Assessment of composing in the lower secondary school in the English National Curriculum

Martin Fautley[1] and Jonathan Savage[2]

[1]Birmingham City University, Centre for Research in Education, Attwood Building, Perry Barr, Birmingham B42 2SU
[2]Institute of Education, Manchester Metropolitan University, 799 Wilmslow Road, Didsbury, Manchester, M20 2RR
martin.fautley@bcu.ac.uk, j.savage@mmu.ac.uk

This article reports the results of research into teachers' practices concerning the assessment of composing at Key Stage 3 in the National Curriculum for Music in England. It finds that many teachers are using NC levels for assessing individual pieces of work, a process for which they were never intended. It also finds that teachers find it difficult to show progress using NC levels, and that many teachers have rewritten the levels into child-friendly language, thus causing further difficulties.

Introduction

In England it is a statutory requirement of the National Curriculum that all pupils study music in the lower secondary school up to the age of 14 years (we are specifying 'English' here as the National Curriculum is different for each member country of the UK). This phase is known in the UK as Key Stage 3 (KS3), and encompasses the part of secondary education which falls between the ages of 11 and 14 years. The contents of the National Curriculum are delineated by statute, and composing is one of the key processes in this. Although a statutory activity, assessment of composing does not take place as a separate component, instead summative assessment of attainment in music is afforded by a single National Curriculum level, for which level statements are documented in the National Curriculum for music itself.

There have not as yet been any major national studies in England of teachers' assessment practices with regard to composing. This study set out to address this by investigating how classroom music teachers undertook assessment of composing. We were interested in both formative and summative aspects of assessment of composing, and set out to inquire as to what sorts of assessment mechanisms and systems were employed by teachers for assessing composing at KS3 in England. Our research was governed by two principal questions:

- What are teachers actually doing when they assess classroom composing?
- How are official structures (e.g. National Curriculum levels) utilised in this process?

Research into composing in the curriculum

Composing as a curricular activity has received increasing attention in recent academic research, and a number of facets of it have been investigated. Burnard has researched a variety of approaches to, and outcomes from, the composing process (Burnard, 2000b, 2006; Burnard & Younker, 2002, 2008), including how learners make meaning from composing and improvising, and the cognitive and practical mechanisms in place whilst undertaking it (Burnard, 2000a, 2002). In the context of the UK National Curriculum (NC), composing often takes place as a collaborative process, and this has formed the locus of investigation for a number of studies (Burland & Davidson, 2001; Fautley, 2004, 2005). Social interaction plays a large part in group composing, and this has also been investigated (Miell & MacDonald, 2000; Burland & Davidson, 2001; MacDonald *et al.*, 2002; Major, 2007, 2008). More recently, what Activity Theory can tell us about group composing has been studied (Burnard & Younker, 2008). Pedagogic attitudes to composing have sometimes been an issue, and this area has been investigated too (Odam, 2000; Paynter, 2000; Berkley, 2001; Byrne & Sheridan, 2001).

Assessment research

Assessment is a key area of interest in contemporary educational discourse. Key distinctions are drawn between summative and formative assessment. Summative assessment is where marking, grading and certification of learning takes place, usually at the conclusion of a learning episode (Harlen, 2005, 2007). Formative assessment, frequently referred to as assessment for learning, has been shown to play a key role in the development of learning, and has been well documented (Black, 1995; James, 1998; Black & Wiliam, 1998, 2006; Assessment Reform Group, 1999, 2002; Black *et al.*, 2003a, 2003b, 2004).

Assessment of composing

Assessment of composing, particularly in the English situation has received less attention, however. There are some general overviews (Stephens, 2003; MacDonald *et al.*, 2006) and some researchers have investigated teacher practice in this area (Byrne & Sheridan, 2001; Byrne *et al.*, 2003). In their discussions concerning the composing pathways undertaken by individuals, Burnard and Younker (2004) mention assessment. Mills (1991) investigated the musical nature of assessment, whilst Brophy (2000) and Colwell (2002, 2007) have both considered a range of other issues. From a classroom perspective, Bray (2000, 2002) and Adams (2000) discuss ways in which teachers can undertake assessment in the classroom, whilst Fautley (2008, 2009) has provided guidance for both serving and trainee teachers, as well as considering assessment issues more widely (Fautley, 2010). From a more general perspective, issues concerned with assessment and the arts have been discussed (Murphy & Espeland, 2007), as have the more specific issues of assessment in music education (Murphy, 2007). Kaschub and Smith, writing from the perspective of the USA, address the issue of grading composing, observing that 'In many cases there are no compelling reasons why grades must be given in composition classes' (Kaschub & Smith, 2009, p. 97). This is a situation that many older music teachers in the UK will recognise from pre-NC days.

The context of schools

In England, considerable investment has been made into developing teacher proficiency in assessment, in term of both resources (DfES, 2002, 2003, 2004a, 2004b) and in-service training. Formal examination-based assessment of composing in schooling happens after the period of investigation of this current study, at Key Stage 4. Assessment of composing at KS3 does take place, as it forms part of the teacher assessment which gives an overall level grade to all pupils at the end of the key stage. It is important to note that teacher assessments of pupils in terms of the NC levels they are awarded are not subject to any form of external verification or moderation. There are number of issues associated with NC assessment practice, both from an academic perspective (Ruthven, 1995; Wiliam, 2003) and from music teachers themselves (Fowler, 2008). Whilst there have been moves to facilitate a common understanding of what might constitute attainment at each specific level, previous research in this area has shown differences in understanding amongst teachers (Harlen, 2005).

