

Assessment issues within national curriculum music in the lower secondary school in England

Martin Fautley, Birmingham City University, UK

Abstract

This paper reports on a meta-analysis of two pieces of research into assessment in generalist class music teaching in England in the lower secondary school. It finds that there are five main issues with the ways in which assessment is employed appertaining to National Curriculum levels. These issues are: Frequency of use of National Curriculum levels; Subdivision of the levels; Linear progression; Assessment getting in the way of music making; and Constructivist and Behaviorist confusions. It finds that schoolteachers are being put under pressure from school leadership to change the results from teacher assessments in order to fit with school-based data prediction, and that the level statements of the National Curriculum are being used in ways for which they were not originally intended.

This paper investigates issues with assessment in England in the compulsory National Curriculum subject of music, as taught in the lower secondary school. The role of assessment in music education in the English situation has not as yet received a great deal of academic attention. There are general overviews (Stephens 2003; MacDonald, Byrne, and Carlton 2006) and some researchers have investigated teacher practice (Byrne, MacDonald, and Carlton 2003; Byrne and Sheridan 2001). In their discussions concerning the composing pathways undertaken by individuals, Burnard and Younker (2004) mention assessment, and Mellor (2000) discusses the views of pupils in composing and assessment. Mills (1991) investigated the musical nature of assessment. From a classroom perspective, Bray (2000, 2002) and Adams (2001) discuss ways in which teachers can undertake assessment in the classroom, whilst Fautley (2008, 2009) has provided guidance for both serving and pre-service teachers, as well as considering assessment issues in music education in the UK more widely (Fautley 2010). From a more general perspective, issues concerned with assessment and the arts have been discussed, (Murphy and Espeland 2007), as have the more specific issues of assessment in music education (Murphy 2007).

Music is a statutory subject in the English¹ National Curriculum for all pupils up to the age of 14. The sorts of lessons that result are what might be termed 'generalist music', and this is taught as a compulsory subject to all students, aged 11 through 14, in the lower secondary school in England. In the UK context this period of formal education is known as Key Stage 3 (KS3). The

¹ England is specified here as each member state of the UK has its own separate and different curriculum. This study is solely concerned with England, although similar issues have been noted elsewhere.

Part 3: Assessment in National, Regional, and State Contexts

statutory program of study which governs this in the English National Curriculum is set out in a curriculum document (QCA 2007). The content of the National Curriculum can be perceived as being eclectic to outside views. All pupils must participate in composing, performing and listening activities as part of their generalist music education. This is often done using classroom instruments, and ICT, keyboards, and guitars. It should also be noted that there is no widespread tradition of programs such as Orff or Kodaly in the UK. The National Curriculum states that pupils should acquire:

knowledge, skills and understanding through the integration of performing, composing and listening ... Performance, composing and listening are interrelated. Pupils should be encouraged, for example, to develop listening skills through performance and composition activities. Knowledge, skills and understanding in each of these areas should be developed interactively through practical music-making. (QCA 2007).

Assessment of pupil attainment in the program of study is done by individual teacher assessment, wherein each student is awarded a National Curriculum level. These levels are defined by level statements, which range from level 1 to level 8, with a further level also being available for exceptional performance. The level statements are arranged such that level 5 is deemed to indicate the average attainment for a 14-year-old pupil. The levels are given by teacher judgment alone; there are no set tests for pupils. Teachers decide on the 'best fit' of the level statements depending on the work the pupils have been doing during the three years of the key stage. There is no external moderation of the awarding of the levels. This leveling process has received a number of critical commentaries (Adams 2001; Fautley 2010; OFSTED 2003a, 2003b; Spruce 2001a).

