Low stakes and high temperatures

A study of music performance assessment in teacher panels in terms of micropolitics

Elizabeth Oltedal, Volda University College/University of Bergen

ABSTRACT

For instrumental teachers, who often have differing views and intentions concerning music education and who teach students individually, assessment by social moderation presents particular challenges. Not only is the assessment meeting an arena in which multiple aspects of teachers' professional identities intersect, including a threefold role of teacher, assessor and colleague, but, like all social interactions, this is a context in which individuals present themselves to each other, making room for various kinds of vulnerability and power. This study draws on theory from micropolitics and ethnomethodology to explore the discussions of instrumental teachers in a Norwegian upper secondary school as they assess student performances on their principal instrument. Results suggest that the teachers were enabled and constrained by a range of different interests with regard to making assessments in grading meetings. They used available situated and situational resources to establish epistemic authority, for example privileged knowledge of student, domain expertise, and projection of the spontaneous, "pure" judgment as unassailable. Assessment discussion is revealed as a layered social interaction in which teachers use available resources strategically to achieve their aims, with implications for student outcomes.

INTRODUCTION

For instrumental teachers, who often have differing views and intentions concerning music education and who teach students individually, assessment by social moderation presents particular challenges. Not only is the assessment meeting an arena in which multiple aspects of teachers' professional identities intersect, including a threefold role of teacher, assessor and colleague, but, like all social interactions, this is a context in which individuals present themselves to each other, making room for various kinds of vulnerability and power. Any attempt to separate the activity of assessment from the relational interaction between participants would therefore be to neglect important dimensions of the situation, yet this has not been a focus in the assessment literature. Drawing on theory of micropolitics (Ball, 1987; Kelchtermans & Ballet, 2002) and Goffman's (1971, 1972b) concepts of face-work, I attempt to provide insights into these issues through an exploration of the discussions of instrumental teachers in a Norwegian upper secondary school as they graded individual student performances on principal instrument.

Research in the field

Recent research in music performance education highlights the ways in which the views and practices of instrumental teachers are informed by different professional and institutional discourses (Angelo, 2012; Gaunt, 2011; Nerland, 2007). Teachers' aims for students range from the development of flexible skills for all-round musicianship, to induction into the specialist skills and dispositions of professional musicians (Angelo, 2012; Ellefsen, 2014), and aspects such as gendered discourses have been problematized (Green, 1997; Rowe, 2008; Zhukov, 2008). The research also reveals student-teacher relationships as close and complex (Ellefsen, 2014; Gaunt, 2011; Schei, 2007).

That assessing music performance is problematic is reflected in the many studies addressing aspects of the validity and reliability of assessment systems, for instance on the merits of using pre-ordained criteria and atomistic or holistic assessment (e.g. Mills, 2005; Stanley, Brooker, & Gilbert, 2002); the construction and usage of criteria and rating scales (e.g. Abeles, 1973; Wesolowski et al., 2017; Wrigley & Emmerson, 2013); and multifarious influences on judgment (e.g. Elliott, 1996; Platz & Kopiez, 2013; Wapnick, Ryan, Lacaille, & Darrow, 2004). Links have been made between student outcomes and teachers' prestige

(Maugars, 2006), and between severity of assessment and the student-teacher relation[1] (Davidson & Da Costa Coimbra, 2001).

Research into music teachers' perceptions about assessing performance reveals ambivalence. While there are concerns about validity, equity and consequences for students' self-confidence and motivation, there are also perceptions that the system of social moderation provides enhanced validity and reliability and enables the development of teachers' assessment literacy (Oltedal, 2018; Vinge, 2014). Studies of authentic contexts using moderation have found that teachers have difficulty in verbalizing and exchanging conceptions of quality, with some criteria remaining implicit (Davidson & Da Costa Coimbra, 2001; Gynnild, 2015; Stanley et al., 2002). While the large body of research in this field concerns Western art music, Zandén (2010) suggested that this problem is even larger for performance in popular genres, linking this hypothetically to resistance to framing popular music in the same conceptual terms as art music, and the notion of opinions about musical quality being a private matter, involving taste.

These studies reveal music performance assessment as a complex and multifaceted field, and social moderation as a context with the potential for both enhancement of teachers' assessment practices and the crisscrossing of various different interests, having implications for student outcomes and collegial relations. Given the evidence in assessment literature about the conflicting interests that can imbue assessment contexts (e.g. Black, Harrison, Hodgen, Marshall, & Serret, 2010; Orr, 2007; Yorke, 2011), there is surprisingly little research on these tensions in the field of music performance education. In the current climate of accountability, gaining more understanding of the inner workings of social moderation in this field seems highly pertinent.

CONCEPTUAL FRAMEWORK
Social moderation as micropolitics

Social moderation is held to be a system assuring assessment quality and professional development (Maxwell, 2010), defined as "a socio-culturally structured collective activity that involves teachers in a coordinated effort to develop shared meaning" (Adie, Klenowski, & Wyatt-Smith, 2012: 225). However, a review by

---

1   In Davidson & Da Costa Coimbra's (2001) study of assessment of vocal performance it is suggested that teachers can be more severe on their own students, while those who know their students pastorally might be more lenient.

