PLOS ONE
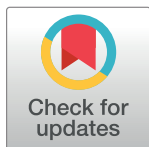
Speech-in-noise, psychosocial, and heart rate variability outcomes of group singing or audiobook club interventions for older adults with unaddressed hearing loss: A SingWell Project multisite, randomized controlled trial, registered report protocol

Chi Yhun Lo[1]*, Benjamin Rich Zendel[2], Deniz Baskent[3], Christian Boyle[4], Emily Coffey[5], Nathan Gagne[5], Assal Habibi[6], Ellie Harding[7], Merel Keijzer[7], Gunter Kreutz[8], Bert Maat[9], Eva Schurig[8], Mridula Sharma[4], Carmen Dang[1], Sean Gilmore[1], Helen Henshaw[10], Colette M. McKay[11], Arla Good[1], Frank A. Russo[1]

1 Department of Psychology, Toronto Metropolitan University, Toronto, ON, Canada, 2 Faculty of Medicine, Memorial University of Newfoundland, St John's, NL, Canada, 3 Faculty of Medicine, University of Groningen, Groningen, GR, Netherlands, 4 College of Nursing and Health Sciences, Flinders University, Adelaide, SA, Australia, 5 Department of Psychology, Concordia University, Montreal, QC, Canada, 6 Brain and Creativity Institute, University of Southern California, Los Angeles, CA, United States of America, 7 Faculty of Arts, University of Groningen, Groningen, GR, Netherlands, 8 Institute of Music, Carl von Ossietzky University of Oldenburg, Oldenburg, NI, Germany, 9 Department of Otorhinolaryngology, University of Groningen, Groningen, GR, Netherlands, 10 NIHR Nottingham Biomedical Research Centre, Hearing Sciences, School of Medicine, Mental Health and Clinical Neurosciences, University of Nottingham, Nottingham, United Kingdom, 11 Bionics Institute, Melbourne, VIC, Australia

* chi.lo@torontomu.ca

Abstract

Background

Unaddressed age-related hearing loss is highly prevalent among older adults, typified by negative consequences for speech-in-noise perception and psychosocial wellbeing. There is promising evidence that group singing may enhance speech-in-noise perception and psychosocial wellbeing. However, there is a lack of robust evidence, primarily due to the literature being based on small sample sizes, single site studies, and a lack of randomized controlled trials. Hence, to address these concerns, this SingWell Project study utilizes an appropriately powered sample size, multisite, randomized controlled trial approach, with a robust preplanned statistical analysis.

Objective

To explore if group singing may improve speech-in-noise perception and psychosocial wellbeing for older adults with unaddressed hearing loss.

Methods

We designed an international, multisite, randomized controlled trial to explore the benefits of group singing for adults aged 60 years and older with unaddressed hearing loss (registered

at clinicaltrials.gov, ID: NCT06580847). After undergoing an eligibility screening process and completing an information and consent form, we intend to recruit 210 participants that will be randomly assigned to either group singing or an audiobook club (control group) intervention for a training period of 12-weeks. The study has multiple timepoints for testing, that are broadly categorized as macro (i.e., pre- and post-measures across the 12-weeks), or micro timepoints (i.e., pre- and post-measures across a weekly training session). Macro measures include behavioural measures of speech and music perception, and psychosocial questionnaires. Micro measures include psychosocial questionnaires and heart-rate variability.

Hypotheses

We hypothesize that group singing may be effective at improving speech perception and psychosocial outcomes for adults aged 60 years and older with unaddressed hearing loss—more so than participants in the control group.

Introduction

Age-related hearing loss and its consequences

Hearing loss is common and currently affects 1.6 billion people worldwide [1]. The consequences of hearing loss have a significant impact on individual quality of life and place a considerable burden on global health and economic systems [2], with a conservative cost-analysis estimated at $981 billion [3]. This is exacerbated due to a rapidly ageing population, with hearing loss projected to extend globally to 2.1 billion people by 2050 [4]. Age-related hearing loss (ARHL, or presbycusis) is a progressive and bilateral sensorineural hearing loss that impacts both central and peripheral auditory functions, and disproportionately affects higher frequencies [5, 6]. ARHL is highly prevalent amongst older adults, with the estimated incidence doubling for every ten-year increase in age [7]. For hearing losses greater than 25 dB (equivalent to a mild hearing loss or worse), the estimated incidence varies from 42.3% to 63.1% for adults aged 70 years and older, depending on test context and population factors [7, 8].

ARHL is associated with communication challenges, which is typically exemplified by difficulties perceiving speech in the presence of noise [9–11]. This is significant given the congregation of humans at social activities and experiences such as dining-out at eateries are typified by noisy sound propagation. Real world noise levels indicate that older adults with mild to moderate hearing loss spend approximately 7.5% of their time in listening situations where noise meets or exceeds speech levels [12]. Importantly, while speech-in-noise (SIN) tasks used in clinical and research contexts are typically receptive in nature (i.e., focussed on individual ability to perceive SIN); the real-world outcomes are far broader, with significant social and emotional impacts that extend to communication partners and beyond [13, 14]. The ability to communicate effectively is fundamental to the development of strong interpersonal relationships and helps counteract loneliness and isolation, which is becoming a growing public health challenge [15]. Mechanistically, it has been posited that ARHL communication challenges reduce active social participation; this withdrawal has a cascading effect leading to social isolation and feelings of loneliness; finally, there are also increased risks to health and mortality [16–19]. While these risk factors are associated with general ageing [20, 21], these risks are significantly exacerbated by hearing loss.

There is currently no way to reverse ARHL and management is generally focused on the provision of assistive listening devices such as hearing aids to amplify sounds. However, despite the noted benefits of hearing aids for communication and broad quality of life outcomes [22], hearing aid adoption rates are low, with global estimates indicating 17% of persons that could benefit from a hearing aid do so [23]. Hence, most older adults with hearing loss have *unaddressed* hearing loss [24], acknowledging this has been referred to synonymously as untreated [17], unperceived [25], or uncorrected [26]. To the best of our knowledge, there is no gold-standard terminology and there may be subtle variations in definition. For the purposes of our study, a participant with unaddressed hearing loss refers to an individual with a mild-to-moderate sensorineural hearing loss that does not currently use an assistive listening device such as a hearing aid.

Benefits of group singing

The importance of social support, social networks, and social participation has been highlighted as a target for audiological rehabilitation, specifically for adults with ARHL [19]. In response to this, music and arts-based interventions for health and wellbeing have come to prominence in recent years with the benefits of being effective for preventing and promoting better health outcomes; helping to manage and treat health; and being engaging and cost-effective [27]. One potential non-clinical intervention showing promise is group singing—a universally practiced music activity that is more common than individual singing in approximately 70% of all societies [28]. There is an intimate and important cultural connection between music making and social context, with group singing being strongly associated with religious practice, dance, games, ceremony, and work [28].

Community choirs are popular, with 3.5 million Canadians having sung in 28,000 choirs in 2016 [29]. Typically consisting of amateur singers, the focus is to provide a supportive and inclusive environment where people can enjoy singing. These activities promote community engagement, cultural enrichment, and social connections among participants. Moreover, a diverse group of older adults that participated in a methodologically robust community choir intervention experienced significant benefits to loneliness and interest in life, compared to controls [30].

A potential benefit of group singing is the enhancement of speech perception, which is particularly relevant as ARHL is primarily associated with difficulties perceiving SIN [9–11]. Here, we provide a brief overview of two key theoretical frameworks that support speech perception: the OPERA hypothesis [31–33], and the Processing Rhythm in Speech and Music (PRISM) framework [34].

The OPERA hypothesis posits that music may engage speech processing networks if certain conditions are met. 1) *Overlap*—sensory and cognitive processing of music and speech is encoded by brain networks that overlap. More commonly referred to as neural sharing, there is ample neurological evidence to support the overlap between music and speech processing [35, 36]. 2) *Precision*—music must place a higher sensory and cognitive demand than speech. For example, pitch perception thresholds in music tend to operate at the semitone interval, whereas the speech intonation curve of a question utterance is much larger, typically extending beyond an octave (or 12 semitones) [37, 38]. 3) Music training that is enjoyable, regularly practiced, and engaging will satisfy the final requirement related to *Emotion*, *Repetition*, *and Attention*.

The PRISM framework suggests that three rhythm processes in both speech and music may lead to improved speech perception. 1) *Precise auditory processing*—both speech and music perception require accurate processing of small deviations. However, the regularity and

precision required in music production (particularly when performing with others) may fine-tune the auditory system, leading to enhancements in perceiving speech processes such as temporal envelope [39], or formant tracking (such as in plosives such as /ba/ and /da/) [40]. 2) *Synchronization/entrainment of neural oscillations to external rhythmic stimuli*—because musical rhythm typically features temporal regularities, music is a useful stimulus to entrain neural oscillations. Entrainment optimizes processes associated with attention and prediction [41], and hierarchical processing (i.e., phonemic, syllabic, phrase level speech perception) [42, 43], playing a critical role in speech and music perception. 3) *Sensorimotor coupling*—auditory and motor networks are linked and implicated in effective processing of both speech and music domains [44, 45]. Several studies have shown that music training has the capacity to strengthen sensorimotor coupling [46, 47], supporting a potential mechanism for the musician advantage.