Many problems have arisen with the use of National Curriculum levels and the ways in which they are awarded. The original intention of the level statements was that they were to be used solely at the end of the key stage, and that development between levels would be monitored by teachers in order to demonstrate that reasonable progress was being made. However, this requirement was soon overtaken by the demands and requirements of schools and what they felt were the demands of what was commonly perceived as being a punitive inspection regime, Ofsted (Office for Standards in Education). These requirements came in the fullness of time to involve awarding National Curriculum (NC) levels to individual pieces of work, so that school leadership teams (SLTs) could monitor progression across all subjects, and music was no exception

The problematic nature of this can be found in the nature of the level statements themselves. As an example, here is the wording of the level statement for level 5, the level which is average performance for a student at the end of the key stage:

Level 5

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

Music Assessment across Cultures and Continents: The Culture of Shared Practice

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)

This attainment statement is typical of the wording of the NC levels. As a general end of key stage statement, for which it was intended, it can be seen to have a use and purpose. However, in order to be able to use level statements on an individual piece of work of basis, many teachers were encouraged by their schools to rewrite the level statements into 'pupil-speak'. These rewritten level statements were then used to provide assessment data for work observed by the pupils on a day-to-day basis.

As there are only 8 levels (plus exceptional performance), many schools found that the NC levels were not fine-grained enough to show progression, and so they rewrote the levels again, dividing them into various numbers of subdivisions, three being a common number, but ten also being noted in some places. Again, these strayed somewhat from the original intention and wording of the levels.

It is against this background that two related pieces of research were undertaken, of which a synthesis view is reported on here. Both pieces of research involved similar methods, both were conducted on a nationwide basis, and involved an e-survey, with subsequent follow-up interviews being undertaken. The significant difference between the two was that in study A a second more refined e-survey was undertaken.

Participant details are:

Study A

Survey 1: 94 respondents

Survey 2: 34 Respondents

Follow-up interviews: 11

Study B

Survey: 57 respondents

Follow-up interviews: 6

Methodology

This research study combined both qualitative and quantitative methodologies. This was undertaken purposively, bearing in mind the notion that "by mixing the datasets, the researcher provides a better understanding than if either dataset had been used alone" (Cresswell and Plano Clark 2007). It is also important to note that this was very much a "...research approach that is 'joined up', where what we do in one part of the investigation affects how we proceed in another" (Newby 2009).

Methods

The initial on-line survey of study A was advertised widely using the NAME (National Association of Music Educators) website, and by direct e-mail to English schools. This resulted in 844 viewings of the survey, of whom 176 started the survey, and 94 completed it. During the course of this initial survey, respondents were invited to participate in a follow-up survey, also

Part 3: Assessment in National, Regional, and State Contexts

undertaken on-line. Those who indicated a willingness were forwarded a link to the survey. During this second survey teachers were asked if they would consider being interviewed, and from those who volunteered, a list of eleven names was drawn up, to cover teachers from a broad spread of geographical and socio-economic areas across England. 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, Manion, and Morrison 2007; Denscombe 2007). In reporting speech from teachers, what they said has been transcribed verbatim and conventionally represented used punctuation to aid meaning (Marshall and Rossman 2006).

Study B, another on-line survey, was advertised on the NAME website, and in the forums of teachingmusic.org.uk, a recently established on-line hub for music educators. It was completed by 57 respondents. Again, respondents were invited to self-identify for follow-up interviews, and here a short list of six was drawn up, this was a more geographically focused list, and was confined mainly to the English Midlands for ease of travel purposes.

The two pieces of research differ in intentionality. Study A was looking at the assessment *practices* of teachers, focusing specifically on composing. Study B was concerned with the external requirements of assessment upon music teachers, and, importantly, how they viewed those requirements.

Data analysis was undertaken in a number of ways. Free text responses were coded using a modified grounded theory approach (Glaser and Strauss 1967) where codings of 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 and Corbin 1990). Items were deemed significant when mentioned by a number of respondents, and free-text responses were repeatedly revisited to try to ensure that the importance of what was being said was suitably captured.

Quantitative data were analyzed in a number of ways. Nominal scales of denotation were employed in order to establish categories of responses. Likert scales for attitudinal responses were used, and had inferences drawn from them. Ordinal scales were also employed, for example to find how often NC levels were required. Two online survey packages were used; study A utilized QuestionPro, while study B used the Bristol on-line survey tool. Both employ similar heuristics in the calculation of responses.

