

# **Analytiske og holistiske tilganger til vurdering – en teoretisk drøfting av vurderingsstrategier og vurderingshandlinger i musikkfaget**

*John Vinge*

## **ABSTRACT**

### **Analytic and holistic approaches to assessment**

#### **– a theoretical discussion of assessment strategies and actions in music education**

*In recent years we have seen an increased emphasis on assessment as a central didactical tool at any level of schooling in the Nordic countries. New “assessment for learning” pedagogies in many forms seek to reform teaching practices in every subject, including music. Yet assessment in music education is an under-communicated theme in Nordic music didactical literature and research. This article investigates assessment theory in the context of music as compulsory subject in Norwegian lower secondary schools. Even though the argument made in this article is situated and exemplified in a specific national educational context, the theoretical discussions may be relevant and recognizable for all Nordic Countries at any level of music education. In this article, assessment is investigated through the concepts analytic and holistic. Empirical data from an ongoing research project is used to exemplify and at the same time develop the theoretical discussion. Epistemological questions underlying assessment theory, such as discourses of objectivity and subjectivity, reliability and validity are addressed. The article proposes an analytic framework for understanding and identifying assessment theory and practices in music education. Moreover, the point is made that the concepts analytic and holistic are not to be recognised as dichotomous concepts, but as an assessment continuum.*

*Key words: General music, assessment theory, analytic assessment, holistic assessment*

## **Innledning**

Musikkdidaktisk litteratur ser ut til å være lite interessert i å formulere noe om evaluering, og i generell didaktisk litteratur er det påtakelig hvor små innholdsdelene om evaluering er i forhold til den plassen som vies spørsmål om undervisningens mål og innhold (Johansen 1994: 54)

Dette skriver Geir Johansen i en artikkel om vurdering i *Musikkpedagogiske perspektiver* fra 1994.<sup>1</sup>

I Norge begynner dette bildet å forandre seg fra begynnelsen av 2000-tallet. Et søk etter vurderingslitteratur på BIBSYS viser et økende omfang av både nasjonal og internasjonal litteratur om emnet utover 2000 tallet, både generell vurderingsteori og praktiske veiledere i vurderingens “hva, hvordan og hvorfor” for de enkelte fagene. Innenfor academia ser vi også en økende interesse for vurderingsproblematikk i ulike forskningsprosjekter og i pågående og avsluttede oppgaver på master- og doktorgradsnivå. I Norge tilbys kurspakker for hele skoler i “vurdering for læring”, og på etterutdanningsprogrammene er elevvurdering det desidert mest populære kurstilbudet. Et økende fokus på vurdering i Norge utover 2000 tallet kan forklares dels ved innføring av de nye målbaserte læreplanene for grunnskolen og videregående opplæring (Læreplanen for Kunnskapsløftet, LK06).<sup>2</sup> Dels kan det forklares med en påvirkning fra internasjonale pedagogiske strømninger (jf. Britiske *Assessment Reform Group*<sup>3</sup>, Black & Wiliam, Sadler, Hattie m.fl.) der sentrale begreper som vurdering for-, av- og som læring, underveisvurdering og sluttvurdering, god feedback- og feedforward praksis, egenvurdering og hverandrevurdering, skriver seg inn i vår nasjonale politiske og pedagogiske diskurs. Men i all hovedsak kan det økende fokuset på vurdering i Norge på 2000-tallet sees på bakgrunn av til da manglende tradisjoner for systematisk vurdering og testing, en undervisningskultur som var preget av overfladiske innholdsaktiviteter uten spesifikke læringsmål i kombinasjon med “urovekkende” resultater fra internasjonale studier (jf. TIMSS, PIRLS og PISA) av norske skoleelever grunnleggende kompetanse i basale ferdigheter (Dale 2010). En god og systematisk vurderingspraksis, der elevvurdering er en sentral og integrert del av både planlegging og gjennomføring av undervisningen, ble en av metodene for å imøtekomme dette problemet. Et skjerpet fokus på grunnleggende ferdigheter og kompetanse med elevvurdering som sentralt didaktisk redskap skulle gjelde for alle fag og disipliner i norsk skole, også for musikkfaget.

I denne artikkelen undersøker jeg hvilke teorier og forståelser som ligger til grunn for ulike vurderingsstrategier i musikkfaget og hvordan disse forholder seg til hverandre. Konteksten er, som beskrevet ovenfor, norsk skole på 2000-tallet, men den teoretiske diskusjonen vil være relevant og gjenkjennbar for de andre nordiske landene også. For å konkretisere og kontekstualisere vurderingsteorien, vil jeg bruke noen empiriske eksempler fra et pågående forskningsarbeid jeg gjør på læreres vurderingspraksis i det obligatoriske musikkfaget i norsk ungdomsskole (8-10 klasse, 13-16 år). I dette forskningsarbeidet er 15 musikk lærere intervjuet. Lærerne er av ulik alder og kjønn, har ulik kompetanse- og ansiennetsbakgrunn og representerer således bredden av musikk lærerstanden i en større kommune øst i Norge. I tillegg til semistrukturerte forskningsintervju (Kvale & Brinkmann 2009), består det empiriske materialet av lærernes lokale årsplaner, semesterplaner og ulike vurderingsverktøy (skriftlige formulerte forventninger til måloppnåelse, prøveeksemplarer med mer). I skrivende stund er ikke dette empiriske materialet gjennomanalysert og eksemplene fra dette materialet, brukt i denne

artikkelen, er derfor ikke ment å fungere som empiriske funn, men snarere som oppdaterte eksempler fra et underrapportert forskningsfelt.

Denne artikkelen er ikke ment å være en “for eller imot-tekst” om vurdering eller karakterer i musikkfaget. Den er heller ikke ment å være en metodisk artikkel for “best practice”. Artikkelen er først og fremst en fagartikkel, ment som et deskriptivt bidrag for å forstå vurderingsfeltet i relasjon til musikkfaget. Vurdering i musikkfaget undersøkes i denne artikkelen som et knippe epistemologiske utfordringer (Dearden 1979), knyttet til spørsmål som hvordan vi kan vite hva eleven kan eller vet, om denne vurderingen kan være sikker eller sikker nok, om denne vurderingen er pålitelig og rettferdig og hvordan vi kan uttrykke denne vurderingen slik at eleven, klassekameratene, foreldrene og andre aktører i feltet kan ha tiltro til denne vurderingen. Artikkelen gir ingen endelige svar på disse spørsmålene, men søker heller å drøfte dette saksforholdet gjennom begrepsparet analytisk og holistisk (Sadler 1989). Før vi går inn på denne drøftingen skal vi først se nærmere på selve vurderingsbegrepet og hvordan dette kan forstås og avgrenses som et didaktisk begrep i forhold til musikkfaget og den musikalske gjenstand.

## Vurderingsbegrepet

Vurdering er et fenomen som kan forstås på forskjellige måter og forekomme på ulike arenaer. Som et grunnleggende allment fenomen, er vurdering knyttet til menneskelig persepsjon, valg og sortering. I en kulturell estetisk kontekst kan vurdering knyttes til våre smakspreferanser i avgjørelser av hva vi liker og ikke liker. I en didaktisk kontekst vil vurdering først og fremst være knyttet til elevens læringsprosess, som veiledende vurderinger for å øke elevens læringsutbytte eller som bedømmelser ved avsluttende prøver og tester. Hanken og Johansen (1998) definerer vurdering som “et strukturert og systematisk arbeid med å bedømme noe i forhold til visse kriterier” (Hanken & Johansen 1998: 117). Dette strukturerte og systematiske arbeidet forstås i denne artikkelen som de handlinger læreren foretar seg for å innhente informasjon om elevenes kunnskaper og ferdigheter, nettopp for å kunne gi veiledning underveis i læringsprosessen eller en “domsavsigelse” til slutt. Når Hanken og Johansen definerer vurdering som et strukturert og systematisk arbeid, vil det være naturlig å lese inn en gradforskjell i dette arbeidet. Som vi skal se, vil vurderingspraksiser der skjemaer for måloppnåelse, rubrikker for avkryssing av måleresultat, fylldige elevnotater som vurderingsdokumentasjon osv. inngår, kunne oppfattes som mer strukturert og systematisk enn praksiser med utpreget bruk av mer intuitive kjennerskapsvurderinger, der lærerens oppfattelse av hans elever lagres så å si i “lærerens eget hode”. Graden av systematikk vil også variere mellom formelle og uformelle vurderinger. I det samme en lærer rettleider eller korrigerer en elev som for eksempel lærer seg å spille gitar, vil det alltid ligge en vurdering til grunn. Hva læreren velger å si til eleven, hvilke tips eller råd, skryt eller skjenn