Group singing also confers a unique opportunity to practice sensorimotor coupling—the dynamic integration of auditory, somatosensory, and motor systems [48, 49], which likely involves a dorsal pathway and its constituent fibre tracts integrating auditory cortices to the inferior frontal gyrus through the posterior parietal cortex [50]. This same dorsal pathway has been implicated in speech perception in noise, particularly as listening conditions become more challenging [51]. However, because a choir will experience tempo drift and micro-perturbations of intonation, a very high demand is placed on brain areas and fibre tracts subserving sensorimotor coupling. Singers in a choir are understandably motivated to coordinate sensory input and motor output to ensure accurate and harmonious vocal performance. Thus, motivated practice over a period of weeks and months may lead to neuroplastic changes (e.g., enhanced motor areas of the brain), which may ultimately support SIN perception [52, 53].

The musician advantage was brought to prominence in 2009 and is supported by empirical evidence that musicians outperformed non-musicians in SIN tasks—with years of consistent musical practice being positively associated with better SIN, working memory, and pitch perception [54]. More recently, a systematic review [55], and meta-analyses [56, 57], provided further support for this claim. Musicians routinely outperformed non-musicians on various SIN tasks, with more pronounced benefits across lower signal-to-noise ratios (SNRs), as well as across different distractor types and target properties, such as sentences, single words, phonemes, and tones. Other studies have shown stronger positive effects of musicianship on SIN performance across more challenging listening situations, suggesting greater benefits when the auditory system is under greater strain. One possible mechanism involves the type of training musicians receive (i.e., focusing on specific sounds in compositions that contain multiple simultaneous sounds), which may help them to better segregate auditory information in noisy environments. Neuroimaging studies have also revealed better auditory discrimination for musicians at the neural level (e.g. stronger spectral amplitudes in the Frequency-Following Response (FFR), more precise timing of transient response peaks, and improved Event-Related Potentials (ERP) waves such as P2 and N400). Studies have also shown a musician pattern of neural recruitment via shared pathways (both ventral and dorsal auditory streams in both hemispheres). The musician advantage in SIN may also be influenced by cognitive factors such as auditory working memory and selective attention, as well as age-related differences, with older musicians showing a greater SIN advantage compared to non-musicians of the same age. Additionally, an early age of onset and more extensive duration of musical training is correlated with stronger neural encoding and better performance on SIN tasks. Overall, these studies revealed a complex interaction between lower-level auditory processing and higher-level cognitive functions, indicating that musical training may enhance SIN perception through neuroplasticity and improved auditory processing. Despite these positive findings, some studies show variability in results, with either no significant advantage observed, or an advantage only under specific experimental conditions. A recent review raises questions

regarding the causality of music training and its transfer to non-musical benefits, such as SIN perception, are often weak or nonexistent [58].

Hence, the role of music training for individuals with hearing loss has been explored with cautious optimism, with the possibility that music training may enhance SIN perception as well as psychosocial outcomes. However, a systematic review of studies that tested the hypothesis that music training can enhance speech understanding in people with hearing loss showed that most of the studies had flaws in their methodologies and design such as no (or an inappropriate) control group, lack of randomisation, failure to account for multiple comparisons, or inadequate limiting of participant or tester bias [59]. Additionally, the four studies with appropriate designs did not show any benefit to speech understanding [60–63].

The strongest evidence that group singing enhances SIN for older adults with hearing loss, was a study involving a choir-singing group and an age- and audiometrically-matched control group, which did not participate in any activities, over a 10-week period [64]. Both groups were carefully matched based on previous musical experience, quantified by the years of formal musical training, thereby enabling causal inferences that are typically unachievable in correlational or cross-sectional studies of musicians' auditory capabilities. The choir participants exhibited significant improvements in SIN perception, pitch discrimination, and the robustness of the neural encoding of speech stimuli (frequency following response, FFR). The enhancement in SIN perception among choir participants was mediated by improved pitch discrimination, which was, in turn, predicted by the strength of the FFR. Although these results support the hypothesis that short-term participation in choir singing serves as an effective intervention for mitigating ARHL, reasonable concerns about participant bias may be raised due to the use of a self-selected choir group and a do-nothing control group.

Several systematic reviews have been conducted exploring the psychosocial benefits of singing (which is predominantly focused on group singing) in a range of populations and health conditions [65–68]. Overall, these findings indicate that group singing is an enriching activity that shows promise, but more research needs to be conducted to understand the therapeutic benefits for psychosocial wellbeing. Group singing was associated with better health-related quality of life, reduced mental distress, reduced symptoms of anxiety and depression, enhanced mood and emotional states, self-efficacy, a sense of purpose and achievement, social connectedness, and a sense of community and belonging. Physiological benefits have also been shown, with coordinated group singing leading to enhanced synchronization of respiration and heart rate variability (HRV) [69, 70].

There are several frameworks that suggest how and why music provides psychosocial benefits. Firstly, the musical and social bonding hypothesis posits that music-making is a coevolved system for social bonding [71]. The musical and social bonding hypothesis suggests that music can be likened to a toolkit, consisting of the following features: rhythm and dance (which supports synchronization and coordination); melody, harmony, and vocal learning (which supports unison and harmony); repetitive structure (which supports prediction); and music and social identity (which supports group identity). Importantly, these are all features common to group singing—or choir singing—which is a popular form of active music participation in the Western world. Notably, the musical and social bonding hypothesis argues that music making is more effective for large-scale bonding than other socially-facilitated contexts such as grooming or language [71].

Music has also been conceptualized in relation to its therapeutic benefits, consolidated into seven capacities in the Therapeutic Music Capacities Model [72]: 1) *Music is engaging*, utilizing a wide range of cognitive functions such as attention and working memory. There is evidence that long-term music training induces neuroplastic changes [73, 74]. Furthermore, while still inconclusive, there is some evidence that suggests music may enhance cognitive reserve [75],

and limit age-related cognitive decline [76, 77]; 2) *Music is emotional*, and can induce a wide variety of emotional states, and heighten arousal and reward [78]; 3) *Music is physical*, with movement and dance a typical component to music engagement. Significantly, physical activity (such as dance) is a known contributor for supporting healthy ageing and cognition [79, 80]; 4) *Music permits synchronization*. The ability of individuals to synchronize to an external rhythm (such as a musical beat) is a fairly unique feature of musical activity. Group synchronization is associated with cooperation, group cohesion, collective identity, enhanced non-verbal communication, and better learning outcomes [81–84]; 5) *Music is personal*, allowing for personal connection, such as the reinforcement or re-evaluation of self. This is particularly important in the context of ARHL, as the construction and negotiation of identity and self may face pressures due to age and hearing loss stigma [85, 86]; 6) *Music is social*, and helps create bonds and connections with greater efficacy compared to other group-based interventions, such as crafting [87]; and 7) *Music is persuasive*. The authors of the Therapeutic Music Capacities Model argue that music holds a unique cultural position in cultural contexts such as religion, advertising, and health, that may reinforce and support positive treatment outcomes. Thus, music has the potential to amplify therapeutic benefit, or enhance adherence to a program.

To summarise, in the context of a rapidly ageing population, untreated hearing loss has significant consequences that affect communication and psychosocial wellbeing at the individual and community level. On the other hand, there is promising but limited evidence that group singing may support both communication and psychosocial outcomes. However, numerous systematic reviews have consistently highlighted limitations that limit generalizability as well as the ability to formally recommend group singing as an intervention with robust psychosocial benefits [65–68]. These include heterogenous populations; small sample sizes; single site studies lacking control groups (or appropriate control groups); a lack of theoretical frameworks; a lack of large and well-designed experimental randomized controlled trials (RCTs); and a level of bias that was typically moderate to high [65–68]. Hence, to address these concerns, this SingWell Project [88] study utilizes the theoretical frameworks of the musical and social bonding hypothesis, Therapeutic Music Capacities Model, OPERA, and PRISM that will guide hypothesis testing; adopted a multisite design to recruit a sufficiently large sample size based on a statistical power analysis; will limit of risk of bias through our recruitment strategy and experimenter blinding; and will use an RCT design with planned statistical analyses.

Aims and hypotheses

The aim of this study is to explore if group singing is beneficial for older adults with unaddressed hearing loss, hypothesizing that group singing may be effective at improving speech perception and psychosocial outcomes. At the macro-level (i.e., across a 12-week period), the outcomes of interest include: 1) speech-in-noise perception, and 2) psychosocial health. At the micro-level (i.e., pre- and post-session effects at week 2, 7, and 11), the outcomes of interest are 1) psychosocial health, and 2) heart rate variability. Note that additional detail for macro and micro-level tests, and the active control group is described in the Materials and Methods section.

Macro-level hypotheses

H1. Group singing may be associated with better music perception such as pitch, rhythm, timbre, and higher-level music skills—more so than participants in the control group.

H2. Group singing may be associated with better speech perception such as speech-in-noise and emotional prosody—more so than participants in the control group.

    a. The mechanism for speech perception enhancement may be due to improvements in music perception (i.e., pitch, rhythm, timbre, or higher-level music skills.)

H3. Group singing may be associated with better psychosocial health—more so than participants in the control group.

    a. The primary outcome of interest is the enhancement of quality of life and social connectedness.

    b. The secondary outcomes of interest are the reduced symptoms of depression and anxiety, and the enhancement of self-esteem.

Micro-level hypotheses

H1. Group singing may lead to increases in positive mood—more so than the participants in the control group.

H2. Group singing may foster greater feelings of social closeness—more so than the participants in the control group.

Exploratory analyses. Finally, in an exploratory manner, we aim to explore the role of HRV and investigate the role of individual differences—for example, are the benefits of group singing specific to participants with poorer baseline outcomes? These findings may help us understand the specific populations that may benefit from group singing, to optimize future interventions.