Issues arising from the research

From the research five key issues were identified, which are discussed here. They are:

- Issue 1: Frequency of use of NC levels
- Issue 2: Subdividing the levels
- Issue 3: Linear progression
- Issue 4: Assessment getting in the way of music making
- Issue 5: Constructivist and Behaviorist confusions

Music Assessment across Cultures and Continents: The Culture of Shared Practice

Issue 1: Frequency of use of NC levels

A common theme throughout both of the studies was the ubiquitous nature of the utilization of NC levels. Despite running counter to the intentions of the original NC, 25% of teachers in study A said they used them to level individual pieces of work, and 49% said they used them at least once per term. In study B 42% of teachers used them at least half-termly, with a similar number using them termly. Study B therefore shows a higher figure than study A, with 84% of teachers saying they use the NC levels at least termly. The highest number of uses was from a teacher who had to provide NC level data for each individual child 9 times per year, working out to 3 times per term. The reasons for this frequency of use usually given by respondents was that it was a requirement of the Senior Leadership Team (SLT) of the school, and was used for data tracking purposes. This issue is discussed further in issue 3 below.

It is an often cited commonplace observance, that “when the only tool you have is a hammer, everything looks like a nail”. This seems to be the case with regard to both employment and utilization of NC levels in this research. In study A, teachers were asked what sorts of assessment modalities they employed. (The responses were grouped by areas mentioned, so totals do not add up to 100%)

Table A
Teacher use of assessment modalities

NC levels	49%
Formative Assessment	16%
Summative Assessment	11%
Criterion Referencing	44%

The use of the terminology ‘NC levels’ was not problematized in any way at all when referring to the ways in which they undertook assessment activities with their pupils. For these teachers, the ‘hammer’ of NC levels was the sole tool used for assessment. The relative infrequency of use of the terminology ‘formative assessment’ was noteworthy, as in England a lot of money and professional development training has been expended on developing formative assessment strategies for teachers (*inter alia* DfES 2002, 2003, 2004). Indeed, in the national training materials for teachers of all subjects it was a music lesson which was chosen to exemplify good formative assessment practice (DfES 2002).

The terminology ‘criterion referencing’ was not employed by the teachers in many cases, but this is what they described. Indeed, an interesting point here is that these criteria themselves were often re-written NC levels. One teacher alone noted that this was the case:

I think the further you get away from the actual wording, whether you like the wording or not ... the more difficult it becomes to use them. Because if you just want to write it in your own words ... that's fine, if you want to have a set of criteria, but then why pretend that that's the National Curriculum? (Teacher, Study A)

Part 3: Assessment in National, Regional, and State Contexts

Issue 2: Subdividing the levels

The level statements in the statutory National Curriculum are discrete and separated. For schools, however, the fine-tuning of attainment cannot be shown easily by the levels as they stand, and so many schools have chosen to adopt subdivided levels in an effort to show progress. In Study B this was true for 96.5% of respondents. The teachers who used these subdivided levels reported that their schools *require* them to do this. The purpose of this sub-division was always said by respondents to be so that SLTs could more finely data-track attainment of the pupils. All the teachers in study B who used subdivided levels used a 3-point scale; however, within this there was no commonality, for some 4a was the highest, for others 4c was, whilst others still converted them to spot-point decimal fractions, such as 4.3. As subdivided levels have no basis in statute, teachers were asked in study B where the subdivided level statements came from. 44% of respondents made them up themselves, with a further 28% of teachers being told to use the 3-point shading to distinguish between them. 22% of teachers worked collaboratively with other music teachers in the same school to write them, whilst only 3.7% of school music departments worked collaboratively with other schools to do this.

The three-point subdivision which the majority of teachers used was often expressed as a variant on the notion of three phrases deriving from early NC terminology:

- or working towards, or basic attainment
- = or working at, or attainment at this level
- + or working beyond, or secure attainment at this level

Some enterprising schools reportedly subdivide into ten, thus giving tenths of a level, this was seen as very problematic, and is not common.

