Nordisk musikkpedagogisk forskning. Årbok 18 2017, 241–270
Nordic Research in Music Education. Yearbook Vol. 18 2017, 241–270

Challenges of assessing music performance: teachers' perceptions

Elizabeth Oltedal

ABSTRACT

In recent years assessment discourse has given prominence to issues of measurement and accountability, despite the fact that it is principles of individually tailored goals and feedback that characterise the paradigm of educational assessment. It seems that teachers may experience conflicting values involving the dual roles of teacher and assessor, particularly in subjects with larger interpretive scope. Performance on main instrument at upper secondary school is a case in point, having strong traditions for individualised teaching, individually selected repertoire, and a broad range of instruments and genres. Instrumental teachers, who often have different views on knowledge and skills, informed by different discursive practices, are nonetheless part of a school-based assessment practice and accountable to the same curricular goals. This article draws on Wenger's (1998) concept of communities of practice, in which he highlights "the inseparable duality of the social and the individual", to explore the perceptions of instrumental teachers concerning assessment of main instrument. Analysis of data from two semi-structured dyadic interviews at Norwegian upper secondary schools reveals tensions between responsibility to the student, loyalties to personally held conceptions of musical quality, and accountability to the professional mandate. While expressing scepticism to assessment because of the difficulties of attempting to make a fair judgement, and the detrimental effects assessment can have on students, the informants have confidence in their own assessment practices on the basis of their professional knowledge and participation in social moderation. They value school-based moderation for its functions of providing quality assurance and development of assessment expertise.

Keywords: music performance, assessment, communities of practice, teacher identities

Introduction

Assessment discourse in recent years has given prominence to issues of measurement and accountability (Stobart, 2008; Torrance, 2007) despite the fact that it is principles such as individually tailored goals and feedback that characterise the paradigm of educational assessment (Gipps, 2012). Assessment is a major factor for students' learning and motivation (Black & Wiliam, 1998; Harlen & Deakin Crick, 2002), and in light of its 'double duty' to fulfil both formative and summative purposes (Boud, 2000), teachers may experience conflicting values involving the dual roles of teacher and assessor (Reinertsen, 2014; Yung, 2001). Where grading is involved, studies indicate that teacher judgements are often based on evidence not only of students' achievement, but also their effort and attitude (Brookhart, 2013). Summative assessment might be particularly challenging for subjects with larger interpretive scope (Prøitz & Borgen, 2010), and a case in point is performance on main instrument at upper secondary school, having strong traditions for individualised teaching, individually selected repertoire, and a broad range of instruments and genres. Although formative assessment is held to be an inherent feature of music teaching (e.g. Colwell, 2003; Nerland, 2003; Swanwick, 1998), requirements of assessment documentation, including summative judgements, place teachers of creative and artistic subject domains in a vulnerable position (Zandén, 2010a). The diversity and unpredictability of such activities make attempting to measure these by concrete learning outcomes problematic (Constantino & Bresler, 2010; McPherson & Schubert, 2004; Sadler, 2009, 2015), yet various alternative approaches to assessment, such as the application of criteria suitable for the individual's work, from a pool of potential criteria (Sadler, 1989); the practice of holistic assessment from a position of connoisseurship (Eisner, 2003); and the consensus of 'appropriate judges' (Amabile & Hogan, 1983), place emphasis on the expertise and autonomy of assessors. Instrumental teachers, while operating in the 'private room' of individual tuition, are nonetheless part of a school-based assessment practice and accountable to the same curricular goals. Assessment, seen as a situated, socio-cultural activity, thus involves "the inseparable duality of the social and the individual" (Wenger, 1998: 14), where local practices are shaped by individual 'participatory identities'. Using this aspect of community of practice (Wenger, 1998) as a conceptual frame, this article explores the perceptions of four instrumental teachers in Norwegian schools about assessing performances on main instrument. With the aim of contributing to knowledge of teachers' perceptions and practices of assessment on main instrument, the following research questions are asked:

- How do teachers perceive the task of assessing performances on main instrument at upper secondary level?
- Based on Wenger's (1998) axis of identity and practice, how might different identities of participation affect assessment?

The Norwegian context

In Norway, elective music programmes at upper secondary level offer a broad range of instruments and genres, and tuition in main instrument is individualised. Objectives for the subject in the National Curriculum for Knowledge Promotion (LK06) (Kunnskapsdepartementet, 2006) range from preservation and renewal of the musical life and heritage of the local and broader community, to the individual's development of "qualities important to master for a musician" (Utdanningsdirektoratet, 2011). While it is competence, and not participation or effort, that is the focus for assessment, the competence goals include performing a varied repertoire, mastering elementary techniques, making independent choices, and developing personal expression (ibid.). These goals reflect values of student growth and the European ideal of Bildung (Klafki, 2001), conceptualising an "entrepreneurial learner-citizen" who is both a producer and a consumer of culture (Finney, 2013: 149). In line with a central aim of the Knowledge Promotion Reform to enhance teachers' awareness of assessment (Hodgson, Rønning, Skogvold, & Tomlinson, 2010), the steering documents provide only general descriptors for quality, placing responsibility for the concretising of criteria and standards, as well as the organising of formative and summative assessment, at the level of the local school (Utdanningsdirektoratet, 2014). Thus teachers assessing performance on main instrument, though accountable to government steering documents, have considerable room for autonomy. Yet assessment of performance on main instrument carries complex challenges for a number of reasons. Instrumental teachers at various levels often have different views on knowledge, skills, and educational intentions, informed by discursive practices (Angelo, 2012; Asp, 2015; Georgii-Hemming, 2005; Nerland, 2003; Schei, 2007). In the context of individualised teaching at upper secondary school, complex relationships are forged between student and teacher, e.g. teachers may have strong empathy for their students but also expectations that students prioritise their main instrument (Ellefsen, 2014), and assessment might involve roles as different as 'companion' and 'policeman' (Yung, 2001). At the same time, students entering the

upper secondary music programme[1] themselves have a range of intentions for the subject as well as different levels of competence in instrumental skills and related subject domains (Ellefsen, 2014). Also, given that assessment is normally school-based until the public examination in the final term[2], there is room for considerable variation in the ways assessment might be locally organised and understood, e.g. in combinations of the cumulative observations made by the main instrument teacher, or judgements of individual performances involving social moderation (Adie, Klenowski, & Wyatt-Smith, 2012).

There is a long tradition in Norway of scepticism to formalised assessment and grading (Lysne, 2004), and in particular for music teachers to rely on experience and routine rather than engaging with new curricula (Johansen, 2003). Nevertheless, there are indications that teachers focus more after the Knowledge Promotion Reform than previously on assessment and documentation of assessment (Hodgson et al., 2010), and this tendency is also noted in music as a subject at lower secondary school (Vinge, 2014). In what ways the subject of main instrument at upper secondary level might be affected, has not been investigated. However, the particular challenges for assessment of performance on main instrument might differ significantly from some of those experienced in other subjects. For instrumental teachers, the problem is not likely to be the requirements of formative assessment, e.g. finding time for individual student-teacher conversations and feedback (Hodgson et al., 2010), but rather, the requirement to provide qualified, comprehensive and comparable evaluations of students' competences in a context of infinite variables.

Perception and assessment of music performance

Perception of music involves a range of 'interpretive moves', involving musical and extra-musical associations simultaneously (Feld, 1984). According to Green (1988, 2005), these 'inherent' and 'delineated' meanings depend on historical and sociocultural mediation, yet music has autonomy as a perceptible object, since it "cannot be *whatever* people say it is" (2005: 90). Tensions inhabit this dialectic understanding

1 MDD – *Musikk dans drama* is a combined performing arts programme for the foundation year, specialising in one subject domain for the remaining two years.

