To be published April 2024 (Volume 41, Issue 4) at Music Perception http://mp.ucpress.edu/

The Activated Singer: Components of an Emerging Singing Identity in Adults Learning to Sing

Karen J. Wise

Guildhall School of Music & Drama, London, United Kingdom

Andrea R. Halpern

Bucknell University

Abstract

Many adults avoid singing participation, even in informal situations. We examined components of singing identity in self-identified non-singers using questionnaires, including a novel Singing Inhibition (SI) scale, among 238 adults volunteering for a training study. Higher levels of Singing Inhibition were predicted by a combination of lower self-reported singing skill, lower Parental/Family Engagement in singing, and stronger belief that singing is a fixed ability. A subsample of 20 self-reported non-singers (aged 23–71) participated in 10 months of singing lessons, and we tracked changes in objective singing competence as well as self-assessments and singing-related attitudes and beliefs at baseline, at six months, and at the conclusion. Among the trainees, some but not all aspects of singing improved. Importantly, we found that after six months, participants showed a significant reduction in Elitist Attitudes and Sensitivity to Social Judgment in singing and viewed singing as more open to improvement rather than a fixed talent. Self-assessment of accuracy, vocal tone, and physical sensations also improved. We view this shift as becoming an Activated Singer, encompassing both skills and attitudes, which is encouragement for even life-long non-singers to begin the journey to becoming a singer.

Received June 5, 2023, accepted December 15, 2023.

Key words: singing attitudes, singing competence, musical identity, training, adults

As research evidence for the benefits of singing for health and wellbeing (e.g., Pentikäinen et al., 2021) accumulates, public awareness and interest in singing has been growing. And so have singing opportunities, including everything from large-scale open singing workshops organized by major arts organizations, to grass-roots community groups, and even Tuneless Choirs.[1] However, many adults who might enjoy singing, desire to do so, and benefit from participation, do not take up such opportunities. For example:

> *Although I have always loved singing in private I am very uncomfortable singing around other people as I feel embarrassed about my voice. I have always wanted to learn to sing but have felt unable to join any singing groups for adults (even when they're open to anyone regardless of ability) as I think I would be too embarrassed to sing in front of others.*

The person quoted above (a potential project participant) feels unable to participate even in opportunities that aim to be inclusive. This quotation also highlights the interplay of negative self-identity, inhibition, and perceived lack of competence that are common in stories of singing avoidance.

We define non-singing (NSi) adults as *people who do not sing, except perhaps in the shower, and have not done so for most or all of their adult life.* Little is known about the needs of this population in singing participation, nor how their skills and identities may change in training. Our "Finding a Voice" project explored the journey of learning to sing for adults taking their first steps since childhood by participation in 10 months of singing lessons. We report here on quantitative data focused on singing-related beliefs and attitudes, self-assessments, and a comprehensive assessment of vocal and musical skills, in order to shed light on identity and development. As well as contributing to the understanding of adult singing development more widely, which has been rarely investigated, we think this information will be of interest to music educators and choir organizers who aim to be more age and status-inclusive in the singing activities they organize, and to NSi adults who may be unsure about joining such activities.

Non-Singing Identity

Research involving self-identifying non-singing (NSi) adults identifies key characteristics of a non-singing identity. The term "non-singer" can be self-applied in a variety of ways, including as a comparison with accomplished singers (i.e., someone who feels able to sing but not well enough to sing alone in front of others), or to express indifference towards singing (Knight, 2019). Some NSi consider themselves unable to sing; their stories commonly feature negative past experiences involving judgment, criticism, exclusion, or silencing. These events were understood by the individual to be incontrovertible proof of their innate, i.e., unmodifiable, inability to sing (Knight, 2019).

This innate "all or nothing" talent account of singing, at least in Western culture (Welch, 2001) contrasts to instrumental mastery, which may be more obviously associated with a large degree of training. Thus, compared to singers, self-identified non-singers have been found to more strongly endorse elitist attitudes towards singing (Bento-Allpress, 2013), corresponding to a talent account. Drawing on Dweck's (2000) Implicit Theory concept we propose that NSi believe singing is a fixed rather than a malleable ability. Given the strong evidence that holding a fixed (or entity) theory of an ability is related to a range of self-defeating behaviors that can impede achievement (Dweck, 2000), one goal of the present study was to measure this belief, and how it might change as people learned to sing.

Not everyone who is NSi feels suppressed in the manner described above (Knight, 2019); some may be indifferent or may never have had the opportunity to sing. In either case, it makes sense to ask to what extent singing is considered enjoyable by the general population. In Chong's (2010) qualitative study of university students' singing enjoyment, most people reported enjoying singing (88.3%). Furthermore, content analysis of participants' open-response answers identified a variety of perceived benefits of singing (Self-expression, Aesthetic Experience, Stress reduction/Mood change, Interpersonal relationships, Identity/empowerment, Spiritual, Self-actualization). However, 8.3% reported they only enjoyed singing when alone, and 3.5% said they had no enjoyment of singing at all. Those who only enjoyed singing when alone reported being sensitive to and fearful of other people's judgments, whereas

those with no enjoyment of singing had negative perceptions of their own voices and/or skills. In relation to fearfulness of others' judgments, Abril (2007) characterized the singing anxiety of the three participating elementary school trainee teachers as a form of social anxiety. Bento-Allpress (2013) also found that NSi were more likely than singers to avoid singing in formal contexts involving social judgment such as choirs and public performances.

Most of this prior research is qualitative, which has been valuable in providing rich detail of the histories and experiences of non-singers. However, there has been no standardized measurement of attitudes and experiences among adult NSi. Thus, this project quantifies some of these constructs.

Singing Skills in Context

This project is informed by a developmental view of singing (see, e.g., Welch, 2006) based on the substantial literature on children's singing. This literature elucidates a normative progression towards singing simple songs of the culture acceptably in tune, mirroring general trends in development from sensorimotor to cognitive. This begins with: 1) awareness that vocal pitch can be voluntarily controlled, followed by 2) accuracy of the melodic contour, 3) production of more accurate intervals within individual musical phrases, and finally 4) adherence to consistent tonality (Welch, 1994a, 1998). At the same time, children's vocal capability and control develops, such that the pitch range they can successfully use in singing increases with age (Rutkowski, 1997, 2015). Providing enhanced opportunities for singing in schools has been shown to increase both song singing accuracy and vocal range (Welch et al., 2011).

Conversely, the singing skills of NSi are likely to suffer from lack of practice, and it has further been suggested that adult NSi may have stalled in their singing development at a child-like stage (Lidmann-Magnusson, 1997; Welch, 1994b). In one study, adults self-defining as "tone deaf" on average showed performance consistent with a lower developmental level in singing than non-tone-deaf controls, characterized by a restricted vocal range and insecure tonality (Wise & Sloboda, 2008). In a cross-sectional study of kindergarten children, 6th graders, and young adults, the 6th graders showed the highest

level of pitch-matching skill, but the young adults performed similarly to the kindergartners (Demorest & Pfordresher, 2015). This led the authors to suggest that the young adults' skills had deteriorated since late childhood. However, there is little research in singing development past school age.

Consistent with the child research, studies of untrained adults' singing have shown heterogeneity in singing skills, but with a broad consensus that the majority of adults can sing in-tune and in-time, at least in favorable circumstances (Dalla Bella et al., 2007; Wise, 2019). Even among adults with the music processing disorder congenital amusia—normally expected to be highly disruptive to performance—singing skill is highly heterogeneous (Anderson et al., 2012; Dalla Bella et al., 2009; Wise, 2009). Singing accuracy is not reliably predicted by musical perceptual acuity, such as pitch discrimination thresholds, and is understood as a complex sensorimotor skill, such as in the model delineated by Pfordresher, Halpern, and Greenspon (2015). However, one cognitive predictor of pitch matching accuracy is auditory imagery, as assessed by both self-report and objective measures (Pfordresher & Halpern, 2013); this makes sense given the need to have an accurate mental target to co-ordinate the motor planning process.

The "occasional singers" in the study by Dalla Bella et al. (2007) sang less accurately than the professional model, but only a small percentage were "poor pitch singers," defined by singing more than a semitone off-pitch on average. A subset of these were invited into the lab, and their pitch accuracy improved once they simply sang the song more slowly. However, the original volunteers agreed to sing solo for the researcher in a public park, and so would be unlikely to meet our criterion of general singing avoidance.

One issue in assessment of adult singing skill is the challenge in separating self-perceived skill from actual skill. Research has shown that adults are sometimes accurate in self-assessment (Wise, 2009; Wise & Sloboda, 2008) but this is not uniformly the case. For example, although all participants in Pfordresher and Brown's (2007) study claimed to be unable to match pitches, when tested only 13% of them met the criterion for poor pitch singer. Less is known about either the skills or the self-assessment accuracy of the

general population of singing-avoidant NSi. It is an open question whether self-assessment is more important in singing identity than actual skills.

Interventions/Practice

Our project involved an intervention in the form of singing lessons. Interestingly, few intervention or practice-based studies of adult singing focus on singing-specific outcomes—the development of singing skills and/or singing identity—as opposed to social or health outcomes, for which singing was simply a tool. Where singing-specific outcomes are reported, studies show that non-singers can develop more of a singing identity through both choral singing (Richards & Durrant, 2003) and lessons in individual and group settings (Numminen, 2014).

Aside from a few case-study accounts of teaching adult novices (e.g., Heresniak, 2004; see also review in Wise, 2019) there is little information available about singing skill development in NSi adults. Turøy (2018) aimed to investigate adult singing skill development in the context of Kindergarten teacher training, analyzing recordings submitted by trainee teachers for their examinations. The focus of the analysis, however, was on understanding poor-pitch singing rather than reluctance to sing, and not all participants received training; they were offered self-study materials and singing mentoring to improve skills, but up-take of the latter was very low. Furthermore, no baseline measure of singing skill was taken as the author deemed it unethical in the context. It was therefore not possible to draw conclusions about skill development. In studies designed as interventions, Numminen (2005, 2014) and Anderson et al. (2012) showed that non-singing adults' skills can improve with training. However, the participants cannot be taken to represent the wider population of singing-avoidant NSi. The participants in Anderson et al. (2012) were people with amusia. The trainee priests in Numminen's (2005; 2014) study, in common with the teachers in Turøy (2018), were people whose future professional roles demanded a certain level of singing competence; they could not avoid singing, and they had extrinsic motivation to develop their skills. Given the shortage of training studies and longitudinal data, but the evidence that training may be

beneficial for both skill development and singing identity, we reasoned that regular, individualized lessons, over a substantial time period, would enable the clearest insight into development of objective and subjective changes in the general population of NSi.