Issue 3: Linear progression

An area that worried a significant minority of respondents was that despite having to use subdivided levels to level individual pieces of work, and to report these level results regularly, pupils' progression *had* to be shown to be linear. In study B the question was asked, "Are you required to show that your pupils have made at least a specified amount of progress using the NC levels (and sub-divisions, if appropriate), or are you free to use your professional discretion as to how much progress has been made by individual pupils?" 58% of respondents said they could use professional discretion, whilst 41% had to show linear progress. Interviewed about this, one teacher said: "I thought I was free to use my professional discretion but at the end of the key stage was told to change the levels to meet the percentage target." (Teacher, Study B)

The amount of required progress varied between schools, but in study B the modal answer was progression was expected to be one NC level per academic year. However, other schools required this in sub-levels and in a smaller time-frame, so two sub-levels per term was a common answer.

Music is deemed to be a creative subject, and many teachers complained about prescription of progress. Some teachers voiced concern about consequences for themselves if they failed to show that their pupils had demonstrated the linear progress required. One observed

Music Assessment across Cultures and Continents: The Culture of Shared Practice

that the result was a “carpeting by the line manager” and told her “schemes of work are not engaging enough” as her pupils had failed to progress. Another teacher observed “I get b*ll*cked” (sic) which was a refreshingly honest response! Other teachers have it placed as a performance management (a formal review of the individual teacher’s performance conducted by the school) issue. This teacher describes his concerns:

*When your professional judgment of a pupils' musical ability is repeatedly questioned, when you have done your best to enable pupils to reach their targets and they have failed but you are told over and over that you must do more to ensure they reach it then it is difficult to see the levels as anything but a big stick being repeatedly used to beat us with.
(Teacher, study B)*

One of the many problems in using the level descriptors to track pupil performance, and then hold the teachers accountable for any shortfall, is that this is twisting the utility of the levels beyond that which is credible. All of the English NC levels present this problem, and as Sainsbury & Sizmur (1998) note in this regard:

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, but rather of a typical pupil working at that level. But why should this collection of performances be typical of such a pupil? The answer is that this is a pupil who has been following the programmes of study of the National Curriculum. By teaching the programmes of study, teachers are to impose order upon the attainment targets. (Sainsbury and Sizmur 1998)

One of the implications of this for music educators is that a curriculum model becomes tacitly espoused wherein assessment leads the taught program, so that teachers can, ‘impose order’. This results in the pedagogic model becoming that which is shown in figure 1.

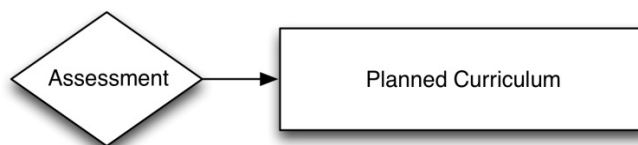


Figure 1. Assessment leads curriculum

The model represented in figure 1 can be seen to be an example of what Popham (2011) calls “assessment-influenced instruction”.

The notion of linear progression in any subject, and, music teachers argue, especially in a subject like music, runs counter to much current educational thinking. As Gary Spruce observes, “The linear model presumes predictable and common stages of development and ignores children’s social and cultural backgrounds which so affect their perception of what music is and *means* to them” (Spruce 2001b, p.20). The outworking of this can be seen in the classroom, where an individual pupil may be highly motivated by, say, a composing project involving

Part 3: Assessment in National, Regional, and State Contexts

songwriting, but much less so by a unit of work based on the Viennese waltz. This hypothetical pupil's attainment level, and the resultant teacher-awarded NC level, is likely to dip between these two units. The relentlessly upward linear progression demanded by SLTs will be difficult to provide here, and teachers will be in trouble. However, one of the findings of this synthesis study is that an unintended consequence of the requirement to demonstrate progression in a linear fashion has been that with the lack of external moderation, teachers simply fabricate level information so as to satisfy demand! These comments represent just a few of the many: "I have to say that we're playing the game [fabricating levels] because we've been met with a brick wall so many times it's not worth it." (Teacher, study A)