2 Although there is the possibility of random selection for external examination of main instrument in year 12, most students have no experience of external assessment in music performance until the final summative performance examination.

of musical perception, since while music has certain characteristics affording meanings that can be agreed upon, to recognise those meanings requires knowledge that is socially situated. The goalposts of musical meaning are therefore negotiable, and this has implications for assessment in light of teachers' professional expertise and identity. Emphasis might be placed on performance affording reproduction of music according to genre-specific criteria, or on the communicative activity of musicking (Small, 1998), where relationships between participants, whether performers or audience, are as meaningful as the sounds produced. A further focus is the understanding that performance affords the mediation of a 'persona' (Auslander, 2006), with possibilities for signifying various forms of authenticity (Gates, 1988; Moore, 2002; Weisethaunet & Lindberg, 2010), and recent research into the physiological, embodied aspects of music performance (Davidson, 2001, 2012, 2014; Leman & Godøy, 2010; Liao & Davidson, 2015) may have made it more admissible to talk about physical and visual elements for *all* genres. Indeed, where there has been a tendency for popular genres to be judged against the norms of classical genres, preserving hegemonies of Western aesthetic values (Danielsen, 2006, 2016; Green, 2005; Middleton, 2000), the pool of potential criteria for performance may not only have expanded, but even shifted in favour of new hierarchical categories (Dyndahl & Nielsen, 2014). Within this broad spectrum, to use universalising criteria (Green, 2014), such as 'expressivity' or 'innovation', can lead to misunderstandings when their meanings within different genres diverge considerably, and when recognising such idiosyncrasies depends on genre-specific knowledge (Danielsen, 2016).

However, the extent to which it is conformity to criteria from within a given musical practice that will be emphasised, or other, more student-centred principles, can be of great importance for assessment. According to Allsup and Westerlund (2012), each of these approaches can hold moral dilemmas. The former, representing values of praxial music education (Elliott, 1995), places the yardstick for assessment with the teacher's expertise and may restrict her imagination to 'accepted' boundaries for the genre in question, while the latter, influenced by Green's work (2008) on informal learning strategies, may give the student motivation and autonomy, but go no further than these "deliberate but limited starting points" (Allsup & Westerlund, 2012: 134). Applying Deweyian concepts, they argue that education should be a moral enterprise where the educational aims are wider than mere disciplinary knowledge, and warning that the arch-enemy of teachers' professionalism is an assumption that musical ends are ends in themselves, unrelated to other values. Amid concerns that the perceived demand for transparency has led to criteria compliance (Torrance, 2007), and a "trivialisation of teaching and learning" (Zandén, 2010a: 140), the aforementioned

competence goals for main instrument in LK06 do not in themselves tie teachers down to technical rationalities, but can be interpreted widely. It is at the level of teacher and teacher community that these goals are operationalised in assessment of the individual student's work, and it is in teachers' actions and choices that values, or the lack of them, are manifested (Allsup & Westerlund, 2012). Defining 'what counts' (Bernstein, 1971) as a valid realisation of the knowledge implicated in curricular goals, thus depends not only on the types of knowledge with which teachers identify, but also on underlying educational values. Furthermore, what counts within a school-based assessment practice, in the absence of detailed criteria or descriptors, must to a large extent depend on the ability of teachers to articulate and share their views (Zandén, 2010a). In this respect, the community of practice (Wenger, 1998) offers a useful frame within which to discuss assessment practices.

Community of practice

According to Wenger (1998), it is in sustained pursuit of a joint enterprise that the practices of a social community are formed and continually renegotiated by its individual members, developing "a shared discourse reflecting a certain perspective on the world" (1998: 125–6). With emphasis on the roles of practice and identity, Wenger describes how members of a community of practice develop 'identities of participation', whereby their experience and competence is shaped by, and can contribute to the continuous shaping of, the practice. Since participation is also a constituent of identity in the broader context of membership in other communities, participation in one community is informed by those memberships, and vice versa. Repertoires of meanings and processes that are formed and negotiated in a community can thus reflect specific backgrounds of particular members, and power relations are implicated within this duality of the social and the individual. For newcomers, participation might initially be peripheral (Lave & Wenger, 1991) or involve a role of brokering between insider and outsider perspectives (Wenger, 1998), affecting the balance of explicit and tacit knowledge (Polanyi, 2009) characteristic of the 'economies of meaning' developed in practices.

In professional life, "it is the collective construction of a local practice that, among other things, makes it possible to meet the demands of the institution" (Wenger, 1998: 46). However, even when a community of practice arises in response to a mandate, the practice will evolve as that community's response to the mandate, and the institution's

power over the practice will only be experienced indirectly (Wenger, 1998). Thus relations of mutual accountability, developing over time through negotiation of the joint enterprise, can lead to particular interpretations of standards, and to certain aspects of the mandate being given higher priority than others. In this article, the concept of community of practice is used to explore assessment practices as described by four Norwegian instrumental teachers. The possibility of a range of educational values and artistic discourses impinging on identities of participation and shared repertoires, makes Wenger's (1998) conceptual approach particularly apt.

Researching teachers' perceptions

The term 'perceptions' is used in this study to include beliefs and conceptions relating to assessment practices. A commonly held view of the individual and subjective beliefs of teachers is that they are both value-laden and relatively stable mental constructs, having significant impact on classroom practice and interpretations of it (Skott, 2015: 19), although there is also considerable research showing discrepancy between beliefs and practice (Fives & Buehl, 2012). While beliefs can be viewed as "the single most important construct in educational research" (Pajares, 1992: 329), both beliefs and conceptions of knowledge are dependent on subjective judgement (Pajares, 1992). Skott (2015) sees a need for research into how teachers' engagement in educational meta-discourses relates to their educational experiences, also concerning "the teacher's participation in a range of other practices at and beyond the school and classroom in question" (p. 26). In the current study, exploration of instrumental teachers' perceptions about assessment can give insights into how different 'participator identities' might contribute to assessment practices and potential challenges.

Relevant research

Much of the research in the field of music performance assessment has involved evaluation of performances by independent raters, focusing on two aspects: the development of tools for assessment in the form of rubrics and descriptive statements; and influences on judgements and inter-judge reliability. A third branch of research with relevance for this article focuses on teachers' perceptions in relation to the contexts for assessment. A summary of relevant findings is given in the following.

Development and validation of assessment tools

The large number of studies identifying variables for performance in classical genres traditionally emphasise aurally perceived factors of expressivity and technical accuracy (e.g.Bergee, 1995; Russell, 2015; Saunders & Holahan, 1997; Wesolowski, 2015). Technical skills are understood to be fundamental to both perception of musical expression and overall performance quality in classical genres (Russell, 2015), and to 'creative development' in jazz improvisation (Smith, 2009). More obviously visual factors are incorporated in studies involving live vocal performance (Coimbra, Davidson, & Kokotsaki, 2001; Davidson & Da Costa Coimbra, 2001; Gynnild, 2015), e.g. posture, charisma, scenic presentation and staging (Gynnild, 2015). Descriptive statements developed for criteria-based assessment across instruments by Stanley et al. (Stanley, Brooker, & Gilbert, 2002) include fidelity to the composer's text, ensemble skills where appropriate, "musical creativity, artistic individuality and effective audience communication" (54). This underlining of performers' autonomy and responsibilities to both co-musicians and audience is also found in criteria involving popular genres that are largely ensemble-based, e.g. jazz (Barratt & Moore, 2005) and rock (Blom & Encarnacao, 2012), bringing a series of 'soft skills' into alignment with the 'hard skills' of technique and stylistic accuracy (ibid.). Research investigating the negotiated development and use of criteria in authentic contexts suggests that implicit knowledge plays an important role in assessment (Davidson & Da Costa Coimbra, 2001; Gynnild, 2010, 2015; Stanley et al., 2002).

