Researching the first year of the *National Singing Programme* in England: an initial impact evaluation

Welch, G.F., Himonides, E., Saunders, J., Papageorgi, I., Rinta, T., Preti, C., Stewart, C., Lani, J., and Hill, J.

Institute of Education, University of London, UK

Researching the first year of the *National Singing Programme* in England: an initial impact evaluation

Welch, G.F., Himonides, E., Saunders, J., Papageorgi, I., Rinta, T., Preti, C., Stewart, C., Lani, J., and Hill, J.

Institute of Education, University of London, UK

Introduction and background

Research suggests that singing behaviours are subject to developmental processes in which individual neuropsychobiological potentiality is shaped (nurtured and/or hindered) by learning experiences within socio-cultural contexts (Welch, 2007; in press; Knight, 2009). Although singing is commonplace, it is also marked by cultural diversity, with development related to opportunity (e.g. Mang, 2007), the prosodic features of indigenous languages (Azechi, 2008), as well as the dominant characteristics of the local musical soundscapes (Welch *et al*, 1997; Welch, 2006a; 2006b; in press).

In many parts of the world, the ability to sing is seen as a mark of an individual's underlying musicality (*cf* Sloboda *et al*, 2005). Consequently, those individuals whose singing development has been hindered in some way are often labelled (including self-labelled) in some absolutist sense under a bi-polar categorisation of 'can'/'cannot' sing, with variations in their ascribed musical identity as a 'non-singer', 'tone-deaf', or 'tone-dumb' being found in virtually all cultures. Yet, as mentioned above, contrary evidence from developmental and neurological studies continues to emerge that singing and musical behaviours are context bound and susceptible to improvement with appropriate experience which can be informal as well as formal (e.g. Brown *et al*, 2004; Koelsch et al 2005; Mang, 2006; 2007; Dalla Bella *et al*, 2007; Kleber *et al*, 2007; Fuchs *et al*, 2007; Mithen & Parsons, 2008; Stewart & Williamon, 2008; Welch *et al*, 2008; see Welch, 2006a for review).

Furthermore, the recent wealth of studies into the neurosciences and music (*cf* Avanzini *et al*, 2003; 2005) continue to amass evidence of the multi-sited representation of musical behaviours in various regions of the brain, including singing (Kleber *et al*, 2007). These and related studies also indicate that there are various other-than-musical benefits that can accrue for the individual from engaging in musical (including singing)

activity, such as related to physical and psychological health and well-being (Clift & Hancox, 2001; Clift *et al*, 2007; Kreutz *et al*, 2004; Welch, 2005), social skill development and social inclusion (Odena, 2007; Portowitz *et al*, 2008) and cognitive development (Schlaug *et al*, 2005).

Therefore, it is perhaps not surprising that music is figuring more significantly in the contemporary educational policies of several of the world's governments. Brazil, for example, has recently (August, 2008) signed into law its first compulsory provision of music education in schools; Italy's regional government for Emilia Romagna has supported a special project (2005-2008) concerning the provision of vocal and choral education in primary schools; and the UK Government has a formal 'Music Manifesto', defined by the Department of Culture, Media and Sport (DCMS) as '…a campaign for improvement in music education. It is about creating more music for more people.' [1]

One major component of the 'Music Manifesto' concerns the *National Singing Programme* (2007-2011). This arose from a recommendation in the 2nd Report of the Music Manifesto group ('Making every child's music matter' October 2006) that singing be provided for all early years and primary-aged children by 2012.

> 'Singing offers the most direct route to providing a music-making experience for all children and young people, so we believe it should be a central element of the universal music offer. As a result, we recommend putting group singing at the heart of all primary school musical activity.' (Music Manifesto Report No 2, 2006:8)

In part, this was because of the opportunity afforded for the development of a cultural programme (2008-2012) that would be linked to the London-based Olympic Games (Education Guardian, 18th October, 2006). But it was also in recognition of the perceived importance of singing as a foundation for all round music education development[2].

---

[1] The Music Manifesto was launched in July 2004 by the then Schools Minister, David Miliband, and the Arts Minister, Estelle Morris, together with sixty plus signatories from the music industry. It is described on the Music Manifesto website as *'the result of a unique collaboration between the DfES and DCMS [two Government Ministries] with music organisations, musicians, teachers and composers, the music industry, broadcasting, teacher and musicians' unions, arts and education charities and Trusts'* (see http://www.musicmanifesto.co.uk/history [retrieved 21 July 2007]). *'At the heart of the Music Manifesto is a desire to see more opportunities in music for more young people – from high quality curriculum tuition to out of school hours youth and garage bands; from composing to live performance, from classical concerts to DJing and gigs. In its final form, the Music Manifesto offers a strategic direction for the future of music education and a common agenda for joint action'* (ibid). The Department for Children, Schools and Families (DCSF – created in 2007) reported that its predecessor, the Department for Education and Skills (DfES) had invested over £500 million in music education between 1999 and 2008, with £95 million proposed for investment in 2007/08 alone.

[2] In the introduction to the 2nd Music Manifesto report, Marc Jaffrey, the 'Music Manifesto Champion' wrote '*Singing has the potential to involve children and young people in music on a scale that we*

The UK Government's then Secretary of State for Education and Skills, Alan Johnson, together with the then Culture Minister, David Lammy, announced in January 2007 the launch of an additional £10m funding package (subsequently confirmed as £10m per year for four years) to support school singing, both in and out of school hours, through a major national singing campaign for primary school aged children, led by the British composer and broadcaster Howard Goodall in a new role as the 'Singing Ambassador' for England (DfES Press Notice, 16th January 2007 - http://www.dfes.gov.uk/pns/DisplayPN.cgi?pn_id=2007_0009).

Following a tendering process, the two Government Departments (DCMS, DfES) jointly appointed a consortium of Youth Music, The Sage Gateshead, Faber Music, and advertising agency Abbot Mead Vickers to lead on the actual provision of the National Singing Programme in 2007-2008 and on through to 2011. Included in the intentions of the Programme are that 'children experience high-quality singing, both within and without their daily school curriculum, on a daily basis' and that 'Every school has a teacher committed to facilitating high quality singing and vocal work for the whole school'.

The *Sing Up* National Singing Programme was launched in November 2007 and a team from the Institute of Education, University of London, led by the first author, were appointed to undertake a research evaluation of key elements of the Programme. Two prime foci were: (i) to undertake an initial baseline audit of children's singing behaviours in randomly selected schools and (ii) to link this baseline data collection to a post-impact evaluation of particular *Sing Up* programme interventions with children and adults (teacher, parents and other professionals involved in promoting singing in community contexts). This paper focuses on the evidence concerning the possible impact of the programme on participant children's singing behaviours during the first year (to July 2008).

Methodology

Participants

The participants for the research (baseline and post-intervention evaluation) were drawn from eighty-one schools located across England. The schools were in major cities and adjacent population centres across the South-East (Greater London), East

---

have not witnessed before. It is the most elemental form of music making, and is within the grasp of all of us, whatever our ability. It is a powerful community activity binding individuals and community together.'

