Comparing vocal emotion perception between singers and instrumentalists

Christine Nussbaum1,2, Jessica Dethloff1,2, Annett Schirmer3,2, and Stefan R. Schweinberger1,2,4

1Department for General Psychology and Cognitive Neuroscience, Friedrich Schiller University Jena, Germany

2Voice Research Unit, Friedrich Schiller University, Jena, Germany

3Institute of Psychology, University of Innsbruck, Austria

4Swiss Center for Affective Sciences, University of Geneva, Switzerland

Correspondence should be addressed to Christine Nussbaum, Department for General Psychology and Cognitive Neuroscience, Friedrich Schiller University Jena, Leutragraben 1, 07743 Jena, Germany. Tel: +49 (0) 3641 945939, E-Mail: christine.nussbaum@uni-jena.de; or to Stefan R. Schweinberger, Department for General Psychology and Cognitive Neuroscience, Friedrich Schiller University Jena, Am Steiger 3/Haus 1, 07743 Jena, Germany. Tel: +49 (0) 3641 945181, Fax: +49 (0)3641 9 45182, E-Mail: stefan.schweinberger@uni-jena.de. Preprocessed data, analysis scripts and supplemental materials can be found in the associated OSF repository (ToDo)

**Word count:** aim 5000

Contents

[*Abstract* 3](#_Toc194687188)

[*Public significance statement* 4](#_Toc194687189)

[1 Introduction: associations between musicality and vocal emotion perception 5](#_Toc194687190)

[1.1 Singers vs. instrumentalists 6](#_Toc194687191)

[1.2 Amateurs vs. professional musicians 7](#_Toc194687192)

[1.3 Rationale, outline etc. 8](#_Toc194687193)

[2 Part I: Comparison of non-professional singers and instrumentalists 9](#_Toc194687194)

[2.1 Hypotheses 9](#_Toc194687195)

[2.2 Method 10](#_Toc194687196)

[2.2.1 Participants 10](#_Toc194687197)

[2.2.2 Stimulus material 11](#_Toc194687198)

[2.2.3 Design 13](#_Toc194687199)

[2.2.4 Data analysis 15](#_Toc194687200)

[2.3 Transparency and openness 16](#_Toc194687201)

[2.4 Results 16](#_Toc194687202)

[2.4.1 Demography, musicality, and personality of participants 16](#_Toc194687203)

[2.4.2 Emotion classification performance 18](#_Toc194687204)

[3 Part II: Comparison of professionals, amateurs and non-musicians 20](#_Toc194687205)

[3.1 Hypotheses 20](#_Toc194687206)

[3.2 Method 21](#_Toc194687207)

[3.3 Results 21](#_Toc194687208)

[3.3.1 Demography, musicality, and personality of participants 21](#_Toc194687209)

[3.3.2 Emotion classification performance 24](#_Toc194687210)

[4 Part III: Correlational analyses 25](#_Toc194687211)

[4.1 Hypotheses 25](#_Toc194687212)

[4.2 Method 25](#_Toc194687213)

[4.3 Results 26](#_Toc194687214)

[5 Discussion 26](#_Toc194687215)

[5.1 Singers vs. instrumentalists 26](#_Toc194687216)

[5.2 Professional musicians vs. amateurs 27](#_Toc194687217)

[5.3 Constraints on generality and future directions 27](#_Toc194687218)

[6 Summary and Conclusion 28](#_Toc194687219)

[7 Acknowledgements 28](#_Toc194687220)

[8 Conflicts of Interests and Funding 28](#_Toc194687221)

[9 Credit Author Statement 28](#_Toc194687222)

[10 Supplementary material 29](#_Toc194687223)

[References 29](#_Toc194687224)

# Abstract

Musicians outperform non-musicians in vocal emotion recognition, presumably due to predisposed differences in auditory sensitivity for melodic patterns that carry emotional meaning. However, the current literature is inconclusive regarding differential effects in subgroups of musicians. Therefore, we focused on two contrasts potentially relevant in the context of vocal emotions: singers (N= 45) vs. instrumentalists (N=43) and professional musicians (N = 40) vs. amateurs (N = 88) vs. non-musicians (N = 38). Importantly, due to consistent evidence against a causal role of formal musical training, we assumed that vocal emotion recognition would be unaffected by the type and amount of musical training. Using both frequentist and Bayesian inference, we found the predicted null effects for singers vs. instrumentalists, and for professionals vs. amateurs. The pattern for amateurs vs. non-musicians was inconclusive. In a subsequent correlational analysis, we replicated the consistent link between vocal emotion perception and auditory sensitivity, especially for melodies. Hence, the current work adds a new perspective to the accumulating evidence that the musicians’ advantage for vocal emotions is associated with differences in auditory sensitivity but not to the type of musical activities or the amount of formal training.

**Keywords:** vocal emotion perception, singers, instrumentalists, amateurs, parameter-specific voice morphing, musicality

# Public significance statement

* ToDo

# Introduction: associations between musicality and vocal emotion perception

The human voice is a prime carrier of emotional information. Therefore, adequate perception of vocal emotions is important for everyday social interaction (Laukka et al., 2016; Schirmer et al., 2025). On average, humans can infer emotion from voices well above chance (Banse & Scherer, 1996; Juslin & Laukka, 2003; Scherer, 2018), but this capacity is subject to great individual variability and seems to be linked to differences in **musicality**. It has been shown repeatedly that musicians outperform non-musician in vocal emotion recognition, although the overall effect size can be considered small to moderate (M. Martins et al., 2021; Nussbaum & Schweinberger, 2021; Schellenberg & Lima, 2024). Several recent works sought to unravel the potential mechanisms underlying this advantage, and collectively tend to emphasize the role of predisposed **acoustic sensitivity** rather than causal effects of musical training: Musicians have more fine-grained basic auditory skills compared to non-musicians, such as pitch and rhythm perception, musical memory, or signal-in-noise discrimination (Baldé et al., 2025; Kraus & Chandrasekaran, 2010). Importantly, this does not seem to depend directly on formal musical training: Correia et al. (2022) found that the link between music training and vocal emotion perception was fully mediated via these auditory perception skills. The association between auditory perception and vocal emotion recognition was even observed in the absence of any formal musical training (Correia et al., 2022; Nussbaum et al., 2024). While these insights are limited by their correlational nature, the presumably strongest evidence is provided by a recent randomized-controlled study in school children, which found no causal effects of musical training on vocal emotion perception performance (Neves et al., 2025). Thus, there is consensus in the literature that the observed performance difference of musicians and non-musicians is due to variations in acoustic sensitivity rather than the result of formal musical education (Schellenberg & Lima, 2024).

In a previous study, we investigated how musicians’ auditory skills promote vocal emotion perception in more detail, by focusing on different auditory cues that transport emotional meaning (Nussbaum et al., 2024). We employed parameter-specific voice morphing to create vocal stimuli that expressed emotion only through fundamental frequency contour (F0), timbre, or both. F0 is linked to dynamic pitch variation (also referred to as voice melody), and timbre is linked to perceived voice quality (i.e. whether it sounds harsh or gentle). Professional musicians outperformed a group of non-musicians when emotions were expressed by F0 and both cues, but not timbre alone. This group difference was complemented by an exploratory correlation between melody perception in music and vocal emotion recognition. Thus, musicians seem to be specifically proficient at exploiting melodic pitch patterns to infer vocal emotions.