Current Study

The present study had two broad aims. Aim 1 was to gain a nuanced understanding of NSi identity via creation and validation of a new measure of Singing Inhibition, administered to a large sample of individuals interested in participating in our project. We characterize this as a measure of the strength of non-singing identity. Aim 2 was to chart the singing journeys of adults taking their first steps in singing since childhood. Specifically, we aimed to examine: a) NSi adults' vocal and musical skills, and how these might develop in training, b) quantifiable changes in key attitudinal, self-concept, and self-assessment factors, and c) the ways in which these attitudinal, self-concept, and self-assessment factors combine with objectively measured skills in comprising an emerging singing identity.

To assess skills (Aim 2a), the present study used a battery of established measures to comprehensively account for the multidimensional nature of singing as a complex sensorimotor skill. This incorporated measures of singing pitch accuracy and vocal facility to objectively assess vocal development, as well as measures of musical perception ability and of auditory imagery. We expected that improvements in vocal facility would be shown, given that singing lessons directly target these skills through technical instruction and research evidence suggests that non-singing adults' skills can improve with training. We left open the question of whether perceptual and auditory imagery skills would also change.

In relation to Aims 2b and 2c, based on previous research we identified four key attitudinal and self-concept aspects of non-singing identity that were quantifiable, and measurable with self-report scales: 1) belief in fixedness of singing ability, 2) Elitist Attitudes, 3) Sensitivity to Social Judgment, and 4) Self-perceptions of Voice. Given narrative accounts of absent or negative experiences of singing in childhood, which provides context for skill acquisition or lack thereof, we also included quantitative measures for

childhood experiences with singing, and for the extent to which singing took place in the childhood home. We expected that attitudes and beliefs would become more positive in general. We were further interested to: 1) examine how attitudes, beliefs, and self-perceptions may be related to objectively measured skills and, 2) whether self-assessment was as important, or perhaps even more important, than objectively measure skills in the journey to becoming a singer. We therefore also aimed to establish what specific components of skills, attitudes/beliefs, self-assessments, and past experiences best predict singing inhibition and singing participation.

Method

The project followed a funnel design, comprising an initial call for volunteers (Stage 1) followed by an invitation to submit more information from among those qualified (Stage 2). In Stage 3, a subset of 40 were invited for an in-person assessment (34 attended) and from among those, 20 were selected for the training project (Stage 4), with pre-and post-training assessments described below.

Stages 1 and 2 (Initial Cohort)

Participants from the London (UK) area were recruited via advertisements in local media and publicity to local community groups and employers. The call asked for adult volunteers who believed themselves to be non-singers and generally avoided singing, and who had not sung for most or all of their adult life, but who would like to explore learning to sing. Expressions of interest (Stage 1) were received from 352 individuals, who responded to items about their singing self-concept and behavior, and musical listening. These included two items designed to screen out potentially severe cases of amusia (severe inability to process pitch patterns): “I can only recognize songs I’ve heard before by the lyrics” and “Music sounds like toneless or unpleasant noise to me.” Participants also gave brief open responses describing their interest in the study. Participants whose responses showed they did not meet the definition of non-singer were screened out. Remaining participants were invited to respond to a background questionnaire that asked about singing participation and self-concepts, musical skills, and

training (Stage 2; more details below). Out of 252 questionnaires submitted, 238 were complete. This cohort comprised 171 females and 65 males (2 preferred not to answer) aged 19–75 (M = 43.64, SD = 13.58).

The background questionnaire (Stage 2; N = 238) covered self-reported musical skill and background, singing participation (not reported in this paper), and the newly created Singing Inhibition (SI) scale. The questionnaire was administered online. Data from the questionnaire were used to characterize the cohort, as well as to identify participants to invite to Stage 3.

Goldsmiths Musical Sophistication Index – Self-report Scales. The Goldsmiths Musical Sophistication Index (Gold-MSI, Müllensiefen et al., 2014) is designed to assess musical skills and training in the general population, incorporating self-report and behavioral tasks. The self-report scales of the measure were included here, comprising five subscales: Musical Engagement, Perceptual Skills, Singing, Emotional response to music, and Training. Questions from the Gold-MSI about general education, musical preferences, and occupational status were also included to inform selection of a balanced sample.

Singing Inhibition. The newly created Singing Inhibition measure comprised 18 items designed to reflect aspects of singing self-concept. Some items were taken from Zienkowska (2013), others newly created. These included behavioral statements (e.g., “I sing on a daily basis”; “I avoid singing”); global self-assessments (e.g., “I feel confident about my singing skill”; “My singing voice is not good enough”); singing identity (e.g., “It is in my nature to sing”; “I consider myself to be a non-singing individual”); and social dimensions (e.g., “I look forward to situations where I know there will be a lot of singing involved”; “I find singing with a group of people stressful”). Participants responded on a 5-point scale of *strongly disagree* (1) – *disagree* – *neither agree nor disagree* – *agree* – *strongly agree* (5). Positive items were reverse scored so that higher scores indicated stronger Singing Inhibition. See Supplemental Materials accompanying the online version of this paper at mp.ucpress.edu for the full scale.

Stage 3 (Candidates for Training Study)

For Stage 3 of selection we invited people to participate who: a) scored high in singing inhibition and/or low in self-reported singing skill, and b) did not report recent formal singing experience. We also maximized diversity in age, gender, and educational background. Forty people were invited to attend an in-person session with the first author for a pre-study singing assessment, comprising the following measures, and 34 completed all measures.

Perception and Auditory Imagery

Gold MSI-Melodic Memory and Beat Alignment Perception (Müllensiefen et al., 2014). In the Gold-MSI Melodic Memory test, participants listen to pairs of novel melodies, then identify whether the second melody is the same as or different to the first and give a confidence rating for their judgment. The second item is always transposed, i.e., presented in a different key to the first, so that the absolute pitches are different even when the melody is the same. Participants therefore must remember the interval structure of the melody.

In the Gold-MSI Beat Alignment Perception (BAT) test, participants listen to short melodies presented with a click track sounding an even pulse. Click tracks either match the tempo of the melody, so that the pulses align with the musical beats, or are paced slightly slower or faster, so that they occur out of time with the melody. Participants indicate whether the clicks were in or out of time and give a confidence rating.

Bucknell Auditory Imagery Scales (BAIS) (Halpern, 2015). The BAIS comprises items asking participants to imagine a variety of environmental, verbal, and musical sounds. In the Vividness scale, participants rate how vividly they are able to imagine the given sound on a scale of 1 (*no image present at all*) to 7 (*as vivid as the actual sound*). For example:

For the next item, consider attending classes.

The slow-paced voice of your English teacher.

In the Change scale, participants are asked to first imagine a sound, then to imagine that sound changing in some way. They then rate the ease with which they could change their auditory image on a scale of 1 (*no image present at all*) to 7 (*extremely easy to change the image*). For example:

> For the next pair, consider the beginning of the song "Happy Birthday."
>
> a. The sound of a trumpet beginning the piece.
>
> b. The trumpet stops and a violin continues the piece.

Vocal Skills

Vocal Range. Vocal range and pitch direction control measures were taken from Wise and Sloboda (2008) and Wise (2009). These tasks were performed after spoken exercises in which a comfort pitch was identified for each participant and used as the starting point for subsequent tasks. Two measures of vocal range were used: Slides and Discrete pitches.

1) Slides. Starting on a comfort pitch, and using the syllable /na/, participants were asked to glide their vocal pitch smoothly downwards as far as possible, and as far upwards as possible. Each was performed twice. The researcher demonstrated a glide. Vocal outputs were analyzed with Praat. The pitches of the highest and lowest measurable waveforms at the end of each glide were extracted, omitting vocal fry (the characteristically creaky, rough, low-pitched vocal sound). The average of the two attempts at each glide was taken.
2) Discrete pitches. Starting on a comfort pitch, and using the syllable /na/, participants were asked to sing single pitches downwards and upwards as far as comfortable, guided by the researcher playing tones on a keyboard. There was no requirement for participants to match the pitches played. The researcher began by playing notes according to a major scale, moving by semitone once it became apparent that the participant's limit might be nearing. The mean pitch of the

highest and lowest sustained pitches were extracted using Praat (Boersma & Weenink, 2019), selecting the steady portion of each pitch (excluding onset and offset).

Non-imitative Vocal Pitch Control. This task was an assessment of voluntary manipulation of vocal pitch without the need for pitch matching. Each time, the participant was asked to start on their comfort pitch (the researcher reminded them of this pitch using the keyboard) and sing on the syllable /na/. The five stages of the task were:

1. Sing the comfort pitch and repeat it, that is, sing /na/ twice, keeping the sound the same
2. Sing the comfort pitch followed by a higher sound, that is, go up on the second /na/
3. Repeat Step 2, making a bigger difference between the pitches
4. Sing the comfort pitch followed by a lower sound, that is, go down on the second /na/
5. Repeat Step 4, making a bigger difference between the pitches

Up to three attempts were allowed, and participants scored one point for each stage successfully accomplished (maximum = 5).

Vocal Pitch Accuracy

Seattle Singing Accuracy Protocol. The Seattle Singing Accuracy Protocol (SSAP) (Pfordresher & Demorest, 2020) is a battery yielding multiple automated measures of pitch imitation accuracy, song singing accuracy, and a pitch discrimination task. The test scores reported in this paper represent pitch imitation accuracy, and familiar song accuracy. For the familiar song, participants are given choice of well-known Western songs (e.g., “Twinkle Twinkle Little Star”) and sing unaccompanied at a pitch of their choosing. The song is sung as a warm-up, and again later in the protocol both with words and without, on the syllable /doo/. The pitch imitation tasks comprise single vocal tones and piano tones, and four-note patterns (vocal tones). Stimuli are presented at a pitch level determined by the participant’s comfort pitch, and using a synthesized vocal model of the same gender. Participants hear each pitch or pattern once, then imitate it unaccompanied.