Thornberg & Jonsson (2015) finds that results are unclear on the effects of social moderation, and Orr (2007: 647) describes the assessment process as a "messy practice" involving contextualized power relations. While a view of the teacher panel as a community of practice (Lave & Wenger, 1991; Wenger, 1998) emphasizes the development of particular interpretations of a mandate through participatory interaction, Wenger (1998) refuted the idea that such communities are democratic arenas in which decisions are reached freely and fairly. Rather, such tasks are likely to involve micropolitical activity, defined as a dynamic process in which power is operationalized and legitimized according to context, depending on the interests, resources and strategies of those involved (Ball, 1987). Previously seen as the unrespectable "dark side" of schools as professional organizations (Hoyle, 1982: 87), a more positive view of micropolitics is that this is not only an inevitable but a natural and necessary part of the life and development of the organization (Achinstein, 2002; Ball, 1987; Blase, 1991).

Three categories of professional interests that might be present simultaneously within the interaction are described by Kelchtermans & Ballet (2002)[2]. *Cultural-ideological interests* denote values, ideals and norms that are acknowledged as legitimate and defined through processes and interactions within the professional culture. For teachers, these involve both pedagogical principles specifically related to a subject domain, and universal ideological principles such as equality and justice (Ball, 1987). *Social-professional interests* concern the quality of interpersonal relationships in the school context, including both teacher-student relations and those between colleagues, and often seem to weigh heavily in teacher decisions (Kelchtermans & Ballet, 2002). *Self-interest* infers the construction of professional identity, as in "the sense of self or identity claimed or aspired to by the teacher" (Ball, 1987: 17), involving self-esteem and framings of self, as well as perception and definition of the tasks of the teacher (Kelchtermans & Ballet, 2002). While more than one category of interest is often at stake in specific interactions, the protection of self-integrity and professional identity is central (ibid.). The interactive character of the assessment meeting puts teachers in a vulnerable position in which their actions can be "perceived, interpreted and judged" (p. 111) by other colleagues. This has particular relevance for a context in which the interaction requires the individual to articulate

---

2   Material interests and organisational interests are also among the categories identified by Kelchtermans and Ballet (2002).

a subjective, intuitive response informed by connoisseurship (Eisner, 2003) and the kind of practitioner knowledge that is always richer in information than verbal descriptions can provide (Schön, 1983: 276). These concepts seem particularly apt for instrumental teachers coming from "private" teaching practices in specialist domains.

Epistemic authority and face-work

At the micro-level, the decision-making work of the teacher panel requires participants to use a variety of resources as they negotiate rights concerning who is entitled to have or evaluate specific types of knowledge relevant to context (Telenius, 2016). The term *epistemic authority* (Heritage & Raymond, 2005) is used in describing the ways a particular view becomes established as more significant or authoritative than another within a moment-to-moment interaction. While legitimate authority is found in the formal structures and roles of the organization, epistemic authority denotes "an incessant situated accomplishment" (Mondada & van Dijk, 2013: 598) which can be challenged and negotiated as participants position themselves through the operationalization and legitimization of various "situated" and "situational" resources (Goffman, 1983) according to context. The former are contingencies of habitus in the form of participants' predispositions and knowledge, formalized status and role which might have implications for the situation, including such factors as gender. The latter are the contingencies of "procedural form" (ibid.), involving opportunities within talk for navigating the situation in one's interests, using domain-specific knowledge and conversational practices such as turn-taking (Sacks, Schegloff, & Jefferson, 1974), preference organization (Pomerantz, 1985), and repair organization (Schegloff, Jefferson, & Sacks, 1977). Viewed in this way, micropolitical activity is a matter of agency based upon the affordances within the situation. However, there are certain principles fundamental to interaction of the individual with the social that micropolitical theory is in danger of glossing over; Goffman's work on social interaction can give important insights to a micropolitical approach.

For Goffman, social intercourse is constrained and conditioned by an inherent moral contract of reciprocal respect between social actors (Treviño, 2003). Using the metaphor of Ritual, Goffman (1972a) explains how the "sacred self" is affirmed through rules of courtesy and deference. The importance of "face" as a sacred object of the self is evidenced in *face-work*, involving the individual's

attempts to guide and control the impressions others have of her, and the cooperation and collusion of others in order to affirm this impression. At every stage of a conversational encounter, participants have obligations to be sensitive to the consequences of the interaction on both their own and others' selves, and to avoid anyone losing face (Williams, 1980: 220). In the event of embarrassment, face-work involves the corrective work of aligning behavior with face, in order to maintain the "ritual equilibrium", either through defensive action to demonstrate self-respect, or protective action, using tact and rules of courtesy in order to save others' face (Goffman, 1972b). Such action, involving the handling or management of talk through patterns of verbal and non-verbal action, is itself expected to be achieved without apparent effort and without becoming a topic of conversation. In this way, ritual elements structure and impinge on interaction, and talk can be seen not only as a medium of communication, but as an activity in which acts and gestures comprise "communication about communication". Concomitantly, while any interaction involves its own particular constellations of power relations, certain ways of talking might have "cultural salience" according to context, for example a tendency for women's conversation to be more collaborative and less competitive than men's (Coates, 2003). These conceptions of interaction enable a deeper understanding of how cultural-ideological and social-professional interests, and self-interest (Kelchtermans & Ballet, 2002) are interrelated and might overlap in a workplace context such as teachers' assessment meetings.