National curriculum levels in music

Statutory assessment of attainment in music occurs, as we have observed, using National Curriculum levels, and takes place at the end of the period of study of compulsory music lessons, normally at the age of 14. In common with all NC subjects there are eight levels, plus one for 'exceptional performance'. The levels holistically delineate the processes of musical learning, including performing, composing and listening. It is important to note that there are not separate levels for each of these activities, rather, the levels are predicated on the conception of integrated practice. There are no assessment tests in music, the judgements for levels are decided by teachers alone, and there is no statutory requirement for external moderation or verification of an individual teacher's NC levels.

The level expected of all pupils at the end of their compulsory period of study is level 5. The level statement for which reads thus:

> Pupils identify and explore musical devices and how music reflects time, place and culture. They perform significant parts from memory and from notations, with awareness of their own contribution such as leading others, taking a solo part or providing rhythmic support. They improvise melodic and rhythmic material within given structures, use a variety of notations, and compose music for different occasions using appropriate musical devices. They analyse and compare musical features. They evaluate how venue, occasion and purpose affect the way music is created, performed and heard. They refine and improve their work. (QCA, 2007)

The composing component is contained within this level statement, but is not delineated in attainment terms, such as 'pupils will be able to ...'. This enables holistic judgements to be made concerning pupil progress, without having unwieldy assessments to undertake on classes of 30 pupils at a time. NC level statements can be viewed as having a constructivist perspective, rather than a behaviour-driven series of outcome statements. This has implications, as we shall see.

The original intention of the NC levels was that they should only be used once, at the end of a key stage, in this instance at age 14+, therefore they were not intended to be used

for individual pieces of work. Anecdotally we understood that this was no longer the case, and teachers were required to provide schools with regular updates of all their pupils using the NC levels. We wanted to find out how much truth there was in these stories, and how the NC levels were used by teachers.

Research methodology

This research involved combining both qualitative and quantitative methodologies (Gorard & Taylor, 2004). The assessment of composing is a complex issue, and employing a mixed methodology enabled us to grapple with the issues raised. As Creswell observes, 'the problems addressed ... are complex, and the use of either quantitative or qualitative approaches by themselves is inadequate to address this complexity' (Creswell, 2009, p. 203). Another factor in the employment of a mixed methods approach was that we wished to combine quantitative responses with qualitative judgements, and were cognisant of the notion that '... research approaches should be mixed in ways that offer the best opportunities for answering important research questions' (Johnson & Onwuegbuzie, 2004, p. 16), and that employing a mixed methodology is '... more than simply collecting and analysing both kinds of data; it also involves the use of both approaches in tandem ...' (Creswell, 2009, p. 4).

With these factors in mind, the research was designed to take place in three phases. The first phase was an on-line survey. The second phase was a follow-up on-line survey, which developed themes from the first survey in greater depth. Materials from both of these surveys were employed to inform the final phase, which was a series of individual semi-structured interviews with teachers. In the initial on-line survey questions were asked which involved attitudinal responses, Likert scale prioritisations, pedagogical procedural questions and free text responses. The methodological justification for this modality was that we were keen to establish an ontological perspective with regard to *what* was taking place with regard to assessment of composing, before moving on to explore reasons *why* this was the case in the survey phase, what Punch (2009, p. 25) refers to as 'question-method connections'. Bearing in mind the observation that 'Rating scales ... are limited in their usefulness to researchers by their fixity of response caused by the need to select from a given choice' (Cohen *et al.*, 2007, p. 328), we wanted to explore using semi-structured interviews the reasons for answers to the surveys.

The initial on-line survey was made available to all secondary school music teachers in England, and was widely advertised in subject association material, and by direct e-mail contact. The survey received a considerable number of views ($n = 866$) and was started by 176 respondents, of whom 94 went on to complete it. There was a range of respondent contexts, from rural to inner city. In this initial survey, respondents were asked if they would be prepared to be involved in a more in-depth follow-up survey, also to be undertaken on-line. This considerably more complex survey received 339 initial views, was started by 34 respondents, and completed by 18. In reporting the data from the various stages of the research, we have included in the analysis details as to the phase of the research the data in question arose from, and the number of respondents. Although not all respondents completed the surveys, the on-line survey tool (QuestionPro) allowed completed questions to be analysed, even if the survey itself was unfinished. On completion of the second

questionnaire, respondents were asked if they would be prepared to be interviewed, and 11 teachers were selected for this. Geographically, the spread of teachers was from the North-East of England, via the North-West and the Midlands, to London and the South East. All but one of the interviews was conducted on a face-to-face basis, the other being done by telephone. Interviews were semi-structured, based on a pre-determined interview schedule, which allowed the possibility for supplementary questioning to take place (Cohen *et al.*, 2007; Denscombe, 2007). In reporting speech from teachers, we have transcribed directly what they said, and used conventionally represented punctuation to aid meaning (Marshall & Rossman, 2006).