Supported Imitation. A pitch accuracy task in supported singing (Wise, 2009; Wise & Sloboda, 2008) was also included. This involves a similar imitation task to the SSAP using vocal stimuli, with single tones and patterns of 2, 3, and 5 pitches. However, after hearing the stimulus, the participant sings along with a repetition of the stimulus. Stimuli are voiced by a model of the same vocal gender, and the pitches lie within the average comfortable untrained adult range.

Vocal Self-assessments: Accuracy, Quality, Physical Discomfort

These measures comprised a set of descriptors, each of which participants rated on a seven-point scale for the extent to which it applied to their own singing, from 1 (*not at all*) to 7 (*completely*). A "don't know" option was also given. Example items for Accuracy: tuneful; pitch-perfect; flat. Negative items were reverse scored. For Quality, participants were asked to rate the descriptors "regardless of whether you sing the right notes". Example items for Quality: clear; strong; wobbly; breathy. Negative items were reverse scored. For the Physical Discomfort measure, participants were asked to rate the descriptors based on their "physical sensations while singing, regardless of how it sounds". Example items for Physical Discomfort: relaxed, tight (in the throat), tight (elsewhere in the body), short of breath. Positive items were reverse scored.

Experiences and Attitudes

Five self-report scales were created or adapted from previous studies to measure facets of past experiences and attitudes towards singing. All scales not previously published can be found in the Supplemental Materials.

Childhood Experiences. The Childhood Experiences measure comprised 14 items designed to indicate how positive or otherwise participants' memories of singing in childhood are. Items covered opportunity (e.g., "Singing activities were frequently available to me as a child"); encouragement or silencing (e.g., "I was told not to sing when I was a child"); and emotion (e.g., "I always felt embarrassed when singing at school"). The response scale was *Completely disagree* (1) – *Strongly disagree* – *Disagree*

– *Neither agree nor disagree* – *Agree* – *Strongly agree* – *Completely agree* (7). Negative items were reverse scored so that higher scores represent a more positive childhood experience. Some items in the Childhood Experiences scale—as well as the Parental and Family Engagement, and Sensitivity to Social Judgement scales described below—were taken from Zienkowska (2013).

Parental and Family Engagement. Four items concerned how often singing occurred in participants' childhood family activities at home, for example, "I often heard my parents sing"; "We hardly ever sang together in my family". The response scale was the same as for childhood experiences. Negative items were reverse scored so that higher scores indicated higher parental and family engagement.

Sensitivity to Social Judgment. This scale consisted of 12 items, for example "I avoid singing when others might hear me"; "I rarely worry that people will dislike my singing". The response scale was the same as for childhood experiences. Positive items were reverse scored so that higher scores indicated higher sensitivity to social judgment.

Elitist Attitudes (Bento-Allpress, 2013). The Elitist Attitudes Scale included 23 items designed to capture the extent of participants' endorsement of an inclusive versus an elitist account of singing, for example, "We are all singers"; "A singer is someone who is born with incredible musical talent and vocal power." Statements also included those relating to a talent account of singing, e.g., "Only a small percentage of people in society are able to sing well"; "Any speaking voice can become a good singing voice"; "It's unacceptable for someone to sing in public unless they have a very good voice." The response scale was the same as for childhood experiences. Positive items were reverse scored so that higher scores indicated stronger endorsement of elitist attitudes.

Implicit Theory of Singing Ability. This scale is an adaptation of Dweck's (2000) Theories of Intelligence Scale – Self form for Adults, substituting singing ability for intelligence. It measures the extent to which participants believe singing ability is fixed (an entity belief) versus malleable (an

incremental belief). In line with Dweck's (2000, p. 176) advice for using the scale in longitudinal studies, in which respondents may drift towards higher endorsement of the more attractive incremental items, we chose to use the entity-only form of the scale. For example: "You have a certain amount of singing ability, and you can't really do much to change it." However, we thought the close similarity in wording among the four entity items might discourage participants from considering their responses, so we retained one incremental item: "You can always substantially change how able you are in singing." This resulted in five items, with the response scale: *Strongly disagree* (1) – *Disagree* – *Disagree somewhat* – *Agree somewhat* – *Agree* – *Strongly agree* (6). The one incremental item was reverse scored, so that higher scores indicate a stronger belief that singing ability is fixed. Item scores were totaled, giving a possible range of 5 to 30.

Stage 4: Training

The Training Group (n = 20), chosen from among the 36 Stage 3 participants, comprised 11 females and 9 males, aged 23–71 (M = 47.85, SD = 13.92). They were chosen to represent diversity in age, initial level of singing skill, and musical and educational background. It was also ascertained that all participants could judge the direction of pitch changes within a threshold of 100 cents, as measured by the Seattle Singing Accuracy Protocol (SSAP, Pfordresher & Demorest, 2020). Practical considerations such as location and availability were also taken into account, to maximize likelihood that participants would be able to sustain involvement in the study.

The Training Group received a combination of individual singing lessons and group lessons delivered over three terms (approximately 12 weeks per term, over 10 months). Eight 40-min individual lessons and two 2-hour group lessons were offered in each of the first two terms. In Term 3, 10 group sessions were offered, and a maximum of five individual lessons. The project finished with an informal, private, celebration workshop co-designed by teachers and participants. Fifteen participants completed the entire duration of the study, participating in up to 14 group sessions and 16–21 individual lessons.

Each participant was paired with one of ten teachers for individual lessons. Five of the teachers took responsibility for group sessions, co-teaching in pairs. Teachers taught in classical, jazz, and drama fields, with institutional positions in tertiary or adult education, and/or private studios. All had particular interests and experience in teaching adult beginners. Teachers were free to follow their own teaching approaches and styles; however, all participants received a combination of vocal technique, musical skill-building, song learning, and creative activities, designed by the teachers to meet individual and group needs and capabilities.

All pre-study assessment measures from Stage 3—i.e., self-report and vocal tasks—were repeated after six months' participation, with the exception of Childhood Experiences and Parental and Family Engagement. Vocal tasks (vocal range, pitch direction control, vocal pitch imitation and song) were additionally repeated after 10 months' participation. All statistical analyses were carried out in JASP (JASP Team, 2022).

Results

Self-Assessed General Musical Ability (Stage 2 and Training Group)

The Goldsmith's Musical Sophistication Index (Gold-MSI v. 1.0) self-report subscales (five factors, plus a General factor) were used to characterize the general musical ability of the initial cohort (n = 237). Comparison of the profile of this cohort with published norms for the general population (Müllensiefen et al., 2013, 2014) showed that non-singers' mean scores fell in the lowest quartile of the norm distribution (Table 1), with particularly low means for both Perceptual Abilities and Singing Abilities (both around the 10th percentile). The range of scores was, however, wide, with some scoring fairly high on each factor. Scores were approximately normally distributed except in, as expected, Musical Training, where the distribution was skewed due to a large proportion of low scores.

Table 1. *Descriptive Statistics for Initial Cohort (n=237) on Six Self-report Subscales of the Gold-MSI and Comparison with Published Norms*

Gold-MSI Factor	Scale Min/Max	Non-singers' Min/Max	Non-singers' *M* (*SD*)	Norm *M* (*SD*)	Norm percentile of non-singers' mean
1. Active Engagement	9 / 63	9 / 60	33.84 (9.97)	41.52 (10.36)	20-21
2. Perceptual abilities	9 / 63	19 / 57	40.54 (7.45)	50.20 (7.86)	9
3. Musical Training	7 / 49	7 / 40	15.70 (8.05)	26.52 (11.44)	21-22
4. Singing Abilities	7 / 49	7 / 42	20.05 (7.03)	31.67 (8.72)	10
5. Emotions	6 / 42	14 / 43	31.60 (4.59)	34.66 (5.04)	20-25
6. General	18 / 126	22 / 84	53.49 (12.74)	81.58 (20.62)	9

Gold-MSI data for the 20 people in the Training group are provided in the Supplemental Materials. The percentile scores for the Training group were a little lower than those for the initial cohort as a whole, with the exception of Perceptual Abilities (which was the same) and Emotions (where the Training Group's percentile was higher).

Singing Inhibition Validation and Structure (Stage 2)

Following internal consistency analysis on complete responses from the Stage 2 cohort (n = 234), three items were removed that did not affect the internal consistency of the scale ("My singing voice is not good enough"; "I find singing in a group a wonderful experience"; "I am not a singing type of person"). This left a final scale of 15 items with a Cronbach's α value of 0.896.

The mean Singing Inhibition score from the Stage 2 cohort was 5.06 (0.91), range = 2.44 to 7. In the Training Group (n = 20), M = 5.63 (0.90), range = 3.50 to 6.94 (higher scores = higher Singing Inhibition). In order to further validate the Singing Inhibition Scale, we examined its correlation in the

Stage 2 cohort with the subscales of the Gold-MSI. In particular, negative correlation with the Singing Abilities factor would provide an indication of external construct validity. As expected, there was a significant negative correlation between Singing Inhibition and (self-reported) Singing Abilities. Singing Inhibition was also negatively correlated with General Musical Sophistication, and significantly but more weakly with Perceptual Abilities and Active Engagement (Table 2). Given the wide age range of the Stage 2 group we checked whether these scale measures were correlated with age, finding that only Active Engagement decreased as age increased (r = -.20, p = .008).

Table 2. *Pearson's Correlations of Singing Inhibition with Each Factor of the Gold-MSI*

	1. Active Engagement	2. Perceptual Abilities	3. Musical Training	4. Singing Abilities	5. Emotions	6. General
Singing Inhibition	*-.16	**-.17	-.09	***-.56	-.07	***-.44

*p < .05; **p < .01; ***p < .001

We also examined the factor structure of the SI scale (see Supplementary Materials), identifying three factors as follows:

Factor 1. Perceived Ability (global judgment). This comprises statements related to participants' overall sense of their level of skill, for example, "I believe I am a good singer"; "I feel confident about my level of singing skill."