(Cambridgeshire, Essex), South-West (Bristol and Gloucester), Midlands (Birmingham, Coventry, Derby), North-West (Manchester) and North-East of England (Newcastle, Gateshead, Durham, York), supplemented by a smaller number of schools in other parts of the country in urban, suburban and rural settings, as well as a group of Cathedral Choir Schools. Contacts were made initially with Local Authority music advisors and university music education colleagues for advice on possible participant schools[3], the intention being to draw on local knowledge to ensure that a diverse range of school singing 'cultures' were accessed. Cathedral Choir Schools were contacted directly.

Within each school, participant children were drawn primarily from two contrasting age groups, 7-year-olds and 10-year-olds, representing the upper and lower age groups within Key Stage 2 of the National Curriculum in England[4]. However, where the prime age focus baseline children were in classes with mixed age groups (such as 10-year-olds with some 9-year-olds), normally all the children in the class were assessed in order to ensure that none felt excluded. This meant that the baseline data also included 8 and 9-year-olds, as did the classes assessed in the post-intervention data collection[5].

The initial assessment phase ran from late September 2007 through to February 2008 and was focused on generating some sense of the commonality and diversity of singing behaviours across pupils in English Primary schools. This phase was termed the Year 1 'Baseline Assessment'. In this baseline phase, n=3,510 children were assessed from 77 schools (see below for the assessment protocol). Of these, 10 schools subsequently were visited again, i.e. one visit during the baseline phase and then again between May to July 2008 after the specific *Sing Up* singing development intervention. These 10 schools accounted for n=495 assessments within the baseline phase and n=324 assessments post-intervention (see Table 1a). The post-intervention number of assessments was smaller because not all of the original baseline children were involved subsequently in the intervention (the classes selected for the intervention were at the

[3] See Acknowledgements.

[4] Previous research (e.g. Rutkowski, 1997; Stadler Elmer, 2002; Welch, 1998; 2006a, 2006b; 2007) had demonstrated that clear developmental differences in singing behaviour by age and sex were likely to be evidenced by the selection of these two age groups. Other recent findings from research into the acoustics of children's singing voices (Sergeant & Welch, in press a & b) and children's vocal health in singing and speaking (Rinta & Welch, 2008; Williams *et al*, 2005) similarly supported such a conception.

[5] All participants (headteachers, teachers and pupils) had the purpose of the assessment explained in advance (and in writing to the school). Under our ethical guidelines, we guaranteed anonymity to all participants and reminded them that they were allowed to withdraw from the assessment process at any time that they felt uncomfortable.

school's discretion). In addition, another n=70 children from 4 schools were assessed only after their *Sing Up* intervention, but not before. Together, this makes a total of 14 schools and 394 assessments in the post-intervention phase. The breakdown of numbers of participants by age and sex for each assessment phase is shown in Table 1b[6].

Table 1a: Numbers of individual participant assessments by school and phase (baseline n=3,510; and post-intervention n=394 [324+70])

	Baseline Phase	**Post-intervention Phase**	Grand Total
School visited twice	495	324	819
School visited during baseline phase only	3015	-	3015
School visited during post-intervention phase only	-	70	70
Grand Total	3510	394	3904

Table 1b: Numbers of participants by age and sex for each phase (baseline n=3,510; and post-intervention n=394 [324+70])

	Boys	**Girls**	Grand Total
Year 2	77	68	145
Year 3	769	703	1472
Year 4	55	56	111
Year 5	170	114	284
Year 6	780	718	1498
Baseline totals	*1851*	*1659*	*3510*
Year 2	41	39	80
Year 3	33	46	79
Year 4	41	44	85
Year 5	38	54	92
Year 6	25	33	58
Post-intervention totals	*178*	*216*	*394*
Grand Total	2029	1875	3904

Notwithstanding the essential class-based organisation of the baseline study and the *Sing Up* interventions, as each child had been assigned a unique reference number within

[6] The total numbers of individual pupils involved across all assessment phases was 3,762.

the database, it was possible to identify subsequently those individual children (n=107) who had been assessed during the baseline phase and again after they had their *Sing Up* input.

Assessment Protocol

As mentioned above, amongst the prime foci of the research evaluation of *Sing Up* in its first year (2007-2008) was (i) to create an initial baseline profile of (a) children's singing and vocal behaviours and (b) attitudes to singing[7] that could be used for comparative purposes subsequently and (ii) evaluate the possible impact of a specific *Sing Up* intervention on participant children. This particular intervention was termed 'Singing Playgrounds' and was provided by members of *Ex Cathedra*, one of the UK's leading choir and Early Music ensembles. 'Singing Playgrounds' is an educational outreach programme designed to develop children's musicianship through singing games. Expert adult singers visit school playgrounds and work with older children – called 'Song Leaders' – who lead their peers in singing games.

> 'Through the use of weekly set tasks, the Song Leaders are encouraged to develop and evaluate their own activities. Equipped with clipboards and stickers to hand out to the younger children for enthusiastic participation in "Jump Jim Joe" and other popular playground hits, the song leaders…are seen as role models throughout the school and are chosen for their enthusiasm.' (retrieved 18 August 2008 from http://www.singup.org/teachers_and_music_leaders/recipes_for_success/Singing_Playgrounds.php)

The research protocol for the assessment of singing and other vocal behaviours drew on established models on singing development from the literature. Previous research indicated that it would be helpful to assess more than one aspect of children's vocal behaviour in order to build a composite, rounded picture. Consequently, the protocol investigated: (i) the children's *habitual speech pitch centre* (by asking each participant to count backwards from twenty and noting the spoken pitch centre in relation to an adjacent piano keyboard); (ii) *comfortable singing range*[8] (by imitative singing of a musical

[7] During the research visit, class teachers also made provision for each child to complete a 45-question survey of their attitudes to singing. This will be reported elsewhere, as the data analyses (including factor analyses) are extensive and beyond the space available in this paper. However, a brief overview is attached in the Annex.

[8] Comfortable singing range, rather than singing range limits, is regarded as a more valid measure of children's customary singing behaviour with regard to song items in their local culture (Welch, 1979).

song fragment at various starting pitches, transposed upwards and downwards with reference to an adjacent keyboard); and (iii) *singing behaviour in two well-known song items* (either 'Twinkle, Twinkle' and 'Happy Birthday' or one or other items that the particular child knew well – on advice from the teacher – if these were unknown). The last of these three elements was assessed against two established rating scales (Rutkowski, 1997; Welch, 1998) (see Figure 1).

