Baseline assessment of singing in English schools

Graham F. Welch & Julia Elsley
University of Surrey, Roehampton Institute

The statutory background to baseline assessment

In January 1996, the UK Government’s Secretary of State for Education and Employment, Gillian Shepherd, instructed the School Curriculum and Assessment Authority (SCAA) to undertake a preliminary survey of current practice and views on baseline assessment in primary schools and to prepare draft proposals for a formal consultation. About half of all English LEAs had already introduced some form of local baseline assessment, but each had developed its own scheme, founded on local educational priorities and agendas. The Government’s action was aimed at generating an overall national policy for all LEAs and so in September 1996, following its survey, SCAA sent out draft proposals to all schools, Local Education Authorities (LEAs), higher education establishments and other interested professional organisations.

In addition to the overt focus on profiling pupil development, the underlying political impetus for baseline assessment was (and is) a perceived need to generate a measure of how each school is ‘adding value’ to the basic academic attainment of pupils from the point of entry to compulsory schooling at age five. Such assessment became politically as well as educationally imperative, particularly for some inner-city LEAs, when national ‘league tables’ for schools and LEAs were introduced in the early nineties. The tables were based on pupil performance in national standard attainment tasks (SATs) at 7, 11 and 14 years and also in GCSEs and ‘A’ Levels at 16 and 18 years respectively. Such tables have been widely criticised by LEAs and university researchers (such as Mortimer, Sammons & Thomas, 1994; Goldstein & Spiegelhalter, 1996; Morrison & Cowan, 1996; Sammons, West & Hind, 1997; Strand, 1997; Gibson & Asthana, 1998) because they are based on raw performance data that takes no account of pupil background. Baseline assessment was seen as a means of determining the effectiveness of a particular school by providing a starting point from which progress through Key Stage 1 (ages five to seven years) and older could be measured. In addition, from the pupils’ viewpoint, baseline assessment was projected to facilitate an early insight into each child’s knowledge, understanding, skills and attitudes so that the school curriculum could be better matched to their individual needs, thus promoting greater likelihood of effective schooling and pupil progress.
The Education Act of 1997 subsequently stated that all maintained (= state) primary schools had to adopt a baseline assessment scheme from 1 August 1998 (DfEE, 1998). Each scheme adopted had to be accredited by the Qualifications and Curriculum Authority (QCA, the successor to SCAA). However, there was to be no single national scheme, but instead each local scheme had to include as a minimum the basic skills of language and literacy (speaking and listening, reading, writing), mathematics and personal and social development (PSD).

In addition, the previous Government had set out a group of educational goals for children to reach on entry to statutory education in the term after their fifth birthday. These ‘Desirable Learning Outcomes’ (SCAA, 1996) were an outcomes-based curriculum statement for pre-school provision (nurseries, kindergarten) that embraced language and literacy, mathematics and PSD (key elements in the subsequent 1998 minimalist baseline requirement), plus knowledge and understanding of the world, physical development and creative development. In 1997, the incoming Labour Government asked the QCA to review these ‘Desirable Learning Outcomes’ alongside the QCA review of the 1995 National Curriculum for schools, seeking to clarify the aims of education for children aged three to five years and to establish universal pre-school education as a ‘Foundation Stage’ in 2000. Thus successive governments had initiated a wide range of (not always complimentary) statutory frameworks with the intention of increasing the quality (and political profile) of pre-school education.

Baseline assessment and music

An independent study carried out in the Autumn term of 1997 for QCA revealed (not surprisingly) much diversity of practice amongst the baseline schemes that they had accredited, but very little emphasis on creative development, even though this was one of SCAA’s original foci in their 1996 consultation, one of the 1996 ‘Desirable Learning Outcomes’ and is also one of the suggested components of the proposed pre-school ‘Foundation Stage’ of educational provision (QCA, 1999[a]). Only a very small minority of schemes have incorporated other areas of children’s development outside the basic core of language and literacy, mathematics and PSD. The predominant additional area is physical development. Of the 90 baseline schemes reviewed in the summer of 1998, only 15 identify creative development. Even where creative development is mentioned, the teacher record is normally at most a very brief comment set alongside a box to be ticked against ‘always, sometimes, never’.