Factor 2. Singing Self. This factor expresses how much singing is part of a person's daily life and sense of self, for example, "I think I've always spent a lot of time singing"; "It is in my nature to sing."

Factor 3. Social Singing. This factor relates primarily to singing with other people and in groups, for example, “If I ever get the chance to sing with other people, I do” and “I find singing with a group of people stressful.”

Correlations between these factors and the factors of the Gold-MSI are provided in the Supplementary Materials. In particular, we note that all three SI factors were significantly correlated with Gold-MSI Singing Abilities, as well as with the General Factor. There were different patterns of relationship among the three SI factors with the remaining Gold-MSI factors, suggesting the value of the SI factors for capturing aspects of NSi identity not reflected in the Gold-MSI.

Training Group (Stages 3 and 4)

We now turn to results from the small training group, comparing pre- to post-training. We first examine auditory imagery, followed by objectively measured skills: perceptual tasks of the Gold-MSI, and vocal tasks from simple to complex. We also include a self-assessment of vocal quality, accuracy, and sensations.

For measures repeated at 6 and 10 months, one-way related ANOVAs were used to analyze changes in skills from baseline. Since we made a prediction of improvement from baseline but not specifically differentiating time points 2 and 3, we carried out contrasts comparing baseline to the combination of times 2 and 3 (as given by Helmert contrasts). As we sometimes had incomplete data from Training participants (due for example to illness or technical difficulties) this allowed a more robust analysis.

Auditory Imagery

The Bucknell Auditory Imagery scale (BAIS) was administered at baseline and at six months. Fourteen participants completed the measure at both timepoints. Student’s related *t*-tests showed no significant change after six months in either the vividness (BAIS-V) or ease of change (BAIS-C) scores,

BAIS-C: $t(13) = -0.24$, $p = .82$; BAIS-C: $t(13) = 1.57$, $p = .14$ (Table 3). This stability is as expected for a trait-based construct.

Table 3. *Mean BAIS Scores at Baseline and Time 2 (Six Months)*

	Baseline M *(SD)*	Time 2 M *(SD)*
BAIS-V	4.83 *(0.57)*	4.89 *(1.09)*
BAIS-C	5.02 *(0.92)*	4.49 *(1.57)*

Musical Perception

Descriptive statistics for the training group on the Gold-MSI Melodic Memory and Beat Alignment Perception (BAT) tests are show in in Table 4, along with published norm data. We report Area Under Curve (AUC) scores, which indicate accuracy taking account of participants' self-rated confidence in their judgments (Müllensiefen et al., 2013). At baseline, the training group's mean Melodic Memory performance was between the 16th and 20th percentiles, while their BAT mean scores fell between the 36th and 40th percentiles. Thus, as a group they performed less well than average, but better in beat than melody perception. The range of scores was large, although with the exception of beat perception AUC, no individual obtained a score above the 80th percentile.

Table 4. *Descriptive Statistics for Gold-MSI Melodic Memory and Beat Alignment Perception in Comparison to Published Norms*

	Scale Min-Max	Training group's Min-Max	Norm Mean *(SD)*	Baseline Mean *(SD)*	Norm percentile of baseline mean	Time 2 Mean *(SD)*	Time 3 Mean *(SD)*

Melodic memory AUC	0.5-1	0.50-0.90	0.79 (*0.17*)	0.63 (*0.12*) **0.63 *(0.13)***	16-20	0.69 (*0.09*) **0.69* *(0.08)***	0.70 *(0.15)* **0.72* *(0.15)***
Beat perception AUC	0.5-1	0.50-1.00	0.74 (*0.13*)	0.70 (*0.16*) **0.72 *(0.17)***	36-40 41-45	0.72 (*0.15*) **0.73 *(0.16)***	0.68 *(0.18)* **0.69 *(0.18)***

Statistics for the Training Group for whom data are available at all three timepoints (n = 13) are shown in **bold**. **p* < .05, T2 and T3 combined, in comparison to baseline.

Melody memory improved over the course of the study. A related one-way ANOVA showed a significant change in the AUC scores, $F(2, 24) = 3.070$, $MS(\text{error}) = 0.008$, $p = .033$. Helmert contrasts showed a significant difference between baseline melody AUC and the two later times: $t(24) = 2.410$, $p = .012$ in the expected direction. According to the published norms, the group's mean AUC scores climbed to the 26–35 percentile range. There were no significant changes in beat perception, $F(2, 24) = 0.342$, $MS(\text{error}) = 0.013$, $p = .714$.

Vocal Skills

Range. The mean vocal range measured by slides increased (Figure 1); a one-way related ANOVA was marginally nonsignificant, $F(2, 16) = 2.90$; $MS(\text{error}) = 9.88$, $p = .08$, $\eta^2 = .27$. However, Helmert Contrasts showed a significant increase between Time 1 ($M = 22.24$, $SE = 1.50$) and the overall mean of Time 2 ($M = 25.08$, $SE = 1.68$) and Time 3 ($M = 25.54$, $SE = 1.65$), $t(16) = -2.39$, $p = .03$.

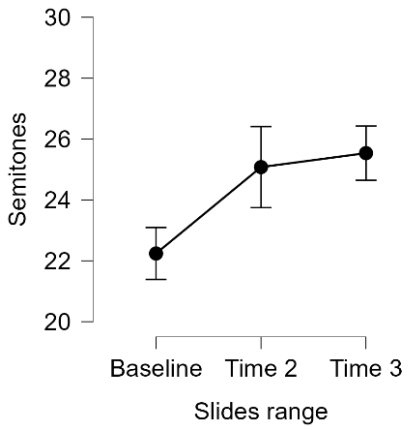

Figure 1. *Slides range means in semitones at three time points (error bars show SE)*

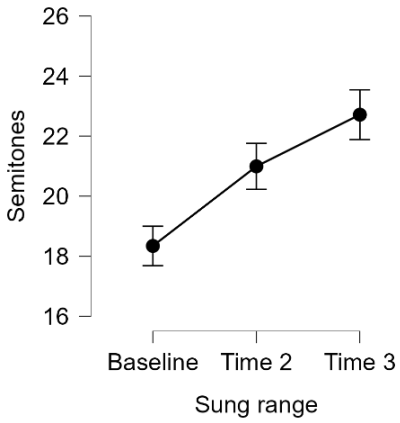

Figure 2. *Sung range means in semitones at three time points (error bars show SE)*

The increase in sung vocal range was significant (Figure 2), as shown by a one-way related ANOVA, $F(2, 18) = 8.54$, $MS(\text{error}) = 5.68$, $p = .002$, $\eta^2 = .49$. Helmert contrasts showed a significant difference between Time 1 ($M = 18.34$, $SE = 1.17$) and the overall mean of Time 2 ($M = 20.99$, $SE = 1.24$) and Time 3 ($M = 22.71$, $SE = 1.09$); $t(18) = -3.80$, $p = .002$. On average, participants expanded their sung range by just over 4 semitones by Time 3.

Non-imitative Vocal Pitch Control. A one-way related ANOVA with Helmert contrasts showed that participants' voluntary (non-imitative) control of pitch direction and size of change in the Up & Down task significantly improved from baseline, $F(2, 18) = 14.22$, $MS(\text{error}) = 0.29$, $p < .001$, $\eta 2 = .61$ (Figure 3). Times 2 ($M = 3.8$, $SE = 0.29$) and 3 ($M = 4.8$, $SE = 0.08$) combined were significantly different from baseline ($M = 3.6$, $SE = 0.28$); $t(18) = -3.352$, $p = .004$. In addition, Time 3 was significantly different from Time 2, $t(18) = -4.147$, $p < .001$. A note of interest is that errors in this task often resulted from perseveration when participants were asked to go from singing an upward change to a downward one. There were no significant changes in consistency of starting pitch in the Up & Down task, $F(2, 18) = 0.03$, $p = .97$ (Baseline $M = 3.2$, $SE = 0.66$; T2 $M = 3.25$, $SE = 0.62$; T3 $M = 3.10$, $SE = 0.55$).

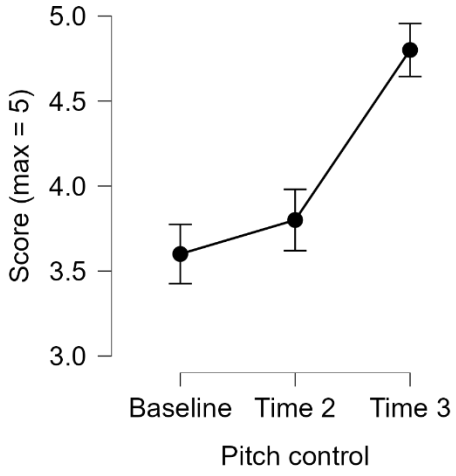

Figure 3. *Voluntary control of pitch direction and size of change (error bars show SE).*

Vocal Pitch Imitation Accuracy and Supported Imitation. Descriptive statistics for overall imitation accuracy (absolute values of deviation from target in cents) across SSAP tasks and supported imitation tasks are shown in Table 5 (note, lower numbers mean more accurate performance). Since data were positively skewed, a Log10 transformation was performed for inferential statistics; the transformed means and *SD*s are also shown in Table 5.