Influences on judgements and inter-judge reliability

Multiple factors are found to have importance for judgement, including performers' gender and race (Davidson & Edgar, 2003; Elliott, 1996), physical appearance and attire (Davidson & Da Costa Coimbra, 2001; Griffiths, 2010; Howard, 2012; Ryan & Costa-Giomi, 2004; Wapnick, Darrow, Kovacs, & Dalrymple, 1997), facial expressions, body movement and gesture (Juchniewicz, 2008; Lehmann & Kopiez, 2013; Platz & Kopiez, 2013). The question of whether holistic judgements or segmented protocols give more dependability for assessment has not been conclusively answered (Bergee, 2007), but different factors may come into play with the two procedures (Ciorba & Smith, 2009). In a study investigating the effects of introducing criteria-based assessment in a conservatoire setting (Stanley et al., 2002), examiners expressed ambivalence to the use of criteria since, while providing useful focus for assessment and feedback, this could impose limitations on examiners and constraints on holistic assessment. Similar findings are reported by Gynnild in a study involving vocal

teachers at a Norwegian conservatoire (2015). Research relating to assessors' own level of expertise and principle instrument seems inconclusive (Hewitt, 2007), but there are indications that both expertise and familiarity with presented repertoire are salient influences on assessment (Kinney, 2009), and that perceptions of identity connected to instrument and previous experience in assessing have importance for assessment practice (Maugars, 2006; Persson, 1994; Vinge, 2014). Research on assessment of performance on main instrument at upper secondary school is scarce, but an exception is Rui's (2010) study involving Norwegian instrumental teachers. Comparisons between individual teachers' allocated grades and written statements concerning video recordings of student performances suggest that, where there is consensus on grades, there are nevertheless discrepancies between what dimensions are valued in assessment. In addition, Rui suggests that expressivity is considered less in performances at a lower technical level, supporting existing research (e.g. Prince & Hallam, 1996; Young, Burwell, & Pickup, 2003), but also that teachers who emphasise expressive aspects of performance are less severe in assessment than those who focus more on technical aspects (Rui, 2010).

Where several assessors are involved, although not needing to be experienced performers on the instrument in question, a background in the same general family of instruments is found to increase reliability (Bergee, 2003, 2007; Fiske, 1975). These findings align with those of Amabile and Hogan (1983) concerning the consensus of 'appropriate judges' in assessment of creative work. In addition, stability is found to improve with increased panel size (Bergee, 2003; Fiske, 1975, 1977), while a panel size of two or three is not recommended (Bergee, 2003). Assessment using social moderation, however, holds complex issues. Despite perceptions that assessment and assessment competence are strengthened by the use of moderation (Vinge, 2014), several studies find that communities of music teachers have difficulties in verbalising and exchanging conceptions of quality, and that implicit understandings can be problematic (Asp, 2015; Davidson & Da Costa Coimbra, 2001; Gynnild, 2015; Vinge, 2014; Zandén, 2010a, 2010b).

Teachers' perceptions and assessment practices

Research investigating various aspects of instrumental teachers' professional identity and practice reveals different values for performance as a subject domain (e.g. Angelo, 2012; Asp, 2015; Georgii-Hemming, 2005; Nerland, 2003). Values of building a positive self-image, personal enjoyment and enabling of active participation in the community are important in general music programmes (Angelo, 2012; Georgii-Hemming, 2005).

Similarly, in Zandén's (2010b) study concerning popular music ensembles in Swedish upper secondary schools, teachers placed high value on authenticity in the form of students' autonomy, enjoyment and physical expressivity (Dyndahl & Nielsen, 2014; Zandén, 2010b). In contrast, Asp (2015) identifies both instrumental skills, breadth of repertoire and deep knowledge of individual genres among learning objectives for ensemble work. Research concerning main instrument at upper secondary and conservatoire level emphasises the appropriation of conventions of performance practice and repertoire in preparation for a professional career (e.g.Angelo, 2012; Ellefsen, 2014; Nerland, 2003; Persson, 1994), and a discourse of 'specialisation' is identified within the Norwegian upper secondary programme, where a trajectory towards expertise at a professional level is assumed (Ellefsen, 2014: 258).

There is little research exploring the perceptions of instrumental teachers about assessment, but one study (Maugars, 2006), investigating the attitudes of music teachers to examination of their students by external jury at French conservatoires, found that teachers criticised the assessment system in light of personal experience as students, yet perpetuated the same system when they themselves became teachers. The possibility that teachers feel prestige reflected in results is suggested since they felt uncomfortable if their students failed, and proud when they succeeded. In Vinge's (2014) study of music teachers in Norwegian lower secondary school, the task of allocating grades is experienced as a particularly challenging aspect of assessment work, and teachers avoid giving low grades to students who do not score highly in tests, but demonstrate effort and participation in class activities.

In summary, diverse studies indicate tensions involving various types of criteria and implicit knowledge in assessment, but there is little research exploring how instrumental teachers themselves perceive their interactive and locally situated practices. Furthermore, none of these studies addresses issues of assessment of the broadly different types of performance regarding genre, repertoire and instrument in the context of upper secondary school. The purpose of this study is therefore to contribute to knowledge on the perceptions of instrumental teachers about assessment of performance on main instrument at upper secondary level, and to gain insight into the interaction of different 'identities of participation' (Wenger, 1998) in assessment situations.

Research design and method

Taking the perspectives and accounts of individuals as a starting point (Ritchie, Lewis, Nicholls, & Ormston, 2014), the present study uses qualitative methods with the intention of exploring phenomena 'from the interior' (Flick, 2009). Central to qualitative research is that meanings and events will always be subject to interpretation, placing agency with the researcher, and making the context in which data is generated, interpreted and presented of crucial importance (Ritchie et al., 2014). However, an interpretative stance does not preclude discussion of realities as more than individual constructions. Drawing on the ontological principles of critical realism (Bhaskar, 2008), this study explores the statements of four instrumental teachers in order to identify possible underlying factors informing their perceptions about assessment as a professional practice. This approach acknowledges the difference between the observable events in empirical data, and structural factors that might underlie them (Danermark, Ekström, Jakobsen, & Karlsson, 2002). Rather than a realist approach, which assumes the existence of an external reality independent of any observer, critical realism assumes such realities initially as a hypothesis in order to 'scaffold' the research, preserving a more or less 'agnostic' position throughout. Nonetheless, qualitative research has the ultimate goal of explanation (Miles & Huberman, 1994), setting out to clarify the nature and interrelationship of different factors contributing to a phenomenon (Ritchie et al., 2014). Bearing in mind that actors can give no more than *accounts* of their experiences, their communication of these using the 'intellectual tools' of their situated practices (Säljö, 2005; Vygotsky, Cole, John-Steiner, Scribner, & Souberman, 1978) is the key for interpretation. To contextualise these accounts within a conceptual frame and a systematic and reflexive procedure conduct is therefore essential, and it is this which gives authority to the findings of the qualitative researcher.

In this study, semi-structured dyadic (i.e. with two informants) interview was considered to be an appropriate instrument for an exploration of teachers' perceptions about assessment. While the asymmetric relationship between researcher and informant is often referred to concerning interviews (Kvale, Brinkmann, & Anderssen, 2009), dyadic interviews are claimed to give informants more control over the situation, allowing them to 'co-construct' their version of the research topic, and to stimulate ideas that might have gone unrecognised or forgotten (Morgan, Ataie, Carder, & Hoffman, 2013). Dyadic interviews are valued "for providing a measure of the depth and detail available in individual interviews at the same time that they provide the interaction present in focus groups" (Morgan et al., 2013: 1283). However, the possibility of multi-faceted power dynamics (Hammersley, 2012) between any combination of actors should be

taken into account, e.g. sympathy or antipathy between colleagues, causing individuals to feel inhibited or censor their participation, or trying to meet the researcher's assumed expectations. This form of interview might also be regarded as "a meeting between professionals of different fields" (Bruun, 2015: 139) where different types of knowledge might be accorded higher status, functioning in its turn as a community of practice. In this study, the researcher's professional background as a teacher educator with classical piano as main instrument, but scant experience from upper secondary school, might be seen in various ways, e.g. as credentials for sharing professional understandings, or as peripheral legitimacy (Lave & Wenger, 1991), placing the informants in a position of expertise from which to initiate the researcher.

