Journal of Experimental Psychology: Human Perception and Performance

Associations Between Musical Expertise and Auditory Processing

Aíssa M. Baldé, César F. Lima, and E. Glenn Schellenberg

Online First Publication, March 3, 2025. https://dx.doi.org/10.1037/xhp0001312

CITATION

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://dx.doi.org/10.1037/xhp0001312

Journal of Experimental Psychology:
Human Perception and Performance

ISSN: 0096-1523

https://doi.org/10.1037/xhp0001312

Associations Between Musical Expertise and Auditory Processing

Aíssa M. Baldé[1], César F. Lima[1], and E. Glenn Schellenberg[1, 2]
[1] Centro de Investigação e Intervenção Social (CIS-IUL), Instituto Universitário de Lisboa (ISCTE-IUL)
[2] Department of Psychology, University of Toronto Mississauga

Many studies have linked musical expertise with nonmusical abilities such as speech perception, memory, or executive functions. Far fewer have examined associations with basic auditory skills. Here, we asked whether psychoacoustic thresholds predict four aspects of musical expertise: music training, melody perception, rhythm perception, and self-reported musical abilities and behaviors (other than training). A total of 138 participants completed nine psychoacoustic tasks, as well as the Musical Ear Test (melody and rhythm subtests) and the Goldsmiths Musical Sophistication Index. We also measured and controlled for demographics, general cognitive abilities, and personality traits. The psychoacoustic tasks assessed discrimination thresholds for pitch and temporal perception (both assessed with three tasks), and for timbre, intensity, and backward masking (each assessed with one task). Both music training and melody perception predicted better performance on the pitch-discrimination tasks. Rhythm perception was associated with better performance on several temporal and nontemporal tasks, although none had unique associations when the others were held constant. Self-reported musical abilities and behaviors were associated with performance on one of the temporal tasks: duration discrimination. The findings indicate that basic auditory skills correlate with individual differences in musical expertise, whether expertise is defined as music training or musical ability.

Public Significance Statement

The human capacity to perceive, play, and remember music requires basic auditory skills, such as the ability to discern differences in pitch, duration, timbre, or loudness. In the present study, we found that individual differences in auditory skills correlate with participants' level of music training, and with their performance on objective tests of music perception. Our findings help to explain the bases of musicality.

Keywords: psychoacoustics, music, ability, expertise, training

Supplemental materials: https://doi.org/10.1037/xhp0001312.supp

Auditory skills are crucial for music perception and musical expertise in general. Basic auditory processing provides the foundation on which acoustic and cognitive analyses operate. To perceive and remember music, a listener must first decode its basic acoustic properties. This initial acoustic analysis involves general auditory mechanisms that also process speech or environmental sounds (Kraus & Banai, 2007; Mueller et al., 2012). Low-level auditory features (e.g., intensity, pitch, timbre) are extracted before acoustic signals are allocated to their sound categories and grouped into increasingly complex structures that are

---

Joel S. Snyder served as action editor.

E. Glenn Schellenberg https://orcid.org/0000-0003-3681-6020

This research was funded in whole, or in part, by the Foundation for Science and Technology of Portugal (FCT) (Grant SFRH/BD/2022.12352; CEECIND/03266/2018). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

This work was funded by the Portuguese Foundation for Science and Technology through a PhD studentship awarded to Aíssa M. Baldé (SFRH/BD/2022.12352) and a Scientific Employment Stimulus grant awarded to E. Glenn Schellenberg (CEECIND/03266/2018). César F. Lima and E. Glenn Schellenberg contributed equally to this work. The authors declare that they have no conflicts of interest.

Aíssa M. Baldé served as lead for project administration and contributed equally to conceptualization. César F. Lima served as lead for project administration, resources, and software, contributed equally to supervision, and served in a supporting role for data curation, formal analysis, funding acquisition, investigation, validation, visualization, writing–original draft, and writing–review and editing. E. Glenn Schellenberg served as lead for supervision, contributed equally to investigation and project administration, and served in a supporting role for funding acquisition, resources, software, validation, and visualization. Aíssa M. Baldé and E. Glenn Schellenberg contributed equally to data curation, writing–original draft, writing–review and editing, and formal analysis. César F. Lima and E. Glenn Schellenberg contributed equally to methodology.