While the available literature paints a fairly consistent picture regarding the link between musicality and vocal emotion perception, a key limitation inherent in most studies targeting group differences is that they treat musicians as one uniform group, whereas they are, in fact, highly heterogeneous. On the one hand, there are quantitative differences regarding levels of expertise. On the other hand, qualitative differences between musicians can relate to a great variety of styles, genres, and forms of expression, within the scope of the Western music system and beyond. A particularly interesting distinction in the context of vocal emotions is the one between **singers and instrumentalists.** Singing is arguably the form of musical expression that is most closely related to vocal emotions (Akkermans et al., 2019; Mithen et al., 2006). When seen against substantial evidence for strong links between expression and perception in auditory communication via the voice (e.g., Frühholz & Schweinberger, 2021), one specific possibility is that, with their supreme expertise in vocal expression, singers might outperform instrumentalists in vocal emotion perception as well. Another interesting debate evolves around differences between **professional musicians and amateurs**. Therefore, the present study targeted these subgroups to explore differential patterns in vocal emotion perception. In what follows, we review current insights and outstanding research gaps, cumulating in the rationale for the present study.

## Singers vs. instrumentalists

Singing and playing an instrument are both fundamental forms of musical expression in humans, but they require very different motor skills and, typically, different amounts of formal musical training (Fisher et al., 2020; Krishnan et al., 2018). This is reflected in vocal performance differences, as singers outperformed instrumentalists in imitation of foreign speech (Christiner & Reiterer, 2015). For auditory perception, however, evidence is less clear. For example, two studies were unable to identify differences between singers and instrumentalists in pitch discrimination and pitch production overall (Nikjeh et al., 2008, 2009).

With regard to vocal emotion perception, evidence is sparse and inconclusive. Intuitively, it may be assumed that singing fosters vocal emotion perception abilities to a larger degree than instrumental activities. In line with this idea, both Correia et al. (2022) and Nussbaum et al. (2024) reported correlations between self-rated singing abilities and vocal emotion perception performance. This was complemented by a study of Greenspon and Montanaro (2023) showing that the same holds for objectively measured singing ability. On a neural level, however, I. Martins et al. (2022) found no differences in electrophysiological response to emotional voices between singers and instrumentalists, suggesting similar profiles of auditors processing. Intriguingly, a music-intervention study reported that singing may even interfere with vocal emotional processing, while keyboard lessons has a positive effect (Thompson et al., 2004). However, the validity of this finding is limited by an extensive drop-out of participants and a small sample size (Schellenberg & Lima, 2024). For the present study, we therefore recruited a well-powered sample of instrumentalists and singers for a comparison of vocal emotion recognition abilities.

## Amateurs vs. professional musicians

Most musicians start with their formal training in childhood, but when they enter adulthood, they pick different paths: some convert their musical activity into a profession, others pursue another career but keep it as a hobby. Interestingly, these groups seem to display several differences with regard to neurocognitive functioning. While amateurs, unsurprisingly, score lower on musical abilities, they show greater cognitive abilities than professional musicians (Vincenzi, 2022). Additionally, musical engagement as leisure activity seems to have a larger protective effect on “brain aging” than engaging in a music-related occupation (Rogenmoser et al., 2018). This also seems to be reflected in general health, which was found to be better in amateurs than professionals (Bonde et al., 2018). Specifically, professional face larger risks for their hearing and mental health (Hake et al., 2024; Loveday et al., 2023; Maghiar et al., 2023). Amateurs may gain more positive outcomes from their musical activity, perhaps because it provides enrichment in addition to their profession, while coming with less noise and less performative pressure. In contrast to this view, it was found that professionals more often experience a state of flow during their musical activity, which is usually considered very enjoyable (Rakei & Bhattacharya, 2024). However, to the best of our knowledge, there are no insights into differences between amateurs and professionals with regard to vocal emotion perception. This gap is addressed with the present study. Because all our newly recruited singers and instrumentalists were amateurs and we used an established experimental procedure, the present study also offered the opportunity to compare findings with our previously recruited groups of professional musicians and non-musicians (Nussbaum et al., 2024).

## Rationale, outline etc.

This study focuses on the comparison between singers and instrumentalists, and thus zooms into possible differences between specific subgroups or amateur musicians, while using the identical protocol as Nussbaum et al. (2024). Accordingly, we report findings in three parts. For Part I, we recruited an original sample of amateur instrumentalists and singers and assessed their vocal emotion recognition, their musical perception performance and self-rated musicality Nussbaum et al. (2024). For Part II, we collapsed all participants recruited in Part I into one group of amateur musicians and compared it to professional musicians and non-musicians from the previous study. In Part III, we explored the correlational links between musicality and vocal emotion recognition. Specifically, we investigated whether the positive relationships we observed in Nussbaum et al. (2024) between music and voice perception would replicate in the newly recruited samples of amateur musicians.

On first sight, it may be intuitive to assume that singers have a larger advantage in vocal emotion perception than instrumentalists, because their form of musical expression is of vocal nature. However, the literature consistently suggests that the link between musicality and vocal emotion perception is not driven by formal training, but rather by predisposed differences in auditory sensitivity. If this is the case, the form of musical engagement should not make a difference. We therefore predicted that singers and instrumentalists would perform equally in our vocal emotion recognition task. This prediction holds for emotions expressed by all available vocal cues, as well as emotions expressed by either F0 or timbre cues in the voice.

A similar logic holds for the comparison for professional musicians and amateurs. We predicted that their vocal emotion recognition performance would be comparable as well (for full emotions, as well as the F0 and timbre condition). However, as the evidence reviewed above suggested that amateurs can differ from professionals in cognitive abilities which could be linked to emotional sensitivity, we also considered the option that amateurs could be more proficient at making emotional inferences than professionals. Compared to our group of non-musicians, we assumed that amateurs would outperform them when emotion were expressed via full emotion cues and F0 cues only, but not timbre, because this is exactly the pattern we observed for professional musicians in Nussbaum et al. (2024). Finally, we aimed to replicate the positive link between music perception, especially sensitivity for melodies, and vocal emotion recognition. The whole study and its hypothesis were preregistered (<https://doi.org/10.17605/OSF.IO/76PV5>).

# Part I: Comparison of non-professional singers and instrumentalists

## Hypotheses

**H1:** We expect NO difference between singers and instrumentalists in overall vocal emotion recognition performance.

**H2:** We expect NO difference between singers and instrumentalists in vocal emotion recognition performance based on timbre and F0 cues only.

## Method

Note that this is a follow-up to the study reported in Nussbaum et al. (2024). Thus, the stimulus material and the design are identical, but we recruited a new sample.

### Participants

According to our preregistered plan, we aimed at a sample size of 40 singers (20 male, 20 female) and 40 instrumentalists (20 male, 20 female), because in our previous study, this sample size allowed us to reveal medium-sized group effects (d =0.81 for the Full and d = 0.56 for the F0 morphing condition) when we compared professional musicians and non-musicians.

