Descriptions and evaluations of "good singing" in the age of *The Voice*

ABSTRACT: What constitutes "good singing" can be hotly contested amongst singing voice pedagogues, yet little is known about what the general public considers to be "good singing". Within a program of research on musical identity and singing self-concept, this mixed-methods pilot study considered how members of the public (N = 52) described and evaluated stylistically different versions of a sung melody to test a hypothesis that reality TV singing may be deemed as "good singing". Participants were exposed to three versions of "Happy Birthday": 1) amateurs singing "as they would normally sing"; 2) professionals performing a "plain" version; 3) the same professionals singing an embellished version in the style of *The Voice* reality TV show. Results indicate that both professional versions were considered "better singing" than the amateur singing. While respondents focused on the technical deficiencies for amateurs, descriptions of the professionals concerned style. Stated exemplars of "good singing" were split between the two professional versions—based on sophistication and creativity ("professional: embellished") or vocal quality ("professional: plain"). While respondents' preferred version largely matched their chosen exemplar of "good singing", participants were more likely to sing along with the "amateur" version. Implications for singing voice pedagogy and engagement in singing activities for wellbeing are considered.

KEYWORDS: singing perception, performance quality, singing style, vocal technique, CCM singing

From *bel canto* to belt, singing teachers of all stripes hold passionate views about what they believe to be "good singing". These views may be based on a wide range of factors such as vocal technique, style, performance parameters, musicality — even the personality of the singer! For some, the answer is self-evident regardless of style or genre: "good singing is good singing" just as "bad singing is bad singing" (Bartlett as quoted in Forbes, 2018, p. 585). Bartlett further states that good singing in any style is "...overlaid on a foundation of genre-appropriate technique" (Forbes, 2018, p. 585). LoVetri (as quoted in Forbes, 2018, p. 583) uses authenticity as a benchmark when she remarks that the appropriate standard for any singing performance is "...to sing [the music] the way it was intended to be sung". Regardless of the position, our views as singing teachers concerning the nature of "good singing" are likely to influence what and how we teach our students as well as the repertoire we perform ourselves.

To best serve our students, singing teachers might also consider prevailing popular attitudes towards singing quality, including an understanding of what is broadly considered by the public to be "good singing". Singing teachers serve members of the public who may be pursuing singing lessons for enjoyment or the sheer personal satisfaction of improving their voice. It is important for singing teachers to consider the influences on students' beliefs around good singing and musical preferences, to both maintain student interest in lessons and ensure that students receive functional training appropriate to their vocal instrument (LoVetri, 2013).

Understanding the general public's beliefs relating to what constitutes "good singing" is also relevant to the field of singing, health, and wellbeing. It is well established that musical and singing engagement can improve wellbeing (see Daykin et al., 2018 for a systematic review of this literature; Krause et al., 2018), but people's beliefs and attitudes towards their own singing voice can act as a barrier to engaging in singing activities (Sloboda et al., 2005; see also Lamont, 2017 on the role of self-concept and musical engagement). These beliefs may be informed, in part, by portrayals in popular culture of "good singing". In an examination of the reasons for dropping out of musical activities, Krause et al. (2020) found that participants made assumptions regarding what participation means, including a belief that a person must possess a certain level of skill or musical ability to participate in musical activities. Therefore, it is important to understand general perceptions of "good singing", as this is the likely benchmark against which people will judge whether they are able to participate in singing activities. Moreover, within the context of singing groups for health and wellbeing, it is important for facilitators to ensure that participants feel comfortable to sing along with the group facilitator, the repertoire, and the style in which the music is presented (see, e.g., Forbes & Bartlett, 2020). Perceptions of "good singing" have

implications for self-concept and identity, which in turn can influence participation in beneficial singing activities.

We undertook this pilot study using a mixed-methods survey to consider whether, and how, people described and evaluated vocal versions of the same song. Here we report the analyses of these data concerning the descriptions of the different versions as well as respondents' comparisons and preference. We conclude the article by discussing the implications of these results for future research, singing voice pedagogy, and engagement in singing activities for health and wellbeing.

In comparing different versions of the same song, we were especially interested to test a hypothesis that reality TV show singing may be commonly deemed "good singing" by members of the public. Since their rise to popularity approximately 20 years ago, global franchises such as *Idol*, *The Voice* and *X Factor* have not only garnered huge television audiences but also received millions of views through social media platforms such as YouTube and Facebook (see Bartlett, 2020 for a summary of these statistics). Bartlett (2020, p. 185) wrote that considering the astounding popularity of these shows, "...it is not a far reach to suggest that reality television talent shows are influencing public attitudes to singing generally and singers in particular."

It is well-known that contestants on these shows are carefully and strategically chosen based on vocal ability and, in many cases, a novel backstory (Cheng, 2017). As is the case with any commercial television, the purpose of these shows is to create compelling viewing to ensure a large audience, and with it, increased advertising revenue (Anderson, 2005). A familiar trope is the contestant with a disability who brings the house down when they perform for the judges and the studio audience (Cheng, 2017). All vocal performances build to a predictable climax, bringing "cheers and tears" from judges and the audience (Cheng, 2017, p. 184). The combination of contestants' dramatic backstories with a victorious, soaring style of singing positions this as *the* winning formula for reality TV singing competitions. As Robinson (2014, p. 587) notes, these whirlwind ascensions to fame and celebrity portray this type of performance as something anyone can do if they are lucky enough to be "plucked from obscurity and placed on a national/international stage and celebrated for his or her newly found vocal prowess".