Table 5. *Overall Imitation Accuracy, Log10 Transformed and Original Data (Cents)*

	Baseline Mean (*SD*)		Time 2 Mean (*SD*)		Time 3 Mean (*SD*)	
	Transformed	Original	Transformed	Original	Transformed	Original
SSAP overall imitation	1.98 *(0.43)*	138.93 *(114.51)*	1.85 *(0.43)*	106.11 (93.15)	*1.75 *(0.43)*	91.80 *(102.27)*
Supported overall imitation	2.09 *(0.37)*	158.03 *(144.47)*	2.04 *(0.42)*	147.57 *(122.19)*	2.05 *(0.43)*	165.95 *(156.69)*

*Significant change from baseline

A one-way related ANOVA (on transformed data) showed a marginally nonsignificant change over time, $F(2, 20) = 2.48$, $MS(\text{error}) = 0.06$, $p = .109$, $\eta^2 = .20$. Contrasts showed a significant improvement in accuracy between baseline and T3, $t(1, 10) = 3.39$, $p = .038$; deviations decreased from an average of almost 1.4 semitones to under a semitone by Time 3. Overall imitation accuracy in supported tasks (singing along) did not significantly improve, $F(2, 18) = 0.08$, $MS(\text{error}) = 0.03$, $p = .46$.

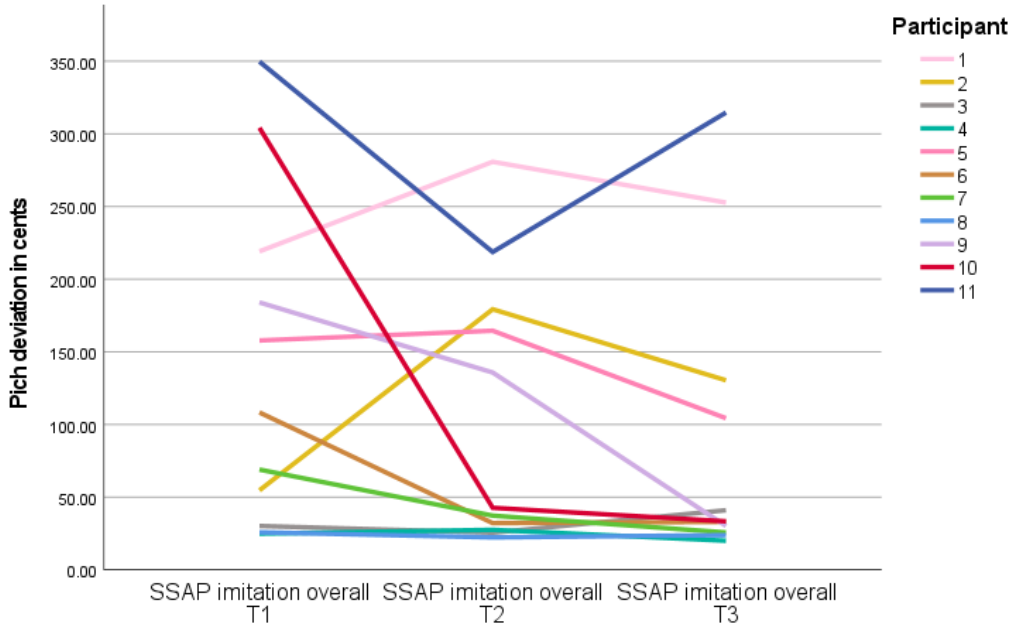

Figure 4. *Individual changes in overall SSAP pitch imitation deviations. Each line is a participant. Lower scores mean more accurate performance.*

Analyzing performance on separate imitation tasks was not informative. This is likely due to the wide variability in participants' performance, an observation that is consistent with previous studies of adult novices and poor-pitch singers (e.g., Dalla Bella et al., 2007; Pfordresher & Brown, 2007; Wise, 2009). Furthermore, there were individual differences in trajectories of change (Figure 4). Although the majority of participants showed an overall improvement, some participants' pitch accuracy worsened at either Time 2 or 3.

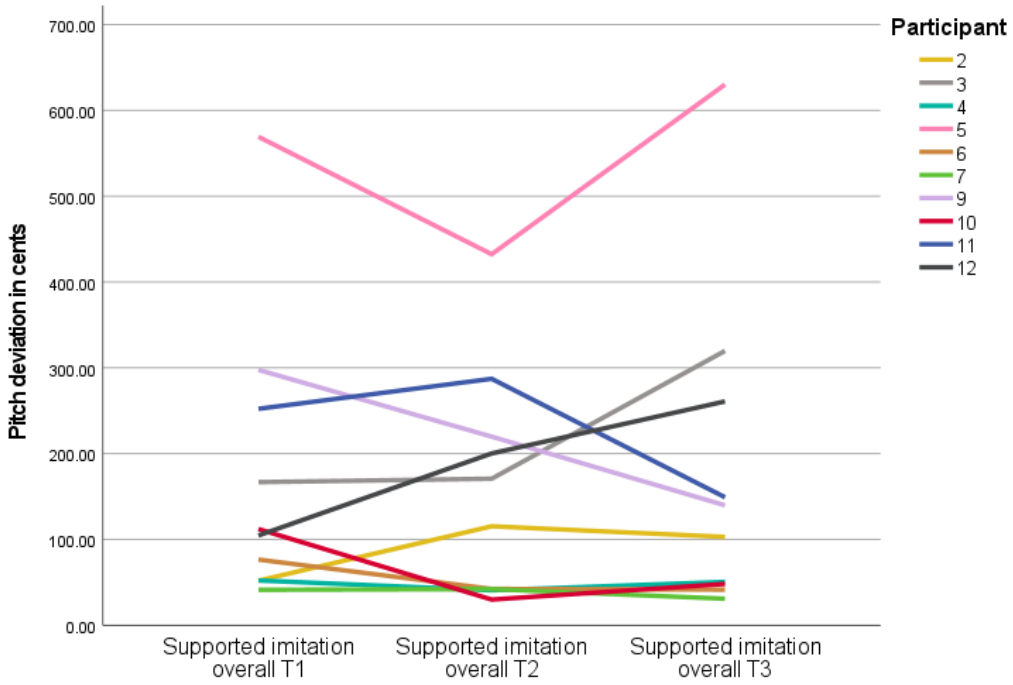

Figure 5. *Individual changes in overall supported imitation deviations. Each line is a participant. Lower scores mean more accurate performance.*

Familiar Song Accuracy. Familiar song singing accuracy estimates, expressed as a percentage, were: baseline *M* = 84.9 (*SD* = 10.59), T2 *M* = 88.25 (*SD* = 9.56), T3 *M* = 90.31 (*SD* = 7.10). For the same songs sung without words to the syllable /doo/, accuracy estimates were: baseline *M* = 87.53, (*SD* = 11.75), T2 *M* = 88.42, (*SD* = 8.71), T3 *M* = 87.46, (*SD* = 12.37). One-way related ANOVA showed no significant change over time in familiar song accuracy either when singing with words, $F(2, 18) = 0.47$, *MS*(error) = 75.881, $p = .63$, or without words, $F(2, 18) = 0.58$, *MS*(error) = 100.03, $p = .57$.

Vocal Self-assessments: Accuracy, Quality, Physical Discomfort

Related *t*-tests showed there were significant differences between baseline and six months in participants' self-assessment of voice on all three measures (Table 6). Self-perceived accuracy significantly increased, $t(1, 14) = -6.16$, $p < .001$. The effect size, as measured by Cohen's *d*, was -1.53,

indicating a very large effect. Self-perceived vocal quality ratings also increased, $t(1, 14) = -2.91$, $p = .011$, $d = -0.86$ (a large effect). Physical discomfort significantly decreased, $t(1,14) = 3.11$, $p = .008$, $d = 0.73$ (a medium effect). Thus, after six months of training, participants perceived improvements in all three areas.

Table 6. *Vocal Self-assessments*

	Baseline Mean *(SD)*	Time 2 Mean *(SD)*
Accuracy	2.42 *(0.98)*	***4.30 *(3.04)*
Quality	3.30 *(4.57)*	*4.31 *(2.53)*
Physical discomfort	4.17 *(1.23)*	**3.03 *(0.94)*

*$p < .05$; **$p < .01$; ***$p < .001$

Past Experiences, Attitudes, and Beliefs

Turning now to participants' singing-related history and attitudes, we first report descriptive statistics for early experiences (measured only once) before examining changes in attitudinal measures.

Early Experiences. Early experiences with singing were measured with the Childhood Experiences and the Parental/Family Engagement scales. Possible scores ranged from 1–7, with higher scores representing more positive childhood experiences and greater parental/family engagement in singing. Descriptive statistics were $M = 2.58$ ($SD = 1.84$) for Parental/Family Engagement and $M = 2.94$ ($SD = 1.01$) for Childhood Experiences. Both means lie within the lower half of the available scale, implying relatively negative experiences and low levels of family-social engagement with singing during childhood.

Singing-related Attitudes and Beliefs. Singing-related attitudes and beliefs were measured at baseline and again at six months, with complete data from 15 participants. Internal consistency for each scale was ascertained using data from all 34 respondents at baseline, with Cronbach's alpha values as

follows: Childhood Experiences: α = .88; Parental/family engagement: α = .89; Sensitivity to Social Judgment: α = .90; Elitist Attitudes: α = .96; Implicit Theory: α = .90. Descriptive statistics for the three scales are shown in Table 7.

Table 7. *Means and Standard Deviations for Measures of Singing-related Attitudes and Beliefs at Times 1 and 2 (n = 15)*

Measure (Possible scores)	**Time 1: M *(SD)***	**Time 2: M *(SD)***
Sensitivity to Social Judgment (1-7)	5.09 *(1.00)*	**4.37 *(1.12)*
Elitist Attitudes (1-7)	3.94 *(0.87)*	**3.29 *(0.63)*
Implicit Theory (5-30)	13.62 *(6.15)*	*9.69 *(2.50)*

*Significantly different $p < .05$. **Significantly different $p < .005$

Significant improvements with medium to large effect sizes were seen in singing-related attitudes and beliefs at Time 2 in comparison to baseline, with Fear of Social Judgment ($t = 3.74$, $p = .002$, $d = 0.97$) and Elitist Attitudes ($t = 3.72$, $p = .002$, $d = 0.92$) reducing, and Implicit Theories becoming less fixed ($t = 2.77$, $p = .017$, $d = 0.77$).

Relationships Between Singing Inhibition and Other Measures

We examined the relationship between Singing Inhibition and the other self-report and behavioral measures, to further characterize non-singing identity. All the participants with complete baseline assessment data ($n = 34$) were included, in order to increase power.