The context of the study

In Norway, the three-year elective upper secondary program, Music, Dance, Drama (MDD) offers a specialized arts subject in combination with subjects qualifying for entry to general higher education. The music program includes principal instrument as a subject domain, traditionally organized as one-to-one tuition. Instrumental teachers, who often work part-time in several educational settings, might see their students only for principal instrument instruction, while others may have more extensive knowledge of students through teaching in several disciplines (Ellefsen, 2014). The National Curriculum for Knowledge Promotion (Kunnskapsdepartementet, 2006) stipulates broad goals for the subject, devolving responsibility for interpreting these to the local level. Term tests are commonly arranged, giving students the opportunity to perform a self-chosen repertoire for a teacher panel, consisting of the teacher of the student

(hereafter termed ToS) and other teachers with appropriate competence, such as similar instrumental expertise. The work of the teacher panel in this context thus contains an inherent asymmetry with regard to the knowledge each teacher has of the student in question, representing interesting challenges. In the present study, I have undertaken an exploration of instrumental teachers' assessment meetings from a perspective of micropolitics, using the above concepts and guided by the following research questions:

- *RQ1: In what ways is epistemic authority American spelling: operationalized in the social moderation of instrumental teachers?*
- *RQ2: In a perspective of micropolitics and social interaction, what challenges can be identified for instrumental teachers participating in school-based social moderation?*

METHODS

The article draws on empirical material from an instrumental case study (Stake, 1995) at an upper secondary school in Western Norway, in which 14 assessment meetings were video recorded between December 2014 and November 2015. Observation of teachers' assessment meetings can give insight into an authentic context of music performance assessment with the potential to provide insight into defined issues for which the case itself is a "functioning specific" (Stake, 2000: 437). The case school was selected through purposive sampling, seeking a range of instruments and a system for social moderation that could yield about five meetings per term, as well as geographical accessibility. Relevant permissions were obtained from the Norwegian Centre for Research Data, the school leadership, music department, and individual teachers across instrument groups. Due to possible third-party identifiability, informed consent was procured from students who were mentioned in the assessment meetings, with data concerning the two students who did not give consent removed from transcriptions.

The 14 meetings yielded 10 hours 37 minutes of video data, during which 87 student performances were assessed, averaging at 7 minutes of discussion time per performance. Because several meetings were scheduled simultaneously, I was not present at all meetings, setting up the recording equipment beforehand, and for two meetings only partial data was procured. These variations in the research situation are acknowledged as limitations. Eighteen teachers participated in the study, 11 males and 7 females, with various Nordic and non-Nordic backgrounds

and expertise variously ranging from orchestral instruments to guitar, electric bass, piano and singing. For all the meetings, panels of 3-5 teachers grouped by similar instruments came directly from the student performances to the meeting-room in order to discuss and grade the performances. While formalized lists of criteria were not used, most teachers had written comments during the performances and brought these to the meeting. At each meeting, one teacher had a fixed role of moderator with responsibility for managing the interaction and ensuring a decision was reached. The moderator role was decided in advance by the teacher groups, favoring teachers with long experience, and the same teacher often had this role over several meetings. Normal meeting structure was for each teacher to give an account of the performance, with the ToS coming last. Whether grades were proposed in this initial account or not, varied from meeting to meeting.

The recordings were transcribed verbatim using elements from Jefferson (2004) and including non-verbal behavior I considered to contribute to mean- ing-making. The transcriptions were sent back to participants for verification in order to ensure that I had heard correctly with respect to diction and pronunci- ation, rapidity or overlapping of speech, and the use of domain-specific termi- nology. According to Oliver, Serovich & Mason (2005), a transcriber hears the recording through her own cultural-linguistic filters, affording her "significant interpretive and representational power that could affect analysis and results." Sending the transcriptions seemed particularly important for two reasons: my own instrumental expertise had its domain-related limitations, and, while a range of geographical accents was represented, Norwegian is not my first language. There was, therefore, also a reciprocity in assuring participants of the authenticity of the data I would be using in my analyses.

My analytic approach was abductive, based on multiple readings of the data and progressively informed by readings of literature in what Layder (1998) termed a "zigzag" process (p. 76). Following an initial coding process, close readings were made of passages I identified as "hotspots", inspired by MacLure's (2013) use of the concept of "glowing data" (p. 661). These were points of the interaction that seemed to hold tensions with regard to epistemic direction, having significance as sites for multiple interpretative framings. The analytic framework uses theory and tools of ethnomethodology and conversation anal- ysis (Gibson & Lehn, 2018). The purpose of this methodology is not to uncover "what this means for this person", but how individuals display to one another

what they think through their use of interaction rights, and how they use these meanings for social purposes (ibid.).

In this article, three excerpts from the study are explored in detail using constructs as described. All data are translated to English by the author, and fictitious names are used for all participants.

FINDINGS

The three excerpts illustrate significant ways in which participants use resources strategically in the management of interactions in the assessment meetings. After a brief introduction explaining the local context, the interaction is presented in numbered utterances with speaker identification, and then discussed.

Excerpt 1: Epistemic positioning (M01)

In this interaction, four teachers of guitar, bass and percussion convene to discuss student performances. Tom, the percussion teacher, is also the moderator of the meeting. The first performances to be discussed at this meeting are those of Adam and Ben, who each performed a side-drum étude, and then played a rock duo together on drum-set. Although meetings generally follow a rule that the ToS waits until the other teachers have given their accounts of a performance, Tom starts the discussion as follows.