Reality singing shows portray contestants as ordinary people, sending a message to viewers that they too could audition for the show and become famous. Yet, at the same time, contestants' performances usually far exceed the singing capabilities of the everyday viewers (Arditi, 2020). Reality TV show singing most commonly involves performing contemporary commercial music (CCM) repertoire (LoVetri, 2013; LoVetri, 2008), with the occasional sensational operatic performance thrown into the mix for variety. Singers on these shows are expected to perform in a manner that is "harder, higher, and louder" (Bartlett, 2020, p. 185). Singing teachers will be all too familiar with the technical aspects of this form of singing, which requires loud, high-pitched, energized singing (usually high belt or chest-mix) which can take the capabilities of the human voice to extremes. For singing teachers, this may result in students presenting with "... expectations of their own singing voice capabilities that may be completely unrealistic and unsustainable" (Bartlett, 2020, p. 185). As these shows have become "a significant part of twenty-first-century contemporary culture" (Butler, 2019, p. 401), they have exposed millions of viewers worldwide to technically challenging and highly stylized singing performances (Bartlett & Naismith, 2020).

Has the elevation of this style of singing through these popular television shows translated into a belief among people that this type of singing is the exemplar of "good singing" in today's culture? Might the pervasiveness of this style in popular television culture lead people to assume that to sing well at all is to sing in such a manner?

The extant literature provides little in the way of answers to these questions, particularly as they relate to lay listeners/viewers and CCM singing styles. Studies on singing evaluation tend to investigate expert evaluations of classical singing (e.g., Wapnick & Ekholm, 1997), the acoustic parameters of good singing (Gupta et al., 2017), expert evaluations and perceptions of belt singing in musical theatre (e.g. Bourne & Kenny, 2016; LeBorgne et al., 2010), or compare expert evaluations of classical singing with acoustic measurements (e.g., Ekholm et al., 1998; Sonninen et al., 2005). The sheer complexity of the task of singing description and evaluation is also due to the ambiguity of terms used to describe singing (Hausknecht, et al., 2021; Mitchell & MacDonald, 2012). In a rare example of evaluation of CCM singing, one study investigated the evaluation of pronunciation using an online corpus of solo karaoke performances of popular songs, but this research was more concerned with developing a strategy for automatic evaluation of sung

pronunciation than with overall evaluations of singing quality (Gupta et al., 2018).

The small body of research concerning evaluations of singing quality is limited in that it is concerned in the main with expert evaluations and/or the acoustic parameters of good *classical* singing or musical theatre belt singing. Other studies on perception of musical performance have investigated the role of visual information in addition to sound information (e.g. Davidson, 1993; Mitchell & MacDonald, 2016). A case study of high profile CCM singer, Annie Lennox, highlights that gesture in singing is crucial for observers to perceive expression (Davidson, 2001). The present study sought to examine how people described and evaluated three stylistically different *audio only* versions of a common melody as sung by amateurs and professional vocalists, including one CCM version performed in the style of *The Voice* reality TV show. Whilst we acknowledge that the visual element of these television shows is an important part of the performance "recipe", for this small-scale pilot, we chose to investigate audio only, to get as close as possible to the singing itself, removed from the usual sophisticated visual production. Again, in keeping with the small scale of this pilot study, we were not concerned with listeners' ability to differentiate between singers or to recognise specific singers (cf. Mitchell & MacDonald, 2011; Mitchell & MacDonald, 2012; Mitchell & MacDonald, 2016), but focused on descriptions and evaluations only. We therefore hypothesized that, due to its dominance within popular culture, reality TV show singing may be described and evaluated as "good singing" by members of the public.

METHOD

Sample and Ethical Approval

A total of 52 Australian residents, aged between 17 and 71 (M = 27.67, Mdn = 23, SD = 12.79), completed an online survey. This convenience sample consisted of 37 (71.20%) people who identified as female, 14 (26.90%) as male, and 1 (1.90%) as non-binary. The sample was largely made up of people with little musical experience/education (36 "hardly ever play/ed", 4 "occasional", 8 "amateur", 2 "semi-pro", 2 "professional"). Participants were recruited via online advertising and a university participation scheme. Participation was voluntary; however, students who participated via the university scheme received course credit. The James Cook University human research ethics committee approved this study (ID: H8209).

Study Design

We used an online, cross-sectional survey (hosted using Qualtrics) that included quantitative and qualitative questions (as outlined below in the Materials section).

Stimuli

Three audio recordings were prepared specifically for this study. The audio clips were 60-second versions of a male and female singing "Happy Birthday" (approx. 30 seconds each) in the key of F major. As the professional singers were located in cities some distance from the research team, they recorded their samples and these were provided to a sound engineer for mixing, together with the amateur versions which were recorded by the same sound engineer. These recordings were mixed to achieve as close a match in average intensity as possible between samples, producing an average intensity level of 45 dB when played on computer speakers at a moderate volume.

The two amateur singers in this study are untrained in singing technique and style. The two professional singers have both had extensive vocal training and both earn a living from their singing.