Data were collected in a pseudonymized format from June 2023 to January 2024. All participants were aged between 18 and 54 years and fluent German speakers. Participants provided informed consent before completing the experiment and received compensation in the form of 12.50 € or course credit upon completion. The experiment was in line with the ethical guidelines of the German Society of Psychology (DGPs) and approved by the local ethics committee of the Friedrich Schiller University Jena (Reg.-Nr. FSV 19/045).

In total, we collected data from 94 non-professional musicians (further also referred to as amateurs) that were divided into singers and instrumentalists. Recruitment criteria specified that participants had to be non-professional musicians (i.e., they held no music-related academic degree or worked professionally as a musician). Singers were required to be currently active in a choir or another singing group but should not play an instrument actively and regularly (i.e., they must not currently be instrumentalists in an orchestra or a band). Instrumentalists, conversely, were required to be currently active in an orchestra or a band, but they should not engage in singing activities actively and regularly (i.e., they must not currently be in a choir or another singing group).

*Singers*

We recorded data from 48 singers, of which three were excluded (N = 2 had > 5 % trials of omission, N = 1 had technical issues during stimulus playback). Thus, data from 45 singers were analyzed (22 female, 22 male, 1 diverse, aged 18 to 53 years [M = 27.02, SD = 8.2]). Mean onset age of musical training was 8 years (SD = 3.08, 5 - 20 years). Twenty-two participants had over 10 years of training, ten between 6-9 years, four between 4-5 years and three between 2-3 years. Six participants had less than one year of training (for more details see ToDo).

*Instrumentalists*

Data from 46 instrumentalists were collected, of which three were excluded. One had technical issues during stimulus playback, one was also active in a choir, one held a master’s degree in music science and was therefore relocated into the group of professional musicians (see Part II). Thus, data from 43 instrumentalists entered analysis (24 female, 18 male, 1 diverse, aged 18 to 54 years [M = 28.51, SD = 10.64]). Mean onset of musical training was 7 years (SD = 2.27, 4 - 14 years). Thirty-five participants studied their instrument for over 10 years, one between 6-9 years and three between 1-2 years. Four participants had less than one year of training (for more details see ToDo).

### Stimulus material

As stimulus material, we used parameter-specific voice morphs that express emotional information either through the fundamental frequency contour only (F0), through timbre only (Tbr) or through a combination of both (Full).

For voice morphing, we selected original audio recordings from a database of vocal actor portrayals, comprised of pseudowords (/molen/, /loman/, /belam/) uttered by eight speakers (four male, four female) with expressions of happiness, pleasure, fear, and sadness. To synthesize the parameter-specific emotional voice morphs, we created morphing trajectories between each emotion and an emotional average of the same speaker and pseudoword, using the Tandem-STRAIGHT software (Kawahara et al., 2013; Kawahara et al., 2008). The averages had been created previously by blending all emotions together and were thus assumed to be uninformative and unbiased with respect to the four emotions of interest. After substantial preprocessing (e.g. manual mapping of time- and frequency anchors in each stimulus), Tandem-STRAIGHT enables voice morphing via weighted interpolation of five independent parameters: (1) F0-contour, (2) timing, (3) spectrum-level, (4) aperiodicity, and (5) spectral frequency; the latter three are summarized as timbre.

We created three types of morphed stimuli (see **Figure 1**). **Full-Morphs** were stimuli with all parameters taken from the emotional version (corresponding to 100% from the emotion and 0% from average), except for the timing parameter, which was taken from the average (corresponding to 0% emotion and 100% average). **F0-Morphs** were stimuli with the F0-contour taken from the emotion, but timbre and timing were taken from the average. **Timbre-Morphs** were stimuli with all timbre parameters taken from the emotion, but F0 and timing from the average. Note that the timing was kept constant in all conditions to allow a pure comparison of F0 vs. timbre. Furthermore, we kept all average stimuli as a further ambiguous reference category. In total, this resulted in 8 (speakers) x 3 (pseudowords) x 4 (emotions) x 3 (morphing conditions) + 24 average (8 speakers x 3 pseudowords) = 312 stimuli (duration M = 780 ms, range 620 to 967 ms, SD = 98 ms). Using PRAAT (Boersma, 2018), we normalized all stimuli to a root-mean-square of 70 dB SPL.

For a more detailed description of the stimulus creation, see Nussbaum et al. (2024). For general information about voice morphing, refer to Kawahara and Skuk (2019).

**Figure 1**

*Morphing matrix for stimuli with averaged voices as reference*



*Note. Figure reprinted from Nussbaum et al. (2024)*

### Design

Data were collected online via PsyToolkit (Stoet, 2010, 2017), but after completion of the study all participants met with the experimenter for a short personal debriefing. This was done to increase commitment and conscientiousness for the experiment.

Participants were required to ensure a quiet environment for the duration of the study and use a computer with a physical keyboard and headphones. As browser, we recommended Google Chrome, and excluded Safari for technical reasons. Prior to the listening tasks, participants could adjust their sound settings to a comfortable sound pressure level

First, participants entered demographic information, including age, sex, native language, profession, and potential hearing impairments such as tinnitus. They then completed an emotion classification experiment, a test on music perception (the Profile of Music Perception Skills) and several questionnaires on musicality, personality and socioeconomic background. Mean duration of the whole online study was about 75 minutes.

Emotion classification experiment

In the experiment, participants classified vocal emotions as happiness, pleasure, fear, or sadness. Each trial started with a green fixation cross presented for 500 ms. Then the sound was played while a loudspeaker symbol was shown on the screen. Subsequently, a response screen showed the emotion labels and participants could enter their response within a 5000 ms time window starting from voice offset. Responses were entered with the left and right index and middle fingers, with random mapping of response keys to emotion categories for each participant, out of four possible key mappings (see Tables on OSF ToDo). In case of a response omission, the final trial slide (500 ms) prompted participants to respond faster; otherwise, the screen turned back. Then the next trial started.

Preceding the experiment, participants completed eight practice trials with stimuli not used during the actual task. Subsequently, all 312 experimental stimuli were presented once in randomized order across six blocks of 52 trials each. Between blocks, participants could take self-paced breaks. The total duration of the experiment was about 25 minutes.

Profile of Music Perception Skills (PROMS)

To measure music perception skills, we used a modular version of the Profile of Music Perception Skills (Law & Zentner, 2012; Zentner & Strauss, 2017), comprised of the four subtests „Melody“, „Pitch“, „Timbre”, and „Rhythm“. Participants completed 18 items per subtest, always preceded by one practice trial. Each trial, participants heard a reference stimulus twice followed by a target stimulus. Then, they indicated whether reference and target were the same or different via a 5-point Likert scale with the labels “definitely same”, “maybe same”, “don’t know”, “maybe different”, and “definitely different”. The duration of the PROMS was about 20 minutes.

Questionnaires

After the PROMS, participants completed several questionnaires: the German Version of the Autism Quotient Questionnaire, AQ, (Baron-Cohen et al., 2001; Freitag et al., 2007), a 30-item Personality Inventory measuring the Big-Five domains (Rammstedt et al., 2018), the Goldsmiths Musical Sophistication Index, Gold-MSI, (Müllensiefen et al., 2014) to assess the participants’ degree of self-reported musical skills, additional questions concerning music experience and musical engagement, their socioeconomic background, and the 20-item version of the Positive-Affect-Negative-Affect-Scale, PANAS (Breyer & Bluemke, 2016; Watson et al., 1988).

