange som styrer figurer samtidig, lydene man sier under framføringen må være tydelig og høy slik at alle får det med seg.»

Noen av eksemplene tyder på at studentene har fått bekreftelse på ting de allerede visste. Dette handler ikke nødvendigvis om å endre sine framtidige handlinger, men at studentene gjennom en kritisk refleksjon og dialogen med de andre, har styrket sin bevissthet om hva som fungerer bra og kan danne grunnlaget for forbedring.

Enkelte utsagn fra loggene kan tyde på at studentene ikke fullt ut forstår den siste delen av GLL-metoden, eller at de ikke klarer å ha et kritisk blikk på sin egen utøvelse, men i stedet bekrefter det de har gjort og lært. Et eksempel som viser dette, er en gruppe som oppsummerer som *lurt* bare det de har opplevd som positivt, og det eneste de har *lært* er: «Vi har lært hvor enkelt det er å lage en sang.»

## Avsluttende betraktninger - vurdering som læring

Et av målene med dette utviklingsarbeidet er kvalitetsutvikling av egen undervisning og bedre tilrettelegging for studentenes læring. Et annet mål har vært å bidra til kvalitetsutvikling av undervisning generelt og av profesjonsutdanninger spesielt, ved å trekke fram og diskutere erfaringer med studenters egenvurdering og hvordan dette kan bidra til læring.

Materialet i utviklingsarbeidet er begrenset og tillater ikke å trekke noen generelle slutninger, men det gir nyttig informasjon for videreutvikling av et studieemne, samt inspirasjon til å fortsette å bruke GLL-metoden i studenters egenvurdering. Det gir også innsikt i utfordringene og mulighetene studentene møter i sin praksis og senere i sin profesjon, og hvordan utdanningen kan forberede dem på disse møtene. Dette er ikke et avsluttet prosjekt, og vi vil fortsette i en eksperimenterings- og læringsprosess, som i likhet med aksjonsforskning er «en syklisk prosess

av konkret eksperimentering og læring» (Levin, 2017, s. 35). I den videre anvendelsen av GLL-metoden i musikk og drama i barnehagen, vil jeg bygge videre på erfaringene fra dette prosjektet. Det betyr også at jeg evaluerer underveis og justerer bruken av metoden slik at egenvurderingen fungerer best mulig, og at jeg samtidig er åpen for å anvende andre måter å gjennomføre egenvurdering på.

De eksemplene som er trukket fram, indikerer at flere av studentgruppene har oppnådd læring gjennom vurderingen av eget arbeid med samlingsstundene. Når studentene har vurdert sitt kunstpedagogiske arbeid i samlingsstundene ved bruk av GLL-metoden, har de reflektert de over sin egen praksis som et viktig ledd i læringsprosessen. Studentene har ikke bare beskrevet det som har fungert godt, men også identifisert og reflektert over elementer som kunne vært bedre, og foreslått endringer for fremtiden. Noen av gruppene har imidlertid kun beskrevet aktivitetene de gjennomførte uten å stille kritiske spørsmål til sitt eget arbeid eller foreslå alternative handlinger eller innhold. Dette kan forstås som at alt fungerte som forventet, men det kan også tyde på manglende innsikt i GLL-metoden. Kanskje har innføringen i metoden vært mangelfull, med for få eller for lite relevante eksempler som grunnlag for fordypning og diskusjon? Og kanskje har studentene for lite trening i å reflektere, for eksempel ved å knytte sine erfaringer til fagteori? En annen forklaring kan være at vurderingskriteriene ikke har vært tydelige nok. Studentene har i stor grad selv bestemt hva som skulle vurderes og hvilke kriterier som skulle gjelde. Muligens har vi som faglærere tatt for gitt at studentenes vurderingskriterier samsvarte med mål og innhold i studieemnets undervisning.

Hva bør endres til neste gang studieemnet gjennomføres for å styrke studentenes evne til å vurdere eget arbeid? Dette leder til spørsmålene jeg stiller meg selv når det gjelder kvalitetsutviklingen av *min* undervisning med utgangspunkt i GLL-metodens nivåer *gjort-lært-lurt*: Hva har jeg *gjort*? Jeg har undersøkt hvordan studentene har brukt GLL-metoden i egenvurdering i sitt drama- og musikkpedagogiske arbeid i barnehagen, og jeg har *lært* at noen studenter har oppnådd læring gjennom vurderingsarbeidet, mens andre hovedsakelig beskriver sine erfaringer med positiv bekreftelse. Hva vil det være *lurt* å gjøre neste gang? En grundigere innføring i GLL-metoden er ett av tiltakene. Studentene trenger trening

i å se muligheter for endring og forbedring, i tillegg til å identifisere hva som har fungert godt. Klarere vurderingskriterier for samlingsstundene er et annet tiltak for å øke læringsutbyttet av studentenes egenvurdering. Disse kan utvikles i fellesskap med studentene, slik at de utvikler mer eierskap til hva som er planen og intensjonene med samlingsstunden. Vurderingen kan slik bli en kritisk og kraftfull refleksjon knyttet både til tankegodt og handling (Dewey, 1910; Søndena, 2004). Å trenes i å vurdere eget og andres arbeid kan på denne måten være et verdifullt bidrag til å utdanne reflekterte profesjonsutøvere.