Two semi-structured dyadic interviews were conducted at upper secondary schools in different geographical regions of Norway. Examples of teacher perceptions were sought, rather than typicality, as well as the possibility of variation at comparable institutions. This was a non-equivalent group design, using stratified purposive sampling (Creswell, 2013) to find schools of similar size offering the elective music programme, and teachers representing several instruments and at least five years' teaching experience at upper secondary level. At each school, teachers were informed of the aims of the study and volunteered to participate. While an advantage of this form of recruitment is the likelihood of finding informants who have particular interest and engagement for the topic of assessment, this must be weighed up against the possibility of imbalances in certain variables. At School A, two female song teachers (A1 and A2) were recruited, with respectively 10 and 19 years' teaching experience. Both of these have a background of studies in classical genres, but teach both classical and popular genres. At School B, recruits were a male teacher of bass guitar (B1) with background in popular genres and five years' experience, and a female teacher of classical piano (B2) with 18 years' experience. The imbalances in gender, instrument and professional experience are acknowledged, yet it is not atypical for the domain studied that e.g. bassists are male, or that singing teachers are a large group, accommodating demand.

Twelve questions were prepared for the interviews, designed to cover key issues for the research questions on the basis of a literature review, and incorporating learning outcomes in LK06 (Appendix). With the purpose of giving more autonomy to the informants, creating an informal tone and reducing the role of researcher, the prepared questions were drawn by informants at random from a box. This method of selection meant that there was no set order of sequence of questions or of which informant might answer first, precluding an organised progression of topics with a

particular thematic chronology. One advantage of this method was giving the informants control of the time used for topics as they accorded them importance, within the time they had available. Nevertheless, the researcher's situated role as interviewer (Kvale, 1996) involved engaging in the conversation with purposes e.g. of clarification, eliciting more information on a topic, or reminding participants of the time remaining. Due to logistical constraints such as timetabling and commuting distances, there were differences in the practical circumstances of the two interviews. For informants from School A, the interview took place at a higher education institution after working hours, lasting 91 minutes. For School B, the interview took place at the school itself between teaching commitments and lasted 51 minutes. Although all twelve questions were drawn at both interviews, answers were longer and there was room for more anecdotal exchanges for informants from School A.

Audio recordings were made of the interviews, giving the possibility of repeated relistenings (Psathas, 1995), and the recordings were transcribed verbatim employing elements of the Jefferson system (2004) in order to document pauses, laughter or other non-verbal elements which might communicate meaning. Acknowledging that the process of listening and transcribing is itself "an act of interpretation and representation" (Bucholtz, 2000: 1463), transcriptions of the interviews were sent back to informants for verification. The subsequent analytical strategy was abductive, involving a zigzag movement of mutually influencing elements of research (Layder, 1998), and coding in Nvivo according to principles of thematic analysis (Attride-Stirling, 2001), using topics from the interview questions as well as open coding. This led to new readings of literature, in particular focusing on the work of Lave and Wenger (1991) and Wenger (1998). Through subsequent readings and revisions of the coded material, three categories were developed in light of Wenger's (1998) concept of participatory identities. These were perceptions of the individuals as (i) teacher-mentor, (ii) instrumentalist, and (iii) participant in a school-based assessment practice. Findings relating to these three aspects are presented and discussed in the following three sections. Numerical codes are employed in order to preserve informants' anonymity. All citations are translated from the Norwegian data to English by the author.

Findings and discussion

Perceptions of assessment as teacher-mentor

All four informants express scepticism to making assessments of music performance, concerning both the fundamental problem of whether trustworthy assessment is possible at all, the difficulties of verbalising implicit knowledge, and the effects of assessment on students. Discomfort in taking the role of assessor is evident, for example, when B2 describes how students sometimes select repertoire which, for various reasons, seems inappropriate, saying:

> it feels as though one is judging somebody's musical taste, in a way, and you shouldn't do that!

A view that formalised learning objectives are inappropriate for performance on main instrument emerges when A1 claims the goals in LK06 are

> first and foremost for the theoretical subjects. Things you can measure on paper.

She perceives the goals as broad, pointing out that a student with "a very narrow talent" cannot achieve a top grade despite outstanding achievement within his specialism. This kind of assessment problem might be felt particularly keenly by instrumental teachers, in view of possible strong mentor relationships to their students (Ellefsen, 2014). It might therefore be equally problematic to make an assessment when a student who struggles with schoolwork generally, nevertheless manages one narrow aspect of music performance quite well. All four informants emphasise that assessment, and in particular grading, can have negative effects on students. It is suggested that the grading system may actually limit a student's progression, particularly for weaker students, causing fear and a narrow focus on specific learning outcomes, as in A2's remark:

> As soon as they've got their first grade, it's as if a hierarchy is established; it's hurtful.

A1 and A2 describe how they try to 'reduce the damage' of normative aspects of assessment by telling students that a test reflects neither a student's all-round competence,

nor the quality of single aspects of performance. As such, they clarify the boundaries of the assessment, but nevertheless also call into question its validity.

These statements are interpreted as demonstrating values of teachers' responsibility to students, not only to be fair, but to nurture students' growth and motivation, giving support to existing findings of tensions in the dual role of teacher-assessor (e.g. Yung, 2001). The informants' descriptions of the negative effects of summative assessment, aligning with established research (Harlen & Deakin Crick, 2002), and their criticism of curricular goals, combine in a picture of the teacher-mentor who wishes to protect the student from detrimental forces, and to encourage student autonomy—even to the extent of ratifying choices considered inappropriate. The task of assessment, in this perspective, is fraught with considerations connected to students' development as autonomous individuals unfettered by curricular constraints.

Perceptions of assessment as instrumentalist

Statements which can be related to types of knowledge, skills and traditions of particular instruments give a different picture of how the informants perceive assessment, and in particular criteria and standards. The notion of what characterises a 'top mark performance' for the two singers, A1 and A2, compared with that of the bass guitarist, B1, has less to do with keeping a rhythmic nerve than with communication and personal expression, but this is not to say that timing and technical issues are not essential, for all three instruments. A fuller description of expertise is given by B2, a pianist:

> You have a superabundance of energy to communicate music well, everything is in place, it is like—that moment when you feel the student is just at one with the material, has made it her own, and can feel it's easy to play, fun to play, goes in for it with her whole self.

Issues of technique are here subsumed by a holistic criterion that to a high degree reflects the praxial music discourse (Elliott, 1995), and its recognition is dependent on the assessor.

Another factor mentioned for demonstrating competence is breadth of repertoire. The singers claim this is important, and there might be several reasons for this. These teachers have a classical background, yet teach across genres. If students have a preference for popular genres, their teachers might nevertheless recommend classical repertoire because of a conviction that working with this will provide the best

grounding for good voice production. Indeed, A1 claims that there are some universally accepted criteria for the craft of singing, governed by current conceptions of healthy technique, which carry more importance than curriculum requirements:

> I don't think singing as a subject will change just because the curriculum does, because there are some criteria for learning to sing, that just are accepted as the craft itself, about what is currently seen as healthy and good.

Nonetheless, the singers also warn that an assessor who lacks broader genre competence might emphasise features of technical mastery in classical genres, missing important aspects like «feeling and timing and all that», which the informants say they have gained through exposure to rhythmic and popular genres. The bassist, too, recommends his students to present a certain breadth of repertoire, but his rationale is different:

> It's to do with what life is really like out there. It's unlikely you'll be playing in a band that's 'gonna make it', that will never play in any other genre. So I stress this, especially for the bassists—don't play four songs in the same genre, right?