Materials and methods

Ethics

Ethical approval to conduct this study was obtained from the Toronto Metropolitan Research Ethics Board (ID: REB 2024–103). A copy of the ethics protocol is available in S1 Appendix. All participants will be required to provide informed written consent.

Study design

A schematic of the study design can be seen in Fig 1. The study is a longitudinal RCT. Initially, participants will undergo an eligibility screening process and complete an information and consent form (described in detail under 'Testing Procedure'). Participants will be randomly assigned to either group singing or audiobook club (control group) intervention for a training period of 12-weeks and will be unaware if they are in the experimental or control group (i.e., participant-blinded). After this, the study has multiple timepoints for testing, that are broadly categorized as macro, or micro timepoints.

The rationale for the use of an audiobook club as a suitable control group is due to several factors. Firstly, listening to audiobooks is a common clinical recommendation for hearing loss and have been used in other RCTs as a control [89, 90]. Secondly, discussion from the audiobook club will generate asynchronous speech, which is a suitable compliment to investigate the targeted benefits of synchronized group singing. Finally, we anticipate that the credibility and expectation between both interventions will be similar given they are both group activities in the auditory domain.

This study was registered as a clinical trial (clinicaltrials.gov, U.S. National Library of Medicine, ID: NCT06580847); and adheres to a Standard Protocol Items: Recommendations for
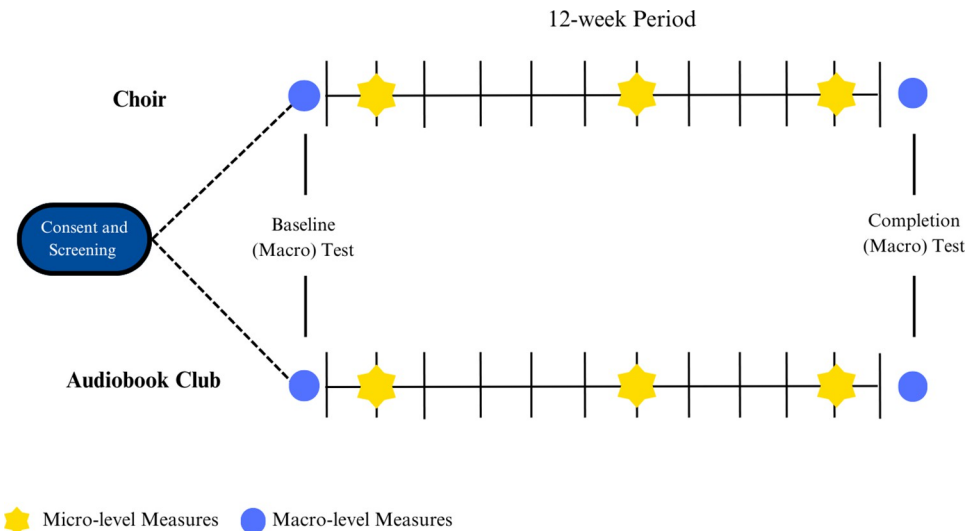
Fig 1. Overview of the study design.

https://doi.org/10.1371/journal.pone.0314473.g001

Interventional Trials (SPIRIT) Checklist with 2022 extension [91–93], which is provided in S2 Appendix. Any modifications to the protocol will require approval from the Toronto Metropolitan Research Ethics Board and will be noted in the subsequent registered report article.

Participants

Participants will be recruited from seven sites around the world: Toronto, Montreal, and St John's (Canada); Los Angeles (United States of America); Groningen (Netherlands); Oldenburg (Germany); and Adelaide (Australia). Each site will aim to recruit a total of 30 participants with a simple randomisation using a 1:1 computer generated allocation (15 randomly assigned to group singing, and 15 randomly assigned to the audiobook club), for a total of 210 participants among all the sites. Participants will not be able to change their allocation.

Recruitment will be variable as each site will have access to different databases, networks, and communities. Hence, recruitment flyers will be sent to a range of organisations such as: deaf and hard-of-hearing organisations, aged-care homes/retirement centres, senior centres, veterans organisations; and to databases such as the Toronto Metropolitan University Auditory Participant Pool and the Toronto Metro Senior Participant Pool. To avoid potential participant bias, flyers will advertise that the study is looking for participants to engage in a creative group activity, rather than group singing or an audiobook club, specifically.

Inclusion criteria. 1) Adults aged 60 years and older; 2) Bilateral mild-to-moderate hearing loss (20–49 dB hearing level, as proposed by the Global Burden of Disease Expert Group [94]), measured using four-frequency pure-tone average across both ears (4FPTA) measured at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz; 3) Unaddressed hearing loss (i.e., participants must not currently use a hearing aid, cochlear implant, or assistive listening device); 4) No significant cognitive impairment, to be assessed with the Montreal Cognitive Assessment for people with hearing impairment (MoCA-H) [95], with participants requiring a score $\geq$ 24; 5) Not use a pacemaker or anti-arrhythmic agents/medications; 6) Not currently participating in regular active music learning (e.g., choir, formal music training) or audiobook clubs within the last year; and 7) Sufficient language capacity to understand and complete the test materials. Note: all materials will be presented written and/or aurally in English at the sites located in Canada, United States of America, and Australia; Dutch at the Netherlands site; and German at the

Germany site. Test material examples corresponding to English, Dutch, and German will be provided where applicable.

Test battery

Cognitive screening. MoCA is a validated brief screening tool designed to detect mild cognitive impairment [96]. MoCA-H is an extension of the MoCA, designed specifically for adults with acquired hearing loss [95]. The cognitive domains assessed include short-term memory recall, visuospatial abilities, executive functioning, phonemic fluency, verbal abstraction, attention, concentration, and orientation.

Demographics, social networks, musical background, credibility, and expectation. Participants will be asked to provide their demographic details such as age, sex, gender, residential status, and education. We will also measure their social network diversity (e.g., connections with friend, families, neighbours, etc.) with a subscale from the Social Network Index [97]. Participant's previous musical training will be measured using the Musical Training subscale from the Goldsmith's Musical Sophistication Index (Gold-MSI) [98]. This 7-item questionnaire captures the participants' history of formal musical training with more points representing more musical training. Participant perspectives on the credibility and expectations in respect to their assigned group will also be recorded with the credibility/expectancy questionnaire [99].

Speech perception materials. The Coordinate Response Matrix (CRM) sentence recognition test will be used to evaluate speech-in-noise (SIN) perception. Our CRM test is based on the original versions that were developed in English by [100–103]. The carrier phrase is a sentence with a call sign of cat or dog and where one of six colors and one of eight numbers are mentioned (e.g., "Show the dog where the green (color) three (number) is."; "Laat de hond zien waar de groene (color) dire (number)" [Dutch]; "Zeige dem Hund, wo die grüne (color) Drei (number) ist" [German]. We use a two-down/one-down adaptive procedure based on correct color and number identification. Initial step sizes are 10 dB, which is set to 3 dB after two reversals. The test ends after an additional four reversals, generating a speech reception threshold (SRT).

EmoHI is a vocal emotion perception test [104]. The stimuli consist of pseudospeech sentences such as 'Koun se mina lod belam' and 'Nekal ibam soud molen' that are not meaningful in any Indo-European languages and are based on the Geneva Multimodal Emotion Portrayal (GEMEP) Corpus materials [105]. Stimuli were spoken by four monolingual native Dutch speakers (two male, two female) without any discernible regional accent, consisting of three emotions: happy, angry, and sad. These were selected for the widest applicability across age ranges [106], and for maximum comparability of the EmoHI test across populations of differing hearing profiles [104]. Stimuli were recorded in an anechoic room at a sampling rate of 44.1 kHz and equalized in RMS. The materials and methods of the EmoHI test are the same as in a previous study [104], but the current interface will not have a child-directed game-like environment. Participants will instead be instructed to listen to the pseudo-sentences and determine whether the voice of the speaker sounded happy, sad, or angry, by clicking on one of three corresponding buttons on the screen labeled "blij"/happy, "verdrietig"/sad, "boos"/angry. For practice, three additional stimuli (one per emotion, one production per speaker) will be used for the training session, but not included in the experimental phase. For the experiment, 36 vocal emotion stimuli are used, comprising the categories happy, sad and angry, with 12 items in each category (4 speakers x 3 utterances). Each emotion category comprises both arousal and valence dimensions: either high or low arousal, or positive or negative valence. Practice items, but not experimental items, will be presented with feedback.

Experiment items will be presented in randomized order in one block of approximately 6 to 8 minutes.

Music perception materials. A Frequency Difference Limen (FDL) task will be used to evaluate pitch perception. Participants will be presented with a three-alternative forced choice (3AFC) paradigm containing two pure tones (200 ms in duration, with 20 ms envelope rise and fall times) at 500 Hz, with the target stimulus adaptively presented at a higher frequency. The participant will identify which corresponding tone is higher on a computer keyboard (i.e., the first, second, or third tone). A pitch discrimination threshold will be calculated with an adaptive staircase procedure with the frequency difference between the target and 500 Hz tones divided by two after three correct responses, or multiplied by two after one incorrect response. After five reversals, the step size will change, with the frequency difference divided by 1.414 after 3 correct responses or multiplied by 1.414 after one incorrect response. Participants will complete two-blocks, with each block ending after 12 reversals. The calculation of their FDL threshold is determined from the mean of the last 10 reversals of each block, which were then averaged between the two blocks.