han gir, er et resultat av en vurdering. Disse uformelle vurderingshandlingene er mer eller mindre intuitive og reflektsive hos den trente lærer og vil kunne forstås som mindre systematiske og strukturerte enn de formelle vurderingene som gjøres ved skriftlige tester, spilleprøver, eksamener, opptaksprøver osv. Begrepet bedømmelse i Hanken og Johansens vurderingsdefinisjon forekommer både underveis i læringsprosessen, som en veiledende vurdering (formativ), og som en avsluttende tellende bedømmelse i slutten av læringsprosessen (summativ).<sup>4</sup> Bedømmelsen kan formidles til eleven verbalspråklig i form av muntlige eller skriftlige kommentarer, eller ved symbolske representasjoner som for eksempel graderte karakterer. Et klapp på skuldra, et smil eller et nikk, eller i negasjon; et sukk eller et rist på hodet, vil også kunne formidle en lærers bedømmelse, uten at vi går nærmere inn på disse gestiske aspektene her. Kriteriebegrepet i Hanken og Johansens definisjon må også undersøkes nærmere. Vi kan, som vi skal se, forstå kriterier som noe uttalt og eksplisitt, noe som er skrevet ned og gjort tilgjengelig for offentligheten. Men kriteriene kan også være personlige, uuttalte og vanskelig å verbalisere.

Begrepet vurderingshandling kan fungere utfyllende til Hanken og Johansens vurderingsdefinisjon. En vurderingshandling forstås som selve vurderingen, dokumentasjonen av denne vurderingen og rapporteringen tilbake til eleven (som en avsluttende karakter eller som en vurdering underveis). Denne handlingen vil da følgelig være mer eller mindre systematisk og strukturert, mer eller mindre intendert og profesjonell i praksis eller i den utøvende lærergjerningen.

## **Den musikalske gjenstand, den musikalske aktiviteten og musikkfaget**

Før vi ser nærmere på hvordan begrepene analytisk og holistisk kan elaboreres i henhold til musikkfagets vurderingsutfordringer, vil det være interessant, men også nødvendig å redegjøre for hvilket fag og hvilke faglige aktivitetsformer som konstituerer basisen for den forekommende drøftingen. Gjenstanden for vurdering i musikkfaget er ikke utelukkende musikken pr. se, det vil si kun musikken som klingende objekt underlagt en kvalitativ verdivurdering (slik vi kjenner det fra musikk-kritikken). Gjenstanden for vurdering i musikkfaget er først og fremst musikk og musikalske aktivitetsformer som *uttrykk* for læring av musikalske kunnskaper og musikkfaglige ferdigheter. For å konkretisere: En musikk-anmelder vil vanligvis ikke være opptatt av den læringsprosess som ligger bak la oss si Rihannas vokale prestasjoner, det er uvedkommende for den kvalitative verdivurderingen anmelderen gjør av hennes låt og hennes plate. Her vil det være musikken og det musikalske uttrykket slik den “kommer ut av høyttalerne” som vil være gjenstanden for vurderingen. Anmelderens vurderingsarbeid handler om hvordan man kan avgjøre denne musikkens verdi jf. respektive funksjonelle-, historisk-kulturelt situerte- og estetiske kriterier (se Nielsen 2002). Men for en lærers vurderingsarbeid vil nettopp det som ligger forut for den klingende musikken være det sentrale. Det er de kunnskaper og ferdigheter, som ledd i en læringsprosess og som muliggjør den klingende musikken,

læreren skal mene noe om. En kompliserende faktor her blir det faktum at en vesentlig kilde for informasjon om denne læringsprosessen allikevel må manifestere seg i den klingende musikken. Oppsummert: Hensikten med vurdering i musikk i en didaktisk kontekst er vurdering av og for elevens læring, men en vesentlig kilde til informasjon for å kunne gjøre denne vurderingen hentes gjennom den musikalske gjenstanden.

Dersom vi godtar det premiss at en vesentlig kilde til informasjon om elevens læringsprosess manifesteres i den klingende musikken, vil det være avgjørende å se hva ved denne musikken som det er mulig for oss å si eller mene noe om, og på hvilken måte. En fenomenologisk analyse av den musikalske gjenstand tegner et bilde av et musikalsk objekt konstituert av ulike meningslag (Nielsen 1998: 133-139). Noen av disse meningslagene er direkte håndgripelige for oss da de er begrepsliggjorte og fremstår som objektivt gitte, f.eks. de akustiske og strukturelle meningslagene (form, sjanger, time, instrument, akkorder, melodiske løp, osv.). En vurdering av disse ytre lagene vil kunne fortone som relativt uproblematisk, ved at musikkens objektive attributter, slik de høres som følge av en musikalsk aktivitet, måles og veies som rett eller galt, presist eller upresist (f.eks. falsk sang, feilspilling, feil form jf. sjangerspesifikke krav osv.). Mens de dypere meningsstrukturer, som musikkens uttrykksfullhet og skjønnhet, vil være tilgjengelige for oss i en mindre språklig forstand og den opplevde kvalitet vil derfor avhenge av lytterens (eller lærerens) erfaring eller preferanser. Selve den musikalske gjenstanden er med andre ord kompleks og mangefasettert. En fullstendig vurdering av dette objektet (det være seg både for deskriptive eller normativ-verdimessige formål) i seg selv vil måtte dra vekslers på analytiske deskriptive tilganger så vel som hermeneutisk fortolkende tilganger.

Når det gjelder musikkfaget som skolefag er også selve kunnskapsgrunnlaget komplekst og mangefasettert, konstituert av musikkvitenskaplige-, håndverksmessige og kunstneriske dimensjoner (jf. Nielsens modell om musikkundervisningens tredimensjonelle basis, Nielsen 1998: 110). Dette kunnskapsgrunnlaget blir i vekslende grad, og med ulik prioritering fra skole til skole, forvaltet gjennom fagets ulike aktivitetsformer (ibid: 295). I den gjeldene norske læreplanen er disse aktivitetsformene organisert i hovedområdene musisere, lytte og komponere, hver med sine respektive kompetansemål (LK 06). Vurdering i musikkfaget er et komplekst fenomen da selve den musikalske gjenstanden er kompleks, aktivitetsformene i faget forholder seg til og tar opp i seg den musikalske gjenstanden på ulike måter, og kunnskap og ferdighetsgrunnlaget består både av teoretiske og verbal-skriftlige dimensjoner samt estetiske erkjennende dimensjoner. Det vil med andre ord ikke være mulig å forene en felles vurderingsstrategi og en form for vurdering som kan brukes gjennomgående i musikkfaget. Vurderingsstrategiene vil til enhver tid måtte forholde seg til den konkrete gitte situasjonen, dennes musikalske gjenstand og utfoldelse. Det er med et slikt komplekst utgangspunkt drøftingen av analytiske og holistiske tilganger til vurdering i musikkfaget i det kommende løftes opp på et mer overordnet og prinsipielt nivå.

## Analytisk-holistisk

I vurderingslitteraturen er det vanlig å stille opp begrepsparet analytisk-holistisk for å argumentere for fordelene ved den ene versus den andre (se f.eks. Sadler 1987, 1989, 2005, 2007, 2009; Fautley 2010; Mills 2005; Sætre & Vinge 2010). I ytterste forstand vil en analytisk vurderingsstrategi søke å bryte læringsmålet ned til enkeltstående vurderingskriterier og gjøre disse eksplisitt kjent, helst skriftlig for eleven i forkant av en vurderings-situasjon. Et kriterium defineres hos Sadler som: “A distinguishing property of characteristic of any thing, by which its quality can be judged or estimated, or by which a decision or classification may be made” (Sadler 1987: 164). I norsk kontekst brukes *kjennetegn på måloppnåelse og vurderingskriterier* om hverandre og betyr det samme (Utdanningsdirektoratet 2007).<sup>5</sup> Kriteriene blir gjerne formulert som mestringsutbytte: eleven skal kunne demonstrere at han kan *det* og *det*, vet *det* og *det* osv. Av formative hensyn vil disse kriteriene også være tilgjengelig for eleven i begynnelsen av selve undervisningsforløpet, som viktige rettesnorer i læringsprosessen og mot den endelige summative vurderingen. Når læreren vurderer elevens læringsutbytte eller læringsresultat, vil han på en teknisk rasjonell måte kunne krysse av på et egnet skjema om hvorvidt eleven har oppnådd det læreren ønsker å måle. Neste steg i denne analytiske prosessen blir å summere opp antall poeng og relatere dette for eksempel til en på forhånd oppsatt poengskala som også angir en karakter. Denne såkalte “bottom-up”-strategien (Fautley 2010) søker å sikre en rettferdig, pålitelig og håndterlig vurdering av alle elevene.