### Data analysis

In line with our preregistered plan, we collapsed data across speakers and pseudowords for analysis. Further, data on emotional averages were excluded because they were not relevant for our hypotheses. Response omissions (~1%) were treated as errors and participants with more than 5% of such omissions excluded from data analysis. Analyses of Variance (ANOVAs) and correlational analyses were performed using R Version 4.3.2 (R Core Team, 2020). Post-hoc tests were Benjamini-Hochberg corrected where appropriate (Benjamini & Hochberg, 1995).

We complemented these classical frequentist analyses with a Bayesian approach, which – in contrast to null hypothesis significance testing - allows a quantification of evidence for null findings (Rosenfeld & Olson, 2021). These analyses were conducted in JASP Version 0.19.3 (Quelle) using default priors. We report the Bayes factor (BF10) as an indicator for the likelihood of the null and alternative hypothesis given the observed data. BF10 > 1 indicate larger evidence for the alternative hypothesis, BF10 < 1 indicate larger evidence for the null hypothesis. For example, a BF10 = 3 means that the alternative hypothesis is three times more likely than the null hypothesis, whereas the reciprocal BF10= .33 means that the null hypothesis is three times more likely than the alternative. Following the guidelines by Jarosz and Wiley (2014), we consider values of BF10 = 1-3 (.1-.33) as anecdotal, BF10 = 3-10 (.33-.10) as moderate, BF10 = 10-30 (.10-.03) as strong, BF10 = 30-100 (.03-.01) as very strong and BF10 > 100 (< .01) as decisive evidence for the alternative hypothesis, and the reciprocal values in parentheses as respective evidence for the null hypothesis.

In alignment with the approach by Nussbaum et al. (2024), we recoded responses in the PROMS from 0 to 1 in 0.25 steps starting with the “definitely” correct option down two the “definitely” incorrect option (thus, “don’t know” was always coded with 0.5) and subtracted 0.5 from the final measure. Thus, a positive score indicates that participants were more correct/confident, whereas a negative score indicates more incorrect/uncertain ratings. For statistical analyses, we used the averaged performance across trials for each subtest.

## Transparency and openness

We specified how we determined our sample size, all data exclusions, all manipulations, and all measures in the associated preregistration ([https://doi.org/10.17605/OSF.IO/76PV5](https://doi.org/10.17605/OSF.IO/76PV5" \t "_self) ). Preprocessed data, analysis scripts and supplemental materials can be found in the associated OSF repository (ToDo). For stimulus examples, refer to Nussbaum et al. (2024).

## Results

### Demography, musicality, and personality of participants

Singers and instrumentalists did not differ significantly in the socioeconomic status assessed via educational level, X2(2, N = 88) = 1.06, p = .588; highest academic degree, X2(7, N = 88) = 9.06, p = .249, and household income, X2(4, N = 88) = 5.23, p = .264 (for more details see Table ToDo). Further, the groups did not differ in age or positive and negative affect (assessed with the PANAS) and were comparable regarding Big Five personality traits and autistic traits. In the Gold-MSI, singers and instrumentalists scored comparatively on the general musicality score, but there were differences on two subfactors: instrumentalists scored higher on the subfactor Formal Education, while singers scored higher on Singing. In the PROMS, both groups performed comparably in all four subtests. Participant characteristics assessed via self-report and music performance in the PROMS are summarized in **Table 1**.

**Table 1**

*Characteristics of participants - Demography, personality, and musicality*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Singers** | **Instrumentalists** |  |  |  |  |  |
|  | M (SD) | M (SD) | t | dfa | p | Cohens d |  |
| Age | 27.02 (8.2) | 28.51 (10.6) | -0.73 | 78.93 | .465 | -0.17 [-0.61, 0.28] |  |
|  |  |  |  |  |  |  |  |
| *PANAS* |  |  |  |  |  |  |  |
| positive Affect | 3.00 (0.68) | 3.11 (0.57) | -0.77 | 84.78 | .446 | -0.17 [-0.59, 0.26] |  |
| negative Affect | 1.53 (0.47) | 1.40 (0.35) | 1.49 | 80.61 | .141 | 0.33 [-0.11, 0.77] |  |
|  |  |  |  |  |  |  |  |
| *Big Five* |  |  |  |  |  |  |  |
| Openness | 4.04 (0.55) | 3.99 (0.51) | 0.46 | 85.96 | .647 | 0.10 [-0.32, 0.52] |  |
| Conscientiousness | 3.47 (0.69) | 3.76 (0.70) | -1.91 | 85.62 | .060 | -0.41 [-0.84, 0.02] |  |
| Extraversion | 3.21 (0.70) | 3.00 (0.73) | 1.44 | 85.3 | .155 | 0.31 [-0.12, 0.74] |  |
| Agreeableness | 3.81 (0.57) | 4.01 (0.60) | -1.61 | 85.29 | .112 | -0.35 [-0.77, 0.08] |  |
| Neuroticism | 2.74 (0.77) | 2.61 (0.78) | 0.80 | 85.75 | .426 | 0.17 [-0.25, 0.60] |  |
|  |  |  |  |  |  |  |  |
| *AQ* |  |  |  |  |  |  |  |
| Total | 18.2 (6.15) | 19.28 (8.55) | -0.68 | 76.04 | .500 | -0.16 [-0.60, 0.30] |  |
| Attention to Detail | 5.4 (2.33) | 5.63 (2.53) | -0.44 | 84.64 | .662 | -0.10 [-0.52, 0.33] |  |
| Social | 12.8 (5.37) | 13.65 (7.53) | -0.61 | 75.65 | .545 | -0.14 [-0.59, 0.31] |  |
| Social Skills | 2.4 (1.94) | 3.09 (2.95) | -1.30 | 72.01 | .200 | -0.31 [-0.77, 0.16] |  |
| Communication | 2.53 (1.94) | 2.44 (2.3) | 0.20 | 82.02 | .842 | 0.04 [-0.39, 0.48] |  |
| Imagination | 2.51 (1.75) | 2.81 (1.88) | -0.78 | 84.86 | .437 | -0.17 [-0.60, 0.26] |  |
| Attention Switching | 5.36 (1.91) | 5.30 (2.23) | 0.12 | 82.69 | .904 | 0.03 [-0.40, 0.46] |  |
|  |  |  |  |  |  |  |  |
| *Gold-MSI* |  |  |  |  |  |  |  |
| General ME | 4.78 (0.85) | 4.75 (0.80) | 0.19 | 85.99 | .866 | 0.04 [-0.39, 0.46] |  |
| Active Engagement | 3.83 (0.82) | 4.21 (1.13) | -1.79 | 76.79 | .078 | -0.41 [-0.86, 0.05] |  |
| Formal Education | **4.39 (1.14)** | **4.95 (0.62)** | -2.85 | 68.31 | .006 | -0.69 [-1.18, -0.20] | \*\* |
| Emotion | 5.50 (0.81) | 5.60 (0.76) | -0.60 | 85.99 | .549 | -0.13 [-0.55, 0.29] |  |
| Singing | **4.98 (0.97)** | **4.19 (1.27)** | 3.25 | 78.56 | .002 | 0.73 [0.27, 1.19] | \*\* |
| Perception | 5.73 (0.82) | 5.77 (1.03) | -0.22 | 80.11 | .825 | -0.05 [-0.49, 0.39] |  |
|  |  |  |  |  |  |  |  |
| *PROMS* |  |  |  |  |  |  |  |
| Pitch | 0.23 (0.08) | 0.24 (0.06) | -0.30 | 82.92 | .766 | -0.07 [-0.50, 0.37] |  |
| Melody | 0.17 (0.10) | 0.14 (0.10) | 1.29 | 85.08 | .199 | 0.28 [-0.15, 0.71] |  |
| Timbre | 0.29 (0.08) | 0.3 (0.09) | -0.59 | 86.00 | .556 | -0.13 [-0.55, 0.30] |  |
| Rhythm | 0.31 (0.09) | 0.30 (0.08) | -0.56 | 85.13 | .577 | -0.12 [-0.55, 0.30] |  |