1	Tom	So, in Year 2 I have one that can get a six (.) So:: and:: the other two won't be getting sixes
2	Max	You're probably talking about Adam
3	Tom	No, I'm talking about **Ben**, about Ben, yes (.) But Ben, he played outside of his comfort zone today (.) He played something he's ehh:: actually quite unfamiliar with (.) so::
4	Max	(hhh)
5	Tom	Yep (.) Now let's see (.) There's the side-drum bit (.) a formality (.) and suchlike and: yes, and that's it (1,0) Or do you want to-
6	Pete	Nah:: I've written:: that ehh (1,0) OK for those (.) There is some flow missing and forward drift in what they're playing, but ((glances at Tom)) OK
7	Tom	Hm, no, Adam could have practiced a bit more (.) but ehh:: but:: it's OK
8	Rob	Ehm:: I think Adam had best flow
9	Tom	Yehyeh ((shrugs shoulders))
10	Rob	<at least on the side-drum, and::

A turn-by-turn description of this episode shows how Tom uses his double authority as both moderator and ToS to give information he has privileged access to as ToS at the outset of the interaction.

Turn 1: Tom produces an evaluative statement predicting potential levels of achievement for three of his students for the next year, presenting this as fact, not opinion, yet not mentioning names. (Whether he is referring to discourses of talent/ability, or students' work ethics, or a combination, is not clear). In this way he sets up a frame of reference, having the first position statement of the sequence, and invites the participants to see one student as having more potential than the others.

Turn 2: Max responds by offering information on his own positive evaluation of Adam's performance, affiliating with what he assumes is Tom's opinion.

Turn 3: Tom makes a negative response, correcting Max and naming Ben as the student he meant, and reinforcing this by repetition. He then produces an evaluative statement presenting information about the conditions for Ben's performance (playing something unfamiliar). This is information he has access to as ToS and is presented as fact.

Turn 4: Max makes a small sound demonstrating that the previous utterance did not go unnoticed.

Turn 5: Tom produces an affirmative, which might be a reinforcement of his previous statement, or an opening for a new topic. Either way, this projects certainty of something, said with the authority of his combined role of ToS and moderator. His next statement is the set phrase "Let's see", an imperative to himself which nevertheless focuses attention on the managerial role. He then presents the side-drum étude as a formality, i.e. not part of the test to be graded. However, after a short pause, he produces an incomplete question to the panel, making the status of the étude in the assessment ambiguous.

Turn 6: Pete's negative is a confirmation that he will not include the étude in the grade. Then, referring to his written notes, he offers an evaluation of the études of the two students, claiming epistemic authority by referring to domain-specific knowledge (on flow and forward drift). His evaluation does not differentiate between the two, and he makes eye contact with Tom as he ends up with a non-committal, though positive evaluation.

Turn 7: Tom's response of "Hm, no" suggests a combination of both alignment and skepticism to Pete's evaluation. He then supplies new information about Adam, referring to perceived lack of practice. In this way, knowledge to which Tom

has access as ToS is used to give negative valence to one student and it is stated as fact, not opinion. He ends the turn by making an evaluative statement ("it's OK") that could be understood to mean either that the performance was adequate, or that Pete's evaluation is adequate. Either way, Tom claims epistemic authority as the person with particular expertise for stipulating requirements for quality.

Turn 8: Rob carefully initiates his first turn with a minimal token. Producing an evaluative statement as an opinion rather than fact, his claim to epistemic authority is thus nuanced by an indirect acknowledgment that others might think differently. Nonetheless, his reference to domain-specific knowledge is a warrant for authority. What he says resonates with Max's indirect evaluation (turn 2) in placing Adam's performance higher than Ben's, but this is the first direct articulation of this standpoint.

Turn 9: Tom aligns to Rob's statement with physical accenting. He is agreeing, but expressing that Rob is only telling him something that is both "old news", and inconsequential to him. In this way, he positions himself again as the expert and downgrades the authority of others.

Turn 10: Rob quickly adjusts his previous statement. By saying "at least on the side-drum", he opens up the possibility that Adam might be considered the better performer also when it comes to the main piece. Having already heard that Tom considers Ben to be a potential top achiever, Rob's comment opens a different scenario here, with the potential for more serious conflict with Tom.

Throughout this excerpt, Tom uses his formalized authority as moderator to take the floor and provide a certain backdrop for the discussion of Adam and Ben's performances. A clear picture is drawn of Adam as the less able and the lazier of the two students, and of Ben as a potentially high-achieving student who has taken the risk of going outside his comfort zone. In micropolitical terms, Tom operationalizes the situated resources of formal authority and expertise, as well as contingent situational resources of conversational practices, with an immediate agenda we can to some extent extract: to influence his colleagues to view Ben in a positive light. In the process, Tom positions himself as having primary rights to evaluate the performances, downgrading others' authority to do so. While his actions can be understood to be in the interests of providing fair judgment for both his students, he is also displaying an understanding of his role as ToS having high epistemic authority and a legitimate agenda of influencing colleagues.

The other participants can be understood to be at a disadvantage in view of Tom's combined role as ToS-moderator, but they nevertheless display their

independence and manage to avoid confrontation (turn 4, 6 and 8), keeping an equilibrium in which face is not overtly threatened for any of the actors.

Interestingly, Pete's reference to his written notes is one of many such allusions in the data corpus, and sets the written account apart as the documentation of a spontaneous response to the performance, a "judgment-in-time". This aspect of separateness or distance gives the possibility for a teacher to use this artefact in different ways, bracketing it as a "pure" judgment made on the basis of professional discernment (Schön, 1983), as well as disclaiming liability for it. This has importance for the next example.

Excerpt 2: Suspension of epistemic claims (M06)[3]

This interaction occurs at the end of a meeting, four months after the previous example, in which Tom, Pete and Max grade five performances on guitar, bass and percussion. From the start of this meeting, Pete has been a stringent marker, commenting that he has "not gone Rambo with the grades", and his proposals have consistently been below the others'. However, when the moderator Tom summarizes the tallied grades at the end of the meeting, Pete understands that he has placed himself in a difficult situation with regard to his own student. A crisis is played out, with the additional pressure of Max needing to get to his next teaching appointment.