Interesting to see people talking about putting them [levels] up just to keep data managers happy. I and my colleagues have done that for years and no member of the Senior team has ever questioned me. I think it's an old, tired system that isn't suited to the more fluid nature of progression that students make within our subject. (Teacher, study B)

Whilst one teacher was very open about what they did: "...to be honest, I make them up as I go along. I have no respect for this system." (Teacher, study A) Which makes the comments of another teacher appropriate here: "NC Levels are so open to abuse." (Teacher, study B)

What did emerge from study B was that for many teachers supplying data had taken the place of knowledge of the pupils. The drive to improve educational practices has given rise to what has become known as the *standards agenda*. The effect of this is to place performativity at the very heart of what teachers and schools do. The most visible manifestation of this is in the form of school league tables, which are published in England in local and national press, and compare school results directly one with another. What this has resulted in are situations where, as Ball puts it, teachers '... find their values challenged or displaced by the terrors of performativity' (Ball 2003). These are very real terrors, and teachers in these studies spoke of how they had been put under a great deal of pressure to ensure that the data was 'right'. One teacher observed: "A child is a [level] 4b and stays there even if absent for nearly 3 months. You have to show progression or lie! Sad world we live in." (Teacher, personal email) This assessment data is not helping pupils learn, nor is it raising standards. It is simply a way of keeping out of trouble! This is not how teachers ought, or want to be behaving.

Issue 4: Assessment getting in the way of music making

In study B teacher were asked to respond using a Likert scale to the statement 'KS3 assessment gets in the way of music making'. 36.8% strongly agreed with this statement, and 35.1% agreed somewhat. Taken together this means that over 70% of teachers have some sympathy with this as a problem. As curriculum time in KS3 is limited, often around 50 minutes a week (although music is a statutory subject, there is no statutory amount of time it should have), the fact that assessment requirements are hindering music making is a significant issue. The comments of this teacher are representative: "Our curriculum is coming down to teaching only what is immediately measurable. True creativity is stifled. The concept of true education has gone out of the window..." (Teacher, study B)

Music Assessment across Cultures and Continents: The Culture of Shared Practice

This is a key issue. Many teachers described how the requirements of data analysis at a whole school level had led to a deterioration in the quality of music teaching and learning: “We are spending more time on assessing and have insufficient time to plan and prepare exciting resources and inspiring lessons.” (Teacher, study B)

Another teacher observed that “...used relentlessly, they stand in the way of music-making and developing social skills.” (Teacher, study B) This was a common cry. One teacher went so far as to observe that their SLT was little concerned with the musical nature of the work they did, so long as the assessment seemed right:

We are now totally driven by numbers on a page. No-one EVER asks to see or hear a kids' work, or shows any interest in what they had to do to gain a particular level, all they care about is whether the 'target level' has been attained or bettered. (Teacher, study B)

What this seems to mean is that a system which was designed to help raise attainment in music is having the opposite effect, and the assessment system is itself an inhibitor towards musical learning and classroom music making.

Issue 5: constructivist and behaviorist confusions

The wording of the level statements themselves is couched in what might be considered to be broadly constructivist terminologies. There are no defined outcome statements which are closed and definitive. The example of the wording of the level 5 statement above evidences this. What is often to be found in the classroom, however, is a reification of the levels rewritten in pupil-speak as what might be considered as behaviorist outcomes. This is evidenced in an observation by Ofsted, “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' (Ofsted 2009, p.31). In a classroom visited during study A, a poster on the wall informed students, “play keyboard with one hand = level 4, play keyboard with two hands = level 5”.