*Note. Descriptive values show mean ratings for the PANAS (Breyer & Bluemke, 2016), the Big-Five Domains (Rammstedt et al., 2018), and the Gold-MSI (Müllensiefen et al., 2014). AQ score were calculated based on Hoekstra et al. (2008) and Baron-Cohen et al. (2001).*

*a Note that original degrees of freedom were 86 but were corrected due to unequal variance.*

### Emotion classification performance

The mean proportion of correct responses was submitted to an ANOVA with Emotion (Happiness, Pleasure, Fear, and Sadness) and Morph Type (Full, F0, and Timbre) as repeated measures factors and Group (singers and instrumentalists) as a between subject factor (see **Table 2**).

The results revealed main effects of **Emotion** and **Morph Type**, which were qualified by an interaction. Crucially, however, we found no main effects or interactions involving **Group** (see **Figure 3**). Planned Bayesian analysis revealed moderate evidence for the null effect of group for overall performance (p = .542, BF10 = 0.265), as well as for Full (p = .392, BF10 = 0.31), F0 (p = .935, BF10 = 0.226), and Timbre morphs (p = .555, BF10 = 0.262) separately. Thus, we found evidence consistent with our hypotheses H1 and H2.

**Table 2**

*Results of the 4 × 3 × 2 mixed-effects ANOVA on the mean proportion of correct responses*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **df1** | **df2** | **F** | **p** | **Ωp2 [95%-CI]** | ***ε*HF** |
| Group | 1 | 86 | 0.38 | .542 | .00 [.00 .01] |  |
| Emotion | 3 | 258 | 72.43 | <.001 | .45 [.36 .53] |  |
| Morph Type | 2 | 172 | 768.93 | <.001 | .90 [.87 .93] | .741 |
| Group x Emotion | 3 | 258 | 2.14 | .095 | .01 [.00 .04] |  |
| Group x Morph Type | 2 | 172 | 0.36 | .635 | .00 [.00 .01] |  |
| Emotion x Morph Type | 6 | 516 | 22.78 | <.001 | .20 [.14 .25] | 827 |
| Group x Emotion x Morph Type | 6 | 516 | 1.33 | .249 | .00 [.00 .01] |  |

**Figure 3**

Mean proportion of correct responses per Morph Type separately for singers and instrumentalists



Note. Whiskers represent 95% confidence intervals. Violin plots represent variation of individual participants. The dotted line represents guessing rate at .25.

Follow-up analysis of the Morph Type effect revealed that performance was best in the Full condition (*M* = 0.75 ± 0.02 SEM), followed by the F0 (*M* = 0.62 ± 0.02) and then the Timbre condition (*M* = 0.41 ± 0.01); Full vs. F0: |*t*(87)| = 23.28, *p* < .001, *d* = 2.50 [2.07, 2.92], F0 vs Timbre: |*t*(87)| = 21.44, *p* < .001, *d* = 2.30 [1.90, 2.70], Full vs Timbre: |*t*(87)| = 33.94, *p* < .001, *d* = 3.64 [3.06, 4.21]). This Morph Type main effect was also found for all emotions separately (all *F*s(2, 174) > 102.44, *p* < .001), although it differed slightly between emotions, as suggested by the interaction (see **Figure 3,** for all post-hoc tests, refer to OSF).

To address our specific interest in the relative importance of F0 and Timbre for the different emotions, we calculated the performance differenceF0-Tbr for each emotion separately. Performance difference was largest for Happiness (M = 0.34 ± 0.02 SEM), followed by Fear (M = 0.21 ± 0.02), Sadness (M = 0.18 ± 0.02), and Pleasure (M = 0.10 ± 0.02; all pairwise comparisons |*t*s(77)| ≥ 2.57, pscorrected ≤ .012, ds ≥ 0.28 [0.06 0.49], except for Fear vs. Sadness (|t(87)| = 1.13, pcorrected = .261). These effects of Morph Type and Emotion therefore present a full replication of the patterns reported in Nussbaum et al. (2024).

**Figure 3**

Mean proportion of correct responses per Emotion and Morph Type



Note. Whiskers represent 95%-confidence intervals. Grey dots represent individual participants’ data. The dotted line represents guessing rate at .25.

ToDo: additional analyses like confusion data or classification of averages on OSF

# Part II: Comparison of professionals, amateurs and non-musicians

ToDo – vllt ne kurze Überleitung

## Hypotheses

**H3:** Amateur musicians outperform non-musicians in vocal emotion perception, in the Full and in the F0 condition.

**H4:** Amateurs perform equally well or better to professional musicians in the Full and the F0 condition.

## Method

For this analysis, we collapsed all participants from Part I into the group of amateur musicians and compared it to the groups of professional musicians and non-musicians reported in Nussbaum et al. (2024). Note that we added one participant to the professional group, because he held a master’s degree in music (see Part I), so numbers slightly diverge from the original publication. All professional musicians reported to have a music-related academic degree or a non-academic music qualification. Non-musicians were required not to play an instrument or engage in any other musical activities. For a more detailed description, please refer to Nussbaum et al. (2024).

In total, we analyzed data from 40 professional musicians (20 male, 20 female, aged 20 to 42 years [*M* = 29.6; *SD* = 5.58]), 38 non-musicians (18 male, 20 female, aged 19 to 48 years [*M* = 30.5; *SD* = 6.54]) and 88 amateurs (40 male, 46 female, 2 diverse, aged 18 to 54 years [*M* = 27.75; *SD* = 9.44].

The stimulus material, design and data analysis were identical to Part I. We focused our analysis on the comparison of amateurs with the other two groups, because the comparison of professional musicians and non-musicians is reported in Nussbaum et al. (2024).

## Results

### Demography, musicality, and personality of participants

Professionals, amateurs and non-musicians did not differ in the socioeconomic status assessed via educational level (χ²(6, N = 166) = 11.11, p = .085) and highest academic degree (χ²(16, N = 166) = 24.04, p = .089). However, there were differences regarding household income (χ²(8, N = 166) = 20.19, p = .01, ϕ = .25), with amateurs reporting higher household income than professionals and non-musicians (for more details see OSF).

The groups were comparable in age as well as in positive and negative affect (assessed with the PANAS). For the Big Five, analyses of variance revealed group differences for extraversion, with slightly higher levels in professionals than in amateurs (|t(82.15)|= 2.91, p = .005, d = 0.64 [0.20, 1.08]; see OSF for a detailed summary). Regarding autistic traits, the three groups did not differ in their overall score, but there were differences on the several subscales. For the detailed pattern and all post-hoc tests, please refer to OSF.