Data analysis was undertaken in a number of ways. Qualitative analysis of free text responses was coded using a grounded theory approach (Glaser & Strauss, 1967) where codings for the data arise from analysis thereof. This was undertaken as an iterative process, with increasingly fine-scaled unique codings arising as a result, in a developed form of axial coding, followed by coding for process (Strauss & Corbin, 1990, p. 163). We assumed items to be significant when mentioned by numbers of respondents, and repeatedly revisited the free-text responses to ensure we were capturing the importance of what was being said.

Quantitative data were analysed in a number of ways. Nominal scales of denotation were employed in order to establish categories of responses, such as composing pedagogies. Likert scales for attitudinal responses were used, and had inferences drawn from them. Ordinal scales were also employed to find how much use was made of assessment techniques. In undertaking this analysis we were mindful of Wright's assertion that '. . . the level of measurement is not an inherent characteristic of a particular variable, but a characteristic that we, as researchers, bestow on it based on our theories of that variable. It is a belief we hold about the variable' (Wright, 2003, p. 127).

For the purposes of our research we did not qualify or define the term 'classroom composing'. We used this consistently, and respondents and interviewees did not question this usage.

Results

In the first survey, we asked the open-text response question: How do you assess composing at Key Stage 3? One hundred and seven teachers responded. From analysis of these free-text responses, we were able to categorise teachers' responses into five areas. These were:

1. NC levels: Almost invariably mentioned directly.
2. Criterion referencing: Some teachers mentioned or discussed some form of criterion referencing in their answers. Sometimes these were of their own devising, but others were taken from NC levels which had been rewritten into 'pupil-speak'.
3. Summative assessment: This included assessment of learning (AofL), and when teachers described their use of summative techniques.
4. Formative assessment: This included assessment for learning (AfL), and when teachers described using AfL strategies, and included self- and peer assessment when mentioned.

Results obtained from this are shown in Table 1.

Table 1 *Respondents' mention of assessment modalities*

	Count	%
NC levels	55	49%
Formative	18	16%
Summative	12	11%
Criterion Referencing	49	44%
Total Respondents	112	

We were able to use these four categories as units of analysis for considering the ways in which teachers were assessing composing at KS3. We shall consider each of them separately.

NC levels

The results shown in Table 1 clearly demonstrate that NC levels were the main modality for teacher assessment at KS3, with 49% of respondents mentioning them directly. The text from which this analysis is derived reveals that, for some teachers, use of the NC levels was felt to be sufficient in and of itself, with short answers such as these being not atypical: 'Use NC Level descriptors'; 'National Curriculum criteria'; 'KS3 levels'; 'NC orders'. This seems to show that in teacher thinking there exists an unproblematic connection between NC levels and assessment of composing.

From a methodological perspective, an issue here is that our analysis of teacher responses only counted when teachers mentioned NC levels *directly* in their answers. Reading the text of respondents it is possible that what teachers were doing was explaining *how* they use NC levels, but not mentioning them. For example, none of these three (different) teacher respondents mention NC levels directly:

> Break task down to several distinct aspects and mark on how well each has been tackled; then a more general mark on how effective the piece was; finally written comments on the complete composition project with comments on what was especially good and what can be done to improve composing in the future.

> Often orally to the group or to the class, mainly through peer assessment (what was good what was not etc.). Traffic lights – set the criteria with the class at start then assess whether it was met. (What makes a good song? X X and X – did they have all three?) Self assessment done in writing – what they feel worked, what didn't, how will they improve next time?

> I look at the four areas of Composing, Performing, Listening and Appraising and this is the structure of my assessment.

To investigate this in more detail, a later question in the first survey asked to what extent teachers used NC levels to make judgements about their pupils. One hundred and four teachers answered this question. Despite the original intentionality concerning the use of

Table 2 *Frequency of NC level use*

	Count	%
To give each piece of work a level	26	25%
To give a level to each child at least once per term	51	49%
To give a level to each child at least once per year	16	15%
To give a level only at the end of a key stage	9	9%
Total Respondents	104	

Table 3 *Reasons for frequency of NC level use*

	Count	%
Yes, I have to provide levels this frequently	37	36%
Partially, and also because I want to keep track of pupils and the NC levels are a useful benchmark	47	45%
No, I do it this way because I want to	11	11%
I don't have to provide levels except at the end of a key stage	7	7%
Total Respondents	104	

Table 4 *Question: Do you find the National Curriculum levels helpful?*

Answer	Count	%
Yes	71	70%
No	31	30%
Total Respondents	102	

NC levels as outlined above, 25% of teachers responded that they use levels to assess individual pieces of work. Only about 9% of teachers use levels in the way which they were originally intended, to report at the end of the Key Stage, with 49% of teachers giving levels to all pupils at least once per term, with about 15% using them on an annual basis in order to review progress. This is shown in Table 2.

What this means is that some 74% of teachers are employing NC levels at least once a term. This frequency of utilisation is an issue for music teachers. We then asked whether teachers were assessing frequently because they had to, or because they wanted to. Again, 104 teachers answered this question. Results from this are shown in Table 3.