Rutkowski (1997) ***Singing Voice Development Measure (SVDM)***

1	"Pre-singer" does not sing but chants the song text.
1.5	"Inconsistent Speaking Range Singer" sometimes chants, sometimes sustains tones and exhibits some sensitivity to pitch, but remains in the speaking voice range (usually a3 to c4 [note: the pitch labels have been altered to bring them in line with modern conventions in which middle C = c4, 256 Hz]).
2	"Speaking Range Singer" sustains tones and exhibits some sensitivity to pitch but remains in the speaking voice range (usually a3 to c4).
2.5	"Inconsistent Limited Range singer" waivers between speaking and singing voices and uses a limited range when in singing voice (usually up to f4).
3	"Limited Range Singer" exhibits consistent use of initial singing range (usually d4 to a4).
3.5	"Inconsistent Initial Range Singer" sometimes only exhibits use of limited singing range, but other times exhibits use of initial singing range (usually d4 to a4).
4	"Initial Range Singer' exhibits consistent use of initial singing range (usually d4 to a4).
4.5	"Inconsistent Singer" sometimes only exhibits use of initial singing range, but other times exhibits use of extended singing range (sings beyond the register lift: b^b4 and above).
5	"Singer" exhibits use of extended singing range (sings beyond the register lift: b^b4 and above).

Welch (1998) ***A revised model of vocal pitch-matching development (VPMD)***

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.
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.
Phrase 3	Melodic shape and intervals are mostly accurate, but some changes in tonality may occur, perhaps linked to inappropriate register usage. Overall, however, the number of different reference pitches is much reduced.
Phase 4	No significant melodic or pitch errors in relation to relatively simple songs from the singer's musical culture.

Figure 1: Two independent measures of Singing Development, each used in the assessment of participants' singing behaviours on two selected song items

Previous research (Mang, 2006) had demonstrated that the two scales could be used alongside each other to investigate complimentary aspects of singing development. The Rutkowski (1997) scale is a measure of singing voice development, whereas the Welch (1998) scale assesses vocal pitch-matching development.

Children were visited at their schools where their singing and vocal behaviours were assessed individually in a quiet space. Each child was taken through the assessment protocol, normally being tested individually within a small group that was drawn from the class. This allowed the other members of the group to observe and see what was required as this had been shown previously to be an appropriate method of accessing better quality responses than individual testing alone (*cf* Plumridge, 1972). To avoid the effects of vocal modelling, no starting pitch was given for the song items and, although the member of the research team provided verbal encouragement to the child, they did not offer any sung prompt (*cf* as advised by Mang, 2006). All children completed the assessments and none were excluded from the study.

Because of the large numbers of participants it was necessary to create a relatively large research team to undertake the fieldwork. Consequently, to promote reliability in the assessment process, this was undertaken initially by moderation, with members of the research team undergoing initial training on sampled items, then undertaking a school visit in pairs prior to making visits on their own. The validity and ease of use of the assessment protocol was established through a short piloting process prior to commencement of the main data collection.

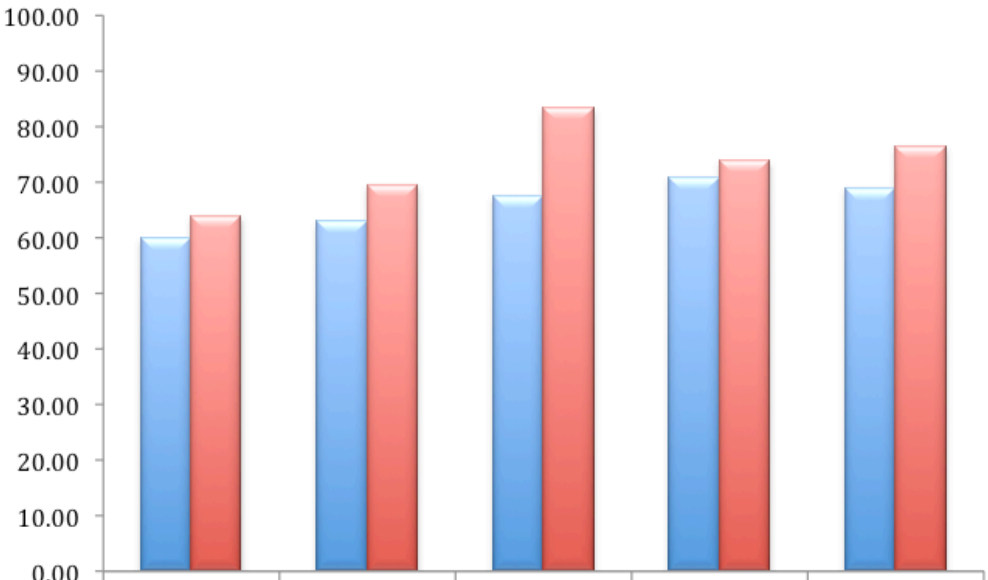
Participants' responses were noted onto individual assessment forms (see Welch *et al*, 2008 for an example) and data were subsequently entered for collation and analysis into a bespoke data entry form that was connected to a structured query language (SQL) based database. Each participant was uniquely coded in order to enable comparative assessment of singing development at a later date as necessary. The database included information on participant demographics (research sites, child's age, year group, sex, ethnicity), unique identification codes for each child, spoken pitch centre, ratings on each of the two independent measures of singing development and normalised singing score (being a conversion of the rated measures into a percentage of the maximum ratings across the combined rating scales[9]).

[9] If a child was rated at the highest possible level on each of the two independent measures for both song items, this would equate to 100% normalised singing score. Slightly lower ratings generate lower percentages and normalises singing scores. This data processing facilitates the possibility of group comparison, such as by class, school, age, sex or ethnicity.

Results

(i) Independent measures of singing development by age, sex and phase of assessment

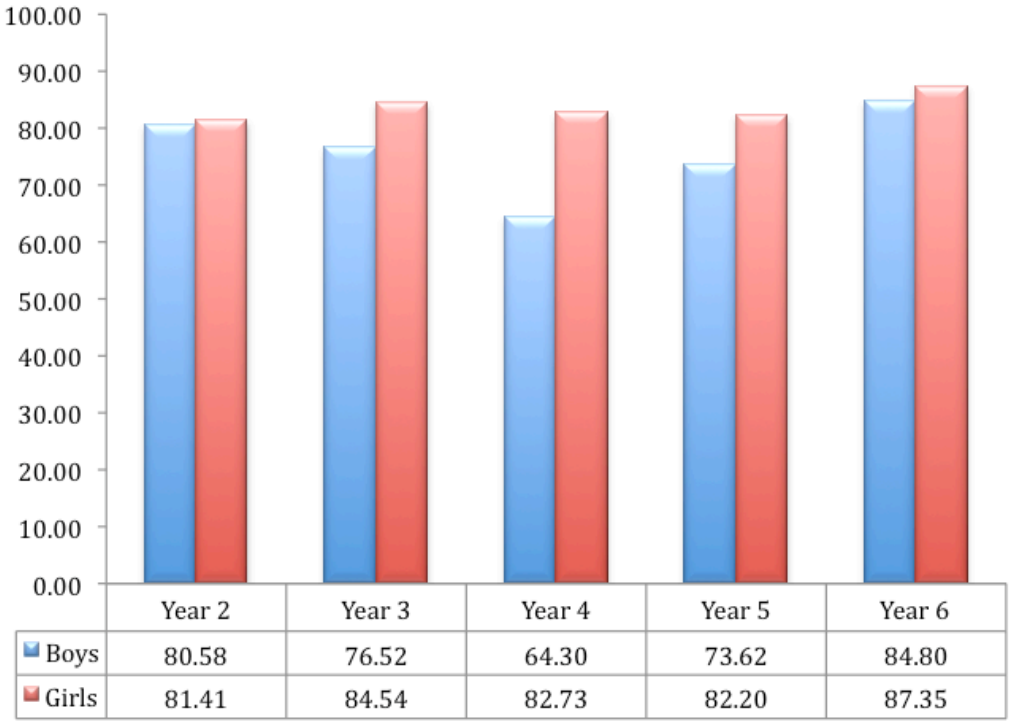
The prime focus for this paper is to explore the evidence of *Sing Up* impact from the 'Singing Playgrounds' intervention on the singing behaviours of participant children[10]. Analyses of the baseline data (n=3,510) indicate that there are age and sex differences in singing development. In general, older children (age 10, Year 6) tend to be rated as more developed on both rating scales than their younger peers (e.g. age 7, Year 3) and girls tend to be rated as more developed than boys in each age group. This is evidenced on both rating scales (Rutkowski; and Welch) separately and also in the combined normalised score for each age group (see Figure 2a). A similar pattern is evidenced in the post-intervention data (n=394), but with higher normalised mean scores (see Figure 2b and Table 3).