Reasoning that students have small chances of a future in a successful band and keeping to one genre, his requirement to present a varied repertoire reflects a discourse that considers future musical enterprise beyond the school context. However, the possibility that the presentation of a varied repertoire may not fulfil intended goals is raised by the pianist, who suggests it is more important to demonstrate a broad 'repertoire of expression' than mastery of several styles. This comment might be a reflection of tensions between views of classical and popular genres (Danielsen, 2006), concerning the possibilities for dynamic and timbral nuance and variation of mood within even a single classical piece. With this in mind, demands for both breadth of repertoire and a certain length of programme might be experienced as different for classical and popular genres.

While *breadth* of repertoire is given importance, conformity to norms for performance practice for *particular* repertoire is discussed for all three instruments. For the pianist, there are limits for how much freedom in tempo and phrasing can be tolerated in music by a composer like Mozart. The singers, on the other hand, while claiming that an unconventional performance might elicit the reaction, "Hallo, you can't sing Mozart like that", concede that there is room to consider the student's own intentions. But it is

not only certain classical genres that are stringently judged according to performance traditions. The bassist, representing popular genres, claims there is a high risk factor involved if students make changes to repertoire that has acquired an 'iconic' position, and that innovations will be tolerated only if the performance is exceptionally good. These views about conformity are connected to assessors' familiarity with repertoire or genre, and seem to be more stringent for the bassist and pianist, than for the singers. In cases where students perform their own material, the singers claim that students' compositions are often "well within [their] own comfort zone", but that a convincing performance is likely to be rated highly. In contrast, the bassist refers to a culture among pop and rock students where there is prestige in composing detailed and challenging music. Students' compositions might thus function as two types of 'shop window' for exhibiting their skills: one where there is comfortable mastery of the material, and the other where virtuosity is stretched to its outer limits.

For assessment, these examples suggest that a complex web of different factors can be claimed to be specific to particular instruments and repertoire, and that instrumental teachers might 'point' assessment in quite different directions. An emphasis on universal criteria of musicking and personal expression over technique and voice quality, as suggested by the singing teachers, stands in contrast to advice given to a bass student to avoid taking risks regarding standard repertoire. The former can be said to reward student autonomy, while the latter focuses not only on performance preserving the 'work' itself, but on a form of criteria compliance (Torrance, 2007), in which future real-life opportunities for performance are envisaged. Individualised teaching contexts throw into relief the professional identity of the teacher as musician (Angelo, 2014; Nerland, 2003), since it is the very specificity of knowledge and skills of the instrument in question that legitimises this form of teaching. In this way, various discourses of music performance, such as the autonomy of certain repertoire or genres, or appropriate vocal technique, might function as mechanisms of gatekeeping as teachers draw up the boundaries for their subject domain, and the informants in this study suggest that these might count more than various learning objectives of the National Curriculum. That these tensions have a bearing on power relations within the duality of social and individual when the qualities of performances are discussed, is likely. Wenger (1998) notes that "different forms of power in a society interact, sometimes reinforcing each other and sometimes creating spaces of resistance" (p. 284). In this way, the question of what types of knowledge teachers emphasise for a particular instrument or repertoire can have great significance for the shaping of assessment in the localised context.

Perceptions as a participant in a school-based assessment practice

Despite the scepticism expressed by the informants to assessment, and to grading in particular, all four seem to perceive themselves as competent assessors with the ability to make independent qualitative judgements of students' performances. They describe how expertise is accumulated over time, enabling them to make decisions in tests or examination situations more rapidly and holistically, and no longer taking personal responsibility for students' results. For example, A1 comments:

> I think I've become less rigid, and it's not because I'm not the same person – but now it's not important for me to have the best student [...] So, in the beginning you almost feel it's you getting the grade.

The importance of the community of practice (Wenger, 1998), both for assuring reliability and for providing experiential knowledge equipping teachers better to assess both within and outside their main instrument domain, is emphasised. A2 describes a learning trajectory through 19 years of teaching and assessing, including various in-service courses, which has involved a shift in focus:

> In the beginning I was much more concerned with [firstly] good technique, and secondly voice material, one's point of departure. But then the performing and personal expression has come in much more, for my part, since we began doing assessments in the team.

Assessing term tests by moderation is practiced at both schools. At School A, it is usual for all fulltime instrumental staff, and the relevant part-time staff, to be present at each performance and assessment discussion. At School B, the teachers assess in teams of 3–5, based on instrument group. At both schools, lists of local criteria for the various instrument groups have been drawn up, but are experienced to have limited value. The two informants at School A value the large team practice they are part of, depicting the development of a holistic assessment practice where, as mentioned, the universal criteria of "feeling and timing" are emphasised ('feeling' here interpreted as appropriate stylistic expression). They present it as a functional and effective vehicle for trustworthy assessment, and as a learning forum where a shared repertoire of implicit knowledge is developed. At the same time, the informants say that when novices or outsiders are introduced into the team there is a healthy incentive to 'unpack' the discussion and articulate meanings more fully. Although admitting that there is a possibility of teachers being swayed against their better judgement,

it seems that the advantages of social moderation outweigh this. When asked by the researcher what they would change, they call for better preparation of undergraduates for assessment work, as well as recommendations that novice teachers should not participate fully in assessment, but first observe and learn.

At School B, the advantages of social moderation are nuanced slightly differently, with more emphasis on aspects of control: here, the informants stress how professional bonds between fulltime and part-time staff are strengthened, and how development of individual "regimes" can be prevented. The use of lists of instrument-relevant criteria is seen as pertinent, both as an assurance of covering prescribed learning objectives, and to avoid basing assessment on "feelings" (here interpreted as 'emotions'), yet teachers' inherent assessment skills have 'the last say'. The informants' criticism of social moderation is that agreement might be reached on very general terms, leaving out important detail or aspects of performance, and in turn leading to inadequate feedback to students. It is acknowledged that teachers' prestige may play a role in assessment discussion, and the call for more thorough guidelines seems to reflect an awareness of challenges to assessment. At both schools it is claimed that there is seldom dissonance between internal and external assessors at the public exam, but examples of exceptions are given in contexts of one examiner and one teacher assessing student performance. While B2 (School B) notes that teachers can experience a battle of prestige over results, A1 (School A) reports having felt the need for support from likeminded members of her team in discussion with the examiner.

Thus, participatory identity connected to the school-based assessment practice, as depicted by the four informants, is nuanced. While the possibility that individual teachers' judgements might be influenced by the group is acknowledged, the informants claim to have gained increased self-efficacy and independence through membership. Bearing in mind that the two informants from School A are both female singers, and probably work in close cooperation, this might have contributed to a certain picture of a harmonious collegium easily finding consensus in assessment. Nevertheless, the basic message from both interviews about social moderation is that this is an advantageous system for assessment, not least for its function of building teachers' assessment competence in the school. For all the informants, gaining experience as assessors seems to have involved a movement towards more holistic assessment where implicit knowledge is foregrounded.

General discussion

The three projected types of participatory identity identified in the analysis were, firstly, the teacher-mentor demonstrating responsibility to individual students, emphasising student autonomy and resisting formalised plans. The second type was the instrumentalist exhibiting loyalties to received instrumental discourses and career paths. The third was the participant in a local assessment practice, emphasising the community of practice as a source of knowledge and a path to consensus. In the statements analysed, it seems the 'goalposts' for assessment might be in traction between these potentially conflicting values. As suggested by Rui (2010), although teachers might come to the same conclusion with regard to grading, assessment might contain quite different rationales, for example using reactive, hegemonic patterns to preserve professional instrumental ideals, or diluted, 'trivialised' discourses (Zandén, 2010a) which stop with the praxial ideal (Allsup & Westerlund, 2012; Elliott, 1995). At the very least, there is potential for localised interpretations of the mandate (Wenger, 1998). While one purpose of the moderated assessment is to provide feedback (Adie et al., 2012; Harlen, 2007) for the individual student and his teacher, if moderation is geared towards effectivity as well as emphasising holistic assessment, the feedback generated to take back to the student is likely to be sparse. Equally, where the application of instrumental and genre-specific criteria are emphasised, certain conforming principles might lead to constraints for students, and a discouragement to go outside the frame of accepted performance practice. Ultimately, if the experience of success in the form of finding consensus reinforces teachers' selection of strategies and perspectives in assessment tasks e.g. which elements of a situation to treat as important or which to ignore, assessment is in danger of becoming strongly biased. Looking at the findings as a whole, the space located between an understanding that reliability depends on attunement within the assessment group (Sadler, 2015), and the concession that teachers experience various tensions concerning their participatory identities, is a black box demanding the exegesis of authentic assessment situations.