To ensure consistency in approach for each recording, the following instructions were given to the singers:

1. ***Professional: plain version*****—**One male and one female professional CCM singer were instructed to sing the melody of "Happy Birthday" without melodic or rhythmic embellishment. Both professional singers are highly experienced CCM vocalists with careers in live and recorded performance.
2. ***Professional: embellished version*****—**The same male and female professional CCM singers from version 1 performed in the style of *The Voice* reality television series—the singers were instructed to use melodic and rhythmic embellishment, melisma, and to build the version to a loud, high climax using belt voice or chest-mix voice.
3. ***Amateur version:*** One male and one female amateur singer were asked to perform "as they would normally sing the song". The amateur singers did not identify as musicians and work in non-music related fields. The only guidance provided to the amateur singers was key and tempo, to ensure consistency across versions.

Materials

Demographic questions asked participants to state their age, gender, and country of residence. Participants also classified their level of musicianship using Kreutz et al.'s (2008) item which involves selecting one of the following options: "hardly ever play/ed", "occasional", "amateur", "semi-professional", "professional", and "other".

To examine participants' evaluations of the audio versions, we used an open-ended question and a set of seven scale items (Forbes, Krause, & Lowe-Brown, 2021). The open-ended question asked participants to provide a short, written response about what they heard using the prompt, "How would you describe what you heard?" Participants used a five-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*) to respond to the seven quantitative items (all items appear in Table 1). A set of three Principal Components Analyses (one for each audio version) were performed to examine the structure of the participants' responses to the seven items. As seen in Table 1, the findings indicated a two-factor structure reflecting an evaluation of quality and the consideration of being able to sing along with the performers in alignment with previous, preliminary analyses (Forbes, Krause, & Lowe-Brown, 2021). Given the item loadings, the two factors are labelled as "singing quality" (e.g., The performer/s had good technique) and factor 2 was labelled as "sing-along ability" (e.g., I would be able to sing along with the performer/s). Two average scores were computed for each participant to be used in subsequent analyses. Two additional author-developed, open-ended questions were designed to address participants' personal responses to the audio versions (after hearing all three versions) — "Which version did you prefer, and why?" and "Which version would you consider to be the best example of 'good singing', and why?" These questions were deliberately broad and their main terms (e.g., "good singing") were undefined to allow participants to interpret the questions freely. We subsequently categorized each response that overtly identified one of the three versions as the participant's preferred version/ exemplar of best singing as either the "professional: plain", "professional: embellished" or "amateur". We note, however, that not all responses were categorized because some people referenced their selection according to the presentation order (which was randomized); therefore, we were unable to label these selections confidently.

Procedure

Participants were directed to the participant information via a direct weblink. After indicating their consent to participate via their response to a yes/no question, individuals completed the survey as a series of webpages. Prior to answering any questions, participants were given the following instructions: "This questionnaire involves listening to some audio. Please ensure that you are wearing headphones at a loud, but comfortable, level". They first answered the demographic questions and were then asked to listen to a series of three short audio clips. Qualtrics randomized the order of the presentation of the audio clips across the sample. After each clip, participants responded to the set of seven quantitative items and provided a short, written description of what they heard. To conclude the survey, individuals responded to the two author-developed, open-ended evaluative questions. Participants were thanked and debriefed via a final webpage.

Data Analysis

SPSS (version 25) was used to conduct the quantitative analyses. In particular, two one-way repeated ANOVAs were used to compare the participants' ratings of the versions regarding singing quality and sing-along-ability. A chi-square analysis was used to consider how participants' chosen preferred version were related to their selection as the exemplar for best singing.

A simple form of qualitative content analysis (Elo & Kyngäs, 2008) was used to produce content summaries of responses to the survey's qualitative prompts. The data were coded inductively, then similar codes were clustered into higher order headings. Finally, content summaries were produced based on code frequency; however, frequency was not assumed to be a proxy for significance (Vaismoradi et al., 2013), in that the summaries also include some outlying responses. Examples of participant quotes are included to support the analyses presented.

RESULTS

Describing and evaluating the versions

Two one-way repeated ANOVAs considered how people evaluated the versions via the quantitative items. The model concerning singing quality was statistically significant, F (2, 96) = 64.614, p <

.001, η_p^2 = .574. Pairwise comparisons indicated that the "professional: embellished" version (*M* = 3.801, *SD* = 0.846) and the "professional: plain" version (*M* = 3.776, *SD* = 0.778) were both rated as significantly higher in singing quality compared to the "amateur" version (*M* = 2.240, *SD* = 0.800). There was no statistical difference between the "professional: embellished" and the "professional: plain" versions.

The model regarding being able to sing along with the performers was also statistically significant, *F* (2, 98) = 40.344, *p* < .001, η_p^2 = .452. The pairwise comparisons indicated that participants were significantly more likely to report being able to sing along with the "amateur" version (*M* = 3.693, *SD* = 0.740) and "professional: plain" version (*M* = 3.320, *SD* = 0.957) compared to the "professional: embellished" version (*M* = 2.407, *SD* = 0.812). There was no significant difference between the "professional: plain" and "amateur" versions.

TABLE 1 to appear about here (tables currently appear at end of the main text)

While the prompt "How would you describe what you heard?" was deliberately neutral so as to elicit how the versions were perceived differently rather than to pre-empt evaluative responses, many of the qualitative responses did contain an evaluation of the performance, some providing different evaluations for the male and female singers within a single version. Across each of the versions, the most common description was the most literal one, for example, "I heard a male and female singing Happy Birthday". Many more nuanced descriptions and evaluations were also provided and are summarized under each version below.