Only about 11% of respondents answered that they assess with this frequency because they wanted to; 36% of respondents were using the NC levels this often because the school required them to. However, 45% of teachers were using the NC levels as a way of monitoring progress, using them for benchmarking so they could track progress over time, and it is to this use of the NC levels that we now turn.

Using levels to show progression

In order to develop our understanding of the ways in which teachers were using the NC levels to show progression in composing, in interviews we asked teachers about this. The NC levels, as was noted above, were not intended to show differentiated development of key processes in music, so in many ways this can be considered to be an artificial question. However, it was designed to uncover teacher thinking in this area.

Answers were broad and varied. For some teachers the levels were not used in this way:

> No. The statements are too broad and sometimes unrealistic as to what can be achieved. They expect too much from students who don't have much actual music teaching.

> I have separate criteria for composition; I don't use the National Curriculum levels.

One teacher had clearly considered the intended nature of the NC levels, and observed that they were not meant to be stand-alone assessment criteria, and so this was not possible:

> Well there are elements that you can pick out of the levels which do relate to composition, there is a strand of composing which runs through them, to show development. Yes I think you can … but I don't think the stranding is particularly thorough, and I don't think it's supposed to be particularly, so you talk about intentions and expressive effects at various different stages, and you can see some sort of differentiated outcomes through those. But I don't think as a set of, well they're not meant to be criteria, and I don't think that as a set of statements they're specific enough really to track progression in all but the vaguest terms.

Some teachers used the NC levels for internal recording purposes, but did not share them with the pupils:

> Yes, we do use them but we use them as teachers. We don't share National Curriculum levels with the students.

A number of teachers had broken down the level statements from the NC, and used these to show progression:

> I think we are able to use them to show development, because the way that I've broken down the levels, in each of the attainments, as a block is performing, appraising and composing. So if you extract from that the different strands, I think you are able to show development through those things. So for example level 5 is one that talks about using chords, the right sort of chords, and then there's the one that's, you are just able to put sounds together. But then if you're able to use chords, and whatever the other thing it says, effectively, then you are showing some form of development. However, within that, I think there's a lot of scope for maybe breaking it down to show more development, because if you're able to do this, it's this, but if you're able to do that, then it's this, but there is quite a lot that happens in between, to go from that, to that. If you follow my, my logic pattern there! So I think you are able to show it, but I think it's quite tricky.

In terms of deconstructing level statements, many teachers talked about the school requiring them to break down the existing level statements into sub-levels, often three. This breaking down is not an officially sanctioned activity, and sub-levels only exist where individual teachers or music departments have re-written them. What this means is that each set of school-specific sub-levels is unique. For example, some schools in our research sub-divided levels into three, in other schools we found sub-division into tenths of a level. The net result of this on a nationwide basis is that lots of different systems are operating simultaneously, with limited or no transferability.

Related to sub-levelling, one teacher we interviewed was having trouble with a school-imposed linear improvement requirement:

> Oh this is a contentious one. Oh well, we're supposed to aren't we? It's just so hard in music. I mean the rule, the rule of the school is that every child should increase by two sub levels per year. And so if we're playing the game, every child increases by two sub levels per year in music.

> *Interviewer:* So how many sub levels are there at this school?

> *Teacher:* Three.

Another teacher was also facing similar issues:

> I find the levels, this is a very difficult one because, I think you can report levels and show progress, but my issue is, how often you are actually reporting the progress. So if you're, you know if your school tells you that you have to report two to three times a year, and says somebody is on a level five and they're still on a level five in the following term, it looks to that student as if they haven't made any progress, and they actually have. They might have fulfilled some actual criteria within the block. What our school's actually making us do on our sub levels, and we're having to do the ABC and breaking these levels down, and I find that really tricky. I don't personally agree with it and I don't like it, and I don't think there's, you know, if you can say, a pupil can display these skills, and they had this knowledge and understanding of *blah blah blah*, I don't see why we had to refer to that so often. I think that actually tells a pupil more, and the parents more, about what they are actually attaining, and achieving, then a number slapped on the head so to speak. That's what it feels like to me, as though we are becoming this number of assessed, and you know, pupils might think, 'oh I'm a level five'. But when you ask them what does that actually mean, because to me, it's more important about what that number means and how they do actually move on and improve. And yes we do, we do, we level twice a year, because we're not allowed to do it any less than that. If it was my choice, it would be much less than that.

These two teachers, and many others besides, talked of how attainment could only be upwards, and *only* in a linear fashion, 'every child should increase by two sub levels per year', the teacher above noted. One teacher, 'off the record', talked of how he had written a spreadsheet routine that did this automatically at the touch of a button! Another teacher was even more disconnected from the process:

> . . . school requires reporting termly, and gets stroppy if no progress is shown, so [NC levels] tend to be made up. The requirement was for end of key stage assessments,

which might have been manageable, but unlikely with 20 classes of 30 or more to teach in one-hour sessions and wide variety of lesson outcomes.

Fautley writes of how assessment in music education is geared to the requirements of three groups of users, teachers, pupils and systems, '. . . there is a clear difference between assessment for classroom purposes which helps with music-making, and assessment for auditing purposes' (Fautley, 2010, p. 69). Many teachers here spoke of assessment being undertaken to fulfil a systemic need for audit, rather than a learning one.