1	Max	You must just say the grades, 'cause I gotta run
2	Pete	Just say it again
3	Tom	Adam, five minus, Ben, six to five. Cass, five plus. Eric, five plus. Felix, five minus (5,5) I feel that Felix has been poorly paid for level of difficulty, actually (.) I think we, we, Felix (.) he performed better than Adam
4	Pete	If, yehyeh, you said Adam, didn't you, then (.) In relation to **Adam**, he performed better than Adam, yeh (.) Ehh:: ((leafs through notes))
5	Max	((to Pete)) But if we think about the combos, I think Felix played very well in Adam's piece (.) and had a very good combo in his own piece, didn't he? But (.) ((to Tom)) let's get him up to a five ((tries to give his notes to Tom, who resists))
6	Tom	No, don't go (.) This here is important, it's important (.) We're looking
7	Max	Ehh, is it going to be corrected (.) up to five, maybe? Just decide, five minus
8	Pete	((to Tom)) I had both of them ehh: on the same grade
9	Tom	OK (.) Hm::

---

3   Three turns are omitted from the excerpt on grounds of relevance.

10		((Max sits down again))
11	Pete	And now Adam has got (.) was it five (.) plus?
12	Tom	Five minus
13	Pete	Five minus (.) And Felix is getting five minus, as well
14	Tom	Well (.) I'm dizzy with all this
15	Max	OK, so if you adjust Adam's grade ehh:: in relation to Felix, just to get that little difference between the two of them?
16	Tom	Mm: ((non-committal))
17	Max	We'll take Felix up to five
18	Tom	**Look**, you have five-to-four, and five-to-four for Adam, and I have five (.) Whereas, for Felix, he has five, five, and five-to-four
19	Max	Should we just::
20	Pete	What?
21	Max	We are **stuck** ((laughs slightly))
22	Pete	Yes, no, I am stingy, stingy today (.) But ehh (.) what happens if I give ehh:: give Adam]ehh::
23	Tom	[No, look (.) but do you feel it's right with a five-minus?
24	Pete	No, I don't really feel it's right for Felix with a five-minus (.) But I haven't given him a better grade here, either, in a way (.) But in relation to the others-
25	Max	((looks at the clock, suddenly stands up)) No, I have]to go
26	Tom	[No:
27	Max	I have to **go**, man. What is the discussion, we're agreed!
28	Tom	No, we aren't agreed, I mean-
29	Max	((sits down again, shakes his head))
30		We are not done
31	Pete	Compared with the others, see, Felix (.) he ehh:: he should maybe (.) he ehh:: he is a candidate who should maybe be over ehh:: yes
32	Tom	Mm (.) Put him one higher? 'Cause really he is a strong-
33	Pete	He is strong (.) Ehm::
34	Tom	I don't think we should pull down another student because of this (1,0) See, I feel we shouldn't put Adam lower because it doesn't feel right with Felix's grade (.) I'd rather we give Felix a five than pull Adam down
35	Pete	((writing)) Yeah, no, give him a five then, and I'll give him up, so he can get himself a five
36	Tom	OK ((writing)) (.) We'll do that.
37	Max	So it's gonna be a five?
38	Tom	It's a five for Felix
39	Max	That's what I said five minutes ago. ((gets up again)). OK, are we done, then? Now I have to go.

In the course of this episode, epistemic authority is suspended in a delicate set of maneuvers in which both Pete and Tom have something at stake. All three teachers express discomfort with an outcome of the same grade for both Adam and Felix, and Tom, using the nuanced epistemic verbs "I think", "I feel", starts the work of persuading Pete to adjust his grade. This work requires Pete to reconcile his conflicting loyalties between his high standards and equity for his own student, Felix. This might not seem so difficult if it were a case of blaming his colleagues for being too lenient. However, after looking in his papers, Pete sees that his initial evaluation was the same grade for the two. To adjust the grades now involves retracting the spontaneous judgment-in-time, something he has valued as a product of reflexive professional practice (Schön, 1983). Tom's comment at turn 9 suggests that he is sensitive to Pete's predicament, but might equally signal concern about another, related matter. Indeed, while Max has been pragmatically offering various solutions to the problem (turn 5, 7, 15, 17), one suggestion is for Pete to lower the grade for Tom's student, Adam. That the grade for Tom's student might be at stake is confirmed at turn 23, when Tom cuts Pete short and appeals to him to consider Felix's grade without the normative comparison. Pete, admitting the grade seems too low, yet still hanging on to the written notes as a "true" representation of his standards for achievement, again links the problem to the difference between students. At turn 31, he falteringly begins a rationale for Felix's grade to be raised, finally making the decision (turn 35), after Tom has reasoned both that Felix is a "strong" student, and that to lower Adam's grade is not a solution he would support.

This episode illuminates problematic issues at the heart of the system of social moderation. Without a common calibration of criteria and standards for quality, assessment processes can run into trouble. The normal bargaining tone of the assessment discussions is replaced in this interaction with a crisis in which something, or somebody, has to "give". The problems are partly due to one teacher's independent stand for his principles, reflecting complex relationships between crossing interests (Kelchtermans & Ballet, 2002), but also giving a glimpse of the sacred self and the attempts of all parties to maintain ritual equilibrium (Goffman, 1972b). In this situation, epistemic authority is suspended, and Pete holds the key to its restitution. Despite the solution-oriented contributions of Max, who is not ToS for either of the students discussed, he offers these only as suggestions. The turning-point comes at turn 31, when

Pete begins to reason that he can raise Felix's grade, and it is at this point Tom resumes his tactics of persuasion through the use of epistemic resources.