	Year 2	Year 3	Year 4	Year 5	Year 6
Boys	59.85	62.84	67.52	70.90	68.96
Girls	63.86	69.53	83.26	73.75	76.46

Figure 2a: Average of normalised singing development ratings by age and sex for baseline participants (n=3510)

[10] Each child also completed a 45-question survey of their attitudes to singing. This will be reported elsewhere as the data analyses (including factor analyses) are extensive.



	Year 2	Year 3	Year 4	Year 5	Year 6
Boys	80.58	76.52	64.30	73.62	84.80
Girls	81.41	84.54	82.73	82.20	87.35

Figure 2b: Average of normalised singing development ratings by age and sex for post-intervention participants (n=394)

Table 3: Normalised singing ratings for age and sex at baseline (n=3,510) and post-intervention (n=394)

	Boys	Girls
Baseline phase (all Year-groups)	66.18	73.05
Year 2	59.85	63.86
Year 3	62.84	69.53
Year 4	67.52	83.26
Year 5	70.90	73.75
Year 6	68.96	76.46
Post-intervention phase (all Year-groups)	75.18	83.45
Year 2	80.58	81.41
Year 3	76.52	84.54
Year 4	64.30	82.73
Year 5	73.62	82.20
Year 6	84.80	87.35

An initial independent samples t-test was conducted in order to compare the normalised singing competency scores (Rutkowski & Welch combined ratings) for all pupils that were assessed during the baseline (n=3510) and post-intervention periods (n=394) in Year 1. The difference between the two (equal variances not assumed) was statistically significant [t(539)=11.2, p<.0005]. The scores for the baseline (M=69.425, SD=20.825) were significantly lower that those for the post-intervention (M=79.714, SD=16.781). Although the magnitude of the difference of the means was small (eta squared=0.031), there is a markedly higher score in assessed singing competency for the post-intervention pupils.

Similarly, an independent samples t-test was conducted in order to compare the normalised singing competency scores (Rutkowski & Welch) of pupils whose schools had been visited both during the baseline (n=495) and post-intervention (n=324) phases. There was a significant difference [t(762)=11, p<.0005, equal variances not assumed] between baseline assessments (Visit 1) (M=65.72, SD=19.792) and post-intervention assessments (M=79.96, SD=16.898). The magnitude of the difference of the means was quite large (eta squared=0.129).

Table 4: Normalised score significant differences between Year-Groups 1.1. (Baseline participants, n=3,510)

Baseline

	Year-2	Year-3	Year-4	Year-5	Year-6
Year-2	-	-2.505*	-5.599***	-5.367***	-5.895***
Year-3	2.505*	-	-4.720***	-4.997***	-8.535***
Year-4	5.599***	4.720***	-	N.S.	N.S.
Year-5	5.367***	4.997***	N.S.	-	N.S.
Year-6	5.895***	8.535***	N.S.	N.S.	-

*p<0.05 **p<0.005 ***p<0.0001

A series of t-tests were then undertaken on each set of normalised singing ratings (baseline and post-intervention) to look at differences within each assessment phase (as illustrated above in Figure 2a and 2b). As can be seen from Table 4 above (re Figure 2a), the analyses of comparative year groups within the baseline data demonstrate significant differences between the youngest children (Years 2 and 3, ages 6 and 7) and their older peers (Years 4, 5 and 6, ages 8, 9 and 10) in normalised scores. In particular, the

differences in ratings between the two prime age focus groups, Years 3 and 6 (ages 7 and 10), are highly significant.

In the between age groups analysis of the post-intervention participants, there are fewer significant differences evident, although the ratings for Years 3 and 6 (ages 7 and 10) continue to demonstrate the same baseline pattern of difference, with older children being more highly rated (and see Figure 2b above).

Table 5: Normalised score significant differences between Year-Groups (Post-intervention participants, n=394)

Post intervention

	Year-2	Year-3	Year-4	Year-5	Year-6
Year-2	-	N.S.	2.447*	N.S.	N.S.
Year-3	N.S.	-	2.836**	N.S.	-2.236*
Year-4	-2.447*	-2.836**	-	N.S.	-4.445***
Year-5	N.S.	N.S.	N.S.	-	-2.972**
Year-6	N.S.	2.236*	4.445***	2.972**	-

*p<0.05 **p<0.005 ***p<0.0001

When the baseline and post-intervention ratings are combined for analysis by age (n=3904, see Table 6), there is an overall pattern of significant difference in the t-tests between the younger pupils (Years 2 and 3, ages 6 and 7) and their older peers (Years 4, 5 and 6, ages 8, 9 and 10), with the latter being significantly more highly rated in their singing development.

Table 6: Normalised score significant differences between Year-Groups (All assessments, n=3904)

Complete dataset (N=3904)

	Year-2	Year-3	Year-4	Year-5	Year-6
Year-2	-	N.S.	-3.113**	-3.024**	-2.975**
Year-3	N.S.	-	-5.185***	-6.472***	-8.407***
Year-4	3.113**	5.185***	-	N.S.	N.S.
Year-5	3.024**	6.472***	N.S.	-	N.S.
Year-6	2.975**	8.407***	N.S.	N.S.	-

*p<0.05 **p<0.005 ***p<0.0001

When comparing the singing development assessment scores of the sexes across the whole data set (n=3904), the boys (n=2029) tended to have a lower rating than the girls (n=1875). An independent samples t-test indicated that there was a significant difference [t(3900)=11.2, p=.000, equal variances not assumed] between male pupils (M=66.667, SD=20.908) and females (M=74.25, SD 19.761). The magnitude of the difference in the means was small (eta squared = 0.031).