Amateur Version. The male and female amateur performances were most often described as "untrained" or "unprofessional", with participants seeming to equate these two terms. Being "untrained", the amateurs were deemed to have "poor technique" which was commonly associated with pitch inaccuracy (e.g., "pitchy", "out of tune") with far fewer responses identifying rhythmic issues (e.g., "out of time slightly") as being relevant to poor technique. Beyond issues of training or professionalism, the amateur singers were positively described in terms of being relatable or sounding like friends or family. Others described the amateurs as sounding uncomfortable or lacking in confidence ("hesitant and unsure"). Overall evaluations ranged from negative ("horrible", "can't sing for shit", "flat and boring") to middling ("they sounded ok", "average voices",) with only one overtly positive evaluation ("good").

Professional: Plain Version. Both the female and male singer in the "professional: plain" version were commonly described as having good technique, tone, or timbre, and as sounding "professional", yet four respondents thought the singers sounded "breathy", "untrained" or "unprofessional". Three other responses also mentioned musical or stylistic elements of the performances ("embellishments", singing "in a fancy manner"). Responses to the "professional: plain" version overall demonstrated a greater likelihood of equating musical elements to style rather than to vocal technique whereas in the amateur version musical elements such as pitch and rhythm were identified as being technically deficient. Indeed, several respondents said these voices were "much better" than the amateur voices and commented that the singing was "very nice" and "on pitch".

Evaluations of the "professional: plain" version were mostly aimed at differentiating the male singer from the female singer, with the female singer being rated more positively (nine respondents overtly preferred the female singer). The female voice was rated as "more complex", "more elegant and pleasant to listen to" and possessing "more volume and vibrato". Unlike the "amateur" version, some respondents described having an emotional response to the female singer in this version ("the female sounded soulful and beautiful"; "The female voice had a greater storytelling ability—there was something deeper being conveyed in their singing").

Professional: Embellished Version. Respondents were less polarized regarding preference for female over male in this version, with only four responses noting gender in their descriptions (with three respondents preferring the female version). The singers were described as "professional", or in terms indicating professionalism, for example, "A performer who knows how to control their voice and singing". One respondent, for instance, described the singers in the "professional: plain" version as "new singers without training" but then described the singers in the "professional: embellished" version as "professional voices"! Implicit in this respondent's response is that the overtly stylistic performance in the "professional: embellished" version was equated with expertise. The most referred to musical and technical element was the ability to sing "high notes" or "high pitch": There was little other description of vocal technique.

Descriptions, however, did focus on the more stylistic elements of the version, much more so than the "amateur" and "professional: plain" versions ("People singing Happy Birthday very stylistically"). One respondent intuitively knew that this was a more stylized version of *Happy Birthday* but struggled to articulate why: "A very souled-up version of happy birthday that showed off some great vocal stylings (but I couldn't name what they are)". Some were able to identify specific stylistic elements including trills and runs and described the singing as "belting". Others said the version was "soulful" or "jazzy". Respondents noted that in this version, the singers made the song "their own" through "their own interpretations and their own vocal intonations to make the song more exciting and interesting"; the singers performed "unconventionally and in their own way". The ability to sing in a highly stylized way appeared to work in the male singer's favor, in that aspects of his technique or tone were no longer the subject of critique by respondents. In this version, aspects of style appear to trump all other considerations in respondents' descriptions.

There were numerous highly positive evaluations of both the male and female singers: "definitely better than the other performances"; "excellent"; "stunning". Other responses described the performances in somewhat critical terms, such as "over the top, trying too hard" and "strangling cats".

Stated Singing Preferences

Table 2 displays the frequencies of those responses that overtly identified one of the three versions as the participant's preferred version. While only two respondents preferred the "amateur" version over either professional version (based on its relatability and "endearing" quality), most respondents preferred the "professional: plain" version, which was described variously as "voices were in tune"; "sounded professional and smooth" (see Table 2 for additional quotes). There was an overall sense that the singers were not trying too hard—as one respondent said, it did not "need to be fancy to be good". Indeed, a comparison was often used in order to reinforce this stated preference for the "professional: plain" version rather than the "professional: embellished" version (e.g., "because 'the voice' style was far too over the top for such a song"; "it was sung with good and clear technique and shows professional performance quality without having to riff and embellish a simple song"). Yet for those who did prefer the "professional: embellished" version, reasons pointed to the embellishments made as demonstrations of what the singers "could do with their voice" (e.g., "they both hit the high notes, lots of energy and sounded exciting").

Exemplifying "Good Singing"

As seen in Table 2, all respondents nominated one of the two professional versions as exemplifying "good singing". The lack of any nominations for the "amateur" version complements the quantitative results where the amateur version was rated the lowest of the three versions in terms of singing quality. In comparison with people's stated preferences, participants were more evenly split on which version exemplified "good singing", though the "professional: plain" version was nominated slightly more times than the "professional: embellished" version.

Responses (see Table 2 for quotes) revealed that preferences for the professional versions were based on sophistication and creativity (for "the professional: embellished" version) and vocal quality (for the "professional: plain" version). These preferences indicate that professional singing is equated with either stylistic sophistication or good technique (whereas amateur singing is perhaps viewed as common and untrained). Implicit in some responses was the idea that "most people" cannot sing well. For example, one nomination for the "professional" embellished" version as best exemplifying good singing was "because very few people can sing that way". In other words, the uncommonness of this version made it "good". Another response captures a similar sentiment:

> The example that would be socially considered to be the best in terms of "good singing" would naturally be the "professional: embellished" version. It would be considered the best as it performs a simple song in a fresh and exciting way, whilst incorporating trained vocals and pitch ranges generally higher than the average person would.