Excerpt 3: Aspersion and reparation (M07)

This excerpt is from a meeting of three female singing teachers, Jen, Kathy and Ingrid. As a panel, they take pride in being a cohesive group who have "tuned into each other" and often agree. This interaction comes at the end of a lengthy discussion about a performance involving the song 'Defying gravity', from the musical *Wicked* (Stephen Schwartz, 2003). While Jen and Kathy have described the performance as somewhat uneven, Ingrid, who is both moderator and ToS, disagrees. She has reasoned that the song is both unusually demanding and is little-known in Norway, adding that it means "incredibly much" to the student performer. Kathy has suggested a grade five-minus for the performance, and Ingrid repeats how difficult this repertoire is.

1	Ingrid	It's difficult to (.) it's very difficult to compare such different demands (.) And to understand some of the challenges (.) It's totally, it's almost im**poss**ible (.) For that, they would have to sing in the same (.) deliver something sort of in the same genre (.) That seems, that seems wrong, to me (.) yes-
2	Kathy	At the same time I'm very concerned that one shouldn't attempt bigger songs than one can master (.) So is, if (.) one shouldn't automatically get a plus just because the piece is so difficult
3	Ingrid	No (.) Yes (.) <that's not what I'm saying, that's not what I'm saying (.) That would be too stupid!
4	Kathy	It's that (.) <no no, I know, I know you don't mean that (.) not **that**, yeh-
5	Ingrid	That's not what I'm talking about
6	Kathy	So, ehh, but, no but, well (.) I can go with a five there
7	Ingrid	Now, I think the clock is beating us
8	Jen	((laughs))
9	Ingrid	Now to [name]. I don't think we'll manage to be finished by quarter to, actually. But it is rather tricky with these (.) I took, I think this forum, it's very important for us
10	Kathy	((nods))
11	Ingrid	Instead of just (0,5) And that we do listen to each other a little, that I think is important (.) [Name], who's student is she? Is she yours, Jen?

In this interaction, Ingrid's interest in her student receiving a higher grade leads to a chain of face-threatening moves, and it is Kathy who must make the repair. Ingrid begins by claiming epistemic authority in reasoning that the premises for

assessment should allow for varied repertoire of different levels of difficulty, yet that understanding and comparing such differences is difficult. In this, she is indirectly saying that particular competence is required in order to assess performances at multiple levels. By the end of the turn, she has oriented her talk to how unfair it would be to limit students to narrower repertoire, giving this argument force by referring to her own sense of justice. Kathy counters this with the argument that students should not be rewarded on the grounds of difficult repertoire, if they cannot master it. In this way, Kathy has claimed epistemic authority to produce a new, altered version of what Ingrid said. Instead of a rationale aknowledging the need for teachers with appropriate assessment competence, Kathy's version assumes such competence, but points out the folly of rewarding poor choice. Ingrid takes defensive action by being offended, and the next four turns comprise denial, retraction and repair as Kathy restores equilibrium by protective action (turn 4) and raising her grade (turn 6). The shifting elements of this interaction seem to illustrate how assessment practice is prey to conflicting interests. Although both teachers base their arguments on cultural-ideological premises, when Kathy adjusts her grade it seems to be in order to pacify Ingrid, rather than because she herself is persuaded of the merits of the case. The epistemic rupture has thus only been superficially repaired, and when Ingrid hurries on to the next topic this prompts a laugh from Jen, who has not taken sides in the matter. Nevertheless, in her legitimate defining power as moderator, Ingrid then shifts focus again, commenting on how she values the assessment forum and the way the teachers "listen to each other a little." In this way, she moves the discourse to a "safer place", in which the potential for social moderation as a forum for the development of assessment literacy seems to be the focus.

DISCUSSION

The aims of this study were to investigate how epistemic authority is operationalized in a context of instrumental teachers undertaking social moderation, and to identify challenges for this type of school-based assessment. From the three excerpts, assessment discussion is revealed as a layered social interaction in which teachers use available situated and situational resources strategically to establish epistemic authority, for example privileged knowledge of student, domain expertise, and projection of the spontaneous "judgment-in-time" as unassailable. Although there is little talk about the music performances themselves, the examples illuminate issues that are central for music teachers'

assessment practices. The *topic* of talk reflects the participants' cultural-ideological interests (Kelchtermans & Ballet, 2002) and a shared aim of making valid assessments and providing equity for their students, yet in the interplay of collegial relations this aim is clouded and diverted, and social-professional interests and self-interest (ibid.) are heavily implicated in the moderation process.

Of particular significance are the roles of moderator and ToS. Excerpt 1 illustrates how high epistemic authority can be established through micropolitical activity, as Tom uses the combined role of moderator and ToS to influence his colleagues. The moderator's responsibility for managing the assessment discussion legitimizes use of situational resources to direct the interaction. Thus, Ingrid, in Excerpt 3 also has the advantage over Kathy first by asserting epistemic authority, then taking offence, and finally defining a new epistemic narrative of the assessment meeting as an arena for teacher development. Nevertheless, the availability of these resources is by no means exclusive, and even within these brief excerpts it is clear that legitimate authority is only half the story – indeed, it is often the other participants who carefully manage the ritual equilibrium (Goffman, 1972b). The importance of such face-work is shown in a particularly tricky situation in Excerpt 2, when Pete experiences a crisis of conflicting loyalties. In the delicate work of moving towards a compromise, epistemic claims are suspended, and it seems that even in a context such as a low stakes term test, there *is* much at stake. Interestingly, while the three examples come from same-sex groups of teachers, the careful face-work of each excerpt reveals both men and women as competent and sensitive handlers of interactions.