Correspondence concerning this article should be addressed to E. Glenn Schellenberg, Centro de Investigação e Intervenção Social (CIS-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Av.ª das Forças Armadas, 1649-026 Lisboa, Portugal, or César F. Lima, Centro de Investigação e Intervenção Social (CIS-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Av.ª das Forças Armadas, 1649-026 Lisboa, Portugal. Email: g.schellenberg@utoronto.ca or cesar.lima@iscte-iul.pt

organized hierarchically (Koelsch, 2011; Krumhansl, 2000; Peretz & Coltheart, 2003).

In the present investigation, we asked whether basic auditory processing is associated with musical expertise. Previous studies of musical expertise have typically focused on music training. The findings reveal that musically trained and untrained individuals differ in brain structure and function (e.g., Criscuolo et al., 2022; Herholz & Zatorre, 2012), and in nonmusical abilities such as general cognition (e.g., Criscuolo et al., 2019) and language (e.g., Neves et al., 2022). Associations between music training and basic auditory skills, however, remain poorly understood. Moreover, musical abilities can differ markedly between two individuals with the same amount of training, just as school performance differs between two students in the same grade. Accordingly, recent studies have focused on aspects of musical expertise other than training, by examining musical abilities in untrained participants (Correia et al., 2023), or by controlling for training in the analyses (e.g., Correia, Castro, et al., 2022).

Observed differences between musically trained and untrained individuals are often interpreted as plasticity and transfer effects, with extensive practice thought to improve nonmusical abilities (e.g., Alain et al., 2014; Engel et al., 2019; Palomar-García et al., 2020; Schneider et al., 2022). Nevertheless, evidence typically comes from cross-sectional designs, which preclude causal inferences (Schellenberg, 2020), and it is well established that musicians and nonmusicians differ in several respects other than training. In addition to differences in personality, cognition, and socioeconomic status (SES) that predict who takes music lessons and for how long (e.g., Corrigall et al., 2013), musical abilities and behaviors have a genetic component (Tan et al., 2014; Wesseldijk et al., 2023). Moreover, associations with general cognitive ability (Swaminathan et al., 2017) are evident whether musical expertise is defined as music training (Corrigall et al., 2013; Schellenberg, 2006) or musical ability (e.g., Correia, Vincenzi, et al., 2022; Slevc et al., 2016). These findings, combined with a lack of clear evidence for a causal role of music training, motivated Schellenberg and Lima (2024) to raise the possibility that nonmusical advantages observed among musicians stem from preexisting factors.

When we think of musical expertise, we are likely to envision skilled performance abilities. Because such skills are typically limited to individuals with music training, equating musical expertise with skilled performance excludes most of the population. Musical performances are also difficult to measure and compare (McPherson & Thompson, 1998). For example, performance evaluations are influenced by irrelevant factors such as visual cues (Tsay, 2013). Musical ability, or musicality (Honing, 2018), is nevertheless considered a universal human trait. In contrast to music performance, musicality is relatively easy to measure with objective tests or self-report questionnaires, which can be administered in the laboratory (or online) to individuals who vary widely in music training. As expected, professional musicians tend to score higher on these measures compared to amateurs, who score higher than untrained adults (Vincenzi et al., 2022).

Musical-ability tests include the Musical Ear Test (MET; Wallentin et al., 2010), the Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al., 2003), the Profile of Music Perception Skills (Law & Zentner, 2012), and the Swedish Musical Discrimination Test (Ullén et al., 2014). These tests are similar to earlier musical-aptitude tests (Gordon, 1967; Seashore et al., 1956), which were designed to measure natural ability and whether a child is likely to benefit from training. On each test trial, listeners hear two tone sequences and judge their similarity. Same/different or yes/no judgments are the most common response format (Peretz et al., 2003; Ullén et al., 2014; Wallentin et al., 2010), but listeners may also rate how confident they are that the sequences are the same or different (Law & Zentner, 2012), whether a sequence is a march or a waltz (Peretz et al., 2003), which tone in a sequence is displaced (Ullén et al., 2014), or whether one tone is higher or lower compared to another tone (Ullén et al., 2014). The tests have different subtests (e.g., melody, rhythm) but they all focus primarily on the ability to perceive, encode, remember, and discriminate tone sequences. In the present study, we assumed that performance on these tests is a marker of musical ability.