In the Supplemental Materials we provide correlational data on the relationships between Singing Inhibition (the overall measure and the subscales) and the behavioral measures of musical perception and vocal skills. None of these relationships was significant. There was a correlation approaching significance between SI-Social subscale and melodic memory accuracy, such that more accurate melodic memory might be related (perhaps counterintuitively) to greater singing inhibition in social situations.

Singing Inhibition scores were, however, significantly related to a number of other self-report measures (Table 8). Overall SI was related to each of the attitudinal measures—Sensitivity to Social Judgment, Elitist Attitudes, and Implicit Theory—in the expected direction. At the subscale level, only SI-Social was associated with an increase in Sensitivity to Social Judgment, while both SI-Social and SI-Ability were related to Elitist Attitudes and Implicit Theory. The SI-Singing Self subscale remained unrelated to the three attitudinal and belief measures.

Looking at past singing experiences, strikingly, Childhood Experiences did not systematically relate to any aspect of SI, while Parental and Family Engagement was significantly negatively correlated with all SI scores except for Ability. Finally, self-assessments of vocal accuracy and physical discomfort were not related to any of the SI measures. Self-assessments of vocal quality were, however, significantly correlated with SI-Social, and narrowly missed being significantly correlated with SI-Ability and overall SI. Thus, participants appear to be distinguishing between vocal quality and accuracy, with vocal quality related to singing inhibition.

We provide further correlational data in the Supplemental Materials. We note that, mirroring the relationship between voice quality and SI, voice quality and Gold-MSI singing ability were also significantly correlated. We also note the general lack of a relationship between self-ratings of voice and objective performance measures.

Table 8. *Pearson's Correlations of Singing Inhibition (SI) with Baseline Self-report Measures*

	Singing related attitudes and beliefs			*Past singing experiences*		*Vocal self-assessments*			*BAIS*	
	Sensitivity to social judgement	Elitist attitudes	Implicit theory	Parental/ family engagement	Childhood	Accuracy	Quality	Physical discomfort	Vividness (V)	Change (C)
SI-Overall	.39*	.50**	.48**	-.51**	-.22	.10	-.30[b]	.17	-.10	-.15
SI-Ability	.29	.54**	.46**	-.27	-.22	-.12	-.34[a]	-.05	-.28	-.34
SI-Singing self	.20	.19	.28	-.44*	-.19	.21	-.07	.29	.04	-.02
SI-Social	.55***	.56***	.42*	-.48**	-.20	.06	-.36*	.03	.05	-.12

*$p < .05$, **$p < .01$, ***$p < .001$, [a]$p = .052$, [b]$p = .083$

Predictors of Singing Inhibition

To further examine attitudes, beliefs, and past experiences in relation to non-singing identity, we identified predictors of Singing Inhibition. Multiple linear regression was used to test whether Singing Inhibition could be predicted by a combination of other variables. To maximize the data available for this analysis we included all Stage 3 participants (N = 34). A backwards stepwise method was used, which enters all predictors into the model and removes those not making a significant contribution, one by one. This is less likely to erroneously exclude contributing variables than the forward stepwise method (Field, 2005). SI-Overall was the outcome variable, with predictor variables of Sensitivity to Social Judgment, Implicit Theory, Elitist Attitudes, Parental and Family Engagement, and Gold-MSI Singing Abilities, chosen because they were previously shown to be significantly correlated with SI.

This process resulted in a model with three predictors. The fitted regression model was: Singing Inhibition = 5.92 + 0.36*(Implicit Theory) + 0.34*(Parental and Family Engagement) + 0.31*(Gold-MSI Singing Abilities). The overall regression was statistically significant, R^2 = 0.48, $F(3, 30) = 9.16$, p = .0002. Each predictor independently made a significant contribution to the model as follows: Implicit Theory, $\beta = 0.36$, p = .01; Parental and Family Engagement, β = -0.34, p = .02; Gold-MSI Singing Abilities, β = -0.31, p = .03.

Thus, Singing Inhibition was significantly predicted by a combination of belief in the fixed nature of singing ability, experience of singing activities in the childhood home, and self-assessed skill, with each making a significant contribution. We present this analysis with caution given the relatively small sample size but offer it as theoretically interesting.

Differences Between Completers and Non-Completers of the Training

We also examined differences between participants who completed the course and those who withdrew. We included this in light of our definition of non-singers as “people who do not sing,” a status

that was immediately challenged simply by the fact of participation in the course. Withdrawal from the course may have signalled a return to non-singing behavior.

Since a small number of participants chose to discontinue the study partway through, or to continue in the study but not to participate in group sessions, participants were classified as having completed the training if they participated in group sessions until the end of the study duration, including one or both final workshop sessions. We included one participant who withdrew soon after training began and was replaced, leaving 14 participants who completed the singing training including group sessions, and seven who did not complete.

We carried out group comparisons using Mann Whitney's *U*, on selected key baseline measures: Singing Inhibition, Implicit Theory, Sensitivity to Social Judgement, Parental/family Engagement, Gold-MSI Singing Abilities, self-assessed vocal quality, and overall vocal imitation accuracy. A significant difference was found in Implicit Theory, $U(19) = 19.00$, $p = .013$, in the expected direction, such that non-completers (Median = 18) held a stronger belief that singing was a fixed skill than completers (Median = 10.5). Non-completers also showed significantly higher Sensitivity to Social Judgment (Median = 6.17), than completers as expected (Median = 5.29), $\underline{U}(19) = 24.50$, $p = .036$. There were no other significant differences.

Discussion

Dimensions of the Non-singing Identity

We began with a behavioral definition of non-singer as someone who generally avoids singing and has not sung for most or all of their adult life (except perhaps in the shower). This population, arguably a sizeable proportion of Western adults, has largely been absent from both practical singing initiatives and research literature. We therefore aimed for a more nuanced understanding of the dimensions of a non-singing identity, and to that end developed and validated a measure of Singing

Inhibition. Singing Inhibition comprised three components that we characterized as: A) “sense of competence”; B) “singing self,” that is, how much singing is part of the respondent’s everyday life; and C) “social,” relating to the context of singing (e.g., in a group). This suggests there are private and public dimensions of singing inhibition combined with a general sense of (a lack of) singing competence.

As expected, Singing Inhibition correlated negatively with self-reported Singing Ability on the Gold-MSI: People who reported more inhibition about singing also reported lower ability. Singing Inhibition also correlated negatively with General Musical Sophistication and more weakly, but still significantly, with reported Perceptual Abilities and Active Engagement. However, significant additional variance in Singing Inhibition remained, suggesting that Singing Inhibition is conceptually distinct from Sophistication as defined in the Gold-MSI. We distinguish the sense of competence from reported singing ability, as items of the Gold-MSI Singing Ability scale refer to specific tasks, such as, “I am able to hit the right notes when I sing along to a recording,” and this scale correlates with objectively measured performance. In contrast, the items in the “sense of competence” component of the Singing Inhibition scale reflect a global concept of being good or bad at singing, and this scale does not correlate with objectively measured performance.

Turning to the Gold-MSI as a measure of musical sophistication, participants’ average levels of musicality fell below adult norms across all dimensions of the Gold-MSI questionnaire, particularly in singing and perception factors, as well as in objectively measured melodic memory and beat perception. In contrast, Numminen (2014) reported musicality in non-singers to be in the average range, although using different measures. The Gold-MSI conceives of musical sophistication more broadly than traditional musicality measures and has a high proportion of self-report, which could account for some differences from Numminen’s study. However, it is also possible that the original Gold-MSI research project (Müllensiefen et al., 2014) disproportionately attracted people with musical interests, despite attempts to mitigate this. And in contrast, our recruitment was designed to attract people who were

particularly dubious about their skills, albeit with an interest in improving them. Thus, we conclude that non-singing identity includes private and public aspects of singing inhibition, and likely a profile of lower engagement with music and less successful music making in relation to the general population.

Singing Development

In line with previous research with novice singers, the training group emerged as heterogeneous in vocal skill levels and pitch matching, with a proportion that could be defined as poor-pitch singers on established criteria. Given that our participants showed lower than average self-assessed skill, and a general perceived lack of competence, we first address whether and how objectively measured skills—including but also extending beyond pitch-matching—changed with training.

As predicted, there were clear improvements over the training period in vocal facility. Participants' vocal range for singing discrete pitches significantly increased and became more commensurate with their range in the physiologically less strenuous pitch slides. On average they expanded their range by just over four semitones, thus they became able to produce more pitches in the manner needed for successfully singing melodies. Non-imitative control of pitch direction and interval size also significantly improved. As a note of interest, errors in this task often resulted from perseveration when participants moved from singing an upward change to a downward one, such that despite intending to sing a downward change they reproduced the upward one (cf. Wisniewski et al., 2013, who observed perseveration of pitch patterns in poor-pitch singers in both spoken and sung imitation). This perseveration often came as a surprise to participants, that is, they could tell the output did not match their intention. By the end of the project these perseverative errors were eliminated, suggesting more reliable coupling of internal representations with motor planning; in Pfordresher and colleagues' (Pfordresher et al., 2015; Pfordresher & Mantell, 2014) MMIA model of vocal pitching this would be a strengthening of the "feed-forward" mechanism by which a movement sequence is planned based on an intended outcome. These simple pitch change and vocal range measures together indicate that vocal development took place.

In this respect, adult learners showed similar developmental markers to children, in their increased voluntary vocal control of pitch and widening vocal range (Welch, 1994a, 1998; and Rutkowski, 1997, 2015, respectively).

In vocal pitch matching, there was also an overall improvement after 10 months in the SSAP imitation tasks, albeit with individual variability in pitch-matching at baseline as well as different trajectories of pitch-matching changes over time. In many cases, pitch matching accuracy did not improve in a linear fashion, with deterioration sometimes occurring before improvement. We hypothesize that temporary deterioration can be understood as an initial disruption of existing sensorimotor pathways for singing, since all participants received technical instruction. This would be analogous to a tennis player working with a new coach who teaches the player a different technique for the backhand to make it more efficient in the long run; initially the player's shot will become less reliable while they learn the new technique. There was, however, no overall change in pitch accuracy in supported singing, suggesting the major changes taking place were to accuracy in self-generated pitch production but not (yet) to matching with an external stimulus.