Conclusion

This study sheds light on how four teachers from two Norwegian upper secondary elective music programmes perceive tasks of assessing performance on main instrument. In light of Wenger's (1998) concept of participatory identities, three types of participatory identity were used to frame interview data: the teacher-mentor with

responsibility to individual students; the instrumentalist with loyalties to received traditions and projected career paths; and the participant in a school-based assessment practice. Bearing in mind that it may only take a few interviews to demonstrate that a phenomenon is more complex or varied than previously thought (Becker in Baker & Edwards, 2012), these data are valuable in giving examples of issues that exist in the current Norwegian education system. Findings from this analysis are instances of challenges for assessment of music performance in the context of Norwegian upper secondary school, where identities of participation are informed by the conflicting value systems of responsibility to the student, loyalties to personally held conceptions of musical quality, and accountability to the professional mandate. While expressing scepticism to assessment because of the difficulties of attempting to make a fair judgement, and the detrimental effects assessment can have on students, the informants indicate that they have confidence in their assessment practices on the basis of their professional knowledge and participation in social moderation. They value school-based moderation for its functions of providing quality assurance and development of assessment expertise.

These inferences carry implications for further study, where investigation should focus not only on what teachers say and intend, but also what they actually do (Pajares, 1992), in authentic assessment contexts. In a field where students need dependable assessment and feedback to help their further development, but where parameters for artistic performance are in constant change, more knowledge is needed on how instrumental teachers meet the challenges of assessment.

References

Adie, L., Klenowski, V. & Wyatt-Smith, C. (2012). Towards an understanding of teacher judgement in the context of social moderation. *Educational Review, 64*(2), 223–240. doi:10.1080/00131911.2011.598919

Allsup, R. E. & Westerlund, H. (2012). Methods and Situational Ethics in Music Education. *Action, Criticism, and Theory for Music Education, 11*(1), 124–148.

Amabile, T. M. & Hogan, R. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology, 45*(2), 357–376. doi:10.1037/0022-3514.45.2.357

Angelo, E. (2012). *Profesjonsforståelser i instrumentalpedagogiske praksiser.* (PhD), Norges teknisk-naturvitenskapelige universitet, Trondheim.

Angelo, E. (2014). *Profesjonsforståelser og kunstpedagogikk : praksiser i musikkfeltet*. Bergen: Fagbokforl.

Asp, K. (2015). *Mellan klassrum och scen – en studie av ensembleundervisning på gymnasieskolans estetiska program.* (PhD), Lund University, Malmö.

Attride-Stirling, J. (2001). Thematic networks: an analytic tool for qualitative research. *Qualitative Research, 1*(3), 385–405. doi:10.1177/146879410100100307

Auslander, P. (2006). Musical Personae. *TDR (1988-), 50*(1), 100–119. doi:10.2307/4492661

Baker, S. E. & Edwards, R. (2012). How many interviews is enough? Methods review paper. Retrieved from http://eprints.ncrm.ac.uk/2273/

Barratt, E. & Moore, H. (2005). Researching group assessment: jazz in the conservatoire. *British Journal of Music Education, 22*(3), 299–314.

Bergee, M. J. (1995). Primary and Higher-Order Factors in a Scale Assessing Concert Band Performance. *Bulletin of the Council for Research in Music Education*(126), 1–14.

Bergee, M. J. (2003). Faculty Interjudge Reliability of Music Performance Evaluation. *Journal of Research in Music Education, 51*(2), 137–151.

Bergee, M. J. (2007). Performer, Rater, Occasion, and Sequence as Sources of Variability in Music Performance Assessment. *Journal of Research in Music Education, 55*(4), 344–358. doi:10.1177/0022429408317515

Bernstein, B. (1971). On the classification and framing of educational knowledge. In M. Young (Ed.), *Knowledge and control*. London: Collier-Macmillan.

Bhaskar, R. (2008). *A realist theory of science*. London: Routledge.

Black, P. & Wiliam, D. (1998). Inside the Black Box: Raising Standards Through Classroom Assessment. *Phi Delta Kappan, 92 (1)*(September 2010 92 (1)), 81–90. Retrieved from https://www.rdc.udel.edu/wp-content/uploads/2015/04/InsideBlackBox.pdf

Blom, D. & Encarnacao, J. (2012). Student-chosen criteria for peer assessment of tertiary rock groups in rehearsal and performance: What's important? *British Journal of Music Education, 29*(1), 25–43. doi:10.1017/S0265051711000362

Boud, D. (2000). Sustainable Assessment: Rethinking Assessment for the Learning Society. *Studies in Continuing Education, 22*(2), 151–167.

Brookhart, S. M. (2013). The use of teacher judgement for summative assessment in the USA. *Assessment in Education: Principles, Policy & Practice, 20*(1), 69–90. doi:10.1080/0969594X.2012.703170

Bruun, H. (2015). The Qualitative Interview in Media Production Studies. In C. Chris Paterson, D. Lee, A. Saha, & A. Zoellner (Eds.), *Advancing Media Production Research: Shifting Sites, Methods, and Politics*. Basingstoke: Palgrave Macmillan.

Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics, 32*(10), 1439–1465. doi:10.1016/S0378-2166(99)00094-6

Ciorba, C. R. & Smith, N. Y. (2009). Measurement of Instrumental and Vocal Undergraduate Performance Juries Using a Multidimensional Assessment Rubric. *Journal of Research in Music Education, 57*(1), 5–15. doi:10.1177/0022429409333405

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. doi:10.1177/1321103x010160010301

Colwell, R. (2003). The Status of Arts Assessment: Examples from Music. *Arts Education Policy Review, 105*(2), 11–18. doi:10.1080/10632910309603457

Constantino, T. & Bresler, L. (2010). Assessment in Schools – Creative Subjects. In P. Peterson, E. Baker, & B. McGaw (Eds.), *The international encyclopedia of education* (3rd ed., pp. 262–267). Amsterdam: Academic Press.

Creswell, J. W. (2013). *Qualitative inquiry & research design : choosing among five approaches* (3rd ed.). Los Angeles: Sage.

Danermark, B., Ekström, M., Jakobsen, L. & Karlsson, J. C. (2002). *Explaining society: critical realism in the social sciences*. Abingdon: Routledge.

Danielsen, A. (2006). Aesthetic Value, Cultural Significance, and Canon Formation in Popular Music. *Studia Musicologica Norvegica, 32*, 55–72.

Danielsen, A. (2016). Nyskapende, sterkt og kompetent! Kvalitetsforståelser i musikkfeltet i lys av tre populærmusikalske sjangre. In K. O. Eliassen & Ø. Prytz (Eds.), *Kvalitetsforståelser : kvalitetsbegrepet i samtidens kunst og kultur*. Oslo: Norsk kulturråd.