In the Gold-MSI, professional musicians scored significantly higher than amateurs (all subscales |*t*s| ≥ 4.08, *p*s < .001), except for the Emotion subscale (|*t*(80.76)|= 2.29, *p* = .025); in turn, amateurs scored higher than non-musicians (all subscales |*t*s| ≥ 2.59 , *p*s < .013). This is a pattern (professionals > amateurs > non-musicians) one would expect for self-rated musicality.

In the PROMS, professionals outperformed amateurs in the Pitch and Melody subtest (Pitch: |t(97.32)| = 2.57, p = .012, d = 0.55 [0.12, 0.98]; Melody: |t(95.24)| = 4.42, p < .001, d = 0.91 [0.48, 1.33]), whereas there were no differences in the Timbre and Rhythm subtests (*p*s ≥ .09). Amateurs performed better than non-musicians in the Pitch, Melody and Rhythm subtest (Pitch: |t(81.21)| = 4.39, p < .001, d = 0.97 [0.51, 1.43]; Melody: |t(91.34)| = 5.65, p < .001, d = 1.18 [0.74, 1.62]; Rhythm: |t(80.84)| = 3.16, p = .002, d = 0.7 [0.25, 1.15]), but not in the Timbre subtest (p = .064). Thus, a clear pattern of professionals > amateurs > non-musicians was only found for melody and pitch. Participant characteristics are summarized in **Table 3.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Professionals** |  | **Amateurs** |  | **Non-Musicians** |
|  | M (SD) |  | M (SD) |  | M (SD) |
| *PANAS* |  |  |  |  |  |
| positive Affect | 3.32 (0.65) |  | 3.05 (0.63) |  | 3.1 (0.67) |
| negative Affect | 1.69 (0.48) |  | 1.47 (0.42) |  | 1.49 (0.69) |
|  |  |  |  |  |  |
| *Big Five* |  |  |  |  |  |
| Openness | 4.12 (0.50) |  | 4.02 [0.53) |  | 3.81 (0.80) |
| Conscientiousness | 3.49 (0.71) |  | 3.61 (0.70) |  | 3.76 (0.72) |
| Extraversion | **3.48 (0.66)** | > | **3.11 (0.72)** |  | 3.38 (0.79) |
| Agreeableness | 3.92 (0.57) |  | 3.91 (0.59) |  | 3.75 (0.66) |
| Neuroticism | 2.95 (0.65) |  | 2.69 (0.77) |  | 2.58 (0.82) |
|  |  |  |  |  |  |
| *AQ* |  |  |  |  |  |
| Total | 15.7 (4.98) |  | 18.73 (7.40) |  | 17.58 (6.41) |
| Attention to Detail | 5.43 (2.04) |  | **5.51 (2.42)** | > | **4.32 (2.01)** |
| Social | **10.28 (4.70)** | < | **13.22 (6.49)** |  | 13.26 (6.51) |
| Social Skills | **1.48 (1.68)** | < | **2.74 (2.49)** |  | 2.61 (2.63) |
| Communication | 1.85 (1.61) |  | 2.49 (2.12) |  | 2.39 (1.73) |
| Imagination | 2.18 (1.52) |  | 2.66 (1.8) |  | 2.87 (1.95) |
| Attention Switching | 4.78 (1.91) |  | 5.33 (2.06) |  | 5.39 (1.92) |
|  |  |  |  |  |  |
| *Gold-MSI* |  |  |  |  |  |
| General ME | **5.68 (0.50)** | > | **4.76 (0.82)** | > | **2.74 (1.07)** |
| Active Engagement | **4.94 (0.81)** | > | **4.02 (1.00)** | > | **2.95 (1.19)** |
| Formal Education | **5.95 (0.56)** | **>** | **4.66 (0.96)** | > | **1.71 (0.68)** |
| Emotion | 5.88 (0.73) |  | **5.55 (0.78)** | > | **4.95 (1.32)** |
| Singing | **5.34 (0.83)** | **>** | **4.59 (1.19)** | > | **2.84 (1.26)** |
| Perception | **6.31 (0.51)** | > | **5.75 (0.92)** | > | **4.22 (1.49)** |
|  |  |  |  |  |  |
| *PROMS* |  |  |  |  |  |
| Pitch | **0.27 (0.06)** | > | **0.24 (0.07)** | > | **0.18 (0.06)** |
| Melody | **0.23 (0.08)** | > | **0.16 (0.10)** | > | **0.07 (0.08)** |
| Timbre | 0.32 (0.08) |  | 0.29 (0.08) |  | 0.26 (0.09) |
| Rhythm | 0.33 (0.08) |  | **0.32 (0.09)** | > | **0.27 (0.08)** |

**Table 3**

*Characteristics of participants - Demography, personality, and musicality*

*Note. Descriptive values show mean ratings for the PANAS (Breyer & Bluemke, 2016), the Big-Five Domains (Rammstedt et al., 2018), and the Gold-MSI (Müllensiefen et al., 2014). AQ score were calculated based on Hoekstra et al. (2008) and Baron-Cohen et al. (2001). Comparison signs (“>” or “<”) indicate significant differences. For a full report of statistical details, please refer to OSF.*

### Emotion classification performance

The mean proportion of correct responses was submitted to an ANOVA with Emotion (Happiness, Pleasure, Fear, and Sadness) and Morph Type (Full, F0, and Timbre) as repeated measures factors and Group (professionals, amateurs, and non-musicians) as a between subject factor (see **Table 4**).

The results revealed the main effects of **Emotion** and **Morph Type**, which were qualified by interaction. Crucially, however, we found no main effect involving **Group,** and only a trend for an interaction with **Morph Type**.

**Table 4**

*Results of the 4 × 3 × 3 mixed-effects ANOVA on the mean proportion of correct responses*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **df1** | **df2** | **F** | **p** | **Ωp2 [95%-CI]** | ***ε*HF** |
| Group | 2 | 163 | 1.96 | .144 | .01 [.00 .06] |  |
| Emotion | 3 | 489 | 130.24 | <.001 | .44 [.38 .49] |  |
| Morph Type | 2 | 326 | 1357.80 | <.001 | .89 [.87 .91] | .829 |
| Group x Emotion | 6 | 489 | 1.17 | .322 | .00 [.00 .00] |  |
| Group x Morph Type | 4 | 326 | 2.14 | .089 | .01 [.00 .04] | .829 |
| Emotion x Morph Type | 6 | 978 | 40.95 | <.001 | .20 [.15 .24] | .875 |
| Group x Emotion x Morph Type | 12 | 978 | 0.74 | .688 | .00 [.00 .00] | .865 |

Planned comparisons between **professionals and amateurs** revealed moderate evidence for the null effect of overall performance (p = .473, BF10 = 0.238), as well as for Full (p = .322, BF10 = 0.286), F0 (p = .435, BF10 = 245), and Timbre morphs (p = .840, BF10 = 0.205) separately. Planned comparisons between **amateurs and non-musicians** revealed inconclusive evidence for overall performance (p = .107, BF10 = 0.546), as well as for Full (p = .044, BF10 = 1.009), and F0 (p = .105, BF10 = 0.576) separately. For Timbre morphs, there was moderate evidence for the null effect (p = .975, BF10 = 0.205). Thus, we found evidence consistent with our hypotheses H4, but inconclusive evidence regarding H3. 