Do teachers find NC levels helpful?

We then asked the simple and stark question: Do you find the National Curriculum levels helpful?

Despite their having been in existence since 2000, and many training courses having been run concerning their use, 30% of teachers reported that they found them unhelpful. Clearly this means that 70% of teachers do find them helpful, but it does seem that the use of NC levels raises a number of questions, not least concerning that of ways of meeting the concerns of this significant minority of teachers.

Translation of levels into 'pupil-speak'

From answers to survey questions, and from anecdotal evidence amongst teachers, we wondered how prevalent the practice of re-writing the wording of the published NC levels into pupil-friendly vocabulary might be. To investigate this we asked interviewees whether they did this. There was a range of responses, from the definite and unequivocal, to those who had not and would not.

This teacher was very definite:

> We don't share National Curriculum levels with the students because we have different kid-speak levels that the students use.

Whereas this teacher was having second thoughts about using the ones he had rewritten:

> I put National Curriculum levels into what I thought was pupil-speak, and we've been using them since when I started, we've tried to augment them every now and again by different things, when people get different ideas. But it's not good, and now I look at it again, and I think that's not what they were really getting out of the National Curriculum. But our kids understand what they were talking about. And now because we just had training days on the new National Curriculum, and now I've had to revisit it all again that five years on from when I first did it, and now I'm thinking I've been marking them all really harshly.

This teacher had done so to a limited extent, but was concerned about the effect that doing this might have:

> To an extent, yes. I'm rather allergic to doing this. I'd rather give them the statement and try to help them to understand it and that's what we do at several points in the year. So we look back at the various statements and say well, if somebody is at such and such a level I'd expect to see most of these things going on, and this is what these

things mean. And so the language that is used in the National Curriculum statements, we do tend to make that part of our teaching language as well. So that those statements are quite accessible. Having said that, there are certain bits of language, for example in level 3, you wouldn't necessarily expect a level 3 child, if there is such a thing, to understand all the terminology in level 3 statement. And so, if you're going to make somebody, help somebody understand that that's the level that they're working at, or working towards or whatever other terminology you want to use, you may have to interpret it.

For some respondents, the prevailing assessment culture in schools meant that teachers found themselves having to do this in order to meet pupil expectations of what was required:

> With the levels, I try and put them into pupil speak. The kids actually like to know what level they're working towards. At first I didn't really use the levels, I sort of gave them an indication, this is what I want it to be, this is like where I want it to be, it's going to be here. So I'm not actually, indicating what the levels were. But then, I was observing and just sort of saying, try actually using the levels with them, as soon as you mention the levels, especially Year 7's they get really, a lot more enthusiastic about stuff if they know they're going to, if they add this they get to a higher level . . .

Criterion referencing

Rewriting level statements into pupil-speak takes us into the territory of criterion referencing. In free-text responses which did not mention NC levels directly, what many teachers did discuss was criterion referencing. Taking this as a unit for analysis, Table 1 showed that 44% of respondents mentioned criterion-referencing of some sort, even if they did not employ the terminology directly. This seems to play an important part in the way teachers think about KS3 assessment.

We were keen to find about the sorts of criteria which were being employed, and so in interviews we asked teachers about this, and found that many of the criteria used were the rewritten pupil-speak NC levels discussed above. This point was made by one teacher:

> We have some things stuck on the wall . . . which has broken down what's said in the various levels into very short sentences, which were designed to be intelligible, and at times I would take certain parts of the statement, ones that refer to a particular topic, and would reword those slightly. But I think the further you get away from the actual wording, whether you like the wording or not, but the further away that you get from it, the more difficult it becomes to use them. Because if you just want to write it in your own words, then, but that's fine, if you want to have a set of criteria, but then why pretend that that's the National Curriculum?

The point made by this teacher is key. The rewritten NC level statements may well be criteria, but they are not the NC levels. Stand-alone criteria of this sort derived from NC levels are possibly a helpful way for teachers to address key aspects of teaching and learning, but as grading criteria they may be problematic. This is because a fundamental issue here is that many of these rewritten pupil-speak criteria were transposed into what might be termed behavioural objectives, such as 'can use an ostinato' or 'plays in time'.

This is problematic, as examples of assessment criteria provided seldom addressed issues of composing-ness. The shift from constructivist levels to their behavioural outworking in classrooms is clearly an issue of more than philosophical concern. This is a point which has been raised by Ofsted:

> . . . in the weakest work seen, they [NC levels] were being misused to assess isolated activities. In one lesson seen, for example, students were told: 'Level 3: clap a 3 beat ostinato; Level 4: maintain a 4 bar ostinato; Level 5: compose an ostinato.' This demonstrated a significant misunderstanding of the expectations inherent in the level description. (Ofsted, 2009, para 101)

What our investigation also reveals is that there is a question as to whether there might be a conflation in teachers' minds between criterion assessment and NC levels. The free-text responses indicate that for a number of teachers, the NC levels themselves are unproblematic examples of criterion-referenced assessments. This runs counter to observation of NC levels generally, that '. . . the level descriptions contain, in themselves, collections of varied attainments that have no necessary unity or coherence. It might be argued that this is a collection of descriptions, not of linked performances' (Sainsbury & Sizmur, 1998, p. 190).