Den motsatte strategien refereres gjerne til som holistisk eller global (Sadler 1989). 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 hentes ut fra lærerens “mentale lager” – en kombinasjon av faglig kompetanse og vurderingserfaring. Denne såkalte “top-down”-strategien (Fautley 2010) søker en kvalitativ, autentisk og helhetlig vurdering.

## Diskusjon

### Analytisk

Ett av hovedargumentene mot den analytiske strategien er at dens atomistiske (ibid.) eller segmenterte (Mills 2005) vurderingsformer ikke passer på alle typer fag og emner. Mange læringsområder er komplekse i den forstand at utbyttet ikke lar seg konseptualisere i “neatly packed units of skills or knowledge” (Sadler 1989: 23). Slike komplekse læringsområder identifiserer Sadler blant annet i performative, kreative og skapende fag. Her vil ikke elevenes læringsresultat, for eksempel manifestert i en komposisjon, nødvendigvis kunne bedømmes som rett eller galt, men snarere som karakteristisk

for en eller annen form av kvalitet. Kvalitet forstås som et integrert konsept, konstituert av potensielt uendelig mange ulike kriterier, og karakteriserer et arbeid som helhet (Sadler 2009). Man sier gjerne at ulike verk, for eksempel ulike elevkomposisjoner, har ulike *kvaliteter*, men at de er like gode – den helhetlige kvaliteten bedømmes likt. Elevene kunne ha fått til ulike ting. Et verk var godt fordi det var rytmisk repeterende og streng i formen, et annet var interessant fordi det fløt og var veldig variert. Fastsatte kriterier for bedømmelse kan derfor umulig dekke alle aspekter ved slike læringsformer eller svare for alle aspekter ved læringsutbyttet.

Et moment som går igjen i mitt datamateriale er forbundet med dette poenget. Utdanningsdirektoratet anbefaler skolene å utarbeide kjennetegn på måloppnåelse på tre nivå, lav, middels og høy, med distinkte kriterier innenfor hvert av nivåene (Utdanningsdirektoratet 2007). Disse kjennetegnene blir gjerne presentert for elevene i en rubrikk eller i et vurderingsskjema (se fig. 1). I mitt forskningsprosjekt forteller mine informanter om utfordringen i å velge vurderingsområder som dekker de områdene man anser som viktige, samt det å formulere disse vurderingsområdene i korte og poengterte vurderingsskjemaer. Her eksemplifisert ved Line:

Line: [...] jeg har nettopp gjennomført en oppgave hvor de har [...] ja de skulle lage en blues – den evinnelige bluesen som er så takknemlig [lattermild]. [...] Og det er en ganske sammensatt greie, for der blir de vurdert på hvordan de hadde jobba for å lage bluesen, altså prosessen, og på komposisjon. Og de ble vurdert på framføringa. Og til slutt så hadde de et refleksjonsnotat hvor de skulle vurdere sin egen læring, og der de hadde sjansen til å bøte på hva som ikke gikk så bra, for å vise at de hadde lært det. Det er verre. Da må man nesten dele den inn [vurderingsskjemaet] i så mange biter at det nesten blir komisk [...]

Når vurderingsskjemaene umulig kan dekke alle aspektene ved læringsaktiviteten, må lærerne foreta noen valg. I dette eksempelet på et vurderingsskjema for gitarspill (fig. 1), ser vi at læreren har valgt akkorder, utførelse og vanskelighetsgrad, som de aspektene som konstituerer kvalitet i gitarspill. Her kunne vi tenke oss andre områder som også kunne være relevante.

Måloppnåelse	Lav	Middels	Høy
<b>Akkorder</b>	Du kan spille akkordene hver for seg	Du kan spille akkordene i en sammenheng	Du kan spille akkordene rent og rytmisk i en sammenheng
<b>Utførelse</b>	Du kan fremføre deler av sangen	Du kan tidvis fremføre sangen uten stopp	Du har god flyt og fremfører sangen uten stopp.
<b>Vanskelighetsgrad</b>	Det er lav vanskelighetsgrad på det du fremfører	Det er middels vanskelighetsgrad på det du fremfører	Det er høy vanskelighetsgrad på det du fremfører

Figur 1:  
skannet originaldokument

### Ekspressive og instruksjonelle mål

Eisners begrep om *expressive objectives* (Eisner 1985a) rommer mye av den samme kritikken som den mot bruk av kriterier i komplekse læringsfenomen. Ekspressive mål spesifiserer ikke det ønskelige utfallet av undervisning, men henspeiler snarere på læringsaktiviteter der elevene utforsker, prøver og feiler, engasjeres skapende, undrende og kreativt og tillates å fokusere på aspekter ved den respektive aktiviteten som er av særlig personlig interesse (ibid.: 54). Eisners begrep om ekspressive mål ble i sin tid lansert som et viktig innspill i debatten om curriculum-design, og da som et supplement til den rådende Tyler-doktrinen med dens krav om klare spesifiseringer av læringsutbytte, såkalte *instructional objectives* (Madhaus & Kellaghan 1992: 130). Vurderingens funksjon i denne mål-middel pedagogikken blir å fange elevens atferdsendring. Læreren instruksjon fører noe inn i eleven (kunnskap og ferdigheter) og læreren måler om dette noe kommer ut av eleven igjen (den samme kunnskapen og de samme ferdighetene). I mange av musikkfagets kreative og skapende områder være umulig, og heller ikke ønskelig, å fastsette klare spesifikke læringsmål eller vurderingskriterier, av den enkle grunn at man ikke kan forutse hva elevene “bringer til bordet”. Fig 2 er et eksempel på et vurderingsskjema for komposisjon. Her skal elevene lage en dansecoreografi i grupper. Vi kan se at læreren har valgt noen vurderingsområder, men han har ikke skrevet inn graderte kriterier eller kjennetegn på måloppnåelse. Læreren forteller at grunnen til at dette feltet står tomt er at han finner det vanskelig å skriftliggjøre sine forventninger til elevene særlig når denne frie oppgaven kan ta ulik retning.

Mål	Lav	Middels	Høy
<b>Eleven:</b>			
Fanger rytme og puls			
Lager bevegelse som uttrykker rytme og puls			
Samspill med gruppa			
Kreativitet			

Figur 2: skannet originaldokument



Man kan styre læringen og læringsutbyttet gjennom en rekke oppgavespesifikasjoner og presise kriterier, men faren ved slike instruksjonelle mål blir at læringsaktivitetens potensial innskrenkes. Hos Sheridan og Byrne (2002) ble kravet om eksplisitte “objektive” kriterier, etter deres syn, negativt regulerende på utformingen av komposisjonsaktivitetene i den skotske videregående skolen. Jeg deltok selv som lærer i *ung.kom*-prosjektet i Oslo-Akershus i 2008.<sup>6</sup> På evalueringen av dette pilotprosjektet ble læreren på en av de videregående skolene som deltok, kritisert for å ha ledet elevene inn på et spor der fokuset var på tradisjonell satsteknikk i kontrast til mer innovative og kreative komposisjonsuttrykk. Forsvaret til denne læreren angikk vurderingshensyn: elevene skulle ha en karakter og grunnlaget måtte vurderes opp mot definerte kriterier for stemmeføring angitt i satslæra.