Davidson, J. W. (2001). The Role of the Body in the Production and Perception of Solo Vocal Performance: A Case Study of Annie Lennox. *Musicae Scientiae, 5*(2), 235–256. doi:10.1177/102986490100500206

Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. *Psychology of Music, 40*(5), 595–633. doi:10.1177/0305735612449896

Davidson, J. W. (2014). Introducing the Issue of Performativity in Music. *Musicology Australia, 36*(2), 179–188. doi:10.1080/08145857.2014.958269

Davidson, J. W. & Da Costa Coimbra, D. (2001). Investigating Performance Evaluation by Assessors of Singers in a Music College Setting. *Musicae Scientiae, 5*(1), 33–53. doi:10.1177/102986490100500103

Davidson, J. W. & Edgar, R. (2003). Gender and Race Bias in the Judgement of Western Art Music Peiformance. *Music Education Research, 5*(2), 169–181. doi:10.1080/1461380032000085540

Dyndahl, P. & Nielsen, S. G. (2014). Shifting authenticities in Scandinavian music education. *Music Education Research, 16*(1), 105–118. doi:10.1080/14613808.2013.847075

Eisner, E. (2003). Educational Connoisseurship and Educational Criticism: An Arts-Based Approach to Educational Evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation* (Vol. 9, pp. 153–166). Dortrecht: Kluwer Academic Publishers.

Ellefsen, L. W. (2014). *Negotiating musicianship. The constitution of student subjectivities in and through discursive practices of musicianship in "Musikklinja".* (PhD), Norges musikkhøgskole.

Elliott, C. A. (1996). Race and Gender as Factors in Judgments of Musical Performance. *Bulletin of the Council for Research in Music Education* (127), 50–56.

Elliott, D. J. (1995). *Music Matters: A Philosophy of Music Education*. New York: Oxford University Press.

Feld, S. (1984). Communication, Music, and Speech about Music. *Yearbook for Traditional Music, 16*, 1–18. doi:10.2307/768199

Finney, J. (2013). *Music Education in England, 1950–2010: The Child-Centred Progressive Tradition* Farnham: Ashgate.

Fiske, H. E. (1975). Judge-group differences in the rating of secondary school trumpet performances. *Journal of Research in Music Education, 23*, 186–196.

Fiske, H. E. (1977). Relationship of Selected Factors in Trumpet Performance Adjudication Reliability. *Journal of Research in Music Education, 25*(4), 256–263.

Fives, H. & Buehl, M. M. (2012). *Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us?* : American Psychological Association.

Flick, U. (2009). *An introduction to qualitative research* (4th ed. ed.). Los Angeles, Calif: Sage.

Gates, H. L. (1988). *The signifying monkey : a theory of Afro-American literary criticism*. New York: Oxford University Press.

Georgii-Hemming, E. (2005). *Berättelsen under deras fötter. Fem musiklärares livshistorier.* (PhD), Örebro Universitet, Örebro.

Gipps, C. V. (2012). *Beyond Testing: Towards a theory of educational assessment.* Abingdon, Oxford: Routledge.

Green, L. (1988). *Music on deaf ears : Musical meaning, ideology, education*. Manchester: Manchester University Press.

Green, L. (2005). Musical Meaning and Social Reproduction: A Case for Retrieving Autonomy. *Educational Philosophy and Theory, 37*(1), 77–92. doi:10.1111/j.1469-5812.2005.00099.x

Green, L. (2008). *Music, Informal Learning and the School: A New Classroom Pedagogy*. Aldershot: Ashgate.

Green, L. (2014). Music as a media art : Evaluation and assessment in the contemporary classroom In L. Green (Ed.), *Music Education as Critical Theory and Practice : Selected Essays* (pp. 165–182). Farnham: Ashgate.

Griffiths, N. K. (2010). 'Posh music should equal posh dress': an investigation into the concert dress and physical appearance of female soloists. *Psychology of Music, 38*(2), 159–177. doi:10.1177/0305735608100372

Gynnild, V. (2010). Kriterier og skjønn i evaluering : en kasusstudie i utøvende musikkutdanning (Vol. 4/2010): Nasjonalt organ for kvalitet i utdanningen.

Gynnild, V. (2015). Assessing vocal performances using analytical assessment: a case study. *Music Education Research*, 1–15. doi:10.1080/14613808.2015.1025732

Hammersley, M. (2012). *What is Qualitative Research?* London: Bloomsbury.

Harlen, W. (2007). Criteria for Evaluating Systems for Student Assessment. *Studies in Educational Evaluation, 33*(1), 15–28. doi:10.1016/j.stueduc.2007.01.003

Harlen, W. & Deakin Crick, R. (2002). *A systematic review of the impact of summative assessment and tests on students' motivation for learning*. Retrieved from London: http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=110

Hewitt, M. P. (2007). Influence of Primary Performance Instrument and Education Level on Music Performance Evaluation. *Journal of Research in Music Education, 55*(1), 18–30. doi:10.1177/002242940705500103

Hodgson, J., Rønning, W., Skogvold, A. S. & Tomlinson, P. (2010). *Vurdering under Kunnskapsløftet: læreres begrepsforståelse og deres rapporterte og faktiske vurderingspraksis* NF-rapport (Nordlandsforskning : online), Vol. 17/2010.

Howard, S. A. (2012). The Effect of Selected Nonmusical Factors on Adjudicators' Ratings of High School Solo Vocal Performances. *Journal of Research in Music Education, 60*(2), 166–185. doi:10.1177/0022429412444610

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the first generation* (pp. 13–31). Amsterdam: John Benjamin.

Johansen, G. (2003). *Musikkfag, lærer og læreplan. En intervjuundersøkelse av læreres fagoppfatning i musikk og en ny læreplans påvirkning på denne.* (PhD), Norges musikkhøgskole, Oslo. (2003:3)

Juchniewicz, J. (2008). The influence of physical movement on the perception of musical performance. *Psychology of Music, 36*(4), 417–427. doi:10.1177/0305735607086046

Kinney, D. W. (2009). Internal Consistency of Performance Evaluations as a Function of Music Expertise and Excerpt Familiarity. *Journal of Research in Music Education, 56*(4), 322–337. doi:10.1177/0022429408328934

Klafki, W. (2001). Kategorial dannelse : bidrag til en dannelsesteoretisk fortolkning av moderne didaktikk. Oslo: Gyldendal akademisk, 2001.

Kunnskapsdepartementet. (2006). *Kunnskapsløftet (LK06) [the National Curriculum for Knowledge Promotion 2006]*. Oslo.

Kvale, S. (1996). *Interviews : an introduction to qualitative research interviewing*. Thousand Oaks, Calif: Sage.

Kvale, S. & Brinkmann, S. (2009). *Det kvalitative forskningsintervju*. Oslo: Gyldendal akademisk.

Lave, J. & Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. Cambridge: Cambridge University Press.

Layder, D. (1998). *Sociological Practice: Linking Theory and Social Research*. London: Sage.

Lehmann, M. & Kopiez, R. (2013). The influence of on-stage behavior on the subjective evaluation of rock guitar performances. *Musicae Scientiae, 17*(4), 472–494. doi:10.1177/1029864913493922

Leman, M. & Godøy, R. I. (2010). *Musical gestures : sound, movement, and meaning*. New York: Routledge.

Liao, M.-Y. & Davidson, J. W. (2015). Enhancing gesture quality in young singers. *Music Education Research*, 1–12. doi:10.1080/14613808.2015.1060953

Lysne, A. (2004). *Kampen om skolen: 1970–2000. Bind 2 – Karakterer og kompetanse* (Vol. B. 2). [Haslum]: AVA forl.

Maugars, C. (2006). Attitudes of music teachers towards final examinations in the French music conservatoires. *International Journal of Music Education, 24*(1), 43–55. doi:10.1177/0255761406063106

McPherson, G. E. & Schubert, E. (2004). Measuring Performance Enhancement in Music. In A. Williamon (Ed.), *Musical Excellence – Strategies and Techniques to Enhance Performance* (pp. 61–82). Oxford: Oxford University Press.

Middleton, R. (2000). Musical Belongings : Western Music and Its Low-Other In G. Born & D. Hesmondhalgh (Eds.), *Western Music and Its Others : Difference, Representation, and Appropriation in Music* (pp. 59–85). Berkeley, CA: University of California Press

Miles, M. B. & Huberman, A. M. (1994). *Qualitative data analysis: an expanded sourcebook*. Thousand Oaks, Calif.: Sage.