Another area in which no change was observed was accuracy of singing familiar songs, which could not be explained by a ceiling effect. Again, we might understand this through a sensorimotor lens, since this may reflect established sensorimotor patterns in over-familiar songs, that are resistant to change. Therefore, the task of singing familiar songs, although a standard feature of singing development research in children, may be less suitable for assessing adults' development. The use of newly taught songs would be more likely to capture progress.

This contrast with children warrants some further discussion, since previous research has raised the question of whether adult non-singers who have pitching difficulties may simply have stalled in their singing development (Lidmann-Magnusson, 1997; Welch, 1994b) or even regressed through lack of practice (Demorest & Pfordresher, 2015). If so, we would expect that their development of pitching

ability might follow a trajectory similar to that of children. As noted, we did see some evidence to support this in improvements of vocal facility (range and pitch control). However, in children, the sequence of vocal pitch matching development parallels their emerging (implicit) musical perceptual and cognitive understanding, from surface (i.e., contour) to deep (i.e., tonality) levels of structure. By contrast, adults' receptive perceptual and cognitive processes for music have been consolidated over many years of listening. In common with many previous studies in non-amusic samples, we found no relationship between musical perceptual and productive skills. However, as a group, participants scored low in melodic memory, which has implications for their ability to learn and retain songs. Observations of group work and qualitative reports of students in this project—which are out of the scope of this paper—highlighted the self-reported and observed difficulties many participants had in learning songs 'by ear'. Melodic memory did improve by the end of the project, but further research would be needed to ascertain the relationship between improvements in melodic memory, and the production skills of sequence imitation and song learning.

The Central role of Self-Assessments in Singing Identity

Given the potential role of sense of competence in the non-singing identity (as captured in the Singing Inhibition scale), self-assessments of skill reflect changes in this area. Self-assessments of voice—accuracy, quality, and physical sensations while singing—significantly improved over the first six months. Interestingly, these self-ratings were not correlated with participants' *overall* singing pitch accuracy and vocal skills. It is tempting to conclude that participants are poor at judging their own skills, but this conclusion may be simplistic because when they are asked to rate their performance on *specific* tasks, non-singers do make accurate judgments (Wise & Sloboda, 2008). In the present study, self-ratings were given for global aspects of singing rather than specific tasks, which may explain the lack of correlation. However, the more important point is that using self-ratings solely as a "reality check" undervalues their importance. We propose that self-assessments may play an important mediating role among the three major components of singing identity: skills and skill development; attitudes and beliefs;

and situated behavior (participation). For example, my actual level of skill and my attitudes and beliefs may influence how likely I am to participate or not in a given singing context, mediated through a self-assessment of how likely I am to be able to meet the requirements. A future study could include a component of group singing (i.e., participation) to see if participation per se or perceived and objective assessment of ability to blend with a group, follow a leader, etc., would explain additional variance in the development of a singer identity.

Another interesting observation is that self-ratings of *accuracy* and *quality* showed different patterns of correlation, suggesting that participants could distinguish between them. This is not always evident in everyday discussions of singing, in which people often do conflate them in the context of general complaints about their singing (or someone else's).

The Influence of Attitudes, Beliefs and Past Experiences

Because prior research has suggested that negative beliefs about singing are also central in non-singing identity and can limit singing development (Knight, 2019), we directly assessed whether those attitudes were subject to change in three key areas: Elitist Attitudes, Implicit Beliefs, and Fear of Social Judgment. After six months participants showed less endorsement of elitist attitudes, greater belief that singing is a malleable rather than a fixed skill, and less fear of social judgment. These quantifiable changes support the narrative research that implicates such beliefs in the genesis and maintenance of non-singing identities and behavior. Welch (2001), among others, has written about widespread (Western) cultural misconceptions manifest in the (false) division between people deemed "musical" and "not-musical," and the dominant discourse of innate talent. That participants were less likely to endorse elitist attitudes after six months can be understood as a change in their underlying belief in a talent account of singing ability. Dweck's (2000) work demonstrates that the belief that an ability is a fixed trait—akin to a talent explanation—causes people to behave in ways that are unhelpful for learning. By contrast, a belief

that a skill is malleable results in learning-oriented behavior. Participants moved significantly towards a more malleable implicit belief, which in turn may have facilitated their skill acquisition.

A belief that singing cannot be learned was one of the three significant predictors of Singing Inhibition that we identified, along with a lack of singing in the childhood home, and a perceived lack of singing skills. These three interacting components represent aspects of belief, past experience, and perceived ability respectively that may inhibit or deactivate individuals' growth into singing adults. If we frame a non-singing status as a "deactivated" singing identity, this leads to a testable hypothesis that the opposite position in these three components should be protective against this deactivation. Future research could test whether a belief that singing is learned, a history of singing in the home (particularly by parents/caregivers), and a sense of competence for everyday singing are indeed associated with singing participation (as opposed to avoidance) and positive singing self-image in adulthood.

Conclusions, Limitations, and Implications: The Activated Singer

Having framed a non-singing identity as a "deactivated" singing identity, we now return to the question of whether non-singing adults are arrested in their vocal development, and therefore whether training can reinstate a pathway of development that is broadly similar to that of children. Beginning with measures directly reflecting competence, the gain in vocal range and voluntary control after training is consistent with the conclusion that these aspects of vocal facility were previously underdeveloped. When it comes to pitch matching, which most people would equate with "good singing", the answer is much more complicated. Although there was an overall improvement in pitch matching, there was also a wide range of results and trajectories among individuals. Furthermore, we did not recruit participants according to pitch matching capability; our definition of non-singer was someone who doesn't sing, not necessarily someone who is a poor singer. This definition is broader than that given for "poor pitch singer" (e.g., Hutchins & Peretz, 2011; Pfordresher & Brown, 2007) or Vocal Pitch Imitation Deficit (Pfordresher & Larrouy-Maestri, 2015), but slightly more restrictive than Dalla Bella et al.'s (2007) "occasional singers";

none of these is defined by the avoidance of singing, although naturally our sample did include poor pitch singers. Thus, one limitation of this study is that we did not have a large enough sample to divide our singers by pitch matching ability, or to examine the differential trajectories of gain and loss, or the different "phenotypes" (Dalla Bella & Berkowska, 2009) of poor singing. However, in characterizing our participants in the way we did, we also allowed for a wider definition of singing which includes—but does not only rest on—pitch accuracy.

We might even question the relative importance of pitch matching ability to singing identity, even though that skill has received so much prior attention in both child and adult singing research. Our results indicated that objective accuracy may be less important than *perceived competence*. However, it would be simplistic to conclude that participants are just poor at judging their own accuracy. Rather, this highlights the dependence on context of both displayed skills and self-assessed skills. While pitch accuracy is affected by musical context, a self-assessment of competence depends both on the musical context and the social context; appropriate singing behavior in Karaoke is quite different to singing in church, for instance. Believing one is a singer might depend on what we could call "functional competence," i.e., how far the individual feels able to produce the behavior that is appropriate in context. A mismatch leads to shame and potential withdrawal from the activity, or even the world of singing. In support of this idea, several participants who would have been classified as poor-pitch singers in other studies were successfully participating in community choirs by the end of the training, even though their pitch matching retained a level of insecurity.

Turning to more global self-assessments, compared to participants who completed the training, participants who did not complete started out with a stronger belief that singing is a fixed skill, as well as higher sensitivity to social judgment of their singing. This has important implications for enabling such people to persist in singing activities. Those who remained in the study for six months and beyond came to believe that singing can be learned and felt less sensitive to social judgment. We suggest that these

shifts may be key to being able to benefit from singing training and participation, respectively. That the influence of these barriers was evident, even though the project was specifically designed with the needs of the most inhibited non-singers in mind, speaks to their powerful influence in the non-singing identity. We suggest therefore that practitioners explicitly discuss these issues with inhibited singers, as this may enable learners to tolerate psychological discomfort and persist long enough for these internal inhibitors to wane.

We close by suggesting that the pattern of objective and self-assessed changes might characterize what we call The Activated Singer. Most people in our initial cohort reported having sung during childhood, with varying degrees of success or enjoyment. Self-identity as being a singer was likely not a foreign concept, but rather an aspect of the self that did not match their assessment of their skill. Unlike for instance, a new and unfamiliar skill like juggling, our participants did sing sometimes (in private), and likely imagined singing music fairly often as their responses to the BAIS were about the same as the general population. Thus the self-image as a singing person likely did not need to be newly created but rather encouraged and enhanced, hence our idea of "Activating" that identity. All participants continued singing after the training ended, in ways that suited them. Most continued singing lessons and/or joined singing groups. One celebrated being able to join in with hymns at church. Others found joy in connecting with family members through informal singing at home, or in singing in private just for themselves. Another limitation of this research is that by definition all our participants had an interest in learning to sing, and therefore our findings may not apply to all non-singers. But the level of interest in our recruitment suggests they are not alone.

A crucial part of activation was instruction and training in how to sing. This is important for a number of reasons. It resulted in objectively measurable skill development and, more importantly, an increased sense of competence. It also allowed participants to see that singing could be broken down into learnable skills, and thus to change their limiting beliefs. Finally, many of our participants reported that

prior to the study they did not know *how* to sing, making even the most ostensibly welcoming and inclusive opportunities inaccessible to them. The element of skill building is often missing from singing opportunities that aim to be inclusive, sometimes under the well-meaning philosophy that "everyone can sing." We therefore recommend to practitioners that modifying this philosophy to "everyone can *learn* to sing," alongside providing specific instruction, might open the door for non-singers to become Activated Singers, even just softly in the last row of the congregation, or loudly but anonymously at the football game.

Author Note

We have no known conflict of interest to disclose.

This project was funded by the UK Arts and Humanities Research Council (AHRC). We thank project coordinators Fanny Martin and Beth Martin, the collaborating singing teachers, and participants.