Den betingelsesløse frie oppgaven som Eisner promoterer i disse ekspressive målene er noe som informantene mine ser verdien av, men flere uttrykker usikkerhet i hvordan de skal inkludere disse i det formelle vurderingsarbeidet. En informant forteller om at han gjerne bruker frie lydformingsaktiviteter (som han også betegner som “ny musikk”), som en del av komposisjonsarbeidet i ungdomsskolen, men at denne læringsaktiviteten ikke teller som karaktergrunnlag.

Mats: [...] så jeg tror at grunnen til at jeg ikke bruker det som karaktergrunnlag i ungdomsskolen, er at det er såpass vanskelig for meg å tydeliggjøre hvorfor dette var et kjempebra resultat. Og så føler jeg meg veldig utsatt for at elever skal kunne si “hva faen er forskjellen på det bråket jeg lagde og det bråket de lagde?”

### **Objektivitet og subjektivitet**

Diskursen om *subjektivitet* og *objektivitet* (Orr 2010), underliggende i eksemplene ovenfor, er et sentralt aspekt ved vurdering som epistemologisk problem og gjennomgående i vurderingslitteraturen: “teachers want their assessments to be objective – when asked, this is the word they use” (Shepard i Abeles 2010: 177). Selv om subjektivitet og objektivitet er grunnleggende begreper innenfor vitenskapsteorien, filosofien og estetikken, vil det for denne artikkelen være fornuftig å forstå begrepene i mer common sense-forstand: En objektiv vurdering henviser til en vurdering som er upersonlig, upartisk og sann eller riktig for oss alle. En subjektiv vurdering vil være 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. Når lærere ønsker at deres vurderinger skal være objektive (jf. Shepard *ibid.*), har det en pragmatisk side like mye som en etisk side. Lærere ønsker å kunne forsvare sine vurderinger overfor eleven og hennes omgivelser, ved å peke på noen konkrete, “objektive”, gjerne målbare kriterier som er uavhengige av lærerens subjektive skjønn.

Vilde: [...] jeg har en mye bedre følelse nå når jeg setter de semesterkarakterene enn det jeg hadde før. Nå vet jeg at det [...] det er lettere når man har sånne gode kjennetegn at det er lettere at det blir objektivt [...]

Rolf: [...] mye av det positive ved å bruke et slikt vurderingsskjema er at du fjerner deg personlig litt i fra den vurderingsbiten som går på inntrykk og sånne ting. Det blir litt mer objektivt.

Una: [...] men det er veldig subjektivt når en elev forteller om en framføring hjemme så kan jeg fortelle om den samme framføringen, men det er jo vanskelig å bli enig om hva som har skjedd.

### **Tingliggjøringer**

Skriftlige vurderingskriterier eller kjennetegn på måloppnåelse, offentliggjort for alle aktører i vurderingsfeltet (elever, lærere og foreldre), fungerer i så henseende som en felles referanseramme for “objektivitet”. Slike skriftlige dokumenter er en tingliggjøring (*reification*) i Etienne Wengers terminologi: “we project our meanings into the world and then perceive them as existing in the world, having a reality of their own” (Wenger i Orr 2010: 7). Av formative hensyn tjener slike tingliggjøringer det formål å gjøre undervisningens objekter kjent, ved å klargjøre for elevene hva de skal strekke seg etter. Av summative hensyn fungerer disse tingliggjøringene som en kontrakt for aktørene i vurderingsfeltet, noe de kan “point to, refer to, strive for, appeal to, and use or misuse in arguments” (ibid.).

Vilde: når de [elevene] har fått et skjema på forhånd “dette skal til”, så er det veldig sjelden man får spørsmål i etterkant når de har fått en vurdering. Jeg opplevde det mye mer før da, også fra foreldre. Holdningen var sånn: “musikk er et fag der alle får gode karakterer” og hvis man ikke fikk det kunne det være mye spørsmål om hvorfor. Og det er det mye lettere å svare på nå da. Det er mye lettere å sette seg ned med skjemaet å spørre: “Fikk du til det som står på høy måloppnåelse her? Gjorde du det? Synes du det selv?”. Som regel så stopper det der da. Da får man ikke noe mer spørsmål om det.

Generelle undervisningsmål, for eksempel kompetansemålene slik de er uttrykt i den norske læreplanen for grunnskolen og videregående opplæring, LK06, er også en tingliggjøring i Wengers forstand, men de er abstrakte og må brytes ned og gjøres eksplisitte hvis de skal fungere som kriterier.

### **Prosedural objektivitet**

I tillegg til klare skriftlige vurderingskriterier, vil det å konstruere oppgaver som kun fokuserer på de aspektene vi kan tilskrive “objektiv” kvantifiserbar kunnskap, være en måte å eliminere “subjektive” vurderinger på. “We seek a objective world, objectively known” (Eisner 1998: 44). Eisner omtaler en slik metodisk fremgangsmåte som *procedural objectivity* (ibid.). En slik prosedyre kan være å innskrenke vurderingene til kun å gjelde de aspekter som kan avgjøres som riktige eller gale. Multiple choice tester er et eksempel på en metode for prosedural objektivitet. Utforming av spesifikke kriterier

forut for en vurderingssituasjon som reduserer menneskelig tolkning, er en annen. Figur 3 er et eksempel på en slik tankegang synliggjort gjennom et vurderingsskjema. Dette er et skjema som følger en komposisjonsoppgave der elevene skal lage musikk ved bruk av et sequencerprogram. Oppgaven er tydelig instruksjonell med en klar bestilling på form og antall takter. I følge dette vurderingsskjemaet kan eleven *noe* om han komponer 30 takter og *noe godt* om han i tillegg har med rett form. Her vil en lærer kunne parentesette alle aspekter som angår musikalske kvaliteter og kun konsentrere seg om “objektive” formalia. Vurderingskriteriet som omhandler det å være tro mot den musikkstilen eleven har valgt, det som vil kvalifisere for det å *kunne noe veldig godt*, kompliserer dette forholdet noe og avviker fra den prosedurale logikken.

Musikk 8-10			
Tema: <b>MUSIKKTEKNOLOGI</b>	<b>Kompetansemål:</b> Bruke digitalt opptaksutstyr og musikkprogram til å manipulere lyd og sette sammen egne komposisjoner		
Lærestoff: Manual og introduksjonsvideo	<b>Oppgaver:</b> Du skal komponere en ”låt” på 30 takter ved hjelp av musikkprogrammet SEQUEL. Du henter ferdige loops fra programmet, oppretter 6 spor og setter sammen komposisjonen etter følgende kriterier: Formen skal være: ABB”A 4 takter intro, 8 takter A, 4 takter B, 4 takter B”, 8 takter A og 2 takter coda Du kan velge musikkstil selv.		
<b>Kjennetegn for måloppnåelse:</b>			
	Kan noe	Kan godt	Kan veldig godt
Form og komposisjon	Du komponerer en låt på 30 takter	Du komponerer en låt på 30 takter og med riktig form	Du komponerer en låt på 30 takter og med riktig form. I tillegg har du vært tro mot musikkstilen du har valgt
Musikkteknologi			

Figur 3: skannet originaldokument

Den sikreste metoden for prosedural objektivitet er skriftlige “true or false”-tester som gjøres og rettes i et dataprogram; *Miles Davis var en kjent jazztrompetist: rett eller galt?* Slike analytiske vurderingsstrategier vil resultere i stor grad av reliabilitet (Stobart 2008: 110) ved å tilby prosedyrer for at like prestasjoner i én og samme klasse vurderes likt av læreren eller at ulike lærere vil kunne vurdere én og samme elev likt. For Eisner (1998) blir likevel et av problemene med slike prosedyrer, slik det også har vært argumentert tidligere i denne artikkelen, at “we often tend [...] to avoid studying what we cannot measure” (ibid.: 45).