The Beat Alignment Test (BAT) will be used to measure musical beat perception and rhythm [107]. The BAT consists of 36 musical excerpts that are played with a superimposed click track in three conditions that are: 1) "On beat"—on the beat; 2) "Off beat"—off the beat with the wrong tempo; and 3) "Phase error"—out of phase. Participants will be provided with practice trials. The task is to identify if the excerpt is on the beat by responding on a keyboard with YES or NO. Participants will be instructed not to tap or move along to the music.

The Spectral-temporally Modulated Ripple Test (SMRT) is a measure of spectral resolution [108], which will be used as a proxy measure of timbre perception. Compared to typical spectral resolution tasks, the SMRT features dynamically changing ripples. Participants will be presented with a three-alternative forced choice (3-AFC) task. The target choice is presented at an initial presentation of 0.5 ripples per octave (rpo), while the other two choices are presented at a reference stimuli of 20 ripples rpo. The target stimuli are modified with a 1-up 1-down adaptive procedure with a step size of 0.2 rpo. After 10 reversals, a threshold is derived from the average of the last six reversals.

The Music-in-Noise Task (MINT) will be used to measure inter-individual differences in hearing-in-noise skills [109]. MINT is utilized as a measure of higher-level musical skills. The MINT uses a match-mismatch trial design, whereby participants first listen to a short instrumental excerpt presented in "multi-music" noise, followed by either a matched or scrambled version of the excerpt presented in silence, and are asked to judge whether the two excerpts are the same. The MINT includes five listening conditions, of which we will utilize three (Baseline, Rhythm, and Prediction [reversed order]), which differ according to the presence or absence of different types of contextual cues, which have been shown to be differentially sensitive to musical and linguistic expertise [109]. The Baseline condition provides no additional cues. In the Rhythm condition, the tones of the instrumental excerpts are all the same pitch within each trial. As such, rhythmic variation between matched and mismatched pairs represents a form of rhythmic cueing. In the Prediction condition, the excerpts are presented in silence first, followed by the noise-masked pair. This predictive cueing helps to anticipate incoming information when the excerpt is presented in noise. Each MINT condition includes 20 trials, with four familiarization trials for each condition prior to the start of the main task. Difficulty is also varied by adjusting the signal-to noise ratio (SNR) between the musical excerpt and the background noise, such that musical excerpts will be more or less hidden in the background noise. Conditions will be tested at four different signal-to-noise levels (0, −3, −6, and −9 dB).

Psychosocial questionnaires. *Macro-level questionnaires.* The 12-item short-form survey version 2 (SF-12v2) from the Medical Outcomes Study (MOS) is a widely used health-related

quality of life questionnaire that measures eight health concepts: 1) Physical functioning; 2) Role physical; 3) Bodily pain; 4) General health; 5) Vitality; 6) Social functioning; 7) Role emotional; and 8) Mental health; aggregated as a Physical Summary and a Mental Summary [110]. We will utilize the 1-week (acute) recall period. Scores range from 0 to 100, where a zero score indicates the lowest level of health and 100 indicates the highest level of health.

The Hospital Anxiety and Depression Scale (HADS) consists of 14-items designed to measure symptoms of psychological distress in medical patients [111]. HADS consists of two scales, one for anxiety (HADS-A) and depression (HADS-D). It is an efficient screening instrument that is commonly used in clinical contexts to identify and quantify symptoms related to anxiety and depression. While the original version was validated with adults aged 18 to 65 years; HADS has also been validated as an assessment for adults aged 65 to 80 years old [112]. Scoring for each item ranges from 0 to 3, with a higher score corresponding to higher levels of anxiety or depression.

The Collective Self-Esteem Scale (CSES) is a 16-item scale that evaluates individual and collective identities [113]. There are four distinct subscales (hence there is no total scale score) measuring: 1) membership self-esteem; 2) private collective self-esteem; 3) public collective self-esteem; and 4) and importance to identity, on a scale between 1 and 7, with higher scores associated with higher levels of self-esteem. The CSES is widely used with good reliability.

The revised Social Connectedness Scale (SCS-R) consists of 20-items that measures the psychological sense of belonging and interpersonal closeness in social contexts [113]. Reponses are measured on a Likert scale ranging from 1 = strongly disagree to 6 = strongly agree on items that represent both positive and negative aspects of social connectedness.

Micro-level questionnaires. The Positive and Negative Affect Schedule (PANAS) is a widely used measure of mood or emotion that consists of 20-items (10 positive and 10 negative affect items) [114]. Participants will be asked to "Indicate to what extent you feel this way right now, that is, at the present moment" on a scale of 1 = very slightly or not at all; 2 = a little; 3 = moderately; 4 = quite a bit; and 5 = extremely; in respect to affective words such as "interested" or "afraid".

The Inclusion of Other in the Self Scale (IOS) is a single item that measures how close an individual feels with another person or group [115]. The tool contains 7-pairs of circles consisting of the word "Self" in one circle, and the word "Community" in the other circle. Each pair of circles differs in proximity, from 1 = no overlap; 2 = little overlap; 3 = some overlap; 4 = equal overlap; 5 = strong overlap; 6 = very strong overlap; 7 = most overlap. Participants will select the image that best describes their relationship with their randomly assigned group.

The Subjective Units of Discomfort Scale (SUDS) consists of a 0 to 100-point scale measuring self-rated anxiety and discomfort [116]. We will utilize descriptive anchor points at 0 = "no anxiety, calm, relaxed"; 25 = "mild anxiety, alert, able to cope"; 50 = "moderate anxiety, some trouble concentrating"; 75 = "severe anxiety, thoughts of leaving"; and 100 = "very severe anxiety, worst ever experienced".

Heart rate variability. Heart Rate Variability (HRV) will be measured using photoplethysmography (PPG) sampled at 135 Hz using Polar Verity Sense heart rate sensors. Each participant in the session will receive their own sensor and all sensors will be time synchronized. The measurement period will include the entire group-singing or audiobook club session. Additionally, we will measure pre- and post-singing/audiobook rest periods, each lasting 60 seconds as a baseline measure. PPG data will be imported to Kubios HRV Scientific software (version 4.1.1) which will allow for computation of Root Mean Square of Successive Differences (RMSSD) between adjacent peaks, a time-domain measurement of HRV. Synchrony of HRV between participants over the course of a session (i.e., between the two rest periods) will be assessed using time-frequency coherence analysis [117]. A pre-post comparison of

HRV during rest periods will also be conducted for each participant to assess global changes in ANS activity over the course of a single session.

Test procedure

Below, we list the three facets to the test procedure, noting that some occur over several sessions. For consistency and logistical pragmatics, we will aim to test all participants within two-week windows, relative to key training timepoints.

1. Consent and Screening—initial session.

2. Macro Timepoint Testing—two sessions in a test booth:

    a. Baseline (up to 2-weeks before training begins).

    b. Completion (up to 2-weeks after training ends).

3. Micro Timepoint Testing—six sessions at the site of training:

    a. Pre- and Post-session testing at Week 2.

    b. Pre- and Post-session testing at Week 7.

    c. Pre- and Post-session testing at Week 11.

Consent, screening, and demographic information.   Data collectors will assess participant's eligibility. This will consist of understanding and signing a participant information and consent form; completion of an audiological assessment to establish their level of unaided hearing loss (pure tone average testing at 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz); and passing the MoCA-H screener with a criterion score ≥ 24. If these inclusion criteria are met, participants will complete their demographic information in a questionnaire.

Macro timepoint testing.   The purpose of the testing from *baseline* to *completion* is to examine any benefit from the training. Macro timepoint testing consists of the following behavioural measures presented in fixed order: CRM, EmoHI, FDL, BAT, SMRT, and MINT; and the following questionnaires: SF-12v2, HADS, CSES, SCS-R, EMO-CheQ.

The macro timepoint testing will occur on-site in an acoustically-treated test booth. The presentation level of behavioural measures will be calibrated to 65 dB with a sound-level meter measured at the participants' position, which will be located 0.5 m directly in front of the loudspeaker. Data collectors conducting the macro timepoint tests will be blinded to the participants' intervention and participants will be asked not to talk about their intervention during testing. Any instances of unintentional blinding (i.e., a participant informing the data collector of their participation) will be recorded. The participants will receive $20 CAD (or equivalent currency) for the consent and screening, and each macro timepoint testing session.

Micro timepoint testing.   The purpose of the micro timepoint testing is two-fold. Firstly, to examine the physiological and psychosocial changes that occur from *pre-* to *post-session*; and secondly, to examine the cumulative longitudinal impacts on physiological measures for each intervention. Micro timepoint testing consists of HRV and the following questionnaires presented in fixed order: PANAS, IOS, SUDS.

The micro timepoint testing will occur on-site where the training occurs at: 1) Pre-session (directly before the intervention begins); and 2) Post-session (directly after intervention ends). Testing will involve biophysical measures and questionnaires (described in detail under the Test Battery and Test Procedure sections). These micro measures will be taken at the training

sessions in Week 2, 7, and 11 and designed to be indicative of pre- and post-session measures corresponding to the start, middle, and end of the training program.

Training procedure

Participants will be randomly assigned to either the group singing or audiobook club intervention. Training will occur once-a-week over a total duration of 12-weeks, with each training session lasting approximately 1.5-hours, including informal networking and discussion. Adherence to the intervention will be reported.