# Part III: Correlational analyses

## Hypotheses

*Correlations with the PROMS*

**H5:** Averaged vocal emotion recognition (VER) performance is correlated with averaged music perception performance.

**H6:** Full-VER and F0-VER are correlated with melody perception in music

*Correlations with the GOLD-MSI:*

**H7:** Averaged-VER and Full-VER are correlated with the General-ME.

**H8:** Averaged-VER and Full-VER are correlated with the Perception Subscale

**H9:** Averaged-VER and Full-VER are correlated with self-rated singing abilities.

## Method

ToDo, Wichtig: nur auf Basis der neuen Daten aus Part I, da die Daten, die in Part II dazugekommen sind ja diese Hypothesen motiviert haben. Die jetzt in die Analyse einzuschließen wäre also unzulässig.

Aus preregistration: - all correlations are controlled for formal musical education

## Results

ToDo

We found evidence for our hypotheses H5 and H6, but not for the hypotheses H7 – H9.

# Discussion

ToDo

## Singers vs. instrumentalists

ToDo

* The distinction between singers and instrumentalists is not as straight-forward, as it may seem
* Singers has a higher performance variance (not surprising, because engaging in singing activities has a lower threshold than playing an instrument, therefor greater variety of people with musical skills)
* Accordingly, structural differences have been observed in the brain, with singers showing increased volume and microstructural complexity in the arcuate fasciculus, a white matter tract connecting regions involved in sound perception and production (Halwani et al., 2011).
* Vocal acotors?
* Tragantzopoulou, P., & Giannouli, V. (2025). A Song for the Mind: A Literature Review on Singing and Cognitive Health in Aging Populations. Brain Sciences, 15(3), 227.
* Choi, W., Ling, C. L. K., & Wu, C. H. J. (2024). Musical Advantage in Lexical Tone Perception Hinges on Musical Instrument: A Comparison between Pitched Musicians, Unpitched Musicians, and Nonmusicians. Music Perception: An Interdisciplinary Journal, 41(5), 360-377.
* Moisseinen, N., Ahveninen, L., Martínez‐Molina, N., Sairanen, V., Melkas, S., Kleber, B., ... & Särkämö, T. (2024). Choir singing is associated with enhanced structural connectivity across the adult lifespan. Human Brain Mapping, 45(7), e26705.
* (Papadaki et al., 2023)

## Professional musicians vs. amateurs

ToDo

* This distinction is also not super straight-forward
* Taylor, A., & Hallam, S. (2011). From leisure to work: Amateur musicians taking up instrumental or vocal teaching as a second career. Music Education Research, 13(3), 307-325.
* Zendel, B. R., & Alexander, E. J. (2020). Autodidacticism and music: Do self-taught musicians exhibit the same auditory processing advantages as formally trained musicians?. Frontiers in Neuroscience, 14, 752.

## Constraints on generality and future directions

ToDo

* Brain measures
* Tervaniemi, M., Castaneda, A., Knoll, M., & Uther, M. (2006). Sound processing in amateur musicians and nonmusicians: event-related potential and behavioral indices. Neuroreport, 17(11), 1225-1228.
* Lotze, M., Scheler, G., Tan, H. R., Braun, C., & Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage*, *20*(3), 1817-1829.
* Oechslin, M. S., Van De Ville, D., Lazeyras, F., Hauert, C. A., & James, C. E. (2013). Degree of musical expertise modulates higher order brain functioning. Cerebral cortex, 23(9), 2213-2224.
* Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., & Lotze, M. (2010). The brain of opera singers: experience-dependent changes in functional activation. Cerebral Cortex, 20(5), 1144-1152.

# Summary and Conclusion

ToDo

# Acknowledgements

This study was conducted by J.D. in partial fulfilment of the requirements for a master’s thesis. The original voice recordings that served as a basis for creating our stimulus material were provided by Sascha Frühholz. We thank Hannah Strauß for support with the PROMS. We are grateful to all participants of the study.

# Conflicts of Interests and Funding

The authors declare no conflicts of interests.

# Credit Author Statement

Christine Nussbaum – Conceptualization, Methodology, Software, Visualization, Formal analysis, Writing - Original Draft, Supervision

Jessica Dethloff - Data collection, Formal analysis, Visualization, Writing - Original Draft

Annett Schirmer – Methodology, Writing - Review & Editing, Supervision

Stefan R. Schweinberger – Conceptualization, Writing - Review & Editing, Supervision

# Supplementary material

Supplemental figures and tables, analysis scripts, and raw data can be found on the associated OSF repository (ToDo).

ADDIN CitaviBibliographyReferences

Akkermans, J., Schapiro, R., Müllensiefen, D., Jakubowski, K., Shanahan, D., Baker, D., Busch, V., Lothwesen, K., Elvers, P., Fischinger, T., Schlemmer, K., & Frieler, K. (2019). Decoding emotions in expressive music performances: A multi-lab replication and extension study. *Cognition & Emotion*, *33*(6), 1099–1118. https://doi.org/10.1080/02699931.2018.1541312

Baldé, A. M., Lima, C. F., & Schellenberg, E. G. (2025). Associations between musical expertise and auditory processing. *Journal of Experimental Psychology: Human Perception and Performance.* Advance online publication. https://doi.org/10.1037/xhp0001312

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol*, *70*(3), 614–636. https://doi.org/10.1037/0022-3514.70.3.614

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J.‑C., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Breyer, B., & Bluemke, M. (2016). *Deutsche Version der Positive and Negative Affect Schedule PANAS (GESIS Panel).* https://doi.org/10.6102/zis242

Christiner, M., & Reiterer, S. M. (2015). A Mozart is not a Pavarotti: Singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience*, *9*, 482. https://doi.org/10.3389/fnhum.2015.00482

Correia, A. I., Castro, S. L [Sao Luis], MacGregor, C., Müllensiefen, D., Schellenberg, E. G., & Lima, C. F. (2022). Enhanced recognition of vocal emotions in individuals with naturally good musical abilities. *Emotion*, *22*(5), 894–906. https://doi.org/10.1037/emo0000770

Fisher, R. A., Hoult, A. R., & Tucker, W. S. (2020). A Comparison of Facial Muscle Activation for Vocalists and Instrumentalists. *Journal of Music Teacher Education*, *30*(1), 53–64. https://doi.org/10.1177/1057083720947412

Freitag, C. M., Retz-Junginger, P., Retz, W., Seitz, C., Palmason, H., Meyer, J., Rösler, M., & Gontard, A. von (2007). Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ) - die Kurzversion AQ-k. *Zeitschrift Für Klinische Psychologie Und Psychotherapie*, *36*(4), 280–289. https://doi.org/10.1026/1616-3443.36.4.280

Greenspon, E. B., & Montanaro, V. (2023). Singing ability is related to vocal emotion recognition: Evidence for shared sensorimotor processing across speech and music. *Attention, Perception & Psychophysics*, *85*(1), 234–243. https://doi.org/10.3758/s13414-022-02613-0

Hake, R., Kreutz, G., Frischen, U., Schlender, M., Rois-Merz, E., Meis, M., Wagener, K. C., & Siedenburg, K. (2024). A Survey on Hearing Health of Musicians in Professional and Amateur Orchestras. *Trends in Hearing*, *28*, 23312165241293762. https://doi.org/10.1177/23312165241293762

Halwani, G. F., Loui, P., Rüber, T., & Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists, and non-musicians. *Frontiers in Psychology*, *2*, 156. https://doi.org/10.3389/fpsyg.2011.00156

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. *J Autism Dev Disord*, *38*(8), 1555–1566. https://doi.org/10.1007/s10803-008-0538-x

Jarosz, A. F., & Wiley, J. (2014). What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *The Journal of Problem Solving*, *7*(1). https://doi.org/10.7771/1932-6246.1167

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol Bull*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Kawahara, H., Morise, M., & Skuk, V. G. (2013). Temporally variable multi-aspect N-way morphing based on interference-free speech representations. *IEEE International Conference on Acoustics, Speech and Signal Processing*.

Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T., & Banno, H. (2008). TANDEM-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation. *IEEE International Conference on Acoustics, Speech and Signal Processing*.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*(8), 599–605. https://doi.org/10.1038/nrn2882

Krishnan, S., Lima, C. F., Evans, S., Chen, S., Guldner, S., Yeff, H., Manly, T., & Scott, S. K. (2018). Beatboxers and Guitarists Engage Sensorimotor Regions Selectively When Listening to the Instruments They can Play. *Cerebral Cortex*, *28*(11), 4063–4079. https://doi.org/10.1093/cercor/bhy208

Laukka, P., Elfenbein, H. A., Thingujam, N. S., Rockstuhl, T., Iraki, F. K., Chui, W., & Althoff, J. (2016). The expression and recognition of emotions in the voice across five nations: A lens model analysis based on acoustic features. *J Pers Soc Psychol*, *111*(5), 686–705. https://doi.org/10.1037/pspi0000066

Law, L. N. C., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the profile of music perception skills. *PLoS One*, *7*(12), e52508. https://doi.org/10.1371/journal.pone.0052508

Loveday, C., Musgrave, G., & Gross, S.‑A. (2023). Predicting anxiety, depression, and wellbeing in professional and nonprofessional musicians. *Psychology of Music*, *51*(2), 508–522. https://doi.org/10.1177/03057356221096506

Maghiar, M. J., Lawrence, B. J., Mulders, W. H., Moyle, T. C., Livings, I., & Jayakody, D. M. P. (2023). Hearing loss and mental health issues in amateur and professional musicians. *Psychology of Music*, *51*(6), 1584–1597. https://doi.org/10.1177/03057356231155970

Martins, I., Lima, C. F., & Pinheiro, A. P. (2022). Enhanced salience of musical sounds in singers and instrumentalists. *Cogn Affect Behav Neurosci.* Advance online publication. https://doi.org/10.3758/s13415-022-01007-x

Martins, M., Pinheiro, A. P., & Lima, C. F. (2021). Does Music Training Improve Emotion Recognition Abilities? A Critical Review. *Emotion Review*, *13*(3), 199–210. https://doi.org/10.1177/17540739211022035

Mithen, S., Morley, I., Wray, A., Tallerman, M., & Gamble, C. (2006). The Singing Neanderthals: the Origins of Music, Language, Mind and Body by Steven Mithen. London: Weidenfeld & Nicholson, 2005. ISBN 0-297-64317-7 hardback £20 & US$25.2; ix+374 pp. *Cambridge Archaeological Journal*, *16*(1), 97–112. https://doi.org/10.1017/S0959774306000060

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), e89642. https://doi.org/10.1371/journal.pone.0101091

Neves, L., Martins, M., Correia, A. I., Castro, S. L [São Luís], Schellenberg, E. G., & Lima, C. F. (2025). Does music training improve emotion recognition and cognitive abilities? Longitudinal and correlational evidence from children. *Cognition*, *259*, 106102. https://doi.org/10.1016/j.cognition.2025.106102

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2008). Hearing of note: An electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians. *PSYCHOPHYSIOLOGY*, *45*(6), 994–1007. https://doi.org/10.1111/j.1469-8986.2008.00689.x

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2009). The relationship between pitch discrimination and vocal production: Comparison of vocal and instrumental musicians. *The Journal of the Acoustical Society of America*, *125*(1), 328–338. https://doi.org/10.1121/1.3021309

Nussbaum, C., Schirmer, A., & Schweinberger, S. R. (2024). Musicality - Tuned to the melody of vocal emotions. *Br J Psychol*, *115*(2), 206–225. https://doi.org/10.1111/bjop.12684

Nussbaum, C., & Schweinberger, S. R. (2021). Links Between Musicality and Vocal Emotion Perception. *Emotion Review*, *13*(3), 211–224. https://doi.org/10.1177/17540739211022803

Papadaki, E., Koustakas, T., Werner, A., Lindenberger, U., Kühn, S., & Wenger, E. (2023). Resting-state functional connectivity in an auditory network differs between aspiring professional and amateur musicians and correlates with performance. *Brain Structure & Function*, *228*(9), 2147–2163. https://doi.org/10.1007/s00429-023-02711-1

Rakei, A., & Bhattacharya, J. (2024). Professional status matters: Differences in flow proneness between professional and amateur contemporary musicians. *Psychology of Aesthetics, Creativity, and the Arts.* Advance online publication. https://doi.org/10.1037/aca0000674

Rammstedt, B., Danner, D., Soto, C. J., & John, O. P. (2018). Validation of the short and extra-short forms of the Big Five Inventory-2 (BFI-2) and their German adaptations. *European Journal of Psychological Assessment.* Advance online publication. https://doi.org/10.1027/1015-5759/a000481

Rosenfeld, J. P., & Olson, J. M. (2021). Bayesian Data Analysis: A Fresh Approach to Power Issues and Null Hypothesis Interpretation. *Applied Psychophysiology and Biofeedback*, *46*(2), 135–140. https://doi.org/10.1007/s10484-020-09502-y

Schellenberg, E. G., & Lima, C. F. (2024). Music Training and Nonmusical Abilities. *Annu Rev Psychol*, *75*, 87–128. https://doi.org/10.1146/annurev-psych-032323-051354

Scherer, K. R. (2018). Acoustic Patterning of Emotion Vocalizations. In S. Frühholz, P. Belin, & K. R. Scherer (Eds.), *The Oxford Handbook of Voice Perception* (pp. 60–92). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.4

Schirmer, A., Croy, I., Liebal, K., & Schweinberger, S. R. (2025). Non-verbal effecting - animal research sheds light on human emotion communication. *Biological Reviews of the Cambridge Philosophical Society*, *100*(1), 245–257. https://doi.org/10.1111/brv.13140

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: do music lessons help? *Emotion*, *4*(1), 46–64. https://doi.org/10.1037/1528-3542.4.1.46

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Zentner, M., & Strauss, H. (2017). Assessing musical ability quickly and objectively: development and validation of the Short‐PROMS and the Mini‐PROMS. *Annals of the New York Academy of Sciences*, *1400*(1), 33–45. https://doi.org/10.1111/nyas.13410