Music is nearly always packaged under a generic arts concept. For example, in one case the assessment wording for music is ‘responds to music through dance’. In another, ‘Explores sound and colour, texture, shape, form and space in 2D and 3D’ (following the SCAA, 1996 guidance). Singing (or the use of voice) is rarely differentiated in the text from other kinds of musical activity. Where singing is mentioned, it is in a general, non-specific, and uncritical fashion that bears no relation to the published research literature (see below). The latest official example is in the wording of the QCA ‘Foundation Stage’ proposals for 2000 that states: ‘By the end of the reception year, most children should be
able to express and communicate their ideas, thoughts and feelings by using ... a variety of songs...' (QCA, 1999[a], p12).

This lack of emphasis on creativity and the arts (including music) in official documents mirrors the similar overall lack of priority within the statutory curriculum. Up until September 1998, music was a statutory requirement for English schools from the ages of five to thirteen years (Key Stages 1 to 3). However, because of greater emphases given to literacy and numeracy by the new Labour Government, the statutory curriculum for music has now been relaxed for two consecutive school years (1998-2000). Schools are still expected to include music in their curricula, but school inspectors (OfSTED) are no longer required to report on the standards achieved. So it is difficult to see how a 'balanced' music curriculum that embraces 'performing (singing and playing), composing, listening and appraising' is currently being ensured for all pupils.

Furthermore, the National Curriculum in England and Wales is currently undergoing its third major revision since 1988, ready for implementation in September 2000. The 'work in progress' draft Order for Music (January 1999) has simplified the wording of the Programme of Study, but the basic elements from the earlier Orders are all still present. For Key Stage 1, the draft has a section on 'singing and playing', with the statements: 'Pupils should be taught how to control sounds made by the voice' and 'Over the key stage, pupils should be taught skills, knowledge and understanding through .... singing a variety of songs' (QCA, 1999[b], p2). Once again, there is no reference (implicit or explicit) to research literature on musical development in general, or to singing development in particular.

The quality of pre-school music teaching

With regard to the quality of pupils' musical experience in music, OfSTED school inspectors have recently reported that:

'Pupils achieve well and make good progress in nine out of twenty nursery and reception classes' [i.e. 45% of all lessons observed]. Lessons frequently focus on singing or repetitive use of simple percussion instruments and give pupils few opportunities to use their imagination..... Within individual schools, pupils' progress across the programmes of study is often uneven. Progress in singing is almost invariably slow when classes are combined, particularly if two or more year groups are taught together' (OfSTED, 1998, p1).

Similarly, Mills (HMI and specialist OfSTED adviser for music; 1998, p60) reports that two 'specific and persistent problems' are 'continuity and progression' and the need to exploit pupils' musical creativity and develop their musical imagination in the nursery and during Key Stage 1. Furthermore, 'good music teaching' is characterised by lessons having a 'clear teaching focus' that caters for the 'full range of pupils' musical development' that 'build[s] on prior learning' (Mills, op.cit. p84).

The reported variability between individual schools (Mills, op.cit. p60) is not confined to any particular age phase and is also evidenced in the pre-school inspection data. As an example, statistical analyses of the under-fives inspection data informing the
judgements contained in 47 published OfSTED school inspection reports on music, conducted by one of the authors [Elsley] from 1992-1999, provide a ‘snap-shot’ of the nature of this variation. OfSTED inspectors are required to rate the quality of observed music education under four headings, namely teaching, response, attainment and progress. Inspectors use a seven-point scale, with '1' being excellent and '7' poor. It is clear from the reported ratings across the 47 inspections that there is considerable variability between schools (ANOVA: $F[3,180] = 4.41; p = 0.005$). However, subsequent two-sample $F$-tests indicate that the quality of observed teaching is not necessarily related to pupils’ response ($F = 2.18; p = 0.005$), nor to progress ($F = 2.85; p = 0.003$). Although there is greater homogeneity between the ratings for teaching and perceived attainment, the difference between the ratings is still statistically significant ($F = 1.81; p = 0.02$). Further scrutiny of the raw data provides an explanation for these findings. Firstly, where the teaching was excellent or very good (rated 1 or 2), this did not necessarily mean that the pupils were making equivalent progress nor attaining at the highest levels. (To many (if not all) teachers, this observation would accord with common sense and their own craft knowledge. Pupils within the same class can vary widely in their current abilities, hence the recognised professional requirement for differentiated teaching.) However, where the teaching was rated less than satisfactory (rated as 5,6 or 7), in almost every case the response, attainment and progress were equally unsatisfactory. Arguably, an appropriate baseline profile could reduce the variability in the quality of pre-school music education provision by promoting more effective (focused, differentiated, developmental) teaching.