This issue is one which, although not articulated directly by the teachers in the research, is one which seems to cause a number of problems. The problem is described by Sainsbury and Sizmur:

With the level descriptions, a more unified approach to describing domains can now be seen, in that they display a consistent degree of complexity, in which specific and general, abstract and concrete are always mixed. This is, therefore, an attempt to represent the nature of the scaled construct itself, rather than trying to derive specific performance descriptions from it. (Sainsbury and Sizmur 1998, p. 191)

And music teachers in the UK have been doing just that – trying to ‘derive performance specific descriptions from it’; performance in both the musical and non-musical sense. Using the NC levels as stand-alone criteria to demonstrate development is concomitantly problematic, as Sainsbury and Sizmur (1998) also observe: “The definition of the level descriptions is indeed

Part 3: Assessment in National, Regional, and State Contexts

dependent upon external understandings, in the sense that they do not consist of unitary and precise performance descriptions, but rather are in need of further interpretation.” (p. 190)

The ‘further interpretation’ is too far removed from the specificity of an individual pupil’s performance to be used reliably to show progress at the micro-level, as many teachers observed, with comments such as this being typical:

...the school line is the line that obviously we sort of have to take but, in reality there are so many different facets to composing, performing, listening, that it’s actually very, very difficult to make accurate assessments; and plus the aesthetic and the creative elements that come on top of that. You know, one person playing ‘Twinkle Twinkle little Star’ and another person can play exactly the same notes and it could be a much more musical and much more creative performance and so, yes technically it shows development, but in reality I think it’s a grey area. (Teacher, study A)

Or, as another teacher observed:

The vagueness with which the level descriptors are written requires a fair amount of knowledge and experience to interpret. After 18 years in the job, I feel that I can make some judgment about the skills and abilities of students, but I still find it very difficult to explain this in a purely objective way (Teacher, study B)

Implications

The English National Curriculum levels are clearly a cause for concern, and this analysis shows widespread mistrust of the system, as well as some questionable practices being undertaken by teachers. It seems to be the case that requirements of the standards agenda have caused a significant misappropriation of National Curriculum levels, and that teachers are struggling to cope with the data collection required of them.

This study is clearly about a specific local issue in a single country, but there are wider lessons that can be learned from it. Measuring progression in generalist music classes requires the teacher to be free to use their professional skill and knowledge to make assessment judgments, formative judgments about what to do to help pupils improve, and summative ones about how well they have done this. Allying this to whole-school accountability issues inappropriately does not help the quality of music making. In England orchestras, bands, and choirs are run voluntarily by teachers outside of core work hours. Many teachers report they no longer have time to do this due to the pressure of performativity and supplying assessment data. Obviously no-one is against lowering standards, but when a system designed to raise standards ends up doing this, then questions need to be asked of such a system.

A further lesson for the international music education community is in the development of learning statements, assessment descriptions, and grade boundaries. If generalized ‘best-fit’ statements of the kind exemplified by the NC level statements are to be employed, then there should not be felt to be an imperative to infer outcome statements measurable in terms of a single behavior. Thus ‘understands cadences’ should not become ‘can play dominant triad followed by tonic on the keyboard’. What should be key in teaching and learning is that assessment should be

linked to learning. The situation in England has become one where assessment drives learning, and all learning is aimed at producing level information. If it cannot be leveled, it does not count. Assessment should not get in the way of music making. The assessment tail is wagging the curriculum dog.