Some previous evidence that musical ability is linked with basic auditory processing comes from atypically developing individuals. For example, the ability of cochlear-implant users to recognize melodies or spoken utterances correlates with their ability to perceive spectral ripple and pitch contour (Won et al., 2010). Impaired auditory abilities also appear to be at the root of amusia (Albouy et al., 2016; Hyde & Peretz, 2004), a music disorder that can be congenital or acquired. In one instance (Hyde & Peretz, 2004), participants heard sequences of five monotonic, isochronous piano tones, and were asked to identify when the fourth tone was displaced in pitch or time. Individuals with amusia had intact detection of temporal displacements, but impaired detection of small pitch changes ($\leq$1 semitone). This pitch-processing deficit can have cascading effects, such as impaired perception of speech prosody (Vuvan et al., 2015) and of emotions in faces and voices (Lima et al., 2016).

For typically developing individuals, auditory skills are often examined in the context of speech and language rather than music. Low-level auditory perception accounts for over 30% of the variance in the recognition of speech prosody (Globerson et al., 2013), and for several aspects of first (L1) and second (L2) language acquisition, including speech-in-noise perception (Boets et al., 2008), vocabulary (Bavin et al., 2010), and phonology (Werker, 2018). According to the auditory precision hypothesis (Saito, 2023), listeners with better auditory processing are better at decoding auditory input, which contributes to their L1 and L2 proficiency. Because individuals vary widely in the precision of their auditory system (Kidd et al., 2007), better acuity may provide an advantage that facilitates language learning broadly.

Although the distinction between auditory and music processing is not absolute (Conzelmann & Süß, 2015), musical-ability tests typically use natural sounding (e.g., piano, woodblock) tones, and sequences comprising multiple tones. By contrast, psychoacoustic tasks use nonmusical stimuli (e.g., pure tones) and focus on a particular acoustic dimension rather than a complex structure. Moreover, psychoacoustic tasks are often adaptive, such that successful responding increases task difficulty, with the acoustic change becoming smaller until the minimum detectable difference between stimuli—the listener's threshold—is determined (Fechner, 1889). Lower thresholds reflect enhanced acuity and less perceptual noise on the dimension being tested. Adaptive tests are also used to measure listeners' ability to discriminate melodies (Harrison et al., 2017) and to determine whether a lead vocal track is in tune with a backing instrumental track (Larrouy-Maestri et al., 2019).

The blurry distinction between tests of auditory and music processing is evident in an older test of musical ability (The Seashore

Measures of Musical Talent), which has subtests that measure listeners' ability to discriminate tones on the basis of pitch, intensity, and duration (Seashore et al., 1956). In more recent studies, tests of pitch or duration discrimination have been used to measure musical ability (Correia, Castro, et al., 2022; Thompson et al., 2012), and musical tasks have been used to measure basic auditory processing (Conzelmann & Süß, 2015; Martin et al., 2018). On the one hand, the failure to distinguish auditory from musical tests is consistent with the proposal that they are associated. On the other hand, a clearer distinction, parallel to the one made for language (e.g., Liberman & Whalen, 2000; Surprenant & Watson, 2001), could serve to provide a mechanistic explanation for the basis of musicality and links between musical expertise and nonmusical abilities, particularly speech.

When musicians are compared with nonmusicians, lower (better) auditory thresholds are evident most often for frequency/pitch (e.g., Bianchi et al., 2016, 2019; Boebinger et al., 2015; Kishon-Rabin et al., 2001; Lima et al., 2019; Micheyl et al., 2006; Schellenberg & Moreno, 2010), but they can also extend to duration (e.g., Jeon & Fricke, 1997; Rammsayer & Altenmüller, 2006; but see Boebinger et al., 2015), timbre (e.g., Lee & Müllensiefen, 2020; Pitt, 1994; but see Shorey et al., 2024), intensity/amplitude (e.g., Endo et al., 2021; Geringer, 1993, 1995), and relative pitch (e.g., Larrouy-Maestri, 2018; Schellenberg & Moreno, 2010). Lower thresholds on tests of gap detection (e.g., Donai & Jennings, 2016, Grassi et al., 2017), backward masking (e.g., Parbery-Clark et al., 2011; Strait et al., 2010; Yoo & Bidelman, 2019), and temporal-interval discrimination (e.g., Banai et al., 2012; Jakobson et al., 2003) suggest that musicians also have enhanced temporal processing. At the lower end of the spectrum, individuals with amusia have higher (worse) psychoacoustic thresholds compared to controls for pitch (e.g., Foxton, 2004; Y. Sun et al., 2017) and amplitude modulation (e.g., Graves et al., 2023; Whiteford & Oxenham, 2017).