## Forfatterbiografi

**Anne Eriksen** er dosent i drama og teater ved UiT Norges arktiske universitet. Hennes forskningsinteresser er nært knyttet til hennes undervisningspraksis i lærerutdanning og inkluderer drama- og teaterpedagogikk, estetiske læringsprosesser og anvendt teater. Eriksen har ledet et internasjonalt samarbeidsprosjekt, publisert en rekke vitenskapelige artikler og vært redaktør for tre antologier.

## Referanser

- Alteren, G. & Eriksen, A. (2020). What do students learn in a cross-cultural negotiation? A transformative approach. *Journal of International Business Education*, 15, 249–270.
- Befring, E. (2015). *Forskningsmetoder i utdanningsvitenskap*. Cappelen Damm Akademisk.
- Berkaak, O. A. (2003). Innledende essay. I P. Freire, *De undertryktes pedagogikk*. Bokklubbenes Kulturbibliotek, De Norske Bokklubbene.
- Boal, A. (2006). *The aesthetics of the oppressed*. Routledge.
- Braun, V. & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, 9(1), 3–26. <https://doi.org/10.1037/qup00001>
- Braun, V. & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. *International Journal of Transgender Health*, 24(1), 1–6. <https://doi.org/10.1080/26895269.2022.2129597>
- Burner, T., Baraas, R. C. & Falkenberg, H. K. (2011). Studentaktive vurderingsformer i norsk lærer- og optometriutdanning. *Uniped*, 34(1), 44–57. <https://doi.org/10.18261/ISSN1893-8981-2011-01-04>

- Dewey, J. (1910). *How we think*. D. C. Heath.
- Duggan, J. & Ofte, I. (2016). Hverandrevurdering som læringsverktøy i lærerstudenters tilegnelse av akademiske skriveferdigheter i engelsk. *Norsk pedagogisk tidsskrift*, 100(3), 207–221. <https://doi.org/10.18261/issn.1504-2987-2016-03-05>
- Earl, L. M. (2003). *Assessment as learning. Using classroom assessment to maximize student learning*. Sage Publications.
- Eriksen, A., Larsen, A. B. & Leming, T. (2015). Acting and reflecting; making connections between theory and practice in teacher education. *Reflective Practice*, 16(1), 73–84.
- Eriksen, A. (2009). Emosjoner i dressur. Forumteater i problemløsning og aksjonsforskning. I R. Jakhelln, T. Leming & T. Tiller (Red.), *Emosjoner i forskning og læring* (s. 179–195). Eureka forlag.
- Eriksen, A. (2020). En samlingsstund tar form – samlingsstundens dramaturgi. I A. Eriksen & B. Isaksen (Red.), *Kunsten å samles*. Universitetsforlaget.
- Eriksen, A., Olsvik, E. & Ryan, S. (2022). Noe om barnehageansattes kreative uro. I B. Isaksen & A. Eriksen (Red.), *Kunsten å være urolig*. Universitetsforlaget.
- Falchikov, N. (2005). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*. Routledge Falmer.
- Freire, P. (1999). *De undertrykte pedagogikk* (2. utg.). Gyldendal.
- Gjølterud, S. (2020). Forskning i egen undervisningspraksis i høyere utdanning. I S. Gjølterud, H. Hiim, D. Husebø & L. H. Jensen (Red.), *Aksjonsforskning i Norge: Bd. 2. Grunnlagstenkning, forskerroller og bidrag til endring i ulike kontekster* (kap. 8, s. 225–251). Cappelen Damm Akademisk. <https://doi.org/10.23865/noasp.121.ch8>
- Hannevig, L., Lundestad, L. & Skogen, E. (2020). *Pedagogisk leder i barnehagen: Samhandling, organisering og dialog*. Fagbokforlaget.
- Illeris, K. (2012). *Transformativ læring og identitet*. Samfundslitteratur.
- Isaksen, B. (2020). Kunsten å undre seg. I A. Eriksen & B. Isaksen (Red.), *Kunsten å samles*. Universitetsforlaget.
- Iversen, L. L. (2014). *Uenighetsfellesskap: Blikk på demokratisk samhandling*. Universitetsforlaget.
- Johnstone, K. (1987). *Improvisation og teater*. Hans Reitzels Forlag.
- Kristiansen, C. W. (2012). *Vurderingens kunst – en diskusjon om vurdering SOM læring i dramafaget* [Masteroppgave]. Universitetet i Agder.
- Levin, M. (2017). Aksjonsforskning som forskning – epistemologiske og metodiske utfordringer. I S. Gjølterud, H. Hiim, D. Husebø, L. H. Jensen, T. H. Steen-Olsen, E. Stjernestrøm, S. Gjølterud, D. Husebø & H. Hiim (Red.), *Aksjonsforskning i Norge: Teoretisk og empirisk mangfold*. Cappelen Damm Forskning. <https://doi.org/10.23865/noasp.17>