The research evidence on singing development in early childhood

The foundations for singing development are established during infancy because of the physical and structural interrelatedness between early infant vocalisation, infant-directed speech (also called ‘motherese’), infant-directed singing, and the existence (and persistence) of songs designed for an infant audience within the maternal culture. During these first few months of life, parents “consistently guide the infant towards vocal expertise in which canonical syllables appear and are treated by parents as ‘protowords’ to which meanings are attributed, being assigned in a declarative manner to the naming of persons, objects and events” (Papousek, 1996 p44-45). This pre-verbal development embraces an adult-child dialogue that is characterised by ‘rich melodic modifications’ during these first months of life (Papousek, op.cit. p48), with repetitions, frequent glides, a prevalence of basic harmonic intervals (3rds, 4ths, 5ths, 8ves) and sometimes dramatic changes in intensity.

Given the bipotentiality of early interactive vocalisation, it is not surprising to discover that the borders between singing and speech are often blurred, particularly for the young child (Davies, 1994; Davidson, 1994; Welch, 1994[a]; Sergeant, 1994; Welch, Sergeant and White, 1996). Such blurring is characteristic of one of the first phases of singing development for many young children (see Annex). For a small minority, this phase can persist into later childhood and result in singing that is labelled as ‘out-of-tune’ because, in the absence of useful information on how to consciously change vocal pitch, the
early vocal behaviour continues to be inappropriately repeated (and rehearsed) in response to singing tasks (Welch, 1985[a], 1985[b]; Kalmar, 1991; Welch (1994[b])). Nevertheless, caution is required in the use of this (or similar) label because ‘in-tune’ singing is culturally defined, perceptually experienced (through categoric perception) and task specific (see evidence below). Overall, early childhood singing development is sequential in character, with certain singing behaviours have primacy over others (Davidson, 1994; Welch, 1994[a]; Welch, Sergeant and White, 1996, 1997; Rutkowski, 1998).

Longitudinal research evidence from a sample of one hundred and eighty-four five-, six- and seven-year-olds in their first three years in London elementary schools (Welch, 1997; Welch et al, 1996, 1997) supports the key features of the sequential model (see Annex) whilst indicating that singing development is multi-faceted and complex, with differences emerging in relation to the children’s age, sex and singing task. Subjects were tested individually in three successive years on their current abilities to sing two songs and also on musical material that was ‘deconstructed’ from the songs and performed without words. This allowed a comparison to be made of the effects of adding words to musical material (such as in songs).

The longitudinal data reveal an explicit hierarchy in children’s developing singing competences. At each age, children were much more accurate in their vocal pitch matching when asked to copy simple glides (uni- or bi-directional, being equivalent to simple melodic contours), single pitches and (from age 6) melodic fragments, compared to complex glides (multi-directional contours) and songs. There was also clear evidence of a general, systematic and statistically significant improvement in pitch matching skills from school year to school year, with the exception of pitch matching in song singing, which was consistently judged as being much poorer by comparison (Welch et al, 1996). In each of the three years of testing, the children’s ability to reproduce the words of the chosen songs were rated extremely highly by the judges. Moreover, in each year, the ratings for word accuracy were significantly better than the ratings for pitch accuracy. In contrast, the ratings for melodic interval pitch accuracy showed no significant improvement until the third year of testing (age seven) and even in this third year, pitch accuracy was still rated as significantly worse than text accuracy (Welch, et al., 1998).