Because existing evidence comes primarily from special groups (i.e., musicians, individuals with amusia), it remains unclear how auditory skills relate to individual differences in musical expertise in the general population. Although associations with other domains (e.g., speech) are sometimes predicted better by musical ability than by training (e.g., Correia, Castro, et al., 2022; Mankel & Bidelman, 2018; Swaminathan et al., 2017), many training studies continue to overlook preexisting links with musical ability (e.g., Benítez-Barrera et al., 2022), let alone basic auditory skills. In our view, good auditory abilities are likely to be a prerequisite for high levels of musical expertise, as they are for language proficiency (Saito, 2023). In the present study, we examined links between auditory skills and musical expertise in a sample unselected for music training. We used psychoacoustic tasks to assess basic auditory skills, including pitch and temporal acuity, discrimination of intensity and timbre, and backward masking. Seven of nine tasks included only pure tones and/or noise bursts. In the other two, complex tones were constructed artificially with five harmonics of equal amplitude. The tasks were selected based on results from previous studies (cited above) indicating an advantage for musicians over nonmusicians.

We defined musical expertise in four different ways. In addition to considering music training, we had two objective tests of musical ability that measured melody and rhythm discrimination. A fourth, self-report measure considered musical activities and behaviors other than training. We also measured demographic variables, personality, and cognitive abilities (abstract reasoning, short-term memory [STM], and working memory [WM]). Because these variables are known to be associated with music training (Corrigall et al., 2013; Kuckelkorn et al., 2021; Schellenberg, 2006) and musical ability (e.g., Correia, Vincenzi, et al., 2022; Swaminathan et al., 2018; Wallentin et al., 2010), we accounted for them in the statistical analyses.

Our general hypothesis was that better auditory acuity (i.e., lower psychoacoustic thresholds) would predict higher musical expertise. We also expected associations to vary depending on how expertise was measured, with pitch-based tasks predicting melody discrimination, temporal tasks predicting rhythm discrimination (Krumhansl, 2000), and musically trained individuals performing better on the musical and auditory tasks (e.g., Correia, Castro et al., 2022; Micheyl et al., 2006; Swaminathan et al., 2021). Because music training has a particularly strong link with melody discrimination (Correia, Vincenzi, et al., 2022; Swaminathan et al., 2021), we hypothesized that music training, like melody discrimination, would be associated most strongly with pitch-based psychoacoustic thresholds (e.g., Boebinger et al., 2015).

Self-reports of musical behaviors other than training served primarily as a control measure and test of discriminant validity, meant to delineate the specificity of the link between basic auditory processing and musical expertise. We expected that auditory skills would be independent of musical behaviors and activities that are likely to be culturally and socially determined, such as writing about music or comparing two performances of the same piece. Nevertheless, the pitch-perception deficit in amusia that leads to cascading deficits in music perception and appreciation could also go in the opposite direction, such that good listening skills lead to cascading benefits in musical behaviors that extend broadly across contexts.

Method

Transparency and Openness

This study was not preregistered. The data, materials, and procedures are available at https://osf.io/gzjea/. Musical-expertise tasks and questionnaires are available at https://app.gorilla.sc/openmaterials/218554. Statistical power was estimated a priori using G* Power (Faul et al., 2009), which determined that 137 participants were necessary to detect a partial correlation of .25 with 85% power, $\alpha = .05$, and six covariates.

Participants

This study was conducted in accordance with the Declaration of Helsinki and approved by the local ethics committee at Iscte, University Institute of Lisbon (reference 24/011). The sample comprised 138 participants (105 women, 32 men, one unspecified), aged 18 to 62 years ($M = 23.02$, $SD = 7.72$). Because of positive skew, age was square-root-transformed for statistical analysis. Most participants were college undergraduates who had completed high school ($n = 82$). Others had a bachelor's ($n = 43$) or postgraduate ($n = 13$) degree. In the statistical analyses, education was coded as a three-level variable based on the highest degree obtained (1 = *high school*, 2 = *bachelor's*, 3 = *postgraduate*). According to self-reports, all participants had normal hearing and were fluent speakers of Portuguese, except for nine who were tested in English. Participants provided written informed consent and received partial course credit or 10€ as compensation for their time. Three additional participants were recruited but excluded because of incomplete responding ($n = 1$) or because their performance was significantly below 50% correct (chance level) on the MET ($n = 2$), and therefore uninterpretable.