A series of t-tests to compare the sexes within each age group (Table 7) indicate that there is no difference in the baseline data for Year 2 (age 6), but that for each successively older age group (with two exceptions), the girls are rated more highly than the boys. The exceptions are in Year 5 in the baseline data and Year 6 in the post-intervention scores. It is hoped that these analyses for Years 2, 4 and 5 will become more robust in future assessment years when they become the prime age foci for extending the baseline. Nevertheless, overall, the post-intervention assessments indicated that boys are generally being rated more highly following their 'Singing Playgrounds' experiences.

Table 7: Normalised comparative ratings for year group and sex in each assessment phase and combined (complete dataset)

	male versus female		
	Baseline	Post-intervention	Complete dataset
Year-2	N.S.	N.S.	N.S.
Year-3	-6.382***	-2.877**	-6.931***
Year-4	-4.720***	-4.742***	-6.690***
Year-5	N.S.	-2.573*	-2.757*
Year-6	-6.931***	N.S.	-7.081***

*p<0.05 **p<0.005 ***p<0.0001

(ii) Evidence of impact on individuals

As reported above in Table 1a, there were n=819 singing behaviour assessments of children in the 10 schools that our team has visited twice. Of those, n=495 assessments were performed during the baseline visit and n=324 assessments were performed during the first follow-up visit. Within these, the number of individual pupils that have been assessed during both visits and whose data can be matched is n=107.

Accordingly, a paired samples t-test was run on the normalised scores. This revealed a statistically significant improvement [t(106)=5.916, p<.0005] between the

baseline phase assessments (M=70.58, SD=16.096) and the post-intervention assessments (M=81.80, SD=15.355). The eta squared (0.25) indicates a large size effect.

(iii) Evidence of impact on sung vocal pitch ranges

A further analysis was undertaken to compare the comfortable sung pitch ranges of the same n=107 children. A paired-samples t-test was applied to evaluate the impact of the *Sing Up* intervention on each individual's comfortable sung pitch range in semitones. There was a statistically significant improvement [t(106)=5.398, p<0.0005] between the baseline phase range (M=10.83, SD=5.614) and range produced in the post-intervention phase (M=13.70, SD, 4.379) of approximately three semitones. The eta squared statistic (0.22) indicated a large size effect.

(iv) Evidence of impact related to participant ethnicity

Schools provided background information on the ethnicity of pupils according to the Department for Children, Schools and Families official classification. An analysis of normalised singing scores for each of these groups indicates that, within and across the two data collection phases, Asian pupils (n=609; M=65.38, SD=21.217) tended to score significantly lower than their White peers (n=2729; M=71.15, SD=20.308) [t(874)=6.12, p.=.000, equal variances not assumed]. Asian pupils also scored significantly lower than their Black peers (n=294; M=72.46, SD=21.458) [t(874)=6.12, p.=.000, equal variances not assumed]. There were no significant differences between White and Black children in either baseline or post-intervention data. (The numbers of pupils in the other ethnicity categories in the Year 1 data were too small for meaningful statistical comparison.)

Table 8: Comparison of baseline and post-intervention data by ethnicity and normalised singing ratings

ethnicity	baseline N	normalised baseline singing score	post-intervention N	normalised post-intervention singing score	Grand Total N	Grand Total singing score
Other	86	70.7	0	0.0	86	70.7
Asian	462	62.1	147	75.6	609	65.4
Black	239	69.9	55	83.4	294	72.5
Chine	14	74.0	5	79.0	19	75.3
Mixed	149	72.8	18	79.0	167	73.5
White	2560	70.4	169	82.2	2729	71.2
Grand Total	3510	69.4	394	79.7	3904	70.5

However, notwithstanding these statistical differences between ethnicities, *all* three major groups have significantly higher normalised singing ratings in their post-intervention assessment data (Asian pupils $t(253)=7.078$, $p<.0005$; Black pupils $t(113)-5.414$, $p<.0005$; White pupils $t(221)=10.478$, $p<.0005$; equal variances not assumed). In the case of Asian pupils, the post-intervention score is also much higher than that for the White and Black pupils at baseline (see Figure 3 and Table 8). Furthermore, when the Year 1 schools are put into a rank order on the basis of their pupils' overall normalised singing scores, schools with Asian pupils are to be found in the upper quartile (scores above M=78), suggesting that they are equally capable of improving their singing abilities as other ethnic groups and, at a school level, of being at an equivalent developmental level.

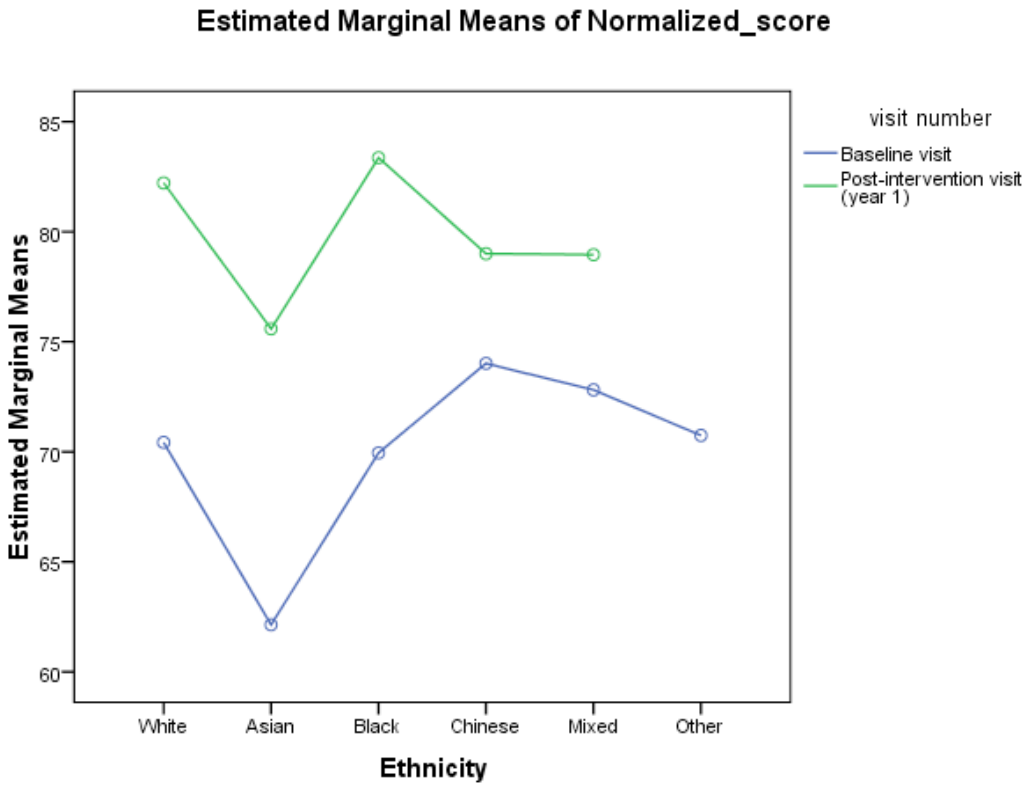


Figure 3: Estimated marginal means for normalised singing scores by ethnicity

Discussion

Some caution is needed in the interpretation of the statistical data because the data set is (of necessity) uneven in its distribution across age groups and the numbers of pupils that were assessed in the post-intervention phase was somewhat smaller than that