In general, the data indicate that children enter compulsory schooling with an ability to learn the words of songs. For many children, this linguistic competence would appear to be significantly in advance of their ability to learn the melodic contour and musical intervals of the same songs (evidenced by the opening year of this longitudinal study when the subjects were assessed in their first year of compulsory schooling and, in many cases, in their first few months). The evidence from this study is that children arrive at school ‘programmed’ to be responsive to words and perhaps ‘biased’ towards words when these are attached to specific melodic contours (as in songs). (Additional evidence of such bias may be found in the perceptual dominance of textual features when pre-schoolers’ invent notations of song (Barrett 1997, 1998).) Furthermore, in each year of the longitudinal study, the children scored more highly for their vocal pitch matching of melodic elements compared to melodies in songs (even though these were identical samples of pitches). Additionally, their general ability to vocally pitch match the musical elements improved across the three years, but there was no similar improvement in relation to song melodies.
The baseline assessment in singing for children entering compulsory schooling

Taking the findings of the longitudinal study (Welch, et al, 1997; 1998), especially in relation to the data on five-year-olds (Welch, et al, 1996), alongside those from other researchers, such as Davidson (1994), Rutkowski (1998) and Levinowitz et al (1998), the research literature indicates that any baseline profile of young children’s singing competences should incorporate a variety of musical tasks. These tasks essentially fall into two main categories; (i) ‘sub-song’ and (ii) ‘song’, with further sub-categories that reflect the constituent and related elements within the whole.

(i) sub-song:
   (a) pitch glides (glissandi)
      - uni-directional (descending or ascending)
      - bi-directional (ascending and descending, descending and ascending)
      - multi-dimensional
   (b) single pitches
   (c) simple pitch patterns (e.g. such as fixed point pitch equivalents of the glides)
      two notes
      three notes
      four notes
      five notes

(ii) song:
   (a) simple songs
      maintain tonality in constituent phrases
      maintain tonality in complete song
      match pitches in constituent phrases
      match pitches in complete song

This schematic is only focusing at present on the pitch aspects of singing. Rhythmic and text accuracies are further sub-categories that could also be assessed (whilst noting that the latter should be assessed separately in its own right, i.e. with regard to accuracy of text reproduction).

Furthermore, each of these particular tasks (‘sub-song’ and ‘song’) can be additionally differentiated in several ways, particularly in relation to pitch ranges. The research literature (cf Welch et al, Davidson, Rutkowski - see Annex) indicates that there are essentially four basic pitch ranges evidenced in samples of young children’s developing singing. These are (i) a ‘spoken’ pitch range (Rutkowski, 1998; Levinowitz et al, 1998), closely allied to the customary speech range (a3 to c4 [middle C]), (ii) a ‘limited’ vocal pitch range (Welch, 1979; Welch, et al. 1991; Welch, 1997) in which pitches are generally accurate within the limited range of a sixth (a3 to a4), linked to a dominance of the lower singing ‘register’ in vocal pitch matching (also termed ‘chest voice’, Wurgler,
1990), (iii) a ‘primary’ vocal pitch range (approximately a3 to c5), so-termed because there is evidence that this is the most common vocal pitch range reported in the literature for elementary school children as a group (Welch, 1979), and (iv) an ‘extended’ pitch range in which vocal pitch accuracy is evidenced beyond the lower singing register and into a higher register (a3 to e5, one and a half octaves) (Davidson, 1994).

When the four basic pitch ranges are combined with the elements of the ‘sub-song’ and ‘song’ categories, a baseline profile begins to emerge (see figure 1) that both celebrates the growing competences of the five-year-olds entering compulsory schooling in the UK and, at the same time, offers a guide to the teacher in the planning of appropriately differentiated activities to support future singing development.

At the very least, it is to be hoped that future music curriculum designers, whether at national or local level, will recognise the complexity of singing development in early childhood. To suggest that young children should be able to ‘express their ideas and communicate their ideas, thoughts and feelings by using .... a variety of songs’ (QCA, 1999[a], p12) without providing any hint of the developmental complexity, not only ignores the research data, but also continues to do young children a disservice. Such weak curriculum framing also perpetuates the likelihood of future generations of adult ‘non-singers’ because of negative experiences in their formative years. If designed appropriately, baseline assessment in singing could be a positive step forward in all young children’s musical development.