## Holistisk

Der en analytisk vurderingsstrategi, i ytterste konsekvens, altså legger absolutte spesifikke kriterier til grunn i forkant av en vurderingssituasjon, og denne vurderingssituasjonen i “beste fall” er regulert gjennom prosedyrer som trakter etter et “objektivt” utfall, vil den holistiske vurderingsstrategien ikke legge slike kriterier eller prosedyrer til grunn i forkant. Men når vurderingen først er gjort holistisk, vil det også her være umulig å begrunne denne vurderingen uten å skjelve til kriterier (Sadler 2005: 179). I begrunnelsen av holistiske vurderinger anlegges kriteriene derfor i etterkant som elementer i en kvalitativ *argumentasjon*. For musikkfaget vil en holistisk vurderingsstrategi derfor være analog med den vurderingen som gjøres i musikk utenfor klasserommet. Vi begrunner vår kvalitative opplevelse av en konsert eller en plateinnspilling etter lytting med å gi uttrykk for hva vi syntes var bra og hva vi eventuelt ikke syntes var bra. Vi møter ikke musikken med en notatblokk i fanget for å krysse av om notene er spilt riktig, om bassen treffer basstromma med millimeterpresisjon, om vokalisten har et liberalt forhold til intonasjon osv. i håp om at summen av alle disse spesifikke kriteriene vil gi oss noe å lytte etter for senere å kunne konstruere et helhetlig kvalitetsbilde.

Sadlers definisjon av kvalitative vurderinger, i motsetning til kriteriebaserte vurderinger, er vurderinger som er “made directly by a person, the person's brain being both the source and the instrument for the appraisal” (Sadler 1989: 124). En kvalitativ vurdering har i følge Sadler noen distinkte kjennetegn (ibid.). For det første anvendes multiple kriterier. Ett enkelt kriterium sees alltid i sammenheng med en større kategori med

tangerende og overlappende kriterier. For det andre så er en del av kriteriene heller uklare (*fuzzy*) enn distinkte. Rett eller feil er distinkte kriterier. Originalitet eller kreativitet er på den andre siden uklare kriterier da alle nivåer mellom fullstendig uoriginalt og originalt, konformt og kreativt, er mulig. For det tredje så kjennetegnes en kvalitativ vurdering av at den kompetente vurdereren appliserer relevante kriterier, som følgelig vil kunne variere mellom elevarbeid: “Professional qualitative judgement consists in knowing the rules for using (or occasionally breaking) the rules” (ibid.: 124). Kvalitative vurderinger “regnes” ikke ut slik som vi har skissert det i en absolutt analytisk tilnærming. Når en kvalitativ vurdering først er gjort, kan denne derfor kun bestrides med en annen kvalitativ vurdering. Med andre ord kan slike vurderinger ikke gjøres til en formel eller en prosedyre som kan adopteres og brukes av en ikke-ekspert. Andre eksperter kan komme frem til liknende kvalitative vurderinger, da de også er bærere av en form for innforstått kunnskap, *guild knowledge*, (ibid.: 126) som kan betraktes som en erfaringskompetanse konstituert av tidligere kvalitative vurderinger gjort av eksperten selv, ofte i samarbeid med andre eksperter i vurderingsfeltet.

### **Connoisseurship**

For å gjøre slike helhetlige kvalitative vurderinger må man altså være en ekspert. En ekspert, eller en *connoisseur*, er en person som gjennom sin levde erfaring, kunnskap og kompetanse har en så gjennomgående kjennskap til sitt fagområde at han har evnen til å “make fine-grained discriminations among complex and subtle qualities” (Eisner 1998: 63). Connoisseurship eller kjennerskap er “the art of appreciation” (ibid.). Kjennerskapsvurderinger er i det “virkelige liv” ofte personlige, intuitive og tause vurderinger. Kjenneren trenger ikke å rettferdiggjøre, forklare eller overtale. Kjenneren vil til og med kunne ha vanskeligheter med å artikulere sine kvalitetsoppfatninger: “I do not know how to define quality, but I know it when I see it” (Dobson 2010: 32). En slik taus kunnskap forstår vi med referanse til Polanyi (2000): “Vi kan vite mer enn vi kan si” (ibid.: 16). Og her ligger mye av hovedutfordringen for holistiske kvalitative vurderinger i en skolekontekst: “Teachers concepts of quality are typically held, largely in unarticulated form, inside their heads as tacit knowledge” (Sadler 1989: 126). For mange av musikkfagets læringsmål blir det vanskelig for lærere å verbalisere hva som konstituerer helhetlig kvalitet, i hvert fall i et språk som kan forstås av ikke-eksperter. I en skolekontekst vil det først og fremst bety eleven, dernest hennes omgivelser: foreldre som vil ha begrunnelser for karaktersettinger og andre elever som kan mene at deres arbeid er vel så bra. Av formative hensyn er det en fare for at lærerens tause kvalitetsoppfatning kan virke uforutsigbar på elevens læringsprosess. I summative vurderinger blir slike kjennerskapsvurderinger ofte prisgitt den enkelte lærer, hennes kvalitetsoppfatninger og med den følgen at det er en fare for at ulike lærere vil kunne vurdere ett og samme arbeid ulikt. Det at kjennerskapsvurderinger kan variere mellom eksperter, sier også noe om at ekspertenes ofte tause kunnskap i stor grad også er en personlig kunnskap.

### Standardreferert vurdering

Hvordan tenker man seg en operasjonalisering av holistiske vurderingsstrategier i en skolekontekst? For det første så ligger premisset til Sadler fast, for mange av skolens emner vil holistiske strategier være den eneste fornuftige måten.<sup>7</sup> Sadler argumenterer for å legge standarder til grunn for kvalitetsvurderinger, i motsetning til detaljerte og presise kriterier som har “[...] been reduced to pea-sized bits to be swallowed one at a time – and for each bit, once only” (Sadler 2007: 390). En standard defineres hos Sadler som: “A definite level of excellence or attainment, or a definite degree of any quality viewed as a prescribed object of endeavor or as the recognized measure of what is adequate for some purpose, so established by authority, custom, or consensus” (Sadler 1987: 194). Denne definisjonen er kilde til noe språklig forvirring, for hva er det som eksakt er forskjellen på en vurdering basert på kriterier og en vurdering basert på standard? En standard i angloamerikansk curriculum-tradisjon er en beskrivelse av det nivået av kunnskap og ferdigheter som alle forventes å nå på gitte trinn, og beskrives derfor i generelle termer. Kriteriene er detaljerte presiseringer av hva som konstituerer kvalitet i disse standardene. Parallelt må vi derfor kunne forstå LK06 kompetansemål som standarder, da de angir hva som eleven forventes å kunne på gitte trinn. I engelsk og amerikansk skole er standardene mer detaljerte enn de norske kompetansemålene, men for både standarder og kompetansemål er presisjonsnivået ikke like spesifikt som i et kriterium.

Standardreferert vurdering er i prinsippet underlagt den samme filosofien som den kriteriebaserte vurderingen. Begge tilbyr et alternativ til normrefererte vurderinger, forstått som en rangering av elevene, der kvalitet blir et innbyrdes relasjonelt fenomen; *en elev er flink på grunnlag av at en annen elev er dårlig*. Og begge kan tilskrives formative hensikter; noe som gir elevenes arbeid en retning. Men der kriteriene, som vi har sett, er detaljerte eksplisitte beskrivelser på hva som konstituerer kvalitet, vil standardene måtte tolkes, forstås og operasjonaliseres. Lærerens forståelse av kvalitet, standarden eller kompetansemålet må ikke forbli taust og uttalt, men *tilbudt* elevene i en form som ivaretar alle aspekter ved faget og er relevant i forhold til læringsmålet. Sadler (1987) mener en kombinasjon av *eksempler* på arbeider som demonstrerer ulik kvalitet (for eksempel tidligere elevbesvarelser eller lærerens egen modellering av gode og dårlige prestasjoner) og *verbale beskrivelser* (som omtaler ulike kvalitative aspekter) er en måte for læreren å gjøre det taust eksplisitt. For veldig mye av musikkfagets performative og skapende sider vil *eksempler* være komplementært med lærerens egen undervisning, mens en verbal beskrivelse, uten å redusere et helhetlige kvalitetsbilde, vil være en større utfordring.<sup>8</sup>

Erik: Altså. Jeg prøver som jeg sa å tydeliggjøre forut for en vurderingssituasjon, [...] altså å modellere så godt jeg kan i forhold til musisering at vi enten lytter til eksempler hvor det er mer avanserte, kreative ting som jeg ønsker at de skal prøve seg på for eksempel.

Nina: [...] nå husker jeg ikke formuleringa, men ordet *flyt* er brukt da for det ordet som beskriver det som skjer når det virkelig sitter da [spill på instrument]. Og jeg tror kanskje ikke alle elevene [...] at du klarer å kommunisere til elevene hva det faktisk betyr, sånn at de faktisk skjønner at vi vurderer det også. Så det tror jeg kan være litt uklart for en del.