available for the overall baseline. Nevertheless, there are some positive early indications that at least one major strand of the *Sing Up* National Singing Programme – the 'Singing Playgrounds' initiative by *Ex Cathedra* – is making a difference to the underlying singing behaviours of the participant children. Taken together with survey evidence from the same children's reported attitudes to singing (to be reported elsewhere – see also Annex below) – that also shows a positive shift towards maintaining engagement in both boys and girls – the initial impression is of an upward trajectory in participants' singing engagement and development during this launch year because of the national programme. Where differences in children's normalised signing behaviours are seen to exist at school level, anecdotal evidence (to be explored more systematically in the coming months) suggests that this relates to school leadership decisions on the relative importance of singing and music in the curriculum compared to other subjects.

As researchers, our prime aims were to establish some form of initial baseline that could act as a comparison data set against which any post-intervention assessments could be measured. The intention in Year 2 (2008-2009) is to undertake a similar number of baseline assessments (somewhere approaching 3,500) from children aged 8 and 9. This will allow a much fuller picture to emerge across ages and sexes. We also intend to follow up the Year 3 (age 7) participants into Year 4 (age 8) to initiate a longitudinal component to the baseline. However, although the numbers of children with data in the post-intervention phase are currently small within the overall total (in part because the *Sing Up* programme did not begin to roll out in schools until late 2007/early 2008), the basic assessment framework appears to be robust and providing useful data on which an independent evaluative assessment of impact can be made. Those children who were tracked from before and after their 'Singing Playgrounds' activities demonstrated a positive development in their singing abilities, both in the underlying available comfortable pitch range (averaging an additional 3 semitones) and in their song singing competency. Although numbers are small (n=107), they are part of a larger group (n=394) who also demonstrate increased singing development behaviours.

The ambition of the UK Government is to develop singing for 3.3 million children aged 5 to 10 years across 17,504 Primary schools and other community settings in England over a four-year period (2007-2011). The challenge is enormous, but this early research data provides encouraging evidence about the possible success of the initial steps on this journey. It also provides evidence that singing development should be

considered as a normal feature of children's musical engagement with the world around them, particularly when they are provided with new and positive singing experiences.

Acknowledgements

The research team wish to thank Maurice Walsh, Senior Vocal Tutor with Manchester Music Service; Ula Weber of *Ex Cathedra*; Dr Penelope Harnett, University of West of England; Dr Liz Mellor, York St John University; and Sarah Kekus and Edward Milner of the *Sage Gateshead* for their invaluable support in identifying participant schools. We are also extremely grateful to all the schools (pupils, teachers and headteachers) for their time and commitment to participate in this national research activity.

References

Avanzini, G., Faienza, C., Minciacchi, D., Lopez, L., & Majno, M. (2003). (Eds.), *The Neurosciences and Music* (Vol. 999). New York: Annals of the New York.

Avanzini, G., Koelsch, S., Lopez, L., & Majno, M. (2005). *The Neurosciences and Music II.* (Vol. 1060). New York: Annals of the New York Academy of Sciences.

Azechi, N. (2008). Young Children's Rhythmic Behaviour in Singing: The Influence of Mother Tongue on Their Development. *Proceedings,* ICMPC10, Sapporo, Japan, 25-29 August 2008.

Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., & Parsons, L. M. (2004). The song system of the human brain. *Cognitive Brain Research 20*, 363 - 375.

Clift S., & Hancox G. (2001). The perceived benefits of singing: Findings from preliminary surveys of a university college choral society. *Journal of the Royal Society for the Promotion of Health, 121*, 248-256.

Clift, S., Hancox, G., Morrison, I., Hess, B., Kreutz, G., & Stewart, D. (2007). Choral singing and psychological wellbeing: Findings from English choirs in a crossnational survey using the WHOQOL-BREF. In A. Williamon & D. Coimbra (Eds). *Proceedings*, International Symposium on Performance Science, Porto, Portugal, 22-23 November, 2007, 201-207.

Dalla Bella, S., Giguère, J-F., & and Peretz, I. (2007). Singing proficiency in the general population. *J. Acoust. Soc. Am. 121* (2), 1182–1189.

Fuchs, M., Meuret, S., Thiel, S., Täschner, R., Dietz, A., & Gelbrich, G. (2007). Influence of Singing Activity, Age, and Sex on Voice Performance Parameters, on Subjects' Perception and Use of Their Voice in Childhood and Adolescence. *Journal of Voice,* [Published on line September 2007]

Kleber, B., Veit, R., Birbaumer, N., & Lotze, M. (2007). Neural correlates of professional classical singing. In A. Williamon & D. Coimbra (Eds). *Proceedings*, International Symposium on Performance Science, Porto, Portugal, 22-23 November, 2007, 335-343.

Knight, S. (2009). *An attributional study of adult 'non-singers': an analysis of ideographic, nomothetic*