- Mezirow, J. (1978). Perspective transformation. *Adult Education*, 28(2), 100–110. <https://doi.org/10.1177/074171367802800202>
- Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass.
- Mezirow, J. (2000). Learning to think like an adult. I J. Mezirow & Associates (Red.), *Learning as transformation* (s. 3–33). Jossey-Bass.
- Mezirow, J. (2003). Transformative learning as discourse. *Journal of Transformative Education*, 1(1), 58–63. <https://doi.org/10.1177/1541344603252172>
- Moxnes, A. R. (2016). Refleksjon i barnehagelærerutdanningen. *Nordisk barnehageforskning*, 12(7), 1–13. <https://doi.org/10.7577/nbf.1563>
- Mæhre, K. S. & Vestly, S. (2012). Om å være i bevegelse mot en profesjonell fagutøvelse. *Uniped*, 35(1), 62–74.
- Postholm, M. (2007). Læreren som forsker eller lærer. *Norsk pedagogisk tidsskrift*, 32(3), 232–244. <https://doi.org/10.18261/ISSN1504-2987-2007-03-05>
- Rønne, N. N. (2017). *Grå*. Magikon forlag.
- Sambell, K., McDowell, L. & Montgomery, C. (2013). *Assessment for learning in higher education*. Routledge.
- Schellekens, L. H., Bok, H. G. J., de Jong, L. H., Van der Schaaf, M. F., Kremer, W. D. J., & van der Vleuten, C. P. M. (2021). A scoping review on the notions of Assessment as Learning (AaL), Assessment for Learning (AfL), and Assessment of Learning (AoL). *Studies in Education Evaluation*, 71, Artikkel 101094. <https://doi.org/10.1016/j.stueduc.2021.101094>
- Søndenå, K. (2004). *Kraftfull refleksjon i lærerutdanninga*. Abstrakt forlag.
- Tiller, T. & Gedda, O. (2017). *Gjort – lært – lurt. Nye verktøy for skolens læringsarbeid*. Universitetsforlaget.
- UiT Norges arktiske universitet. (2024). *MUS-2004 Musikk og drama i barnehage – 30 stp*. Hentet 4. april 2024 fra [https://uit.no/utdanning/emner/emne?p\\_document\\_id=841642](https://uit.no/utdanning/emner/emne?p_document_id=841642)
- Utdanningsdirektoratet. (2022). *Aksjonslæring*. <https://www.udir.no/kvalitet-og-kompetanse/utvikle-praksis/ideer-til-levende-moter-og-hvordan-utvikle-praksis/aksjonslaring/ring>
- Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Learning*, 37(1), 3–14. <https://doi.org/10.1016/j.stueduc.2011.03.001>
- Westrheim, K. (2004). Kritisk pedagogikk og multikulturalisme i lys av Freiretradisjonen. Noen sentrale perspektiv. *Nordisk Pedagogikk*, 24(3), 212–226.

I antologien *Vurderingsformer og tilbakemelding i kunstfaglig utdanning* finner du tekster som belyser ulike sider ved vurderingsformer og tilbakemelding i utdanning av musikere og kunstfagspedagoger. Forfatterne har utforsket og forsket på hva vurderingsformer og -situasjoner innebærer for studentene, lærerne, enkeltfagene, utdanningene og kunsten selv. Artiklene tar opp temaer som: hva som vurderes og hvordan det vurderes, læringspotensialet i vurderingspraksiser, opptaksprøver som vurderingsarena, elevmedvirkning som tilbakemelding i praksisveiledning, sensorveiledningene og læreplanenes funksjon, studenters egenvurdering, og hvordan man kan utvikle nyskapende vurderingspraksiser i kunstfaglig utdanning i dag.

Antologien svarer på et økende behov for å drøfte og ivareta kunstens særpreg og uforutsigbarhet i et utdanningsfelt der målbarhet og tydelige vurderingskriterier har en sentral plass. Leseren får forskningsbasert kunnskap om hva slags landskap vurderinger og tilbakemelding i kunstneriske fag kan innebære. I tillegg gir artiklene et godt grunnlag for større bevissthet og kritisk refleksjon rundt kunst og kvalitet.

ISBN 978-82-02-89535-8



www.cda.no