Group singing will be facilitated by a choir master with at least one-year of experience leading a choir, a piano accompanist, and support from up to two research assistants who will sing with the participants. The broad goal of the group singing is to provide synchronous vocal pitch training in a supportive and casual environment. The beginning of each session will involve approximately 15-minutes of warm-up exercises, and the repertoire will be decided through discussion between the choir master and participants. Given the choir master will have some flexibility and need to adapt the sessions depending on group ability, a full outline of each lesson will be provided as supplementary materials in the final report. Participants will also be assigned an individual online 'homework' task each week—one-hour of using pitch-based games in Theta Music Trainer (https://trainer.thetamusic.com), an application designed for ear training and music theory [118]. The goal of the homework task is to maximize pitch-based perception. Theta Music Trainer provides logging capability that will allow us to monitor participants' weekly homework tasks. Email reminders will be sent to encourage participant adherence.

Participants in the audiobook club will listen to a 1-hour length audiobook segment prior to each audiobook club session. Participants will be provided with a pool of 10 audiobooks with a brief description on genre and plot. Each audiobook club will collectively select which audiobook/s they would like to listen to during their 12-week intervention. Given they are expected to listen to ~1-hour per week; the audiobook/s must total approximately 12-hours in total duration. The pool has been selected to cover a wide range of genres and interests, and are available in English, German, and Dutch. These sessions will feature a facilitator that will constructively guide the participants through a series of open-ended questions developed by the research team that has been designed to elicit open and constructive discussion. Up to two research assistants will also be present and contribute to the discussions. Email reminders will be sent to encourage participant adherence to the weekly audiobook segment.

Dissemination plan

The main findings will be published in the subsequent registered report article. Given the multisite nature of this study, individual sites may choose to develop their own specific hypotheses that extend beyond this protocol, and may choose to develop additional outputs beyond the associated registered report article. The SingWell Project is also committed to knowledge mobilization activities involving key stakeholders such as choirs, physicians, hearing health professionals, and community centres. Full details can be found on their website (https://www.singwell.ca/) and from the SingWell Project protocol [88].

Data analysis

Statistical power analysis. An a priori power analysis was conducted to determine the sample size needed to detect the smallest effect size of interest. We determined the number of participants needed to detect the cross-level interaction of group (i.e., group singing vs. audiobook club) and time (i.e., *baseline* vs. *completion*) on SIN perception. The power analysis was

conducted using traditional approaches where the resulting sample size refers to the level 2 clusters–in this case, the number of participants.

*G*Power* version 3.1.9.4 was used to conduct the analysis for an ANOVA within-between factor interaction [119]. For a conservative estimate, we assumed a small effect size of, $\eta^2 = .01$, corresponding to an effect size of f = 0.1. We estimated the correlation among the repeated measures to be $r = .62$, based on a previous group-singing controlled trial that reported an intraclass correlation (ICC) of .62 [64]. Accounting for a 20% attrition rate [64], the power analysis (f = .10, $\alpha = .05$, $r = .62$) suggests a total sample size of $N = 206$ participants, approximately $n = 30$ per site for 85% power.

In addition to the power analysis, we also calculated the reliability of our repeated measurements to assess for sufficient reliability for each participant based on [120] (1):

$$n_{min} = \frac{\rho(1 - ICC)}{(1 - \rho)(ICC)} \tag{1}$$

where *n* min is the number of time points and ρ is the population correlation across all the timepoints. Given our two timepoints and estimated ICC of .62, we have adequate-to-good reliability of our timepoints, $\rho = .77$.

Planned statistical analyses

Tests of equivalence. Given the sample size and random assignment to group, we assume that covariates such as age, gender, education, musical training/background, credibility, and expectation will be evenly distributed across the participants. Thus, we expect equivalence between participants and their randomly assigned intervention. To test this assumption, we will use a well-known technique to test for equivalence, the Bayesian *t*-test [121, 122]. Each Bayesian *t*-test will produce a coefficient called the Bayes Factor (BF), a measure of the likelihood of the alternative hypothesis (i.e., that the two samples differ). BF of 1 or less, will be interpreted as strong evidence for the null hypothesis [123]. Upon discovery of a BF larger than 1 we will conclude that the groups are not equivalent on this variable, and this variable will be used as a covariant in subsequently modeling of measures of interest.

Difference score and outlier rejection. Outliers due to technical errors will be removed. Univariate outliers on continuous independent and dependent variables will be rejected based on a boxplot method, wherein a data point is an outlier if it falls 1.5 times outside of the interquartile range. Because each group may have unique distributions, we plan to reject outliers based for each group individually. For each of the micro-level dependent measures, a difference score will be computed (i.e., *pre-* to *post-session*).

Modelling and inferential statistics. Due to the nature of this research design, there are various levels at which the data will likely show high levels of covariance (e.g., participant, group and site level), also known as clusters. Not accounting for this covariance can lead to inaccurate estimates of standard error, leading to inaccurate hypothesis testing (Mass & Cox, 2004). A well-known modeling technique that can account for this is mixed-effects linear modeling (MLM) and is also well suited to handle missing data in case of attrition or outliers (Pinheiro & Bates, 2000). MLM accounts for covariance at various levels using random effects, which allow for variance in the slope, or intercept. These random effects are defined within the model parameters based on the clustering in the research design. For our purposes, we will define a random-intercept for individual participant, to account for covariance in the repeated measures, and a random-slope for group, to account for the group level covariance.

The third level of nesting (i.e., site) is not accounted for in the model because we do not have hypotheses relating to the different sites. Modeling the third level will result in an

unnecessary overly complicated model, but ignoring the third level of clustering could result in heteroscedasticity of model residuals, underestimated standard errors and biased *p*-values. To address this, we will plot and visually assess the proposed model's residuals. If there is heteroscedasticity of the model residuals, we will apply a robust standard error estimator to account for the clustering [124]. All modeling will be done in the *lme4* package [125] in R.

For each micro-level dependent variable, three separate MLMs will be defined: 1) an intercept-only model; 2) a main-effect only model (DV_{change} ~ *group* + *time*) and 3) an interaction (DV_{change} ~ *group* * *time*) model. For micro-level models, the time variable will reflect the time point of the micro-level observation and will allow us to model changes between the group across time. Similarly, three separate MLMs will be defined for macro-level dependent variables: 1) an intercept-only model; 2) a main-effect only model (DV ~ *group* + *session*) and 3) an interaction (DV ~ *group* * *session*) model. For both macro and micro-level models we will use a hierarchical modeling technique to determine which model is the best fit to the data. Model fit parameters will be compared between the three models and results of the model with the best fit will be reported on. Upon the discovery of a significant interaction, a follow-up simple slope analysis will be used to examine the differences in time (e.g., *baseline* vs. *completion*) across groups. For all hypothesis tests, alpha values will be set at 0.05.

Mechanisms and contributions to speech perception. If a statistically significant benefit from group singing is found for speech perception, we will examine if changes in musical perception (i.e., FDL, BAT, and SMRT) are potential mechanisms or contributors for this change in speech perception (i.e., CRM and EmoHI). To do this, a hierarchical method will be used to compare a model which predicts speech perception difference score from BAT, FDL, and SMRT difference scores to a model which also contains an interaction term between speech perception variables and the group variable. The model which is a better fit for the SIN change scores will be reported on, and any significant interactions will be decomposed using simple slope analysis.

Supporting information

S1 Appendix. Ethics protocol.
(PDF)

S2 Appendix. SPIRIT checklist.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Chi Yhun Lo, Benjamin Rich Zendel, Arla Good, Frank A. Russo.

Formal analysis: Carmen Dang, Sean Gilmore.

Funding acquisition: Frank A. Russo.

Methodology: Chi Yhun Lo, Benjamin Rich Zendel, Deniz Baskent, Emily Coffey, Assal Habibi, Ellie Harding, Gunter Kreutz, Mridula Sharma, Carmen Dang, Sean Gilmore, Helen Henshaw, Colette M. McKay, Arla Good, Frank A. Russo.

Project administration: Chi Yhun Lo.

Supervision: Chi Yhun Lo, Frank A. Russo.

Writing – original draft: Chi Yhun Lo.

Writing – review & editing: Benjamin Rich Zendel, Deniz Baskent, Christian Boyle, Emily Coffey, Nathan Gagne, Assal Habibi, Ellie Harding, Merel Keijzer, Gunter Kreutz, Bert Maat, Eva Schurig, Mridula Sharma, Carmen Dang, Sean Gilmore, Helen Henshaw, Colette M. McKay, Arla Good, Frank A. Russo.

References

1. Haile LM, Kamenov K, Briant PS, Orji AU, Steinmetz JD, Abdoli A, et al. Hearing loss prevalence and years lived with disability, 1990–2019: Findings from the Global Burden of Disease Study 2019. The Lancet. 2021; 397: 996–1009. https://doi.org/10.1016/S0140-6736(21)00516-X PMID: 33714390

2. Huddle MG, Goman AM, Kernizan FC, Foley DM, Price C, Frick KD, et al. The economic impact of adult hearing loss: A systematic review. JAMA Otolaryngology—Head and Neck Surgery. American Medical Association; 2017. pp. 1040–1048. https://doi.org/10.1001/jamaoto.2017.1243 PMID: 28796850

3. McDaid D, Park A La, Chadha S. Estimating the global costs of hearing loss. Int J Audiol. 2021; 60: 162–170. https://doi.org/10.1080/14992027.2021.1883197 PMID: 33590787

4. World Health Organization. World report on hearing. World Health Organization. Geneva; 2021.

5. Bowl MR, Dawson SJ. Age-related hearing loss. Cold Spring Harb Perspect Med. 2019;9. https://doi.org/10.1101/cshperspect.a033217 PMID: 30291149

6. Robert Frisina D, Frisina RD. Speech recognition In noise and presbycusis: relations to possible neural mechanisms. Hear Res. ELSEVIER; 1997.