With such evaluations in mind, the average person may never be considered to sound "good" when singing!

While the embellishments were equated with "good singing" for some respondents, these embellishments were also judged as "showing off" or "over-doing" the song, leading to other respondents' nominating the "professional: plain" version as exemplifying "good singing". For these participants, it was the performers' talent and technique via their control and command of their

voices that were deemed impressive. Though the response to the embellishments was divisive, it seemed to be stable because a statistically significant chi-square test of contingencies, χ^2 (1, N = 39) = 21.559, p < .001, ϕ = .744, demonstrated that 13 of the 14 people who preferred the "professional: embellished" version and 21 of 25 people who preferred the "professional: plain" version stated that their preferred version was also the exemplar of "good singing". In other words, people's preferences largely matched what they thought was indicative of "good singing".

TABLE 2 to appear about here

DISCUSSION

This pilot study sought to examine how people described and evaluated three stylistically different versions of a common melody as sung by amateurs and professional vocalists, including one version performed in the style of *The Voice* reality TV show. We hypothesized that due to the prominence and popularity of reality singing TV shows such as *The Voice*, the singing style championed in these shows would be considered "good singing" by members of the public.

Quantitative and qualitative results overwhelmingly indicate that both professional versions were considered "better singing" than the amateur singing, but there was little evaluative difference between the "professional: plain" and "professional: embellished" versions. What was apparent in the qualitative descriptions, however, was the focus on technical deficiencies for amateur singers, and the absence of discussion of vocal technique for both professional versions. Respondents made mention of pitch concerning amateur performances, which positions pitch accuracy as a type of threshold criterion for good singing in line with previous studies on classical singing (Gupta et al., 2017; Wapnick & Ekholm, 1997). In contrast, people's descriptions of the professional versions (especially the "professional: embellished" version) concerned style. This suggests that the professional singers were perceived as performing in a stylistic manner "overlaid on a foundation of genre-appropriate technique" (Bartlett as quoted in Forbes, 2018), with their good technique being largely "invisible" to the listeners. In other words, technical considerations such as singing in tune were not remarked on precisely because the professional singers were able to sing with pitch accuracy. Interestingly, the quantitative results indicated that the "professional: plain" version singers had slightly better technique than those in the "professional: embellished" version, yet the "professional: embellished" version was undoubtedly the technically more demanding to sing. This further supports the idea that to the lay listener, technique becomes invisible and is only noticeable in its deficiency rather than proficiency. Qualitative results indicate that the overtly "stylistic" singing of the "professional: embellished" version is strongly equated with professionalism. In contrast, no responses referred to aspects of style in the "amateur" version, demonstrating that the lay listener can detect stylistic authenticity, even if they are unable to articulate specifics.

People's preferences for "good singing" were split between the two professional versions, with people focused on sophistication and creativity (for the "professional: embellished" version) or vocal quality (for the "professional: plain" version). Respondents' preferred version was largely also their chosen exemplar of "good singing". That respondents equated their preferred version with "good singing" supports the notion that in the voice studio, students usually pursue learning repertoire they like, and that this repertoire exemplifies what *they* consider to be "good singing". In the case of the inexperienced singer seeking to learn a vocally challenging song like "Defying Gravity", such a scenario presents a teachable moment in the studio, opening the door for a discussion on personal preference and the implications of this for the student's vocal development and learning journey. In many cases, students' preferences for "good singing" may not, in fact, align with the functional capabilities of their own voices. An early discussion with new students about what they consider to be "good singing" can, therefore, help clarify expectations and the setting of realistic vocal goals.

Contrary to our initial hypothesis that the popularity of reality TV singing shows would (strongly) influence public perceptions of good singing, there was not an overwhelming preference for the "professional: embellished" version nor did that version achieve a clear evaluative "win" over the "professional: plain" version. Rather, respondents were equally enthused about the "professional: plain" version, with its "invisible" solid vocal technique and unadorned singing style. Both professional versions were assessed equally as examples of "good singing". The fact that people nominated one or the other professional version as "good singing" as opposed to the "amateur" version is perhaps not surprising;

however, as suggested by Bartlett (2020), these findings do give some indication as to the benchmarks against which the average person judges their own singing ability.

We agree with Bartlett (2020) that that these benchmarks — particularly for the "professional: embellished" version — are likely to be unrealistic performance indicators upon which to judge singing ability for most people. However, we would argue further that notions of "good" and "bad" as perceived by respondents to be exemplified in the professional and amateur versions respectively are not helpful when used in relation to non-trained or amateur singers. Those who train and identify as professional singers may justifiably remain open to scrutiny or evaluation (particularly when singing commercially), but the same standards should not apply to all who sing. When considering participation in musical activities such as group singing for health and wellbeing (which is intended for a broad participation base), both technical and stylistic aspects of singing are largely irrelevant. Thus, equating good singing with either technical proficiency, stylistic sophistication, or both, may act as a barrier to participation in community music and music for wellbeing activities.