Measures

Musical Expertise

Two of our four measures of musical expertise came from an objective test and two from a self-report questionnaire. The objective test—the MET (Wallentin et al., 2010)—evaluated music-perception abilities. The MET provided separate scores for its melody and rhythm subtests (presented in that order). It has good reliability and validity in both its original (Swaminathan et al., 2021) and online (Correia, Vincenzi, et al., 2022) versions. In the present sample, good internal reliability was evident for both subtests (melody, Cronbach's $\alpha = .78$; rhythm, $\alpha = .73$).

Both MET subtests had 52 trials preceded by two practice trials. On each trial, listeners heard two short piano (melody subtest) or woodblock (rhythm subtest) sequences and made a yes/no judgment about whether the sequences were identical. On half of the trials (26 of 52), standard and comparison sequences were identical. On the other half, the comparison had one or more tones altered in pitch (melody), or the duration of interonset intervals was changed (rhythm). Feedback was provided only for practice trials. Scores were the number of correct responses, converted to percentages for ease of interpretation. We used the online version of the MET (Correia, Vincenzi, et al., 2022), created with Gorilla Experiment Builder (Anwyl-Irvine et al., 2020; available at https://app.gorilla.sc/openmaterials/218554). The test took approximately 20 min to complete.

Participants also completed the online version (Correia, Vincenzi, et al., 2022) of the Goldsmiths Musical Sophistication Index (Gold-MSI; Müllensiefen et al., 2014; Portuguese translation: Lima et al., 2020). The Gold-MSI had 38 items that asked about a variety of musical skills and behaviors, which participants rated on 7-point scales. The items were grouped and averaged to form five subscales: active engagement (nine items, e.g., I enjoy writing about music, for example on blogs and forums), perceptual abilities (nine items, e.g., I am able to judge whether someone is a good singer or not), music training (seven items, e.g., I can play [0–6 or more] musical instruments), singing abilities (seven items, e.g., I am able to hit the right notes when i sing along with a recording), and emotions (six items, e.g., Pieces of music rarely evoke emotions for me).

The music training subscale (possible range: 1–7) served as our measure of music training. In addition to an item that asked specifically about years of training (Item 36), six additional items indexed practice duration and frequency, music theory, number of instruments played, compliments received on performances, and whether participants self-identified as musicians. Internal reliability in the present sample was high (Cronbach's $\alpha = .87$). For our control measure of musical abilities and behaviors (other than music training), we extracted the principal component from the other four subscales. Each loaded highly onto the principal component (lowest: emotions, $r = .642$; highest: perceptual abilities, $r = .789$), which explained 55.4% of the variance in the original four subscales. This measure included everyday musical activities and behaviors, commensurate with the view that musical expertise is multifaceted (Ullén et al., 2016).

Psychoacoustic Tasks

We administered nine psychoacoustic tasks from the Psychoacoustics toolbox for Matlab (Soranzo & Grassi, 2014), simplifying some task names for clarity. Stimuli were initially presented binaurally at 75 dB sound pressure level through Mars Gaming MH4X headphones connected to the computer. Tones included 10-ms onset and offset ramps to prevent audible clicks. Thresholds were estimated with a two-down, one-up adaptative staircase procedure, which identified the minimum acoustic change a listener required to detect a difference between standard and comparison stimuli with 70.7% accuracy (Levitt, 1971; Soranzo & Grassi, 2014). Each task took approximately 5 min to complete. Before each task, the experimenter provided oral instructions and participants completed a training trial. They were allowed to adjust the sound level to a comfortable volume.

Eight of nine tasks had a three-alternative forced-choice response format. On each trial, two standard (identical) stimuli were presented with a comparison stimulus. The position of the comparison stimulus (first, second, or third) varied randomly across trials, and participants were asked to indicate its position. The magnitude of the difference between the standards and the comparison was adjusted in steps, becoming smaller or larger after correct or incorrect responding, respectively. Initially, the difference decreased by a factor of 2 after two correct responses and increased by a factor of 2 after one incorrect response (i.e., the stimulus difference was divided or multiplied by 2). After four reversals, a smaller step size ($\sqrt{2}$) was used for the remaining eight reversals to fine-tune the threshold estimate, which was the average of the difference between standard and comparison stimuli for the eight final reversals. Because lower thresholds indicated better performance, negative correlations were expected between thresholds and musical expertise.