Jeg: Hvordan jobber du med det da?

Nina: Jeg prøver jo, når jeg synes en del har kommet dit at man kan snakke om de tingene, å vise litt hva jeg mener. Men det er vanskelig å vise selv forskjellen på bra og dårlig, for det blir på en måte tull og latter og fjas når man skal illustrere noe dårlig. Det er vanskelig å finne [...] ja jeg leter litt etter sånne måter å vise eksempler på. Hva er bra og hva er ikke fullt så bra.

Rolf: [...] jeg simplifiserer det også gjennom for eksempel å spille en låt for dem på en god måte og på en litt verre måte.

### **Kritikken som didaktisk redskap**

Kvalitetsvurderinger i form av *connoisseurship* eller kjennerskapsvurderinger, som vi har sett, er heller ikke operasjonelt i didaktisk forstand i *seg selv*. For at kjennerskap skal gå fra en personlig og uuttalt vurdering til å bli funksjonell i didaktisk forstand må vi i følge Eisner (1998) henvende oss til kritikken. Vi kjenner kritikkbegrepet fra kunstanmeldelsen som anmelderens bedømmelse *av* eller dom *over* et kunstprodukt, det være seg en spillefilm, en roman, en utstilling eller en konsert eller en innspilling. Der kritikens funksjon i det profesjonelle kunstfeltet som oftest likner det vi tidligere har omtalt som summativ, vil kritikken som didaktisk redskap innenfor utdanningsfeltet også ha formative hensikter. Når Eisner bruker kritikkbegrepet i hans vurderingsteori er det de faglige kvaliteter og faglige prosesser som ligger forut for selve bedømmelsen som han mener er verdifulle for pedagoger og utdanningsforskere.<sup>9</sup>

Det er kritikerens evne til å *se, høre og verbalisere* som er det sentrale. I følge Eisner er kritikerens primære oppgave å oversette “the essentially ineffable qualities constituting works of art into a language that will help others perceive the work more deeply” (Eisner 1985b: 217). Kritikerens oppgave er derfor pedagogisk i den forstand at han hjelper andre som ikke er kjennere eller eksperter, for eksempel elever, til å øke oppfatningsevnen og utdype forståelsen. Denne oversettelsen fra et modus, for eksempel det klingende materialet, til et annet, i vårt tilfelle, språk, er ingen *en-til-en-* korrespondanse. Det finnes ingen ekvivalensregler, derfor vil enhver oversettelse være en rekonstruksjon (Eisner 1998: 86). Denne rekonstruksjonen tar form som et narrativ med argumenter som aldri er uimotsagte; det finnes alltid alternative tolkninger (ibid.). En kritikk er med andre ord ikke objektiv i den forstand at alle kjennere vil konkludere likt, men kritikken er heller ikke forutsetningsløs. Hos Eisner består kritikken av fire dimensjoner: beskrivelse, tematisering, tolkning og vurdering.<sup>10</sup> *Beskrivelsen* består i å verbalisere det en har hørt eller sett i et språk som er analogt med det en søker å beskrive. I og med at slike

oversettelser ikke er ekvivalente, brukes et rikt og metaforisk språk slik at mottakeren blir i stand til å høre det kritikeren har hørt. *Tematiseringen* finner sted når kritikeren summerer opp det han mener er de viktigste eller de mest interessante forholdene ved det som er beskrevet. Det er disse aspektene eller temaene han ønsker å fokusere videre på. *Tolkningen* søker å forklare betydningen av disse temaene. Her trekker kritikeren inn både kontekst, kunnskap om og erfaring med liknende fenomen som nå blir vurdert, og tilbyr derfor mottakeren en sammenheng mellom teori og praksis. Praksis berikes og tydeliggjøres ved adekvat teori. Den konkluderende  *vurderingen* skjer når kritikeren offentliggjør sin mening gjennom å bruke kriterier som er relevante i forhold til det som blir vurdert. Med tanke på slike vurderingsforløp i en skolekontekst argumenterer Eisner for personbaserte vurderinger (ibid.: 102). Her vurderes ikke eleven opp mot forhåndsoppstilte kriterier eller i sammenlikning med andre elever, men i forhold til elevens egen kunnskap og ferdighetsutvikling: “To make such comparisons requires the ability to appraise the qualities of the student’s work and to have some sense of the direction in which it is going” (ibid.).

### **Autentisk vurdering**

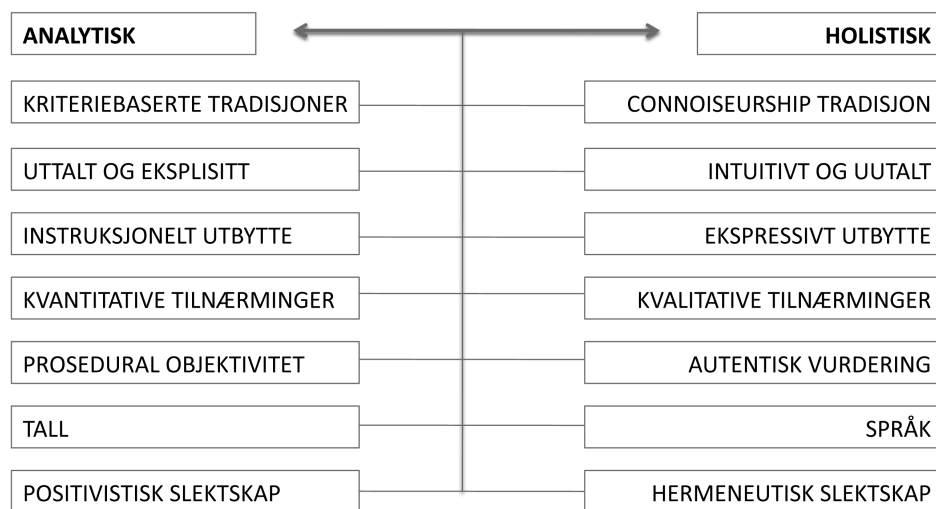
Sadler argumenterer gjennom sitt forfatterskap for holistiske standardbaserte vurderingsstrategier. Disse er av særskilt relevans for komplekse læringsfenomen. Eisner argumenterer for ekspressive læringsmål og fremhever personbaserte vurderinger basert på kjennerskap og kritikk. Vurderingsformene som brukes i disse holistiske strategiene må være relevante for, og tro mot, de objektene man ønsker å vurdere. Autentisk vurdering (se f. eks. Stobart 2008) er vurderingsformer som kan sies å ivareta validitetsaspektet av vurderingen i den forstand at den vurderer det den sier den skal vurdere og at vurderingsformen er “fit to purpose”. Man vurderer ikke nødvendigvis spilleferdighetene hos en elev ved å be eleven holde en powerpoint-presentasjon om Snoo Dog; “someone with a grade in music ought to have played something to someone” (ibid.: 109). Utfordringer med kvalitative holistiske vurderinger og kjennerskapsvurderinger er, som tidligere drøftet, reliabilitetsaspektet forstått som likebehandling av alle elever uavhengig av lærerens subjektive skjønn. Susan Orrs Bourdieu-inspirerte forskning (2010) viser hvordan autentiske vurderinger i kunstutdanning gjøres med bakgrunn av slike “subjektive” skjønnsmessige vurderinger. I hennes studie er lærernes subjektive vurderinger basert på kjennerskap aldri grunnløse eller tilfeldige, men formet av lærerens deltakelse i et praksisfellesskap med andre lærere. Grunnlaget for vurderingens “objektivitet” er at den er konstruert og regulert innenfor rammene av en bestemt sosial, kulturell og politisk diskurs (ibid.: 16). I relasjon til kjennerskapsvurderingens personlige aspekt (se avsnitt 4.2.1) vil vi kunne anta at der eksperter har en lang felles historie, for eksempel gjennom å være kolleger på en skole, vil hver og ens personlige kunnskap bidra konstituerende for en felles sosial kunnskap – *dette mener vi her*. Kvalitetsoppfatninger er derfor ikke private og uuttalte, men forhandlet, formet og gjort innforstått og virksomme i praksisfellesskapet. Dette betyr ikke at gode kolleger og kjennere



innenfor et felt ikke kan være uenige, men at de til syvende og sist taler med en tunge. Å nå konsensus må være målet for kvalitetsdiskusjoner ved formelle vurderinger, men diskusjonen som leder frem til enighet vil være både nødvendig og fruktbar.