and socio-cultural perspectives. Unpublished PhD Thesis, Institute of Education, University of London.
Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *NeuroImage 25*, 1068– 1076.
Kreutz G., Bongard S., Rohrmann S., Hodapp, V., & Grebe, D. (2004). Effects of choir singing or listening on secretory immunoglobulin A, cortisol and emotional state. *Journal of Behavioral Medicine, 27*, 623-635.
Mang, E. (2006). The effects of age, gender and language on children's singing competency, *British Journal of Music Education, 23,* 161-174.
Mang, E. (2007). Effects of Musical Experience on Singing Achievement, *Bulletin of the Council for Research in Music Education*, 174, 75-92.
Mithen, S., & Parsons, L. (2008). Singing in the brain. *New Scientist,* 23 February, p28.
Music Manifesto http://www.musicmanifesto.co.uk/ (Retrieved 24 August 2008).
Odena, O. (2007). *Music as a way to address Social Inclusion and Respect for Diversity in early childhood*. Study Paper for the Bernard van Leer Foundation. Belfast: NFER at Queen's.
Plumridge, J.M. (1972). *The Range and Pitch Levels of Children's Voices, in relation to Published Material for Children's Voices*. Unpublished Diss. Dip. Adv. Study of Ed, University of Reading.
Portowitz, A., Lichtenstien, O., Egorov, L., & Brand, E. (2008). Underlying mechanisms linking music education and cognitive modifiability. In S. Malbran & G, Mota (Eds). *Proceedings*. 22nd International Seminar on Research in Music Education, Porto, Portugal, 13-18 July, 2008.
Rinta, T., & Welch, G.F. (2008). Should Singing Activities Be Included in Speech and Voice Therapy for Prepubertal Children? *Journal of Voice. 22*(1), 100-112.
Rutkowski, J. (1997). The nature of children's singing voices: Characteristics and assessment. In: B.A. Roberts (Ed.), *The Phenomenon of Singing* (pp. 201-209). St. John's, NF: Memorial University Press.
Schlaug, G., Norton, A., Overy, K., & Winner, E., (2005). Effects of Music Training on the Child's Brain and Cognitive Development. *Ann. N.Y. Acad. Sci. 1060*: 219–230
Sergeant, D.C., & Welch, G.F. (in press). Age-related changes in Long-Term Average Spectra of children's voices. *Journal of Voice.* [Published online July 2007]
Sergeant, D.C., & Welch, G.F. (in press). Gender differences in Long-Term-Average Spectra of children's singing voices. *Journal of Voice.* [published online May 2008].
Sing Up http://www.singup.org/ (Retrieved 24 August 2008).
Sloboda, J. A., Wise, K. J., & Peretz, I. (2005). Quantifying tone deafness in the general population, *Ann. N.Y. Acad. Sci.* 1060, 255–261.
Stadler Elmer, S. (2002). *Kinder singen Lieder: Über den Prozess der Kultivierung des vokalen Ausdrucks*. Berlin: Waxmann.
Stewart, L., & Williamon, A. (2008). What are the implications of neuroscience for musical education? *Educational Research, 50* (2), 177-186.
Welch, G.F. (1979). Vocal range and poor pitch singing. *Psychology of Music*, *7*(2), 13-31.
Welch, G. F. (1998). Early childhood musical development. *Research Studies in Music Education, 11*, 27-41.
Welch, G.F. (2005). Singing as Communication. In: D. Miell, R. MacDonald, & D. Hargreaves (Eds.), *Musical Communication.* (pp. 239-259). New York: Oxford University Press.
Welch, G.F. (2006a). Singing and Vocal Development. In: G. McPherson (Ed.) *The Child as Musician: a handbook of musical development.* (pp. 311-329). New York: Oxford University Press.

Welch, G.F. (2006b). The musical development and education of young children. In: B. Spodek & O. Saracho (Eds.), *Handbook of Research on the Education of Young Children*. (pp. 251-267). Mahwah, N.J.: Lawrence Erlbaum Associates Inc.

Welch, G.F. (2007). Addressing the multifaceted nature of music education: an activity theory research perspective. *Research Studies in Music Education*, *28*, 23-38.

Welch, G.F. (in press). Culture and gender in a cathedral music context: An activity theory exploration. In M. Barrett (Ed.), *A Cultural Psychology of Music Education*. New York: Oxford University Press.

Welch, G.F., Sergeant, D.C. & White, P. (1997). Age, sex and vocal task as factors in singing 'in-tune' during the first years of schooling. *Bulletin of the Council for Research in Music Education, 133*, 153-160.

Welch, G.F., Himonides, E., Saunders, J., Papageorgi, I., Rinta, T., Stewart, C., Preti, C., & Lani, J. (2008). The *National Singing Programme* for Primary schools in England: An Initial Baseline Study. In W. Sims (Ed.). *Proceedings,* International Society for Music Education 28th World Conference, Bologna, Italy, 20-25 July, 2008, 311-316.

Welch, G.F., Himonides, E., Saunders, J., Papageorgi, I., Rinta, T., Stewart, C., Preti, C., Lani, J. & Hill, J. (under review). The *National Singing Programme* for Primary schools in England: An initial baseline study.

Williams, J., Welch, G.F. & Howard, D.M. (2005). An exploratory baseline study of boy chorister vocal behaviour and development in an intensive professional context. *Logopedics Phoniatrics Vocology, 30*(3/4), 158-162.

Annex: Evidence of impact on children's attitudes to singing

All children completed a 45-question survey of their attitudes to different aspects of singing in school, at home and elsewhere. The data analyses reveal significant age and sex differences (Welch *et al*, under review)[11]. On average, (i) younger children were more positive about singing than older children and (ii) girls tend to be more positive about singing than boys in each age group (Figure 4). These findings are somewhat surprising when set against the earlier data on children's increasing singing competency with age (see Figure 2a). Overall, there is an inverse relationship between children's singing development and their attitudes to singing. Whilst the older children are more competent singers, both girls and boys are less positive about singing.

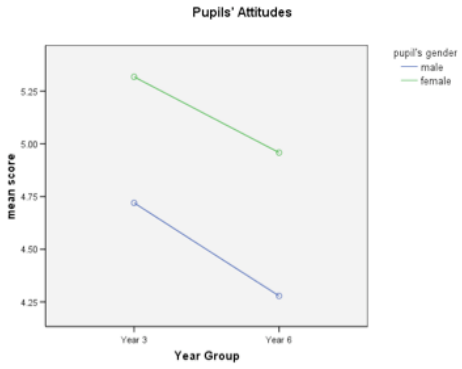


Figure 4: Attitudes to singing by age and sex

In order to explore this finding in more detail, a further statistical analysis was undertaken of children's responses within the 45 questions to see if there were any particular groupings evident in the way that the children had answered[12]. Six clusterings

[11] Taken across the whole participant population, including all year groups, a multivariate analysis of variance indicated that there were significant differences between male and female participants in their attitudes towards singing (F (6, 3337) = 231.796, p<.0001, partial eta squared = .294). There were also significant age differences (F (6, 3337) = 127.79, p<.0001, partial eta squared = .189). A separate ANOVA for sex and age that focused only on the Year 3 (n=1352) and Year 6 (n=1523) pupils confirmed the impact of these variables (sex: F(3, 2871) = 384.53, p<.0001, partial eta squared = .118; age: F(3,2871) = 150.847, partial eta squared = .050), but with no interaction between them.

[12] A principal components analysis was conducted on the 45 questions in order to investigate whether children's attitudes to singing could be summarised in subgroups on the basis of their responses. The

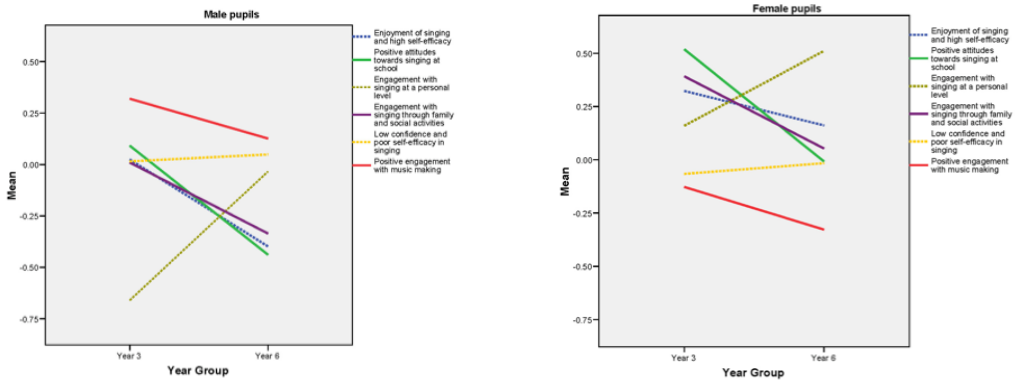
emerged, embracing (1) enjoyment of singing and high self-efficacy, (2) positive attitudes to singing at school, (3) engagement with singing at a personal level, (4) engagement with singing through family and social activities, (5) low confidence and poor self-efficacy in singing; and (6) positive engagement with music making. Sex differences were evidenced in all six factors. Overall, females tended to be more positive than males towards singing and more self-confident (Factors 1, 2, 3, and 4). Males tended to have lower confidence about singing and poorer singing self-efficacy (Factor 5). Yet, in terms of the more creative aspect of the curriculum, males had a higher positive engagement than females with music making (Factor 6).