Three tasks measured pitch perception. The pitch task measured the smallest change in frequency that produced a perceptible difference in pitch. On each trial, listeners heard three pure tones (250 ms) and identified the highest. Standard tones had a frequency of 1000 Hz. The comparison was always higher. The initial difference between the standards and comparison was 100 Hz but became smaller as the task progressed. A second pitch-perception task, contour speed, measured the minimum tone duration listeners required to form a stable mental representation of pitch, such that two tones could be judged in terms of contour (i.e., which tone was higher in pitch). Listeners heard three sequences of four consecutive tones (i.e., no gaps between tones) and determined when the contour of the second and third tones (low–high or high–low) differed. The first and fourth tones were fixed at 650 Hz and 100 ms. For standard (low–high) sequences, the second tone was 550 Hz and the third tone was 710 Hz. For comparisons (high–low), frequencies were inverted. The initial duration of the second and third tones was 200 ms but became shorter as the task progressed.

The third pitch-perception task—scale mistuning—measured relative pitch and the minimum difference required to detect a change in interval size (i.e., pitch distance between tones). Its yes/no response format differed from the other tasks by relying on long-term memory for the intervals of a familiar tone sequence, specifically the ascending equal-tempered major scale (do, re, mi, …). On each trial, listeners heard the scale starting on middle C ($C_4 = 261.6$ Hz) and indicated whether it was in tune or out of tune. Individual tones were 500-ms complexes of five harmonics. On the first trial, the fifth note (sol) was mistuned upward by 80 cents, which increased and decreased the pitch distance from fa and la, respectively. Although correct responding (out of tune) led to a decrease in mistuning, the scale was never presented completely in tune. Rather, the mistuning eventually became imperceptible, such that participants made an error, which caused a reversal and the task to proceed like the others. The overall goal was to determine how close to exact equal temperament (700 cents higher than do) sol needed to be for listeners to perceive it as equidistant between fa (200 cents lower) and la (200 cents higher).

Three tasks measured temporal perception. The duration task determined the minimum discernible difference in tone duration. On each trial, listeners heard three consecutive 1000-Hz pure tones and identified the longest. The interonset interval was fixed at 500 ms. The initial difference in duration between standards (250 ms) and comparison (350 ms) was 100 ms. In the complex duration task, listeners heard three sequences of six 20-ms tones (1,000 Hz). Each comprised three tone pairs. In standard sequences, pairs were separated by 120 ms of silence, but tones within pairs were separated by 40 ms. In the comparison sequence, the separation decreased between pairs (initially by 60 ms) but increased by the same extent within pairs such that total duration of standard and comparisons sequences was constant. The third test of temporal perception was gap detection, a low-level test of temporal acuity that is particularly relevant for speech perception (Phillips, 1999). On each trial, listeners heard three bands of Gaussian noise and identified the one with a silent gap at its temporal center. Standards comprised 750 ms of continuous noise. The comparison was identical but with a silent gap, which was initially 60 ms.

The intensity task measured the minimum perceptible difference in tone intensity. Listeners heard three consecutive pure tones (1,000 Hz, 250 ms) separated by 500 ms of silence between tones. The task was to identify the loudest. Standard tones were always attenuated from base level (75 dB sound pressure level) by −30 dB. The comparison was always louder. The initial difference between the standards and the comparison was +10 dB.

In the backward masking task, listeners heard three 300-ms sounds of bandpass noise (400–1,600 Hz), one of which was preceded immediately by a 20-ms pure tone of 1,000 Hz (i.e., no temporal gap between tone and noise). The initial sound level of the tone was attenuated by −10 dB but became quieter over time. The task was to identify which noise band was preceded by the tone. Backward masking occurred when the noise made the preceding tone imperceptible. This phenomenon is evident in vision (Breitmeyer & Ogmen, 2000) and audition (Elliott, 1971) whenever the perception of a stimulus is obscured by a subsequent stimulus.

Timbre (or tone color) refers to the relative amplitude of individual harmonics in complex tones and how they change over time (McAdams, 2013). For example, tones with the same duration, amplitude, and pitch that are produced by a piano or guitar differ only in timbre. The timbre task measured the minimum difference that allowed listeners to detect a change in timbre. On each trial, they heard three complex tones. Standards comprised five harmonics with the same amplitude (−40 dB). The intensity of the third harmonic was increased for the comparison. The initial difference was +20 dB. To ensure that the task was measuring listener's sensitivity to timbre changes (rather than differences in loudness), the overall level of the standards and comparison was constant within each trial.