Despite the lukewarm reception of the amateur version in people's qualitative responses, the quantitative results showed respondents would feel most comfortable singing along with the "amateur" version. Drawing on the data, some people found the relatable and endearing nature of amateur singing an invitation to join in. It could be that by being more familiar, such a style of singing is also more accessible. Thus, there are implications regarding the facilitation of singing groups and activities. Again, expectations that an elevated level of ability is required due to the "unachievable for most" component of "good singing" may again act as a barrier to broad community participation in singing activities. This corresponds to previous research findings concerning people's barriers to continued musical engagement. Krause et al. (2020, p. 412) found that people's "assumption of what a music participant *is* or *should be* included the perception that an individual must possess specific qualities and/or skills in order to take part". Assumptions included a certain level of musical ability, and as Krause et al. (2020) argued, music educators and facilitators must challenge held assumptions and ideologies to remove such barriers to musical participation.

Concerning musical engagement, the present findings have further implications for research on singing in the health and wellbeing context. Whilst Forbes and Bartlett (2020) found that a background in singing voice pedagogy was advantageous for facilitators of singing groups for people with Parkinson's, little is known about what is required *vocally* for the effective facilitation of community singing groups including those for health and wellbeing. Given respondents' (seemingly) counter-intuitive preference for singing along with the amateur version, future research might seek to evaluate people's judgments of facilitators' leading singing styles, examining not only judgments of quality but also how their style might influence group participants' sense of feeling welcomed and encouraged to sustain participation. Given the growing body of evidence concerning how singing is associated with positive health and wellbeing benefits (Clift, et al., 2010; Daykin, et al., 2018; Fancourt & Finn, 2019; Krause, et al., 2018; Lewellen, Meyere, & van Leer, 2020; Więch, et al., 2020), such work will extend the consideration of how singing group facilitators can best teach and support groups comprised of un-trained singers.

This exploration of lay opinion of singing style and quality is not without its limitations. Beyond the use of a small, convenience sample, the study only made use of three versions of a very familiar song. Because our hypothesis related to reality TV shows which predominantly portray a particular brand of CCM style, classical singing as well as other CCM styles were not included. This limits the generalizability of our findings. A similar point can be made for the use of only one overtly male and female voice. Given the comments made differentiating between the male and female performer within an audio version, future research is also needed to tease out descriptions and evaluations which relate to gender differences in sung performance. Further, the use of "Happy Birthday" may have both minimized and exaggerated some of the results. For instance, as one participant noted, "Happy birthday is such a well-known song and it is sung for others to join in and sing along to" such that the embellishments may have been judged inappropriate for this song, but not others. Even though the "professional: embellished" version was more technically and stylistically demanding, the near-even split in people's evaluations and preferencing of the two professional versions suggests that future research should draw on more performers singing more songs in a wider variety of styles to gain a nuanced view of the perceived elements indicative of good singing.

Additionally, we recognize that participants were not asked about their familiarity or engagement with reality singing television and that the audio versions were not repeated to check for intra-rater reliability. The audio versions were presented in the online study without any performative context. Performance contexts vary (consider, for instance, a community choir versus a TV talent show), and associated expectations may influence what is expected, and thus, judgments of quality. Moreover, musical performances often include a visual component which has been shown to influence people's evaluations of the performance (Platz & Kopiez, 2012; Thompson, et al., 2005; Tsay, 2013). Studies that have investigated the full performance experience including sound and visual information have reported that visuals are crucial to perceiving performance manner and expression (Davidson, 1993; Davidson, 2001). Therefore, future investigations should consider the influence of reality TV contestants' movements and the use of audiovisual materials on lay considerations of singing quality.

CONCLUSION

This pilot study has provided insight into the public's descriptions and evaluations of a limited number of singing styles including the style most commonly portrayed in reality television singing shows. Whilst the sample size was small, it is clear that respondents did not view amateur singing favorably (unless it was to sing along to) and equated "good singing" with "professionalism". As researchers who are interested in everyday musical engagement, these findings give us pause. If the public considers good singing to be the exclusive domain of the professional, the flipside is that non-professionals cannot be considered good singers. Self-concept plays an important role in the extent to which people participate in music (Lamont, 2017). If we do not think we are good at something, we are less likely to participate. Respondents' views confined "good singing" to the professional domain as a performative act rather than as something *to be experienced in everyday life*. Thus, countering this attitude may be one of the great challenges for broadening participation in singing activities.

REFERENCES

Anderson, S. (2005). Regulation of television advertising. In *Virginia Economics Online Papers* (Vol. 363). https://ideas.repec.org/p/vir/virpap/363.html

Arditi, D. (2020). *Getting signed: Record contracts, musicians, and power in society.* Palgrave MacMillan. https://doi.org/10.1007/978-3-030-44587-4

Bartlett, I. (2020) Crossing style borders: New inroads in training teachers of singing. *Voice and Speech Review, 14*(2), 184-195. https://doi.org/10.1080/23268263.2020.1695878

Bartlett, I., & Naismith, M. (2020). An investigation of Contemporary Commercial Music (CCM) voice pedagogy: A class of its own? *Journal of Singing, 76*(3), 273–282.