References

- Adams, P. 2001. Assessment in the Music Classroom. In C. Philpott (ed.). *Learning to teach music in the Secondary School (1st Edition)*. London: RoutledgeFalmer.
- Ball, S.J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy* 18 (2):215-228.
- Bray, D. (2000). *Teaching Music in the Secondary School*. Oxford: Heinemann Educational.
- . 2002. Assessment in Music Education. In G. Spruce (ed.). *Aspects of Teaching Secondary Music*. London: RoutledgeFalmer.
- Burnard, P., and Younker, B. A.. 2004. Problem-solving and creativity: insights from students' individual composing pathways. *International Journal of Music Education* 22 (1):59-76.
- Byrne, C., Sheridan, M.. 2001. The SCARLATTI papers: development of an action research project in music. *British Journal of Music Education* 18 (2):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 (3):277-290.
- Cohen, L., Manion, L., Morrison, K. (2007). *Research methods in education*. 6th ed. London, New York: RoutledgeFalmer.
- Cresswell, J.W., Plano Clark, V. L. (2007). *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage.
- Denscombe, M. (2007). *The good research guide: for small-scale social research projects*. Maidenhead, Berkshire: Open University Press.
- DfES. (2002). *Training Materials for the Foundation Subjects*. London: DfES.
- . (2003)*Key Messages - Pedagogy and Practice*. http://www.standards.dfes.gov.uk/keystage3/downloads/ks3_km2003.pdf London: DfES.
- . 2004. *Assessment for Learning: Whole School Training Materials*. Norwich: HMSO.
- Fautley, M. 2008. Assessment in Music Education - Questions and Answers. Matlock, Derbyshire.: NAME (National Association of Music Educators).
- . 2009. Assessment for learning in music. In J. Evans and C. Philpott (ed) *A Practical Guide to Teaching Music in the Secondary School*., Abingdon: Routledge.
- . 2010. *Assessment in Music Education*. Oxford: Oxford University Press.
- Glaser, B , and A Strauss. 1967. *The discovery of grounded theory: strategies for qualitative research*. NY: Aldine de Gruyter.
- MacDonald, R., Byrne, C., Carlton, L. (2006). Creativity and flow in musical composition: an empirical investigation. *Psychology of Music*.
- Marshall, C., Rossman, G. B. (2006). *Designing qualitative research*. Thousand Oaks, CA: Sage.
- Mellor, L. (2000). Listening language and assessment: The pupils perspective. *British Journal of Music Education* 17 (3):247-263.
- Mills, J. (1991). Assessing musical performance musically. *Educational Studies* 17 (2):173-181.

Part 3: Assessment in National, Regional, and State Contexts

- Murphy, R. (2007). Harmonizing assessment and music in the classroom. In *International Handbook of Research in Arts Education*, edited by L. Bresler. Dordrecht, Netherlands: Springer.
- Murphy, R., Espeland, M. (2007). Making connections in assessment and evaluation in arts education. In *International Handbook of Research in Arts Education*, edited by L. Bresler. Dordrecht, Netherlands: Springer
- Newby, P. (2009). *Research Methods for Education*. Harlow, Essex: Pearson Education Ltd.
- Office for Standards in Education (2003a). Good assessment in secondary schools. London: Ofsted.
- — —. (2003b). Good assessment practice in music (HMI Document 1479). London: Ofsted.
- — —. 2009. Making more of music. London: Ofsted.
- Popham, WJ. (2011). *Classroom assessment: What teachers need to know (6th Ed)*. Boston MA: Pearson.
- QCA. (2007). Music: Programme of Study for Key Stage 3. Retrieved December 30, 2011 from <http://curriculum.qcda.gov.uk/key-stages-3-and-4/subjects/key-stage-3/music/index.aspx>.
- Sainsbury, M, Sizmur, S. (1998). Level Descriptions in the National Curriculum: what kind of criterion referencing is this? *Oxford Review of Education* 24 (2):181-193.
- Spruce, G. (2001a). Music assessment and the hegemony of musical heritage. In *Issues in music teaching*, edited by C. Philpott and C. Plummeridge. London: RoutledgeFalmer.
- — —. (2001b). Planning for music teaching and learning. In *Aspects of Teaching Secondary Music: Perspectives on Practice*, edited by G. Spruce. London: RoutledgeFalmer.
- Stephens, J. (2003). Imagination in education: Strategies and models in the teaching and assessment of composition. In *Why and how to teach music composition: a new horizon for music education.*, edited by M. Hickey. Reston, USA: MENC.
- Strauss, A , Corbin, J. (1990). *Basics of qualitative research: grounded theory procedures and techniques*. Newbury Park: Sage.