General Cognitive Ability

Abstract Reasoning. The Matrix Reasoning Item Bank (MaRs-IB; Chierchia et al., 2019; available at https://app.gorilla.sc/openmaterials/36164) provided an 8-min online test of nonverbal, abstract reasoning modeled after Raven's advanced progressive matrices (Raven, 1965). On each trial, participants saw a 3 × 3 matrix. Eight of nine cells contained abstract patterns that varied systematically on four dimensions (color, size, shape, and location) but the bottom-right cell was empty. Participants chose one of four alternatives to fill the missing cell, based on logical progressions (left-to-right, top-to-bottom) among other cells. The order of the test's 80 trials was the same for all participants, who were unaware of task duration but informed that they had a maximum of 30 s to respond to each trial. If all 80 trials were completed in less than 8 min, the test restarted in the same order, but responses from repeated trials were not added to scores, which were calculated as the proportion of correct responses (number correct/trials attempted). Responses faster than 250 ms were excluded, and proportions were logit-transformed for statistical analyses (as in Correia et al., 2023; Correia, Vincenzi et al., 2022; Vincenzi et al., 2022).

Auditory Short-Term and WM. The forward and backward portions of the digit span subtest from the Wechsler Adult Intelligence Scale–Third Edition (Wechsler, 2008) measured the capacity of auditory STM and WM. The experimenter read aloud a list of numbers, which participants were asked to reproduce verbally, either in the same or reverse order, in the forward and backward parts of the test, respectively. After every second trial, the number of digits increased. Testing was discontinued when participants failed to recall correctly on two lists with the same number of digits. For both forward and backward portions of the test, a participant's score was the maximum number of digits recalled correctly.

Personality

The online version (Correia, Vincenzi, et al., 2022) of the Portuguese translation (Brito Costa et al., 2016) of the Big Five Inventory (BFI; John et al., 2008) assessed the Big-Five Dimensions of Personality (McCrae & Costa, 1987; McCrae & John, 1992). The questionnaire had 44 items describing the participant (e.g., I am someone who is full of energy), who agreed or disagreed on 5-point scales. Items were averaged and internal reliability was calculated for each of five traits: extraversion ($\alpha = .90$), agreeableness ($\alpha = .63$), conscientiousness ($\alpha = .80$), neuroticism ($\alpha = .88$), and openness-to-experience (hereafter openness, $\alpha = .80$).

Procedure

Participants were tested individually in a single 90-min session in a university laboratory. Tasks were administered in the following order: psychoacoustic battery and digit span, Gold-MSI, BFI, MaRs-IB, and MET. The psychoacoustic battery was divided into two blocks and implemented on Matlab r2022b, Version 9.13.0. The first block had five tests (pitch, intensity, duration, complex duration, and contour speed); the second block had four (backward masking, gap detection, timbre, scale mistuning). Digit span was administered between blocks. The order of tests within blocks was randomized separately for each participant. Other tests were administered with Gorilla software (Anwyl-Irvine et al., 2020). The Gold-MSI and BFI were followed by the MaRs-IB and finally the

MET. After completing the experiment, participants received summary feedback about their scores on the BFI, Gold-MSI, and MET.

Data Analysis

Analyses included both standard frequentist and Bayesian statistics conducted with JASP Version 0.19.1 (JASP Team, 2024; default priors). Whereas frequentist statistics evaluate the likelihood of the observed data if the null hypothesis is true, Bayes factors (BF_{10}, reported here with three-digit accuracy) quantify the evidence that the data provide for the alternative compared to the null hypothesis. BF_{10} values greater than 1.00 provide evidence favoring the alternative hypothesis; values less than 1.00 favor the null hypothesis. This value is readily interpretable such that when $BF_{10} = 5.00$, for example, the observed data are 5 times more likely under the alternative than the null hypothesis. Conversely, when $BF_{10} = 0.200$ (1/5), the data are 5 times more likely under the null. By convention (Jarosz & Wiley, 2014; Jeffreys, 1998), evidence is considered weak or anecdotal when $0.333 \leq BF_{10} \leq 3.00$, but when $BF_{10} >$ 3.00 (or <0.333), substantial evidence is provided for the alternative (or null) hypothesis. When $BF_{10} > 10.0$, 30.0, or 100 (or <0.100 or 0.033, or 0.010), the data provide strong, very strong, or decisive evidence, respectively, for the alternative (or null) hypothesis. We considered results to be reliable only when $BF_{10} > 3.00$ for positive results, or $BF_{10} < 0.333$ for null results. This criterion is typically more conservative than $\alpha = .05$, but it can be less conservative that some methods of correcting for multiple tests.