7. Lin FR, Thorpe R, Gordon-Salant S, Ferrucci L. Hearing loss prevalence and risk factors among older adults in the United States. Journals of Gerontology—Series A Biological Sciences and Medical Sciences. 2011; 66 A: 582–590. https://doi.org/10.1093/gerona/glr002 PMID: 21357188

8. Löhler J, Walther LE, Hansen F, Kapp P, Meerpohl J, Wollenberg B, et al. The prevalence of hearing loss and use of hearing aids among adults in Germany: a systematic review. European Archives of Oto-Rhino-Laryngology. Springer Verlag; 2019. pp. 945–956. https://doi.org/10.1007/s00405-019-05312-z PMID: 30737583

9. Moore DR, Edmondson-Jones M, Dawes P, Fortnum H, McCormack A, Pierzycki RH, et al. Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. PLoS One. 2014;9. https://doi.org/10.1371/journal.pone.0107720 PMID: 25229622

10. Vermeire K, Knoop A, Boel C, Auwers S, Schenus L, Talaveron-Rodriguez M, et al. Speech recognition in noise by younger and older adults: Effects of age, hearing loss, and temporal resolution. Annals of Otology, Rhinology and Laryngology. 2016; 125: 297–302. https://doi.org/10.1177/0003489415611424 PMID: 26466858

11. Plomp R, Mimpen AM. Speech-reception threshold for sentences as a function of age and noise level. Journal of the Acoustical Society of America. 1979; 66: 1333–1342. https://doi.org/10.1121/1.383554 PMID: 500971

12. Wu YH, Stangl E, Chipara O, Hasan SS, Welhaven A, Oleson J. Characteristics of real-world signal to noise ratios and speech listening situations of older adults with mild to moderate hearing loss. Ear and Hearing. Lippincott Williams and Wilkins; 2018. pp. 293–304. https://doi.org/10.1097/AUD.0000000000000486 PMID: 29466265

13. Roberts SD, Delich NA. From Isolation to Communication: Connecting Adults Who Have Hearing Loss With Their Communication Partners. JADARA. 2020; 53: 88–106. Available: https://repository.wcsu.edu/jadara/vol53/iss2/5

14. Barker AB, Leighton P, Ferguson MA. Coping together with hearing loss: a qualitative meta-synthesis of the psychosocial experiences of people with hearing loss and their communication partners. International Journal of Audiology. Taylor and Francis Ltd; 2017. pp. 297–305. https://doi.org/10.1080/14992027.2017.1286695 PMID: 28599604

15. US Department of Health and Human Services. Our Epidemic of Loneliness and Isolation: The U.S. Surgeon General's Advisory on the Healing Effects of Social Connection and Community. 2023.

16. Shukla A, Harper M, Pedersen E, Goman A, Suen JJ, Price C, et al. Hearing Loss, Loneliness, and Social Isolation: A Systematic Review. Otolaryngology–Head and Neck Surgery. 2020; 162: 622–633. https://doi.org/10.1177/0194599820910377 PMID: 32151193

17. Jayakody DMP, Wishart J, Stegeman I, Eikelboom R, Moyle TC, Yiannos JM, et al. Is There an Association Between Untreated Hearing Loss and Psychosocial Outcomes? Front Aging Neurosci. 2022;14. https://doi.org/10.3389/fnagi.2022.868673 PMID: 35663574

18. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and Social Isolation as Risk Factors for Mortality: A Meta-Analytic Review. Perspectives on Psychological Science. 2015; 10: 227–237. https://doi.org/10.1177/1745691614568352 PMID: 25910392

19. Pichora-Fuller MK, Mick P, Reed M. Hearing, Cognition, and Healthy Aging: Social and Public Health Implications of the Links between Age-Related Declines in Hearing and Cognition. Semin Hear. 2015; 36: 122–139. https://doi.org/10.1055/s-0035-1555116 PMID: 27516713

20. Taylor HO, Taylor RJ, Nguyen AW, Chatters L. Social Isolation, Depression, and Psychological Distress Among Older Adults. J Aging Health. 2018; 30: 229–246. https://doi.org/10.1177/0898264316673511 PMID: 28553785

21. Lara E, Caballero FF, Rico-Uribe LA, Olaya B, Haro JM, Ayuso-Mateos JL, et al. Are loneliness and social isolation associated with cognitive decline? Int J Geriatr Psychiatry. 2019; 34: 1613–1622. https://doi.org/10.1002/gps.5174 PMID: 31304639

22. Ferguson MA, Kitterick PT, Chong LY, Edmondson-Jones M, Barker F, Hoare DJ. Hearing aids for mild to moderate hearing loss in adults. Cochrane Database of Systematic Reviews. John Wiley and Sons Ltd; 2017. https://doi.org/10.1002/14651858.CD012023.pub2 PMID: 28944461

23. Orji A, Kamenov K, Dirac M, Davis A, Chadha S, Vos T. Global and regional needs, unmet needs and access to hearing aids. Int J Audiol. 2020; 59: 166–172. https://doi.org/10.1080/14992027.2020.1721577 PMID: 32011190

24. World Health Organization. Global costs of unaddressed hearing loss and cost-effectiveness of interventions: A WHO Report, 2017. 2017.

25. Ramage-Morin PL, Banks R, Pineault D, Atrach M. Unperceived hearing loss among Canadians aged 40 to 79. Health Rep. 2019; 30: 11–20. https://doi.org/10.25318/82-003-x201900800002-eng PMID: 31454408

26. Arlinger S. Negative consequences of uncorrected hearing loss—a review. Int J Audiol. 2003; 42: S17–S20. https://doi.org/10.3109/14992020309074639 PMID: 12918624

27. Fancourt D, Finn S. What is the evidence on the role of the arts in improving health and well-being? A scoping review. Copenhagen; 2019.

28. Shilton D, Passmore S, Savage PE. Group singing is globally dominant and associated with social context. R Soc Open Sci. 2023;10. https://doi.org/10.1098/rsos.230562 PMID: 37680502

29. Hill K. 3.5 million choral singers in Canada. Hamilton; 2017 Jul. Available: http://www.cbc.ca/news2/interactives/sports-junior/,

30. Johnson JK, Stewart AL, Acree M, Nápoles AM, Flatt JD, Max WB, et al. A Community Choir Intervention to Promote Well-Being among Diverse Older Adults: Results from the Community of Voices Trial. Journals of Gerontology—Series B Psychological Sciences and Social Sciences. 2020; 75: 549–559. https://doi.org/10.1093/geronb/gby132 PMID: 30412233

31. Patel AD. Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. Front Psychol. 2011; 2: 142. https://doi.org/10.3389/fpsyg.2011.00142 PMID: 21747773

32. Patel AD. The OPERA hypothesis: assumptions and clarifications. Ann N Y Acad Sci. 2012; 1252: 124–8. https://doi.org/10.1111/j.1749-6632.2011.06426.x PMID: 22524349

33. Patel AD. Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. Hear Res. 2014; 308: 98–108. https://doi.org/10.1016/j.heares.2013.08.011 PMID: 24055761

34. Fiveash A, Bedoin N, Gordon RL, Tillmann B. Processing Rhythm In Speech And Music: Shared Mechanisms And Implications For Developmental Speech And Language Disorders. Neuropsychology. 2021; 35: 771–791. https://doi.org/10.1037/neu0000766 PMID: 34435803

35. Hausen M, Torppa R, Salmela VR, Vainio M, Särkämö T. Music and speech prosody: A common rhythm. Front Psychol. 2013; 4: 1–16. https://doi.org/10.3389/fpsyg.2013.00566 PMID: 24032022

36. Peretz I, Vuvan D, Lagrois M-E, Armony JL. Neural overlap in processing music and speech. Philosophical Transactions of the Royal Society B: Biological Sciences. 2015; 370: 20140090–20140090. https://doi.org/10.1098/rstb.2014.0090 PMID: 25646513

37. Chatterjee M, Peng SS-CS. Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. Hear Res. 2008; 235: 143–56. https://doi.org/10.1016/j.heares.2007.11.004 PMID: 18093766

38. Holt CM, McDermott HJ. Discrimination of intonation contours by adolescents with cochlear implants. Int J Audiol. 2013; 52: 808–815. https://doi.org/10.3109/14992027.2013.832416 PMID: 24053225

39. Kraus N, Chandrasekaran B. Music training for the development of auditory skills. Nat Rev Neurosci. 2010; 11: 599–605. https://doi.org/10.1038/nrn2882 PMID: 20648064

40. Lo CY, McMahon CM, Looi V, Thompson WF. Melodic Contour Training and Its Effect on Speech in Noise, Consonant Discrimination, and Prosody Perception for Cochlear Implant Recipients. Behavioral Neurology [Special Issue]. 2015. https://doi.org/10.1155/2015/352869 PMID: 26494944

41. Large EW, Jones MR. The Dynamics of Attending: How People Track Time-Varying Events. Psychol Rev. 1999; 106: 119–159.