Bourne, T., & Kenny, D. (2016). Vocal Qualities in Music Theater Voice: Perceptions of Expert Pedagogues. *Journal of Voice*, *30*(1), 128.e1-128.e12. https://doi.org/10.1016/j.jvoice.2015.03.008

Butler, A. (2019). The judges' decision is final: Judgement in music talent reality TV and school music education. *Journal of Popular Music Education, 3*(3), 399–415. https://doi.org/10.1386/jpme_00003_1

Cheng, W. (2017). Staging overcoming: Narratives of disability and meritocracy in reality singing competitions. *Journal of the Society for American Music, 11*(2), 184–214. https://doi.org/10.1017/S1752196317000062

Clift, S., Nicol, J., Raisbeck, M. Whitmore, C., & Morrison, I. (2010). Group singing, wellbeing and health: A systematic mapping of research evidence. *UNESCO Refereed E-Journal, Multi-disciplinary Research in the Arts, 2,* 1-25.

Daykin, N., Mansfield, L., Meads, C., Julier, G., Tomlinson, A., Payne, A., Grigsby Duffy, L., Lane, J., D'Innocenzo, G., Burnett, A., Kay, T., Dolan, P., Testoni, S., & Victor, C. (2018). What works for wellbeing? A systematic review of wellbeing outcomes for music and singing in adults. *Perspectives in Public Health*, *138*(1), 39–46. https://doi.org/10.1177/1757913917740391

Ekholm, E., Papagiannis, G. C., & Chagnon, F. P. (1998). Relating objective measurements to expert evaluation of voice quality in western classical singing: Critical perceptual parameters. *Journal of Voice*, *12*(2), 182–196. https://doi.org/10.1016/S0892-1997(98)80038-6

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Fancourt, D., & Finn, S. (2019). *What is the evidence on the role of the arts in improving health and wellbeing? A scoping review (2019)*. World Health Organization.

http://www.euro.who.int/en/publications/abstracts/what-is-the-evidence-on-the-role-of-the-arts-in-improving-health-and-wellbeing-a-scoping-review-2019

Forbes, M. (2018). A tale of two pedagogues: A cross-continental conversation on CCM. *Journal of Singing*, *74*, 579–584.

Forbes, M., & Bartlett, I. (2020). 'It's much harder than I thought': Facilitating a singing group for people with Parkinson's disease. *International Journal of Community Music*, *13*(1), 29–47. https://doi.org/10.1386/ijcm_00009_1

Forbes, M., Krause, A. E., & Lowe-Brown, X. (2021, October). *The influence of vocal performance on singing self-concept.* Presented at the International Symposium on Performance Science (ISPS21). Montreal, Canada.

Gupta, C., Li, H., & Wang, Y. (2017). Perceptual evaluation of singing quality. *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference* (APSIPA ASC), 577-586. https://doi.org/10.1109/APSIPA.2017.8282110

Gupta, C., Li, H., & Wang, Y. (2018). Automatic pronunciation evaluation of singing. *Interspeech*, 1507-1511. Retrieved from https://smcnus.comp.nus.edu.sg/wp-content/uploads/2018/07/Chitralekha_Interspeech2018.pdf

Hausknecht, J. B., Murdaugh, K. M., Nagl, E., & Herbst, C. T. (2021). Global inventory and similarity rating of singing voice assessment terms used at English speaking academic institutions. *Journal of Voice*, advance online publication. https://doi.org/10.1016/j.jvoice.2021.08.020

Krause, A. E., & Forbes, M. (2021, July). *Socio-cultural influences on Australian adults' musical identities*. 16th International Conference on Music Perception and Cognition jointly organised with the 11th Triennial Conference of ESCOM (ESCOM-ICMPC 2021), Australia. https://sites.google.com/sheffield.ac.uk/escom2021/home

Krause, A. E., Davidson, J. W., & North, A. C. (2018). Musical activity and wellbeing: A new quantitative measurement instrument. Music Perception, 35(4), 454-474. https://doi.org/10.1525/MP/2018.35.4.454

Krause, A. E., Kirby, M. L., Dieckmann, S., & Davidson, J. W. (2020). From dropping out to dropping in: Exploring why individuals cease participation in musical activities and the support needed to re-engage them. *Psychology of Aesthetics, Creativity, and the Arts, 14*(4), 401-414. http://dx.doi.org/10.1037/aca0000268

Kreutz, G., Schubert, E., & Mitchell, L. A. (2008). Cognitive Styles of Music Listening. *Music Perception: An Interdisciplinary Journal*, *26*(1), 57–73. https://doi.org/10.1525/mp.2008.26.1.57

Lamont, A. (2017). Musical identity, interest, and involvement. In R. MacDonald, D. J. Hargreaves, & D. Miell (Eds.), *The handbook of musical identities* (pp. 176–196). Oxford University Press. doi: 10.1093/acprof:oso/9780199679485.001.000

LeBorgne, W. D., Lee, L., Stemple, J. C., & Bush, H. (2010). Perceptual findings on the broadway belt voice. *Journal of Voice*, *24*(6), 678–689.

Lewellen, R., Meyer, D., & van Leer, E. (2020). The effects on acoustic voice measures and the perceived benefits of a group singing therapy for adults with Parkinson's disease. *Australian Voice*, *21*, 39–48.

LoVetri, J. (2008). Contemporary commercial music. *Journal of Voice, 22*(3), 260-262.

LoVetri, J. (2013). The necessity of using functional training in the independent studio. *Journal of Singing, 70*(1), 79–86.