In criterion-referenced assessment we would normally expect a construct to be linked to an assessment criterion. What we found was that teachers were moving readily between atomistic task criteria based outcome assessment and the NC level statements. In one instance, the assessment criterion observed was: 'can play the keyboard melody with more than one finger'. This was equated with performance at NC level 4, the text of which is:

> Pupils identify and explore the relationship between sounds and how music reflects different intentions. While performing by ear and from notations, they maintain their own part with awareness of how the different parts fit together and the need to achieve an overall effect. They improvise melodic and rhythmic phrases as part of a group performance and compose by developing ideas within musical structures. They describe, compare and evaluate different kinds of music using an appropriate musical vocabulary. They suggest improvements to their own and others' work, commenting on how intentions have been achieved. (QCA, 2008)

There is no mention here of keyboard playing, nor of how keyboard playing equates to composing. In the school in question, the outworking of the composing process was enacted in assessment terms by a performance-related criterion, which in itself has marginal relevance to the NC level statement to which it is linked in the teacher's assessment. This has implications for the validity and reliability of NC assessments nationally.

Summative assessment

Summative assessment was the assessment modality which was mentioned least by respondents. When it was, the way summative assessment was described was that for many respondents a composing unit would end with an assessment lesson. In the assessment lesson pupils performed their composed pieces to the rest of the class, and the teacher

graded it. Criteria for this grading remained unclear. Typical comments include, from different teachers:

> The summative assessment is a final performance that takes place in front of the class.

> . . . in that lesson, we will have recorded onto camcorder the students performing their compositions, then we will watch the recording on the television. I will write a teacher assessment, they will do peer assessments as they watch, and they will also complete a self assessment of their own work, either during the lesson if there's time, or as a home work.

Interestingly what happens in many cases is that these assessment lessons seem to assess performance of the finished composition, rather than of the process of composing itself. The resultant grade seems to be based on performance of the composition, rather than the process of composing. We noted that the *process* of composing was not normally subject to summative assessment, but rather to formative assessment, and so it is to that we now turn.

Formative assessment

Table 1 showed formative assessment was only mentioned by 16% of respondents. In analysing responses involving it, many teachers did not feel the need to expand on what formative assessment involved. Responses such as this were unusual in that they provide examples of formative assessment in action:

> Most assessment experiences in my classroom are designed to be formative. I do very little formal marking of compositions as end products, but I give lots of feedback and suggestions to students on how they might reflect on their own compositions and what the next steps in developing them might be . . . Occasionally students ask me what Level they're at, and we do use the National Curriculum levels in helping us to assess progress . . .

The formative assessment of composing was more often described as 'informal' and 'regular', but we need more research to find out what is really going on in the formative assessment of composing. A number of teachers described things which could more properly be called the formative use of summative assessment (Black *et al.*, 2003a), rather than formative assessment *per se*. Our suspicion, therefore, for which more research is needed, is that what many music teachers are calling formative assessment is in fact the formative use of summative assessment, and this, whilst useful, is not true formative assessment.

Conclusion

NC levels are a summative assessment tool, designed for use at a single point in the education system, at the end of a key stage. Possibly because no other tool exists, they have been subverted for use in ways for which they were never intended. Teachers are trying to assess composing using NC levels, a tool which was never intended for the task. As the folk saying goes, 'when the only tool you have is a hammer, everything looks like a nail'. Music teachers are busy banging at everything with the hammer of NC levels.

There are seven main issues with regard to assessment of composing at KS3 that arise from our research:

1. NC levels are the most frequent modality of composing assessment at KS3.
2. NC levels alone are not sufficient to show progress . . .
3. . . . but, possibly, because they exist, they are frequently used for this purpose.
4. The language of NC levels is problematic for teachers to use in the classroom, and many have rewritten them into 'pupil-speak'. This creates inconsistency between schools.
5. Many schools have had to invent sub-levels. This again leads to inconsistency, and teachers are unsure as to what these mean in practice.
6. Teachers are using NC levels because they need to, rather than because they are considered useful.
7. The role of formative assessment in the assessment of composing is not clear.

Teachers are having to provide assessment data for school systems with what some see as excessive frequency. This coupled with a drive towards linear progression for all is causing problems for some teachers, whilst others are taking a cavalier attitude, and producing assessment data of limited validity. This is a key area of concern. Whilst we recognise that standards do need to be addressed, we are concerned at the amount of time from music lessons which is being taken to fulfil the systemic needs of assessment data collection which do not help learning directly.

References

ADAMS, P. (2000) Assessment in the music classroom. In C. Philpott (Ed.), *Learning to Teach Music in the Secondary School*. 1st Edition. London: RoutledgeFalmer.

Assessment Reform Group (1999) *Assessment for Learning – Beyond the Black Box*. Cambridge: University of Cambridge School of Education.

Assessment Reform Group (2002) *Assessment for Learning: 10 Principles*. ARG. http://k1.ioe.ac.uk/tlrp/arg/publications.html

BALL, S. J. (2003) The teacher's soul and the terrors of performativity. *Journal of Education Policy*, **18**, 215–228.

BERKLEY, R. (2001) Why is teaching composing so challenging? A survey of classroom observation and teachers' opinions. *British Journal of Music Education*, **18**, 119–138.