42. Jones MR. Time, Our Lost Dimension: Toward a New Theory of Perception, Attention, and Memory. 1976.

43. Poeppel D, Assaneo MF. Speech rhythms and their neural foundations. Nature Reviews Neuroscience. Nature Research; 2020. pp. 322–334. https://doi.org/10.1038/s41583-020-0304-4 PMID: 32376899

44. Hickok G, Houde J, Rong F. Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. Neuron. 2011. pp. 407–422. https://doi.org/10.1016/j.neuron.2011.01.019 PMID: 21315253

45. Ross JM, Balasubramaniam R. Time Perception for Musical Rhythms: Sensorimotor Perspectives on Entrainment, Simulation, and Prediction. Frontiers in Integrative Neuroscience. Frontiers Media S.A.; 2022. https://doi.org/10.3389/fnint.2022.916220 PMID: 35865808

46. Sharda M, Tuerk C, Chowdhury R, Jamey K, Foster N, Custo-Blanch M, et al. Music improves social communication and auditory–motor connectivity in children with autism. Transl Psychiatry. 2018;8. https://doi.org/10.1038/s41398-018-0287-3 PMID: 30352997

47. Li Q, Wang X, Wang S, Xie Y, Li X, Xie Y, et al. Musical training induces functional and structural auditory-motor network plasticity in young adults. Hum Brain Mapp. 2018; 39: 2098–2110. https://doi.org/10.1002/hbm.23989 PMID: 29400420

48. Loui P. A dual-stream neuroanatomy of singing. Music Percept. 2015; 32: 232–241. https://doi.org/10.1525/mp.2015.32.3.232 PMID: 26120242

49. Russo FA. Motor System Involvement in the Perception of Singing. The Routledge Companion to Interdisciplinary Studies in Singing. Routledge; 2020. pp. 276–288. https://doi.org/10.4324/9781315163734-21

50. Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. Nature Neuroscience. 2009. pp. 718–724. https://doi.org/10.1038/nn.2331 PMID: 19471271

51. Peelle JE. Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. Ear Hear. 2018; 39: 204–214. https://doi.org/10.1097/AUD.0000000000000494 PMID: 28938250

52. Zendel BR, West GL, Belleville S, Peretz I. Musical training improves the ability to understand speech-in-noise in older adults. Neurobiol Aging. 2019; 81: 102–115. https://doi.org/10.1016/j.neurobiolaging.2019.05.015 PMID: 31280114

53. Fleming D, Belleville S, Peretz I, West G, Zendel BR. The effects of short-term musical training on the neural processing of speech-in-noise in older adults. Brain Cogn. 2019;136. https://doi.org/10.1016/j.bandc.2019.103592 PMID: 31404817

54. Parbery-Clark A, Skoe E, Lam C, Kraus N. Musician enhancement for speech-in-noise. Ear Hear. 2009; 30: 653–61. https://doi.org/10.1097/AUD.0b013e3181b412e9 PMID: 19734788

55. Coffey EBJ, Mogilever NB, Zatorre RJ. Speech-in-noise perception in musicians: A review. Hearing Research. Elsevier B.V.; 2017. pp. 49–69. https://doi.org/10.1016/j.heares.2017.02.006 PMID: 28213134

56. Maillard E, Joyal M, Murray MM, Tremblay P. Are musical activities associated with enhanced speech perception in noise in adults? A systematic review and meta-analysis. Current Research in Neurobiology. 2023; 100083. https://doi.org/10.1016/j.crneur.2023.100083 PMID: 37397808

57. Hennessy S, Mack WJ, Habibi A. Speech-in-noise perception in musicians and non-musicians: A multi-level meta-analysis. Hear Res. 2022;416. https://doi.org/10.1016/j.heares.2022.108442 PMID: 35078132

58. Schellenberg EG, Lima CF. Annual Review of Psychology Music Training and Nonmusical Abilities. 2024. https://doi.org/10.1146/annurev-psych-032323

59. McKay CM. No Evidence That Music Training Benefits Speech Perception in Hearing-Impaired Listeners: A Systematic Review. Trends Hear. 2021;25. https://doi.org/10.1177/2331216520985678 PMID: 33634750

60. Good A, Gordon KA, Papsin BC, Nespoli G, Hopyan T, Peretz I, et al. Benefits of Music Training for Perception of Emotional Speech Prosody in Deaf Children With Cochlear Implants. Ear Hear. 2017; 38: 455–464. https://doi.org/10.1097/AUD.0000000000000402 PMID: 28085739

61. Fuller CD, Galvin JJ, Maat B, Başkent D, Free RH. Comparison of Two Music Training Approaches on Music and Speech Perception in Cochlear Implant Users. Trends Hear. 2018; 22: 233121651876537. https://doi.org/10.1177/2331216518765379 PMID: 29621947

62. Bedoin N, Besombes AM, Escande E, Dumont A, Lalitte P, Tillmann B. Boosting syntax training with temporally regular musical primes in children with cochlear implants. Ann Phys Rehabil Med. 2018; 61: 365–371. https://doi.org/10.1016/j.rehab.2017.03.004 PMID: 28506442

63. Chari DA, Barrett KC, Patel AD, Colgrove TR, Jiradejvong P, Jacobs LY, et al. Impact of Auditory-Motor Musical Training on Melodic Pattern Recognition in Cochlear Implant Users. Otology and Neurotology. 2020; 41: e422–e431. https://doi.org/10.1097/MAO.0000000000002525 PMID: 32176126

64. Dubinsky E, Wood EA, Nespoli G, Russo FA. Short-Term Choir Singing Supports Speech-in-Noise Perception and Neural Pitch Strength in Older Adults With Age-Related Hearing Loss. Front Neurosci. 2019;13. https://doi.org/10.3389/fnins.2019.01153 PMID: 31849572

65. Clift S, Nicol J, Raisbeck M, Whitmore C, Morrison I. Group singing, wellbeing and health: A systematic mapping of research evidence. UNESCO Observatory. 2010;2.

66. Irons JY, Sheffield D, Ballington F, Stewart DE. A systematic review on the effects of group singing on persistent pain in people with long-term health conditions. European Journal of Pain (United Kingdom). Blackwell Publishing Ltd; 2020. pp. 71–90. https://doi.org/10.1002/ejp.1485 PMID: 31549451

67. Williams E, Dingle GA, Clift S. A systematic review of mental health and wellbeing outcomes of group singing for adults with a mental health condition. Eur J Public Health. 2018; 28: 1035–1042. https://doi.org/10.1093/eurpub/cky115 PMID: 29982515

68. Clark I, Harding K. Psychosocial outcomes of active singing interventions for therapeutic purposes: A systematic review of the literature. Nord J Music Ther. 2012; 21: 80–98. https://doi.org/10.1080/08098131.2010.545136

69. Müller V, Lindenberger U. Cardiac and respiratory patterns synchronize between persons during choir singing. PLoS One. 2011;6. https://doi.org/10.1371/journal.pone.0024893 PMID: 21957466

70. Hemakom A, Powezka K, Goverdovsky V, Jaffer U, Mandic DP. Quantifying team cooperation through intrinsic multi-scale measures: Respiratory and cardiac synchronization in choir singers and surgical teams. R Soc Open Sci. 2017;4. https://doi.org/10.1098/rsos.170853 PMID: 29308229

71. Savage PE, Loui P, Tarr B, Schachner A, Glowacki L, Mithen S, et al. Music as a coevolved system for social bonding. Behavioral and Brain Sciences. 2021;44. https://doi.org/10.1017/S0140525X20000333 PMID: 32814608

72. Brancatisano O, Baird A, Thompson WF. Why is music therapeutic for neurological disorders? The Therapeutic Music Capacities Model. Neurosci Biobehav Rev. 2020; 112: 600–615. https://doi.org/10.1016/j.neubiorev.2020.02.008 PMID: 32050086

73. Merrett DL, Peretz I, Wilson SJ. Moderating variables of music training-induced neuroplasticity: A review and discussion. Frontiers in Psychology. 2013. https://doi.org/10.3389/fpsyg.2013.00606 PMID: 24058353

74. Olszewska AM, Gaca M, Herman AM, Jednoróg K, Marchewka A. How Musical Training Shapes the Adult Brain: Predispositions and Neuroplasticity. Frontiers in Neuroscience. Frontiers Media S.A.; 2021. https://doi.org/10.3389/fnins.2021.630829 PMID: 33776638

75. Wolff L, Quan Y, Perry G, Forde Thompson W. Music Engagement as a Source of Cognitive Reserve. Am J Alzheimers Dis Other Demen. 2023;38. https://doi.org/10.1177/15333175231214833 PMID: 37993973

76. Alain C, Zendel BR, Hutka S, Bidelman GM. Turning down the noise: The benefit of musical training on the aging auditory brain. Hearing Research. Elsevier; 2014. pp. 162–173. https://doi.org/10.1016/j.heares.2013.06.008 PMID: 23831039

77. Zendel BR, Alain C. Musicians experience less age-related decline in central auditory processing. Psychol Aging. 2012; 27: 410–417. https://doi.org/10.1037/a0024816 PMID: 21910546

78. Blood AJ, Zatorre RJ. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. Proc Natl Acad Sci U S A. 2001; 98: 11818–11823. https://doi.org/10.1073/pnas.191355898 PMID: 11573015

79. Quan Y, Lo CY, Olsen KN, Thompson WF. The effectiveness of aerobic exercise and dance interventions on cognitive function in adults with mild cognitive impairment: an overview of meta-analyses. Int Rev Sport Exerc Psychol. 2024. https://doi.org/10.1080/1750984X.2024.2332989

80. Daskalopoulou C, Stubbs B, Kralj C, Koukounari A, Prince M, Prina AM. Physical activity and healthy ageing: A systematic review and meta-analysis of longitudinal cohort studies. Ageing Research

Reviews. Elsevier Ireland Ltd; 2017. pp. 6–17. https://doi.org/10.1016/j.arr.2017.06.003 PMID: 28648951