Mitchell, H. F., & MacDonald, R. A. R. (2011). Remembering, recognizing and describing singers' sound identities. *Journal of New Music Research*, *40*(1), 75–80. https://doi.org/10.1080/09298215.2010.545130

Mitchell, H. F., & MacDonald, R. A. R. (2012). Recognition and description of singing voices: The impact of verbal overshadowing. *Musicae Scientiae*, *16*(3), 307–316. https://doi.org/10.1177/1029864912458849

Mitchell, H. F., & MacDonald, R. A. R. (2016). What you see is what you hear: The importance of visual priming in music performer identification. *Psychology of Music*, *44*(6), 1361–1371. https://doi.org/10.1177/0305735616628658

Platz, F., & Kopiez, R. (2012). When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Perception: An Interdisciplinary Journal, 30*(1), 71-83. https://doi.org/10.1525/mp.2012.30.1.71

Robinson, D. K. (2014). TV talent shows: Navigating the minefield. *Journal of Singing*, *70*(5), 585–589.

Sloboda, J. A., Wise, K. J., & Peretz, I. (2005). Quantifying tone deafness in the general population. *Annals of the New York Academy of Sciences*, *1060*(1), 255–261.

Sonninen, A., Laukkanen, A.-M., Karma, K., & Hurme, P. (2005). Evaluation of support in singing. *Journal of Voice*, *19*(2), 223–237. https://doi.org/10.1016/j.jvoice.2004.08.003

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica, 156*(1/4), 203–227.

Tsay, C.-J. (2013). Sight over sound in the judgment of music performance. *PNAS, 110*(36), 14580-14585. https://doi.org/10.1073/pnas.1221454110

Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences*, *15*(3), 398–405. https://doi.org/10.1111/nhs.12048

Wapnick, J., & Ekholm, E. (1997). Expert consensus in solo voice performance evaluation. *Journal of Voice, 11*(4), 429–436. https://doi.org/10.1016/S0892-1997(97)80039-2

Więch, P., Sałacińska, I., Walat, K., Kózka, M., & Bazaliński, D. (2020). Can singing in a choir be a key strategy for lifelong health? A cross-sectional study. *Journal of Voice*, advance online publication. https://doi.org/10.1016/j.jvoice.2020.11.010

BIOGRAPHY

Melissa Forbes is Senior Lecturer (Contemporary Singing) at the University of Southern Queensland. Melissa is singing practitioner-researcher who researches singing experiences across a broad range of contexts, from elite, professional singers to community groups. Her research uses qualitative methods to position singing as a unique health and wellbeing practice from which all can benefit.

Amanda Krause is a Lecturer (Psychology) in the College of Healthcare Sciences at James Cook University. She is interested in the social and applied psychology of music, and her research examines everyday music interactions, with an emphasis on considering how everyday music experiences influence wellbeing.

Xanthe Lowe-Brown Xanthe Lowe-Brown is a Bachelor of Music graduate from the University of Melbourne. Xanthe completed a major in contemporary guitar performance and a minor in psychology, which allowed her to hone her guitar skills and delve into her passion for music psychology. Xanthe aspires to pursue research in this area and make an impact by discovering more about how we can use music to enhance our health and wellbeing.

Table 1.
Promax Rotated Structure of the Principal Components Analyses Concerning the Seven Item Questionnaire for Each Version

	Professional - the voice		Professional - plain		Amateur	
Item	1	2	1	2	1	2
The quality of the performance was high	0.925		0.951		0.915	
The performer/s sounded like a professional to me	0.908		0.892		0.929	
The performer/s had good technique	0.846		0.872		0.888	
The performer/s has had a lot of training	0.800		0.837		0.889	
The song would be hard to sing	0.579	-0.392		-0.835		-0.572
I would be able to sing along with the performer/s		0.845		0.849		0.840
I relate to the sound the performer/s was making	0.362	0.817		0.754		0.777
Eigenvalue						
% of variance						
Cronbach's alpha	0.903	0.535	0.908	0.749	0.925	0.566

Note. Factor 1 label = Singing quality, Factor 2 label = Sing-along ability.

Table 2.
Frequencies and Responses Concerning Participants' Version Selections

	Nominated Version		
	Professional - *The Voice*	Professional - plain	Amateur
Version preference (*n* = 43)	14 (32.60%) "it sounded more sophisticated and was more emotionally impactful." "it was the one that caught my attention the most" "it was unique and it was evident that they had been trained on how to shape their voices into the perfect version"	27 (62.80%) "the vocal quality was good and I could hear the singers individual subtle styles without needing to embellish" "it sounded the most natural yet polished in terms of vocal quality" "it was much easier/more soothing to listen to and sing along with compared to the other clips"	2 (4.70%) "honestly, overall I did prefer the Amateur over the rest. possibly due to its relatability and familiarity" "the amateur version was my preferred option …[it] was most endearing"
Exemplar of "Good singing" (*n* = 40)	18 (45.00%) "because very few people can sing that way" "they sounded like they knew how to sing as the voices where more skilled at reaching higher more complicated notes." "the embellished, the singers had been trained and didn't follow the original version of the song, the song was a mystery, you couldn't predict how it would sound and that is how interesting good music is meant to be to captivate its audience." "they were able to showcase what their voices could do and show how good of a singer they were"	22 (55.00%) "The female professional singer in the professional plain version. She sounded like she had good control of her voice but wasn't overdoing it with the flourishes in the professional embellished version, which almost sounded like a parody of "good singing"!" "it flowed well and showcased the musical talent of the singers. It demonstrated that you don't have to go 'full out' in a performance to impress others." "the singers displayed their vocal talents with precise diction and were consistent with the tone of the song." "their voices were soft, experienced, and pleasant to listen to."	0 (0%)