Nevertheless, for both males and females, the majority of responses tended to be less positive for children in the oldest age group, but with a few exceptions. In general, older children of both sexes reported less enjoyment and engagement with singing compared with their younger peers, whether in school or at home with the family. This

---

suitability of the data for factor analysis was first investigated with the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity (Field, 2000). Both of these tests confirmed the suitability of the data (the KMO measure was .938 and Bartlett's Test of Sphericity was statistically significant, p<.0001). The Varimax Rotation method was selected to ensure that the extracted components were uncorrelated and to aid interpretation of the extracted factors.
Six components were extracted, explaining 41.8% of the variance. The factors were interpreted as follows: (1) Enjoyment of singing and high singing self-efficacy; (2) Positive attitudes towards singing at school; (3) Engagement with singing at a personal level; (4) Engagement with singing through family and social activities; (5) Low confidence and poor self-efficacy in singing; and (6) Positive engagement with music making. Multivariate analysis of variance was conducted on the six extracted components to investigate the possible effects of sex and age and their interaction on pupils' attitudes to singing. With regard to sex, female pupils scored higher on most of the positive components, such as 'enjoyment of singing and high self-efficacy' (F(6,3337) = 163.99, p<.0001, partial eta squared = .047; males $\bar{x}$ = -.21, females $\bar{x}$ = .23), 'positive attitudes towards singing at school' (F(6,3337) = 170.99, p<.0001, partial eta squared = .049; males $\bar{x}$ = -.20, females $\bar{x}$ = .22), and 'engagement with singing at a personal level' (F(6,3337) = 467.89, p<.0001, partial eta squared = .123; males $\bar{x}$ = -.31, females $\bar{x}$ = .20). Male students scored higher on 'positive engagement with music making' (F(6,3337) = 177.60, p<.0001, partial eta squared = .05; males $\bar{x}$ = .21, females $\bar{x}$ = -.24). They also scored higher on the negative component 'low confidence and poor self-efficacy in singing' (F(6,3337) = 4.39, p<.05, partial eta squared = .001; males $\bar{x}$ = .03, females $\bar{x}$ = -.04). With regard to age, year group differences were observed in five out of six components; the exception was 'low confidence and poor self-efficacy in singing' which was relatively stable across age groups. Younger pupils score higher in 'enjoyment of singing and high self-efficacy' (F(6,3337) = 76.07, p<.0001, partial eta squared = .022; younger $\bar{x}$ = .16, older $\bar{x}$ = -.13), 'positive attitudes towards singing at school' (F(6,3337) = 259.00, p<.0001, partial eta squared = .072; younger $\bar{x}$ = .29, older $\bar{x}$ = -.23), 'engagement with singing through family and social activities' (F(6,3337) = 102.96, p<.0001, partial eta squared = .030; younger $\bar{x}$ = .19, older $\bar{x}$ = -.15) and 'positive engagement with music making' (F(6,3337) = 34.05, p<.0001, partial eta squared = .010; younger $\bar{x}$ = .11, older $\bar{x}$ = -.09). In contrast, older pupils scored higher in 'engagement with singing at a personal level' (F(6,3337) = 239.64, p<.0001, partial eta squared = .067; younger $\bar{x}$ = -.28, older $\bar{x}$ = .23).

was particularly marked for the boys[13]. Similarly, in line with other research literature, these older children were less positive about the creative process of music making. However, older children of both sexes were more positive about singing at a personal level (see Figure 5a and 5b), particularly the boys, indicating that it was singing associated with both school and social activities that had become less attractive.



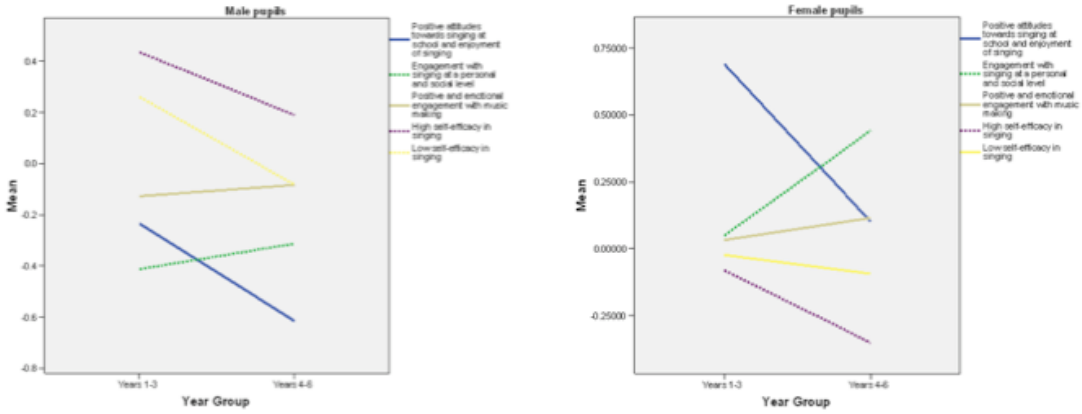
Figures 5a and 5b: Factor analyses of children's attitudes to singing by age and sex (Baseilne)

With regard to children's attitudinal responses post-intervention (Figures 6a and 6b), two key findings emerge in contrast to the baseline data set:

- Both sexes have a more positive attitude to music making following their 'Singing Playgrounds' experiences;
- Older boys and girls post-intervention tend <u>not</u> to report themselves as having low self-efficacy in singing following their 'Singing Playgrounds' experiences.

Taken together with the singing behaviour data reported in the main body of this paper, the overall impression is that *Sing Up* (at least in terms of this focused intervention) has had a positive impact on children's attitudes to singing as well as to their actual singing behaviours.

[13] Interactions between sex and age were observed in two components. These were 'enjoyment of singing and high self-efficacy' ($F(6,3337) = 15.18$, $p<.0001$, partial eta squared = .005) and 'engagement with singing at a personal level' ($F(6,3337) = 18.95$, $p<.0001$, partial eta squared = .006).



Figures 6a and 6b: Factor analyses of children's attitudes to singing by age and sex (post-intervention)