## Oppsummering og konklusjon

Den teoretiske utlegningen i denne artikkelen kan summeres i følgende modell:



Figur 4

En oppsummerende konklusjon på denne fremstillingen vil være at i rendyrket form vil styrkene til den analytiske tilgangen være vurderingens reliabilitet. Tilsvarende kan det virke som at vurderingens validitet blir best ivaretatt i den holistiske tilgangen. Svakheterne til strategiene er det motsatte forholdet: Den analytiske strategien vil i rendyrket form ha utfordringer ved å ivareta validitetsaspektet for flere av musikkfagets læringsaspekter. Tilsvarende vil holistiske vurderinger forde upartiske, rettferdige og etterrettelige lærere for å ivareta reliabilitet innenfor en og samme klasse, samt et utstrakt praksisfelleskap med andre lærere for å sikre reliabilitet i vurderinger av elever på tvers av ulike skoler. Betingelsene for dette saksforholdet er, som vi har sett, den analytiske strategiens streben etter objektivitet, mens den holistiske strategien vil til syvende og sist fremstå som subjektive vurderinger basert på personlig og til dels taus kunnskap.

Dobson (2010) mener at diskursen om vurderingspraksis pendler mellom analytiske og holistiske tilganger. Når det blir et fokus og en overvekt på den ene praksisen, blir det et økt krav om mer av det andre og vice versa. I vurderingslitteraturen, slik vi har sett, blir også disse to tilgangene behandlet som ytterpunkter, der den ene argumenterer mot den andre ved å peke på feil og mangler. En slik dikotom fremstilling, som jeg også har anlagt i denne artikkelen, er problematisk uten forbehold. Slik jeg vurderer det, vil forbeholdene gjelde en erkjennelse av at virkeligheten ikke er like kategorisk som den fremstilte modellformen ovenfor. Et annet forbehold vil være faren for at en slik kategorisk oppstilling vil kunne oppfattes som at jeg fremstiller en tilgang som mer fordelaktig for musikkfaget enn den andre.

For å ta det første først. En utpreget bruk av vurderingsskjemaer med kjennetegn på måloppnåelse og vurderingskriterier, slik vi ser tendenser til i norsk skole i dag (se innledningen), tyder på en vurderingspraksis som nærmer seg en kriteriebasert tradisjon i en analytisk konseptuell forstand. Men der den analytiske tilgangen har som sin fremste "styrke" en parentessetting av subjektet og det subjektive, vil, så langt jeg kan se, ingen av de vurderingsskjemaene som er presentert for meg og som vi har sett eksempler på i den artikkelen, kunne fungere uten en videre forklaring eller presisering. Ordlyden i disse kjennetegnsformuleringene er heller ikke presis nok for å unngå alternative tolkninger eller diskusjoner. Videre: Modellen illustrerer en analytisk tilgang der tall har forrang over språk, mens det motsatte er tilfelle i den holistiske tilgangen. Dette er heller ingen absolutt fremstilling. For eksempel vil nettopp kvalitative kjennerskapsvurderinger i kunstnerisk kommisjonsarbeid eller i opptakskomiteer gjøres ved å plassere kandidatene i forhold til hverandre i en nummerert rekkefølge eller tildele kandidaten en poengskåre. Det er oftest først ved uenighet innad i kommisjonen eller ved uttalte begrunnelser at språket benyttes. Et tredje eksempel på modellens begrensning er den oppsatte dikotomien prosedural objektivitet vs. autentisk vurdering. For eksempel vil prøveformer som rytme- og melodidiktater etter mitt skjønn også kunne betegnes som autentisk vurdering. Selv om disse ofte er basert på en prosedural logikk i den forstand at eleven enten noterer rett eller galt og der besvarelsen til eleven også ofte relateres direkte til en på forhånd oppsatt poengskåre, fanger metoden det den faktisk er ment å vurdere: om elevene kan skrive ned med konvensjonell noteskrift det som blir avspilt for dem eller ikke. Videre kan man selvsagt diskutere hvorvidt prøveformen likner på situasjoner man kan komme opp i utenfor skolen, men her kan man argumentere for metodens autenticitet da ferdigheten med å transkribere klingende lyd i skrift er en vesentlig del av en musikerprofesjon på et eller annet nivå. Disse er bare noen eksempler på praktiske forhold som kompliserer modellen og vi kunne kanskje ønske oss en modellform som tar inn over seg alle mellomposisjoner og alle eventualiteter. Vel vitende om dette saksforholdet mener jeg allikevel at denne modellen kan bidra til å identifisere, teoretisere og sette ord på hovedlinjene i vurderingsteorien.

Når det gjelder om selve fremstillingen kan oppfattes som normativ, ved at en tilgang kan fremstå som mer fordelaktig enn den andre, tenker jeg at det ville være

fornuftig å behandle relasjonen mellom de to tilgangene som et kontinuum i stedet for å betrakte disse som to adskilte tilganger. Verdien av en strategi vil da avgjøres av den konkrete undervisningssituasjonen den er ment å undersøke, en oppfatning av vurderingsteori som også argumenteres for av flere (se bl.a. Fautley 2010; Mills 2005; Sætre & Vinge 2010). I et vitalt og flerfasettert musikkfag vil ikke alle læringsaspekter være like komplekse, men tydelig definerte ferdighets- og kunnskapsaspekter. Og jo tidligere man befinner seg i utdanningsforløpet, jo mer oversiktlige og konkrete vil disse aspektene være. På den andre siden vil heller ikke musikkfaget være vitalt og komplekst uten det usagte, uforutsigbare, personlige og intuitive, koblet til opplevelse og personlig erfaring. Læringsinnholdet og læringsaktiviteten må derfor bestemme vurderingsstrategi.

Videre virker det fornuftig å legge lærerens kompetanse til grunn i forståelsen av vurderingspraksis. Desto mindre *connoisseur* man er med tanke på musikkfaglig kompetanse og lærererfaring, desto større blir behovet for å *objektiverere* vurderingene. Desto større erfaring og faglig oversikt man har, desto mer kan man stole på ens egen personlige, innforståtte kunnskap og evne å gjøre denne transparent for elevene. En faglig *connoisseur* vil også ha en større formening om hvilke deler av faget som egner seg for de respektive vurderingsformene, og vil i så henseende kunne gjøre relevante vurderinger langs hele kontinuumet. Om denne læreren vurderer alene eller sammen med en annen vil også kunne virke inn på valg av vurderingsform: “Consensus breeds confidence in the objectivity of the judgements rendered” (Eisner 1998: 45). En vurderingshandling er ikke tilfeldig, men forholder seg mer eller mindre intendert til et planlagt læringsinnhold og et ønsket læringsutbytte. Vurdering som didaktisk fenomen vil derfor være relasjonelt forbundet med andre sentrale didaktiske kategorier. Det vil dreie seg om mål, arbeidsmåter, elevforutsetninger, rammefaktorer som tid, antall elever, undervisningsmaterieell osv. Lovverk og forskrifter vil også legge sentrale føringer på vurderingshandlingen, det vil også skolens *doxa* eller inneforståtte læringskultur og tradisjoner kunne gjøre. Slik avtegner det seg et bilde av et komplekst saksforhold med ulike ytre og indre strukturer som alle vil kunne virke regulerende både på valg av musikkfaglig innhold og på vurderingsform.