Correspondence concerning this article should be addressed to Karen Wise, Research and Enterprise, Guildhall School of Music & Drama, Milton Court, 1 Milton Street Barbican, London EC2Y 9BH. E-mail: karen.wise@gsmd.ac.uk

References

Abril, C. R. (2007). I have a voice but I just can't sing: A narrative investigation of singing and social anxiety. *Music Education Research*, *9*(1), 1–15. https://doi.org/10.1080/14613800601127494

Anderson, S., Himonides, E., Wise, K., Welch, G., & Stewart, L. (2012). Is there potential for learning in amusia? A study of the effect of singing intervention in congenital amusia. *Annals of the New York Academy of Sciences*, *1252*(1), 345–353.

Bento-Allpress, R. (2013). *Effects of group singing on well-being: Empirical findings and methodological considerations.* [Doctoral Thesis, Canterbury Christchurch University].

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* [Computer software]. http://www.praat.org/

Chong, H. J. (2010). Do we all enjoy singing? A content analysis of non-vocalists' attitudes toward singing. *The Arts in Psychotherapy*, *37*(2), 120–124. https://doi.org/10.1016/j.aip.2010.01.001

Dalla Bella, S., & Berkowska, M. (2009). Singing proficiency in the majority. *Annals of the New York Academy of Sciences*, *1169*(1), 99–107. https://doi.org/10.1111/j.1749-6632.2009.04558.x

Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2007). Singing proficiency in the general population. *Journal of the Acoustical Society of America*, *121*, 1182–1189.

Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2009). Singing in congenital amusia. *The Journal of the Acoustical Society of America*, *126*(1), 414–424. https://doi.org/10.1121/1.3132504

Demorest, S., & Pfordresher, P. (2015). Singing accuracy development from K-adult: A comparative study. *Music Perception*, *32*, 293–302. https://doi.org/10.1525/MP.2015.32.3.293

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality and development*. Psychology Press.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Sage.

Halpern, A. R. (2015). Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 37–47. https://doi.org/10.1037/pmu0000081

Heresniak, M. (2004). The care and training of adult bluebirds: Teaching the singing impaired. *Journal of Singing*, *61*(1), 9–25.

Hutchins, S., & Peretz, I. (2011). A frog in your throat or in your ear? Searching for the causes of poor singing. *Journal of Experimental Psychology. General*, *141*, 76–97. https://doi.org/10.1037/a0025064

JASP Team. (2022). *JASP* (Version 0.16.3 [Computer software]) [Computer software]. University of Amsterdam. https://jasp-stats.org/

Knight, S. (2019). Addressing the needs of the adult 'non-singer' ('NS'). In G. F. Welch, D. M. Howard, & J. Nix (Eds.), *The Oxford handbook of singing* (pp. 621–649). Oxford University Press.

Lidmann Magnusson, B. (1997). Factors influencing singing development in poor pitch singers. *Proceedings of the Third Triennial ESCOM Conference*, 339–343.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLOS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0089642

Müllensiefen, D., Gingras, B., Stewart, L., & Musil, J. (2013). *Goldsmiths Musical Sophistication Index (Gold-MSI) v1.0: Technical report and documentation* (Revision 0.3). https://media.gold-msi.org/test_materials/GMS/docs/gms_documentation_en.pdf

Numminen, A. (2005). *Numminen, A. (2005). Laulutaidottomasta kehittyväksi laulajaksi. Tutkimus aikuisen laulutaidon lukoista ja niiden aukaisemisesta. [Helping adult poor pitch singers learn to sing in tune. A study of stumbling blocks confronting developing singers and means of surmounting them.]* [Doctoral Thesis]. Sibelius-Academy.

Numminen, A. (2014). ‘Singing is no longer forbidden to me – it’s like part of my human dignity has been restored.’ Adult non-singers learning to sing: An explorative intervention study. *The European Journal of Social & Behavioural Sciences*, *12*(1), 1660–1674. https://doi.org/10.15405/ejsbs.149

Pentikäinen, E., Pitkäniemi, A., Siponkoski, S.-T., Jansson, M., Louhivuori, J., Johnson, J. K., et al. (2021). Beneficial effects of choir singing on cognition and well-being of older adults: Evidence from a cross-sectional study. *PLOS ONE*, *16*(2), e0245666. https://doi.org/10.1371/journal.pone.0245666

Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of ‘tone deafness’. *Music Perception*, *25*(2), 95–115. https://doi.org/10.1525/mp.2007.25.2.95

Pfordresher, P. Q., & Demorest, S. M. (2020). Construction and validation of the Seattle Singing Accuracy Protocol (SSAP): An automated online measure of singing accuracy. In F. A. Russo, B. Ilari, & A. J. Cohen (Eds.), *The Routledge companion to interdisciplinary studies in singing: Volume. 1. Development* (pp. 322–333). Routledge.

Pfordresher, P. Q., & Halpern, A. R. (2013). Auditory imagery and the poor-pitch singer. *Psychonomic Bulletin & Review*, *20*(4), 747–753. https://doi.org/10.3758/s13423-013-0401-8

Pfordresher, P. Q., Halpern, A. R., & Greenspon, E. B. (2015). A mechanism for sensorimotor translation in singing. *Music Perception*, *32*(3), 242–253. https://doi.org/10.1525/mp.2015.32.3.242

Pfordresher, P. Q., & Larrouy-Maestri, P. (2015). On drawing a line through the spectrogram: How do we understand deficits of vocal pitch imitation? *Frontiers in Human Neuroscience*, *9*. https://www.frontiersin.org/articles/10.3389/fnhum.2015.00271

Pfordresher, P. Q., & Mantell, J. T. (2014). Singing with yourself: Evidence for an inverse modeling account of poor-pitch singing. *Cognitive Psychology*, *70*, 31–57. https://doi.org/10.1016/j.cogpsych.2013.12.005

Richards, H., & Durrant, C. (2003). To sing or not to sing: A study on the development of `non-singers' in choral activity. *Research Studies in Music Education*, *20*(1), 78–89. https://doi.org/10.1177/1321103X030200010501

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). Memorial University Press.

Rutkowski, J. (2015). The relationship between children's use of singing voice and singing accuracy. *Music Perception*, *32*(3), 283–292. https://doi.org/10.1525/mp.2015.32.3.283

Turøy, A. K. W. (2018). Once a poor pitch singer, always a poor pitch singer? A bottom up study of factors that may support singing development. *British Journal of Music Education*, *35*(1), 91–103. https://doi.org/10.1017/S026505171700016X

Welch, G. F. (1994a). The assessment of singing. *Psychology of Music*, *22*, 3–19.

Welch, G. F. (1994b). Onchi and singing development: Pedagogical implications. In G. Welch & T. Murao (Eds.), *Onchi and singing development: A cross-cultural perspective* (pp. 82–94). The Roehampton Institute / David Fulton Publishers.

Welch, G. F. (1998). Early childhood musical development. *Research Studies in Music Education*, *11*(1), 27–41.

Welch, G. F. (2001). *The misunderstanding of music*. Institute of Education University of London.

Welch, G. F. (2006). Singing and vocal development. In G. E. McPherson (Ed.), *The child as musician: A handbook of musical development* (2nd ed., pp. 441–461). Oxford University Press.

Welch, G. F., Himonides, E., Saunders, J., Papageorgi, I., Rinta, T., Preti, C., et al. (2011). Researching the first year of the National Singing Programme Sing Up in England: An initial impact evaluation. *Psychomusicology: Music, Mind and Brain*, *21*, 83–97. https://doi.org/10.1037/h0094006

Wise, K. (2019). Defining and explaining singing difficulties in adults. In G. F. Welch, D. M. Howard, & J. Nix (Eds.), *The Oxford handbook of singing* (pp. 333–368). Oxford University Press.

Wise, K. J. (2009). *Understanding 'tone deafness': A multi-componential analysis of perception, cognition, singing and self-perceptions in adults reporting musical difficulties*. [Doctoral Thesis, Keele University].

Wise, K. J., & Sloboda, J. A. (2008). Establishing an empirical profile of self-defined "tone deafness": Perception, singing performance and self-assessment. *Musicae Scientiae*, *12*, 3–26.

Wisniewski, M. G., Mantell, J. T., & Pfordresher, P. Q. (2013). Transfer effects in the vocal imitation of speech and song. *Psychomusicology: Music, Mind, and Brain*, *23*(2), 82–99. https://doi.org/10.1037/a0033299

Zienkowska, A. (2013). *Self-reported singing difficulty: Exploration of possible predictors.* [Unpublished Masters Thesis]. City, University of London.

Footnotes

[1]Tuneless Choirs are a UK network of "choirs for those who love to sing, but can't necessarily get all the right notes in the right order." (www.tunelesschoir.com)

Table 8.

Pearson's Correlations of Singing Inhibition (SI) with Baseline Self-report Measures

	Singing related attitudes and beliefs			*Past singing experiences*		*Vocal self-assessments*			*BAIS*	
	Sensitivity to social judgement	Elitist attitudes	Implicit theory	Parental/ family engagement	Childhood	Accuracy	Quality	Physical discomfort	Vividness (V)	Change (C)
SI-Overall	.39*	.50**	.48**	-.51**	-.22	.10	-.30[b]	.17	-.10	-.15
SI-Ability	.29	.54**	.46**	-.27	-.22	-.12	-.34[a]	-.05	-.28	-.34
SI-Singing self	.20	.19	.28	-.44*	-.19	.21	-.07	.29	.04	-.02
SI-Social	.55***	.56***	.42*	-.48**	-.20	.06	-.36*	.03	.05	-.12

*$p < .05$, **$p < .01$, ***$p < .001$, [a]$p = .052$, [b]$p = .083$

Figure Captions

Figure 1. Slides range means in semitones at three time points (error bars show *SE*).

Figure 2. Sung range means in semitones at three time points (error bars show *SE*).

Figure 3. Voluntary control of pitch direction and size of change (error bars show *SE*).

Figure 4. Individual changes in overall SSAP pitch imitation deviations. Each line is a participant. Lower scores mean more accurate performance.

Figure 5. Individual changes in overall supported imitation deviations. Each line is a participant. Lower scores mean more accurate performance.