**figure 1: a schematic for baseline assessment in singing**

<table>
<thead>
<tr>
<th>vocal pitch ranges</th>
<th>spoken</th>
<th>limited</th>
<th>primary</th>
<th>extended</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a3-c4</td>
<td>a3-a4</td>
<td>a3-c5</td>
<td>a3-e5</td>
</tr>
</tbody>
</table>

(i) sub-song:
- (a) pitch glides (glissandi)
  - uni-directional
  - bi-directional
  - multi-dimensional
- (b) single pitches
- (c) simple pitch patterns
  - two notes
  - three notes
  - four notes
  - five notes

(ii) song:
- (a) simple songs
  - maintain tonality in constituent phrases
  - maintain tonality in complete song
  - match pitches in constituent phrases
  - match pitches in complete song

(iii) text:
- matches text in part (phrases), or whole (complete song)
Annex: a model of vocal pitch-matching development

Phase 1 The words of the song appear to be the initial centre of interest rather than the melody, singing is often described as 'chant-like', employing a restricted pitch range and melodic phrases. In infant vocal pitch exploration, descending patterns predominate (e.g. Young, 1971; Moog, 1976; Fox, 1982; Dowling, 1984; Goetze, 1985; Welch, 1986; Rutkowski, 1987, 1998; Ries, 1987; Levinowitz, 1989; Davies, 1992; Thurman and Klitzke, 1993; Davidson, 1994).

Phase 2 There is a growing awareness that vocal pitch can be a conscious process and that changes in vocal pitch are controllable. Sung melodic outline begins to follow the general (macro) contours of the target melody or key constituent phrases. Tonality is essentially phrase based. Self-invented and 'schematic' songs 'borrow' elements from the child's musical culture. Vocal pitch range used in 'song' singing expands from speaking range, through to a less restricted range but still essentially within one vocal register (e.g. Moog, 1976; Davidson et al, 1981; Fyk, 1985; Welch, 1986; Hargreaves, 1986; Rutkowski, 1987, 1998; Fujita, 1990; Minami and Umezawa, 1990; Welch et al, 1991; Davies, 1986, 1992, 1994; Davidson, 1994; White, Sergeant and Welch, 1996; Hargreaves, 1996; Papousek, 1996).

Phase 3 Melodic shape and intervals are mostly accurate, but some changes in tonality may occur, perhaps linked to inappropriate singing register usage. Overall, however, the number of different reference pitches is much reduced (e.g. Davidson et al, 1981; Dowling, 1984; Fyk, 1985; Welch, 1986; Wurgler, 1990).

Phase 4 No significant melodic or pitch errors in relation to relatively simple songs from the singer's musical culture (e.g. Davidson et al, 1981; Fyk, 1985; Welch, 1986; Rutkowski, 1987; Welch, 1994[a]).

Reference list


243


Endnotes

1 The existence of individual ‘pathways’ in development (Welch, 1998) permits some children to become relatively skilled in the dominant musical song genre(s) at a very early age. At present, there is insufficient evidence as to whether such children proceed through the developmental phases of the model much more quickly, whether some phases are missed out or collapsed together, or whether they experience an alternative developmental process. Davidson (1994 p114), for example, suggests that more developed singers have a wider variety of melodic contour schemes available at an earlier age. Nevertheless, the basic model draws on a wide range of existing research data and is thought to be applicable to the majority of young children. In addition to the London-based research (detailed in the chapter) other studies by Davidson (1994) and Rutkowski (1998) offer useful perspectives of the nature of the developmental sequence, each taking complimentary aspects.

2 The term ‘register’ is being used as a label for a particular configuration of the intrinsic laryngeal muscles. Where the vocalis predominates over the cricothyroid muscle, the vocal folds have a thicker configuration and sung pitches are relatively low (often termed ‘chest register’); where the cricothyroid muscle predominates over the vocalis, the vocal folds have a thinner configuration and sung pitches are relatively high (often
termed 'head register') (Thurman and Welch, 1997, p239).

It should be noted that the latest version of the suggested model of singing development (above) is based on a research literature that is firmly rooted in Western musical genres. Singing development in relation to non-Western musics may be different because differing musical traditions and structures shape auditory perception and, potentially, vocal production (e.g. Walker, 1994; Imada, 1994; Addo, 1995, 1998; Chen-Haftek, 1996, 1998, 1999). Although singing is a commonplace human activity, it is also culturally diverse.