## Noter

- 1 Begrepene vurdering og evaluering brukes ofte om hverandre og betyr det samme. Men innenfor vurderingsteori er det en vanlig praksis i nyere tid å bruke vurderingsbegrepet på individnivå, f. eks. vurdering av en elevs kunnskap og ferdigheter, mens evaluering brukes på vurderinger av hele programmer eller prosjekter, f. eks. evaluering av L97. Vurderingsbegrepet slik det anvendes i denne artikkelen samsvarer med den engelske termen *assessment*. Evalueringsbegrepet i Johanssens tekst fra 1994 forstås som vurdering på individnivå.
- 2 Kunnskapsløftet er en utdanningsreform som ble igangsatt i Norge i 2006. Fra 2009 er nye læreplaner (LK06) innført i alle fag i grunnskolen og i videregående opplæring. Utdanningsreformen har fokus på grunnleggende ferdigheter i alle fag og innfører tydelige kompetansemål for

- elevenes læringsutbytte. I følge Erling Lars Dale (2010) markerer Kunnskapsløftet som reform overgangen fra *industriksamfunnets* skole til en skole i tråd med *kunnskapsamfunnet*. Forandring av læring og undervisning i skolen med en større vekt på resultatvurdering av elevene, er sentrale kjennetegn på denne overgangen.
- 3 <http://www.aiaa.org.uk/afl/assessment-reform-group/>
  - 4 Formativ og summativ vurdering (distinksjonen først innført av Michael Scriven i 1967, se Pinar 1995: 734) er to grunnleggende begreper som omhandler formålet ved en vurdering. En vurdering underveis i undervisningsforløpet som brukes av læreren for å modifisere egen undervisning og for eleven og læreren for å utforme det videre læringsopplegget, kalles for en formativ vurdering. Her handler det om å identifisere mulige gap mellom tilstand og formulerte mål og sette inn tiltak for å minske dette gapet. En avsluttende oppsummerende vurdering etter endt undervisningsforløp kalles for en summativ vurdering, og er gjerne forbundet med eksamen og karaktersetning. Det er min mening at utfordringene knyttet til vurdering i musikkfag på et epistemologisk nivå, som er denne artikkelens fokus, vil gjelde uavhengig av vurderingens funksjon. Det tilfaller derfor ikke denne artikkelen å diskutere disse grunnleggende begrepene mer i detalj. (Se f.eks Black & Dylan 2006 eller Gardner 2006 for videre utlegging av begrepene og oppdaterte diskusjoner om begrepene didaktiske utforming).
  - 5 Denne synonyme begrepsbruken er diskutert i Norge blant annet av Throndsen et al som hevder at det hadde vært mer riktig å si at kriterier er det som brukes som målestokk ved vurdering av elevbesvarelser, mens kjennetegn brukes om det som er karakteristisk for en besvarelse (Throndsen et al. i Dobson 2010: 30).
  - 6 Prosjekt i regi av Norsk Komponistforening og Oslo Filharmoniske orkester, der musikklinjene i Oslo-Akershus komponerer for en sinfonietta-besetning, og tilslutt får komposisjonen sin fremført i Oslo konserthus. <http://www.oslofilharmonien.no/lang/no/ung-arena/ung-kom/>
  - 7 Det må nevnes at Sadler skriver om vurdering i høyere utdanning der en rekke grunnleggende kunnskaper og ferdigheter allerede foreligger for anvendelse på komplekse områder.
  - 8 Sætre og Vinge (2009) presenterer *eksempler* på slike holistiske beskrivelser av elevkompetanse i musikkfaget. Disse vil kunne fungere både som kvalitative guider for læring for elevene, men også som eksempler på en kvalitativ argumentasjon som er kompatibel med det presisjonsnivået som forventes av musikkklæreren i formelle klagesaker.
  - 9 Eisners begreper om *connoisseurship* og *criticism* er først og fremst utviklet i relasjon til utdanningsforskning, nærmere bestemt til evaluering av skoleprogram. Men disse enhetene for analyse lar seg etter mitt syn fint applisere på individnivå. (For en mer utførlig utlegning av Eisners kritikk-dimensjoner og dens betydning på individnivå se Sætre & Vinge 2010: 172-173).
  - 10 I utlegningen av kritikerdimensjonene i Eisner 1985b (229-238) er tematisering som dimensjon utelatt. Denne dimensjonen er tatt med i revisjonen av teorien i Eisner 1998.

## Litteratur

- Abeles, Harold F. (2010). Assessing music learning. In: H. F. Abeles & L. A. Custodero (eds), *Critical issues in music education: contemporary theory and practice* (pp. 167-193). Oxford: Oxford University Press.
- Black, P. & Dylan, W. (2006). Assessment for Learning in the Classroom. In: J. Gardner (ed), *Assessment and Learning* (pp. 9-25). London: Sage.
- Dale, Erling Lars. (2010). *Kunnskapsløftet: på vei mot felles kvalitetsansvar?* Oslo: Universitetsforl.

- Dearden, R. F. (1979). The Assessment of Learning. *British Journal of Educational Studies*, 27(2), 111-124.
- Dobson, Stephen. (2010). Najonale og internasjonale utfordringer innen elev- og lærlingvurdering. In: S. Dobson & K. R. Engh (eds), *Vurdering for læring i fag* (pp. 24-36). Kristiansand: Høyskoleforl.
- Eisner, Elliot W. (1985a). *The Art of educational evaluation: a personal view*. London: Falmer Press.
- Eisner, Elliot W. (1985b). *The educational imagination: on the design and evaluation of school programs*. New York: Macmillan.
- Eisner, Elliot W. (1998). *The enlightened eye: qualitative inquiry and the enhancement of educational practice*. Upper Saddle River, N.J.: Merrill.
- Fautley, Martin. (2010). *Assessment in music education*. Oxford: Oxford University Press.
- Gardner, John. (2006). *Assessment and learning*. London: Sage.
- Hanken, Ingrid Maria, & Johansen, Geir. (1998). *Musikkundervisningens didaktikk*. Oslo: Cappelen akademisk forlag.
- Johansen, Geir Rosenløv. (1994). Evaluering av musikkundervisning *Musikkpedagogiske perspektiver* (s. 49-61). Oslo: Ad notam Gyldendal.
- Kunnskapsdepartementet (2006). *Læreplanverket for Kunnskapsløftet (LK06)*. Oslo: Utdanningsdirektoratet.
- Kvale, Steinar, & Brinkmann, Svend. (2009). *Det kvalitative forskningsintervju*. Oslo: Gyldendal akademisk.
- Madhaus, George F., & Kellaghan, Thomas. (1992). Curriculum evaluation and assessment. In: P. W. Jackson (ed), *Handbook of research on curriculum: a project of the American Educational Research Association* (pp. 119-154). New York: Macmillan.
- Mills, Janet. (2005). *Music in the school*. Oxford: Oxford University Press.
- Nielsen, Frede V. (1998). *Almen musikdidaktik*. København: Akademisk Forlag.
- Nielsen, Frede V. (2002). Quality and Value in the Interpretation of Music from a Phenomenological Point of View – a Draft. *Zeitschrift für Kritische Musikpädagogik*, 1-19. Retrieved from <http://www.zfkm.org/>
- Orr, Susan. (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), 5-19.
- Pinar, William F. (1995). *Understanding curriculum: an introduction to the study of historical and contemporary curriculum discourses*. New York: P. Lang.
- Polanyi, Michael. (2000). *Den tause dimensjonen: en innføring i taus kunnskap*. Oslo: Spartacus.
- Sadler, D. Royce. (1987). Specifying and Promulgating Achievement Standards. *Oxford Review of Education*, 13(2), 191-209.
- Sadler, D. Royce. (1989). Formative Assessment and the Design of Instructional Systems. *Instructional Science*, 18(2), 119-144.

- Sadler, D. Royce. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), 175-194.
- Sadler, D. Royce. (2007). Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education: Principles, Policy & Practice*, 14(3), 387-392.
- Sadler, D. Royce. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159-179.
- Sheridan, Mark, & Byrne, Charles. (2002). Ebb and flow of assessment in music. *British journal of music education*, 19(2), 135-143.
- Stobart, Gordon. (2008). *Testing times: the uses and abuses of assessment*. London: Routledge.
- Sætre, Jon Helge, & Vinge, John. (2009). *Vurdering i musikk på ungdomstrinnet: veiledere*. Oslo: Oslo kommune, Utdanningsetaten.
- Sætre, Jon Helge, & Vinge, John. (2010). Musikk og vurdering. In: S. Dobson & K. R. Engh (eds), *Vurdering for læring i fag* (pp. 166-177). Kristiansand: Høyskoleforl. Utdanningsdirektoratet. (2007). *Et Felles løft for bedre vurderingspraksis: en veiledning*. Oslo: Utdanningsdirektoratet.

PhD student, John Vinge  
Hedmark University College  
Faculty of Teacher Education  
Music Department  
Holsetgt. 31, N-2318 Hamar, Norway  
john.vinge@hihm.no