81. Good A, Choma B, Russo FA. Movement Synchrony Influences Intergroup Relations in a Minimal Groups Paradigm. Basic Appl Soc Psych. 2017; 39: 231–238. https://doi.org/10.1080/01973533.2017.1337015

82. Wiltermuth SS, Heath C. Synchrony and Cooperation. Psychol Sci. 2009; 20: 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x PMID: 19152536

83. Overy K. Making music in a group: Synchronization and shared experience. Ann N Y Acad Sci. 2012; 1252: 65–68. https://doi.org/10.1111/j.1749-6632.2012.06530.x PMID: 22524341

84. Schiavio A, Stupacher J, Parncutt R, Timmers R. Learning music from each other: Synchronization, turn-taking, or imitation? Music Percept. 2020; 37: 403–422. https://doi.org/10.1525/MP.2020.37.5.403

85. Hindhede AL. Negotiating hearing disability and hearing disabled identities. Health N Hav. 2012; 16: 169–185. https://doi.org/10.1177/1363459311403946 PMID: 21540252

86. Wallhagen MI. The stigma of hearing loss. Gerontologist. 2010; 50: 66–75. https://doi.org/10.1093/geront/gnp107 PMID: 19592638

87. Pearce E, Launay J, Maccarron P, Dunbar RIM. Tuning in to others: Exploring relational and collective bonding in singing and non-singing groups over time. Psychol Music. 2017; 45: 496–512. https://doi.org/10.1177/0305735616667543

88. The SingWell project protocol: the road to understanding the benefits of group singing in older adults. Public Health Panorama.

89. Humes LE, Skinner KG, Kinney DL, Rogers SE, Main AK, Quigley TM. Clinical Effectiveness of an At-Home Auditory Training Program: A Randomized Controlled Trial. Ear Hear. 2019; 40: 1043–1060. https://doi.org/10.1097/AUD.0000000000000688 PMID: 30575602

90. Rao A, Rishiq D, Yu L, Zhang Y, Abrams H. Neural Correlates of Selective Attention With Hearing Aid Use Followed by ReadMyQuips Auditory Training Program. Ear Hear. 2017; 38: 28–41. https://doi.org/10.1097/AUD.0000000000000348 PMID: 27556531

91. Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ. 2013;346. https://doi.org/10.1136/bmj.e7586 PMID: 23303884

92. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jeric K, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials DEVELOPMENT OF THE SPIRIT 2013 STATEMENT. Ann Intern Med. 2013. Available: www.annals.org

93. Butcher NJ, Monsour A, Mew EJ, Chan AW, Moher D, Mayo-Wilson E, et al. Guidelines for Reporting Outcomes in Trial Protocols: The SPIRIT-Outcomes 2022 Extension. JAMA. American Medical Association; 2022. pp. 2345–2356. https://doi.org/10.1001/jama.2022.21243 PMID: 36512367

94. Stevens G, Flaxman S, Brunskill E, Mascarenhas M, Mathers CD, Finucane M. Global and regional hearing impairment prevalence: An analysis of 42 studies in 29 countries. Eur J Public Health. 2013; 23: 146–152. https://doi.org/10.1093/eurpub/ckr176 PMID: 22197756

95. Dawes P, Reeves D, Yeung WK, Holland F, Charalambous AP, Côté M, et al. Development and validation of the Montreal cognitive assessment for people with hearing impairment (MoCA-H). J Am Geriatr Soc. 2023; 71: 1485–1494. https://doi.org/10.1111/jgs.18241 PMID: 36722180

96. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. 2005.

97. Cohen S. Social Supports and Physical Health: Symptoms, Health Behaviors, and Infectious Disease. In: Cummins EM, Greene AL, Karraker KH, editors. Life-span Developmental Psychology. New York: Psychology Press; 1991.

98. Müllensiefen D, Gingras B, Musil J, Stewart L. The musicality of non-musicians: An index for assessing musical sophistication in the general population. PLoS ONE. Public Library of Science; 2014. https://doi.org/10.1371/journal.pone.0089642 PMID: 24586929

99. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/ expectancy questionnaire. J Behav Ther Exp Psychiatry. 2000. https://doi.org/10.1016/s0005-7916(00)00012-4 PMID: 11132119

100. Moore T. Voice communication jamming research. Advisory Group for Aerospace Research and Development Conference Proceedings. 1981.

101. Brungart DS. Evaluation of speech intelligibility with the coordinate response measure. J Acoust Soc Am. 2001; 109: 2276–2279. https://doi.org/10.1121/1.1357812 PMID: 11386582

102. Hazan V, Messaoud-Galusi S, Rosen S, Nouwens S, Shakespeare B. Speech Perception Abilities of Adults With Dyslexia: Is There Any Evidence for a True Deficit? Journal of Speech, Language, and

Hearing Research. 2009; 52: 1510–1529. https://doi.org/10.1044/1092-4388(2009/08-0220) PMID: 19635940

103. Welch GF, Saunders J, Edwards S, Palmer Z, Knight J, Mahon M, et al. Using singing to nurture children's hearing? A pilot study. 2015;16.

104. Nagels L, Gaudrain E, Vickers D, Matos Lopes M, Hendriks P, Başkent D. Development of vocal emotion recognition in school-age children: The EmoHI test for hearing-impaired populations. PeerJ. 2020;2020. https://doi.org/10.7717/peerj.8773 PMID: 32274264

105. Bänziger T, Mortillaro M, Scherer KR. Introducing the Geneva Multimodal expression corpus for experimental research on emotion perception. Emotion. 2012; 12: 1161–1179. https://doi.org/10.1037/a0025827 PMID: 22081890

106. Widen SC, Russell JA. The relative power of an emotion's facial expression, label, and behavioral consequence to evoke preschoolers' knowledge of its cause. Cogn Dev. 2004; 19: 111–125. https://doi.org/10.1016/j.cogdev.2003.11.004

107. Iversen JR, Patel AD. The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population. Department of Psychology, Hokkaido University; 2008.

108. Aronoff JM, Landsberger DM. The development of a modified spectral ripple test. J Acoust Soc Am. 2013; 134: EL217–EL222. https://doi.org/10.1121/1.4813802 PMID: 23927228

109. Coffey EBJ, Arseneau-Bruneau I, Zhang X, Zatorre RJ. The Music-In-Noise Task (MINT): A Tool for Dissecting Complex Auditory Perception. Front Neurosci. 2019;13. https://doi.org/10.3389/fnins.2019.00199 PMID: 30930734

110. Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. Med Care. 1996; 34: 220–233. Available: https://doi.org/10.1097/00005650-199603000-00003 PMID: 8628042

111. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand. 1983; 67: 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x PMID: 6880820

112. Djukanovic I, Carlsson J, Årestedt K. Is the Hospital Anxiety and Depression Scale (HADS) a valid measure in a general population 65–80 years old? A psychometric evaluation study. Health Qual Life Outcomes. 2017;15. https://doi.org/10.1186/s12955-017-0759-9 PMID: 28978356

113. Luhtanen R, Crocker J. A Collective Self-Esteem Scale: Self-Evaluation of One's Social Identity. Pers Soc Psychol Bull. 1992; 18: 302–318. https://doi.org/10.1177/0146167292183006

114. Watson D, Clark LA, Tellegen A. Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. J Pers Soc Psychol. 1988. https://doi.org/10.1037//0022-3514.54.6.1063 PMID: 3397865

115. Aron A, Aron EN, Smollan D. Inclusion of Other in the Self Scale and the structure of interpersonal closeness. J Pers Soc Psychol. 1992; 63: 596–612. https://doi.org/10.1037/0022-3514.63.4.596

116. Wolpe J. The Practice of Behavior Therapy. Michigan: Pergamon Press; 1969.

117. Orini M, Bailon R, Mainardi LT, Laguna P, Flandrin P. Characterization of dynamic interactions between cardiovascular signals by time-frequency coherence. IEEE Trans Biomed Eng. 2012; 59: 663–673. https://doi.org/10.1109/TBME.2011.2171959 PMID: 22155936

118. Theta Music Technologies Inc. Theta Music Trainer.

119. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007; 39: 175–191. https://doi.org/10.3758/bf03193146 PMID: 17695343

120. Snijders TA, Bosker RJ. Multilevel analysis: An introduction to basic and advanced multilevel modeling. London: Sage Publications; 2011.

121. Kruschke JK. Bayesian estimation supersedes the t test. J Exp Psychol Gen. 2013; 142: 573–603. https://doi.org/10.1037/a0029146 PMID: 22774788

122. Lakens D, McLatchie N, Isager PM, Scheel AM, Dienes Z. Improving Inferences about Null Effects with Bayes Factors and Equivalence Tests. Journals of Gerontology—Series B Psychological Sciences and Social Sciences. 2020; 75: 45–57. https://doi.org/10.1093/geronb/gby065 PMID: 29878211

123. Stefan AM, Gronau QF, Schönbrodt FD, Wagenmakers EJ. A tutorial on Bayes Factor Design Analysis using an informed prior. Behav Res Methods. 2019; 51: 1042–1058. https://doi.org/10.3758/s13428-018-01189-8 PMID: 30719688

124. McNeish D, Stapleton LM, Silverman RD. On the unnecessary ubiquity of hierarchical linear modeling. Psychol Methods. 2017; 22: 114–140. https://doi.org/10.1037/met0000078 PMID: 27149401

125. Bates DM. lme4: Mixed-effects modeling with R Springer. 2010.