BLACK, P. (1995) Can teachers use assessment to improve learning. *British Journal of Curriculum and Assessment*, **5** (2), 7–11.

BLACK, P. & WILIAM, D. (1998) *Inside the Black Box: Raising Standards Through Classroom Assessment*. London: School of Education, King's College.

BLACK, P. & WILIAM, D. (2006) Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and Learning* (pp. 9–25). London: Sage.

BLACK, P., HARRISON, C., LEE, C., MARSHALL, B. & WILIAM, D. (2003a) *Assessment for Learning: Putting it into Practice*. Maidenhead: Open University Press.

BLACK, P., HARRISON, C., LEE, C., MARSHALL, B. & WILIAM, D. (2003b) The nature and value of formative assessment for learning. *Improving Schools*, **6** (3), 7–22.

BLACK, P., HARRISON, C., LEE, C., MARSHALL, B. & WILIAM, D. (2004) Working inside the Black Box. *Phi Delta Kappan*, **86** (1), 8–21.

BOUD, D. (2000) Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, **22**, 151–167.

BRAY, D. (2000) *Teaching Music in the Secondary School*. Oxford: Heinemann Educational.

BRAY, D. (2002) Assessment in music education. In G. Spruce (Ed.), *Aspects of Teaching Secondary Music* (pp. 79–93). London: RoutledgeFalmer.

BROPHY, T. S. (2000) *Assessing the Developing Child Musician: A Guide for General Music Teachers*. Chicago: GIA Publications.

BURLAND, K. & DAVIDSON, J. W. (2001) Investigating social processes in group musical composition. *Research Studies in Music Education*, **16**, 46–56.

BURNARD, P. (2000a) Examining experiential differences between improvisation and composition in children's music-making. *British Journal of Music Education*, **17**, 227–245.

BURNARD, P. (2000b) How children ascribe meaning to improvisation and composition. *Music Education Research*, **2**, 7–23.

BURNARD, P. (2002) Investigating children's meaning-making and the emergence of musical interaction in group improvisation. *British Journal of Music Education*, **19**, 157–172.

BURNARD, P. (2006) The individual and social worlds of children's musical creativity. In C. McPherson (Ed.), *The Child As Musician: A Handbook of Musical Development* (pp. 353–374). Oxford: Oxford University Press.

BURNARD, P. & YOUNKER, B. A. (2002) Mapping pathways: fostering creativity in composition. *Music Education Research*, **4**, 245–261.

BURNARD, P. & YOUNKER, B. A. (2004) Problem-solving and creativity: insights from students' individual composing pathways. *International Journal of Music Education*, **22**, 59–76.

BURNARD, P. & YOUNKER, B. A. (2008) Investigating children's musical interactions within the activities systems of group composing and arranging: an application of Engeström's Activity Theory. *International Journal of Educational Research*, **47**, 60–74.

BYRNE, C. & SHERIDAN, M. (2001) The SCARLATTI papers: development of an action research project in music. *British Journal of Music Education*, **18**, 173–185.

BYRNE, C., MACDONALD, R. A. R. & CARLTON, L. (2003) Assessing creativity in musical compositions: the importance of flow. *British Journal of Music Education*, **20**, 277–290.

COHEN, L., MANION, L. & MORRISON, K. (2007) *Research Methods in Education*. 6th edition. London: RoutledgeFalmer.

COLWELL, R. (2002) Assessment's potential in music education. In R. Colwell & C. P. Richardson (Eds), *The New Handbook of Research on Music Teaching and Learning: a Project of the Music Educators National Conference*. Oxford: Oxford University Press.

COLWELL, R. (2007) Music assessment in an increasingly politicized, accountability educational environment. In T. S. Brophy (Ed.), *Assessment in Music Education. Integrating Theory and Practice. Proceedings of the 2007 Symposium on Assessment in Music Education. University of Florida*. Chicago: GIA Publications.

CRAFT, A. (2003) The limits to creativity in education: dilemmas for the educator. *British Journal of Educational Studies*, **51**, 113–127.

CRAFT, A. (2005) *Creativity in Schools – Tensions and Dilemmas*. Abingdon: Routledge.

CRAFT, A., CREMIN, T., BURNARD, P. & CHAPPELL, K. (2007) Teacher stance in creative learning: a study of progression. *Thinking Skills and Creativity*, **2**, 136–147.

CRESWELL, J. W. (2009) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 3rd Edition. Thousand Oaks: Sage.

DENSCOMBE, M. (2007) *The Good Research Guide: For Small-Scale Social Research Projects*. Milton Keynes: Open University Press.

DfES (2002) *Training Materials for the Foundation Subjects*. London: DfES.

DfES (2003) *Key Messages – Pedagogy and Practice*. London: DfES.

DfES (2004a) *Assessment for Learning: Guidance for Senior Leaders*. Norwich: HMSO.

DfES (2004b) *Assessment for Learning: Whole School Training Materials*. Norwich: HMSO.

FAUTLEY, M. (2004) Teacher intervention strategies in the composing processes of lower secondary school students. *International Journal of Music Education*, **22**, 201–218.

FAUTLEY, M. (2005) A new model of the group composing process of lower secondary school students. *Music Education Research*, **7**, 39–57.