Moore, A. (2002). Authenticity as authentication. *Popular Music, 21*(02), 209–223. doi:10.1017/s0261143002002131

Morgan, D. L., Ataie, J., Carder, P. & Hoffman, K. (2013). Introducing Dyadic Interviews as a Method for Collecting Qualitative Data. *Qualitative Health Research, 23*(9), 1276–1284. doi:10.1177/1049732313501889

Nerland, M. (2003). *Instrumentalundervisning som kulturell praksis* (Vol. no. 24). Oslo: Unipub forl.

Pajares, M. F. (1992). Teachers' Beliefs and Educational Research: Cleaning up a Messy Construct. *Review of Educational Research, 62*(3), 307–332.

Persson, R. (1994). Concert musicians as teachers: On good intentions falling short. *European Journal for High Ability, 5*, 79–91.

Platz, F. & Kopiez, R. (2013). When the first impression counts: Music performers, audience and the evaluation of stage entrance behaviour. *Musicae Scientiae, 17*(2), 167–197. doi:10.1177/1029864913486369

Polanyi, M. (2009). *The tacit dimension*. Chicago: University of Chicago Press.

Prince, V. & Hallam, S. (1996). *Cognitive, technical and expressive music skills: pupils' and teachers' perceptions of the relative emphasis given to each in instrumental music lessons.* Paper presented at the BPS Education Section Conference, Reading, UK.

Prøitz, T. S. & Borgen, J. S. (2010). *Rettferdig standpunktvurdering – det (u)muliges kunst?: læreres setting av standpunktkarakter i fem fag i grunnopplæringen*. Oslo: NIFU STEP.

Psathas, G. (1995). *Conversation analysis: the study of talk-in-interaction.* Thousand Oaks, Calif.: Sage.

Reinertsen, A. B. (2014). Outdoor Becoming Thinking With Data Dewey Derrida Deleuze. *Qualitative Inquiry, 20*(8), 1022–1032. doi:10.1177/1077800413489529

Ritchie, J., Lewis, J., Nicholls, C. M. & Ormston, R. (2014). *Qualitative research practice : a guide for social science students and researchers*. London: Sag.

Rui, E. (2010). *Vurdering i hovedinstrument – Hvordan begrunner lærere karakterer de setter for hovedinstrumentfremføringer?* (Master), Norwegian Academy of Music, Oslo.

Russell, B. E. (2015). An empirical study of a solo performance assessment model. *International Journal of Music Education, 33*(3), 359–371. doi:10.1177/0255761415581282

Ryan, C. & Costa-Giomi, E. (2004). Attractiveness Bias in the Evaluation of Young Pianists' Performances. *Journal of Research in Music Education, 52*(2), 141–154. doi:10.2307/3345436

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144.

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education, 34*(2), 159–179. doi:10.1080/02602930801956059

Sadler, D. R. (2015). Backwards assessment explanations: implications for teaching and assessment practice In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in Music Education: from Policy to Practice* (pp. 9–19). Cham: Springer.

Saunders, T. C. & Holahan, J. M. (1997). Criteria-Specific Rating Scales in the Evaluation of High School Instrumental Performance. *Journal of Research in Music Education, 45*(2), 259–272.

Schei, T. B. (2007). *Vokal identitet.* (PhD), Griegakademiet, Institutt for musikk, Universitetet i Bergen, Bergen.

Skott, J. (2015). The promises, problems, and prospects of research on teachers' beliefs. In H. Fives & M. G. Gill (Eds.), *International Handbook of Research on Teachers' Beliefs* (pp. 13–30). New York: Routledge.

Small, C. (1998). *Musicking: the meanings of performing and listening*. Hanover, N.H.: University Press of New England.

Smith, D. T. (2009). Development and Validation of a Rating Scale for Wind Jazz Improvisation Performance. *Journal of Research in Music Education, 57*(3), 217–235. doi:10.1177/0022429409343549

Stanley, M., Brooker, R. & Gilbert, R. (2002). Examiner Perceptions of Using Criteria in Music Performance Assessment. *Research Studies In Music Education, 18*(1), 46–56. doi:10.1177/1321103x020180010601

Stobart, G. (2008). *Testing times: the uses and abuses of assessment*. London: Routledge.

Swanwick, K. (1998). The Perils and Possibilities of Assessment. *Research Studies In Music Education, 10*(1), 1–11.

Säljö, R. (2005). Læring, kunnskap og sosiokulturell utvikling: Mennesket og dets redskaper. In I. Bråten (Ed.), *Læring i sosialt, kognitivt og sosialtkognitivt perspektiv*, pp. 31–57. Oslo: Cappelen Akademiske Forlag.

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. doi:10.1080/09695940701591867

Utdanningsdirektoratet. (2011). *Programme for Music, Dance and Drama – Music Studies. Curriculum for instrument, choir and ensemble*. Retrieved from https://webcache.googleusercontent.com/search?q=cache:NgPU9_9HGToJ:https://www.udir.no/Upload/larerplaner/Fastsatte_lareplaner_for_Kunnskapsloeftet/english/Music_dance_drama/5/MDD_common_and_optional_programme_subjects.rtf+&cd=3&hl=no&ct=clnk&gl=no.

Utdanningsdirektoratet. (2014). *Veiledning i lokalt arbeid med læreplaner [Guidelines for work with local curricula]*. Retrieved from http://www.udir.no/Lareplaner/Veiledninger-til-lareplaner/Veiledning-i-lokalt-arbeid-med-lareplaner/2-Lareplanverket-for-Kunnskapsloftet-LK06-og-LK06S/Lokalt-arbeid-med-lareplaner-etter-LK06/.

Vinge, J. (2014). *Vurdering i musikkfag: en deskriptiv, analytisk studie av musikklæreres vurderingspraksis i ungdomsskolen* (Vol. 2014:1). Oslo: Norges musikkhøgskole.

Vygotsky, L. S., Cole, M., John-Steiner, V., Scribner, S. & Souberman, E. (1978). *Mind in society : the development of higher psychological processes*. Cambridge, Mass: Harvard University Press.

Wapnick, J., Darrow, A. A., Kovacs, J. & Dalrymple, L. (1997). Effects of Physical Attractiveness on Evaluation of Vocal Performance. *Journal of Research in Music Education, 45*(3), 470–479.

Weisethaunet, H. & Lindberg, U. (2010). Authenticity Revisited: The Rock Critic and the Changing Real. *Popular Music & Society, 33*(4), 465–485. doi:10.1080/03007761003694225

Wenger, E. (1998). *Communities of practice: learning, meaning, and identity*. Cambridge: Cambridge University Press.

Wesolowski, B. C. (2015). A facet-factorial approach towards the development and validation of a jazz rhythm section performance rating scale. *International Journal of Music Education*. doi:10.1177/0255761415590524

Young, V., Burwell, K. & Pickup, D. (2003). Areas of Study and Teaching Strategies Instrumental Teaching: a case study research project. *Music Education Research, 5*(2), 139–155. doi:10.1080/1461380032000085522

Yung, B. H. W. (2001). Examiner, Policeman or Students' Companion: teachers' perceptions of their role in an assessment reform. *Educational Review, 53*(3), 251.

Zandén, O. (2010a). Professionalise or perish: a case for heightened educational quality through collegiate cooperation. *Nordisk musikkpedagogisk forskning: Årbok, 12*, 117–134.
Zandén, O. (2010b). *Samtal om samspel: Kvalitetsuppfattningar i musiklärares dialoger om ensemblespel på gymnasiet.* (PhD), Gothenburg.

Senior Lecturer
Elizabeth Oltedal
Volda University College
P.O. Box 500
6101 Volda
Norway
oltedale@hivolda.no