Several challenges emerge from this analysis related specifically to the assessment of music performance. In an understanding that assessment involves connoisseurship and the invoking of professional discernment (Eisner, 2003; Schön, 1983), the individual's spontaneous judgment can become a problematic "black box" for the assessment discussion. This can involve distancing oneself from a judgment and projecting it as an independent resource produced more purely and spontaneously in response to the performance, superior to a recollection and reconstruction of the experience. Conversely, such judgments can be devalued to something less trustworthy than a reflection made after the event, indeed, after the whole sequence of performances.

In addition to this, the asymmetric relations within an assessment panel represent challenges, for example when there is pressure to acquiesce to a dominant view. Orr (2007) described assessment dialogue in terms of a "dance" or

a "skirmish", noting that teachers with seniority seldom changed their grade proposals, while others seemed to "reckon the cost" in connection to the chances or of succeeding. While the context of an assessment meeting may have its own procedural form and rules for politeness, the present study reveals instrumental teachers using the contingencies of the situation with care, attempting to avoid direct conflict and disagreement. Particularly challenging is the possibility that tensions between cultural-ideological interests, socio-cultural interests and the maintaining of "face" might lead to the adjustment of students' grades.

A third general challenge is the practical parameters of time limits for the assessment discussion itself. The particular construction of the assessment meeting as social interaction in an institutional context involves the imperative to arrive at a result, as illustrated in the mounting pressure "to be finished" in Excerpt 2 in which Max has another appointment to meet. As communities of practice, the assessment teams have a task to fulfil from which they cannot reasonably abdicate.

While there might be a "truth regime" where assessment ostensibly is practiced according to "rules", the study demonstrates that there are openings in the restricted space of the assessment discussion for multiple alternative functions. Indeed, while a panel might develop particular axiomatic principles for assessment, the social interaction between teachers is of crucial importance for assessment outcomes.

IMPLICATIONS

What can these findings mean for the practice? They demonstrate that there is a need to think seriously about the claims made for social moderation, and the challenges this poses for instrumental teachers who work in variegated contexts, often without membership in the larger community of teachers in school. Although the assessment meeting is an important arena for both the development of assessment literacy (Adie et al., 2012) and of collegial relations in a profession marked by isolation, there are challenges which should be considered. In the climate of accountability, the ability of music teachers to verbalize conceptions of quality is seen as an urgent matter, yet equally it is feared that a "doctrine of verbal clarity" may lead to trivialization and the undermining of music subjects in schools (Zandén, 2013). However, it may be that in focusing on inadequate vocabulary and articulation of quality, we miss something more fundamental: by glossing over the power relations implied in the interactions

of social moderation, important issues that we need a better understanding of, go unaddressed. These issues have relevance not only for the field of music education and music teacher education, but also for teachers in general.

REFERENCES

Abeles, H.F. (1973). Development and validation of a clarinet performance adjudication scale. *Journal of Research in Music Education, 21*(3), 246-255. doi:10.2307/3345094.

Achinstein, B. (2002). Conflict amid community: the micropolitics of teacher collaboration. *Teachers' College Record, 104*(3), 421-455.

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.

Angelo, E. (2012). *Profesjonsforståelser i instrumentalpedagogiske praksiser [Philosophies of work in instrumental music education]*. (PhD thesis), Norges teknisk-naturvitenskapelige universitet (NTNU), Trondheim.

Ball, S.J. (1987). *The micro-politics of the school: Towards a theory of school organization*. New York, NY: Routledge Press.

Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2010). Validity in teachers' summative assessments. *Assessment in Education: Principles, Policy & Practice, 17*(2), 215-232. doi:10.1080/09695941003696016

Blase, J.J. (1991). *The politics of life in schools: Power, conflict and cooperation*. London: Sage Publications.

Coates, J. (2003). *Men Talk: stories in the making of masculinities*. Oxford: Blackwell.

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

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 thesis). Oslo: 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.

Gaunt, H. (2011). Understanding the one-to-one relationship in instrumental/vocal tuition in higher education: comparing student and teacher perceptions. *British Journal of Music Education, 28*(2), 159-179. doi:10.1017/S0265051711000052.

Gibson, W., & Lehn, V. (2018). *Institutions, Interaction and Social Theory*. London: Palgrave.

Goffman, E. (1971). *The presentation of self in everyday life*. Harmondsworth: Penguin Books.

Goffman, E. (1972a). *Interaction ritual: essays on face-to-face behavior*. Harmondsworth: Penguin Books.

Goffman, E. (1972b). On face-work: an analysis of ritual elements in social interaction. In *Interaction ritual: Essays on face-to-face behavior*. Harmondsworth: Penguin Books.

Goffman, E. (1983). The interaction order: American Sociological Association, 1982 Presidential Address. *American Sociological Review, 48*(1), 1-17.

Green, L. (1997). *Music, gender, education*. New York, NY: Cambridge University Press.

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

Heritage, J., & Raymond, G. (2005). The Terms of Agreement: Indexing Epistemic Authority and Subordination in Talk-in-Interaction. *Social Psychology Quarterly*, *68*(1), 15-38. doi:10.1177/019027250506800103.

Hoyle, E. (1982). Micropolitics of Educational Organisations. *Educational Management & Administration, 10*(2), 87-98. doi:10.1177/174114328201000202.