FAUTLEY, M. (2008) *Assessment in Music Education – Questions and Answers*. Matlock: NAME (National Association of Music Educators).

FAUTLEY, M. (2009) Assessment for learning in music. In J. Evans & C. Philpott (Eds), *A Practical Guide to Teaching Music in the Secondary School* (pp. 63–72). Abingdon: Routledge.

FAUTLEY, M. (2010) *Assessment in Music Education*. Oxford: Oxford University Press.

FELDHUSEN, J. F. & GOH, B. E. (1995) Assessing and accessing creativity: an integrative review of theory, research, and development. *Creativity Research Journal*, **8**, 231–247.

FOWLER, A. (2008) Assessment – A view from the classroom. *NAME (National Association of Music Educators) Journal*, **23**, 10–12.

GLASER, B. & STRAUSS, A. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine de Gruyter.

GORARD, S. & TAYLOR, C. (2004) *Combining Methods in Educational and Social Research*. Milton Keynes: Open University Press.

HARLEN, W. (2005) Teachers' summative practices and assessment for learning – tensions and synergies. *The Curriculum Journal*, **16**, 207–223.

HARLEN, W. (2007) *Assessment of Learning*. London: Sage.

JAMES, M. (1998) *Using Assessment for School Improvement*. Oxford: Heinemann Educational.

JEFFREY, B. (2006) Creative teaching and learning: towards a common discourse and practice. *Cambridge Journal of Education*, **36**, 399–414.

JOHNSON, R. & ONWUEGBUZIE, A. (2004) Mixed methods research: a research paradigm whose time has come. *Educational Researcher*, **33** (7), 14–26.

KASCHUB, M. & SMITH, J. (2009) *Minds on Music*. Lanham: Rowman & Littlefield.

MACDONALD, R., MIELL, D. & MITCHELL, L. (2002) An investigation of children's musical collaborations: the effects of friendship and age. *Psychology of Music*, **30**, 148–163.

MACDONALD, R., BYRNE, C. & CARLTON, L. (2006) Creativity and flow in musical composition: an empirical investigation. *Psychology of Music*, **34**, 292–306.

MAJOR, A. (2007) Talking about composing in secondary school music lessons. *British Journal of Music Education*, **24**, 165–178.

MAJOR, A. (2008) Appraising composing in secondary-school music lessons. *Music Education Research*, **10**, 307–319.

MANSELL, W., JAMES, M. & Assessment Reform Group (2009) *Assessment in Schools. Fit for purpose? A Commentary by the Teaching and Learning Research Programme*. London: Economic and Social Research Council, Teaching and Learning Research Programme.

MARSHALL, C. & ROSSMAN, G. B. (2006) *Designing Qualitative Research*. Thousand Oaks: Sage.

MIELL, D. & MACDONALD, R. (2000) Children's creative collaborations: the importance of friendship when working together on a musical composition. *Social Development*, **9**, 348–369.

MILLS, J. (1991) Assessing musical performance musically. *Educational Studies*, **17**, 173–181.

MURPHY, R. (2007). Harmonising assessment and music in the classroom. In L. Bresler (Ed.), *International Handbook of Research in Arts Education* (Vol. 16, Part 1, pp. 361–380). Dordrecht: Springer.

MURPHY, R. & ESPELAND, M. (2007) Making connections in assessment and evaluation in arts education. In L. Bresler (Ed.), *International Handbook of Research in Arts Education* (Vol. 16, Part 1, pp. 337–446). Dordrecht: Springer.

NEWTON, P. E. (2007) Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy and Practice*, **14**, 149–170.

ODAM, G. (2000) Teaching composing in secondary schools: the creative dream. *British Journal of Music Education*, **17**, 109–127.

Ofsted (2009) *Making More of Music*. London: Ofsted.

PAYNTER, J. (2000) Making progress with composing. *British Journal of Music Education*, **17**, 5–31.

PUNCH, K. (2009) *Introduction to Research Methods in Education*. London: Sage.

QCA (2007) *Music – Programme of Study: Key Stage 3, QCA*. www.qca.org.uk/curriculum.

QCA (2008) *Music Key Stage 3*. http://curriculum.qca.org.uk/key-stages-3-and-4/subjects/music/index.aspx.

RUTHVEN, K. (1995) Beyond common sense: reconceptualizing National Curriculum assessment. *Curriculum Journal*, **6**, 5–28.

SAINSBURY, M. & SIZMUR, S. (1998) Level descriptions in the national curriculum: what kind of criterion referencing is this? *Oxford Review of Education*, **24**, 181–193.

STEPHENS, J. (2003) Imagination in education: strategies and models in the teaching and assessment of composition. In M. Hickey (Ed.), *Why and How to Teach Music Composition: A New Horizon for Music Education* (pp. 113–138). Reston: MENC.

STRAUSS, A. & CORBIN, J. (1990) *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*: Newbury Park: Sage.

WILIAM, D. (2003) National curriculum assessment: how to make it better. *Research Papers in Education*, **18**, 129–136.

WRIGHT, D. (2003) Making friends with your data: improving how statistics are conducted and reported. *British Journal of Educational Psychology*, **73**, 123–136.