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 Benjamins.

Kelchtermans, G., & Ballet, K. (2002). The micropolitics of teacher induction. a narrative-biographical study on teacher socialization. *Teaching and Teacher Education, 18*(1), 105-120. doi:10.1016/S0742-051X(01)00053-1.

Kunnskapsdepartementet. (2006). *Læreplanverket for Kunnskapsløftet [National curriculum for knowledge promotion in primary and secondary education and training]* (82-486-0397-0). Oslo: Utdanningsdirektoratet Retrieved from https://www.udir.no/laring-og-trivsel/lareplanverket/

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 Publications.

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

Maxwell, G.S. (2010). Moderation of student work by teachers. In (pp. 457-463).

Mills, J. (2005). *Music in the school*. Oxford: Oxford University Press.

Mondada, L., & van Dijk, T.A. (2013). Displaying, contesting and negotiating epistemic authority in social interaction: Descriptions and questions in guided visits. *Discourse Studies*, *15*(5), 597-626. doi:10.1177/1461445613501577.

Nerland, M. (2007). One-to-one teaching as cultural practice: Two case studies from an academy of music. *Music Education Research, 9*(3), 399-416. doi:10.1080/14613800701587761.

Oliver, D.G., Serovich, J.M., & Mason, T.L. (2005). Constraints and opportunities with interview transcription: towards reflection in qualitative research. *Social Forces, 84*(2), 1273. doi:10.1353/sof.2006.0023

Oltedal, E. (2018). Challenges of assessing music performance: teachers' perceptions. In *Nordisk musikkpedagogisk forskning Årbok 18* (Vol. 2017, pp. 241-270). Oslo: Norges musikkhøgskole.

Orr, S. (2007). Assessment moderation: constructing the marks and constructing the students. *Assessment & Evaluation in Higher Education, 32*(6), 645-656. doi:10.1080/02602930601117068.

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.

Pomerantz, A. (1985). Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes. In J. M. Atkinson (ed.), *Structures of Social Action* (pp. 57-101). Cambridge: Cambridge University Press.

Rowe, V.C. (2008). *Patterns and consequences of gender interactions in instrumental music lessons* (PhD thesis). London: University of Roehampton.

Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4), 696-735. doi:10.2307/412243.

Schegloff, E.A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*(2), 361-382. doi:10.2307/413107.

Schei, T.B. (2007). *Vokal identitet [Vocal identity]* (PhD thesis). Bergen: Griegakademiet, Universitetet i Bergen.

Schön, D.A. (1983). *The reflective practitioner: how professionals think in action*. New York, NY: Basic Books.

Stake, R.E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.

Stake, R.E. (2000). Case studies. In N.K. Denzin & Y.S. Lincoln (eds.), *Handbook of qualitative research* (2nd ed., pp. 435-454). Thousand Oaks, CA: Sage Publications.

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

Telenius, J. (2016). *Sensemaking in meetings* (PhD thesis). Helsinki: Aalto University. Retrieved from http://urn.fi/URN:ISBN:978-952-60-7014-8

Thornberg, P., & Jonsson, A. (2015). Sambedömning for ökad likvärdighet? [Social moderation for increased equity?]. *Educare, 2015*(2), 179-205.

Treviño, A.J. (ed.) (2003). *Goffman's legacy*. Lanham, MD: Rowman and Littlefield.

Vinge, J. (2014). *Vurdering i musikkfag: En deskriptiv, analytisk studie av musikklæreres vurderingspraksis i ungdomsskolen [Assessment in music: A descriptive, analytical study of music teachers' assessment practice in lower secondary school]* (PhD thesis). Oslo: Norges musikkhøgskole.

Wapnick, J., Ryan, C., Lacaille, N., & Darrow, A.-A. (2004). Effects of selected variables on musicians' ratings of high-level piano performances. *International Journal of Music Education, 22*(1), 7-20. doi:10.1177/0255761404042371

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

Wesolowski, B.C., Amend, R.M., Barnstead, T.S., Edwards, A.S., Everhart, M., Goins, Q.R., … Williams, J.D. (2017). The development of a secondary-level solo wind instrument performance rubric using the multifaceted rasch partial credit measurement model. *Journal of Research in Music Education, 65*(1), 95-119.

Williams, R. (1980). Goffman's sociology of talk. In J. Ditton (ed.), *The view from Goffman* (pp. 210-232). London: Macmillan.

Wrigley, W.J., & Emmerson, S.B. (2013). Ecological development and validation of a music performance rating scale for five instrument families. *Psychology of Music, 41*(1), 97-118. doi:10.1177/0305735611418552.

Yorke, M. (2011). Summative assessment: Dealing with the "measurement fallacy." *Studies in Higher Education, 36*(3), 251-273. doi:10.1080/03075070903545082.

Zandén, O. (2010). *Samtal om samspel: Kvalitetsuppfattningar i musiklärares dialoger om ensemblespel på gymnasiet [Discourses on music-making: Conceptions of quality in music teachers' dialogues on upper secondary school ensemble playing]* (PhD thesis). Gothenburg: Göteborgs universitet.

Zandén, O. (2013). *Till en kritik av tydlighetsdoktrinen [Towards a critique of the doctrine of clarity]*. Paper presented at the Nordic Network for Research in Music Education (NNMPF), Bergen 2013. http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-26083

Zhukov, K. (2008). Exploratory study of approvals and disapprovals in Australian instrumental music teaching. *International Journal of Music Education, 26*(4), 302-314. doi:10.1177/0255761408096072.