For musical expertise variables other than training, we conducted analyses with and without holding music training constant to ensure that any observed associations with psychoacoustic thresholds were not the consequence of musically trained individuals performing better on the psychoacoustic and expertise measures. For example, associations between melody perception and pitch thresholds could reflect advantages of musically trained individuals in both contexts, rather than a direct link between them.

Before statistical analysis, psychoacoustic thresholds were log-transformed to reduce skew. This transformation is common (Globerson et al., 2013; Micheyl et al., 2006) because thresholds tend to follow a logarithmic distribution (Moore, 2003). After excluding outliers (scores ≥ 3 *SD*s below the group mean), data were missing from one participant for the pitch, intensity, and duration tasks. Data were also missing for complex duration ($n = 2$), gap detection ($n = 2$), timbre ($n = 3$), scale mistuning ($n = 6$), and contour speed ($n = 8$). Thus, sample size varied across analyses depending on which variables were included.

Results

Table 1 provides descriptive statistics for musical expertise, psychoacoustic thresholds (raw data), cognition, and personality. The four measures of musical expertise are referred to as music training (music training subscale from the Gold-MSI), melody (MET-melody), rhythm (MET-rhythm), and informal musicality (self-reported musical abilities and behaviors from the Gold-MSI excluding music training). The data provided decisive evidence that performance was above chance levels ($M = 0.50$) for melody, $t(137) = 15.26$, $p < .001$, $BF_{10} > 100$, and for rhythm, $t(137) =$ 18.92, $p < .001$, $BF_{10} > 100$. (There was no "chance level" for music training or informal musicality.)

Table 1

Descriptive Statistics for All Tasks

Task	*N*	*M*	*SD*	Range
Musical expertise				
MET-melody	138	65.87	12.22	42.31–94.23
MET-rhythm	138	68.59	11.54	36.54–92.31
Music training	138	2.90	1.42	1.00–6.57
Informal musicality	138	0.00	1.00	−2.97–2.30
Psychoacoustic thresholds				
Pitch (Hz)	137	30.84	40.30	1.63–300.13
Intensity (dB)	137	1.93	1.09	0.48–5.46
Duration (ms)	137	33.07	15.21	7.54–87.72
Complex duration (ms)	136	122.95	677.56	3.48–7,784.31
Contour speed (ms)	130	124.73	86.49	13.58–531.07
Backward masking (dB)	138	47.31	18.92	101.63–12.25
Gap detection (ms)	136	2.45	0.79	1.05–5.27
Timbre (dB)	135	6.95	4.29	0.93–24.29
Scale mistuning (cents)	132	141.50	137.17	1.28–659.41
Cognition				
Abstract reasoning	138	63.36	15.78	26.92–95.00
Short-term memory	138	6.08	1.05	3.00–9.00
Working memory	138	4.88	1.17	2.00–8.00
Personality				
Extraversion	138	3.10	0.89	1.13–5.00
Agreeableness	138	3.83	0.49	2.11–4.78
Conscientiousness	138	3.33	0.67	1.78–4.78
Neuroticism	138	3.24	0.79	1.13–4.88
Openness	138	3.75	0.61	2.30–5.00

Note. For backward masking, numbers are absolute values of the attenuation (expressed in dB) applied to the tone preceding the noise. MET = Musical Ear Test.

Pairwise correlations among the expertise variables are provided in Table 2. Evidence for a positive association with music training was decisive for melody and informal musicality and substantial for rhythm. The association between melody and rhythm was also decisive. For informal musicality, there was strong evidence for an association with melody, but for rhythm the evidence was anecdotal. To test whether music training was driving associations among the other three expertise variables, we retested the correlations with music training held constant. The data continued to provide decisive evidence for a partial association between melody and rhythm, $r_p = .538$, $p < .001$, $BF_{10} > 100$. For informal musicality, however, there was now substantial evidence for null associations

Table 2

Associations Among Musical-Expertise Variables

Variable	Music training	Melody	Rhythm
Melody			
r	**.410**		
p	**<.001**		
BF_{10}	**>100**		
Rhythm			
r	**.247**	**.576**	
p	**.003**	**<.001**	
BF_{10}	**7.24**	**>100**	
Informal Musicality			
r	**.559**	**.275**	.199
p	**<.001**	**.001**	.019
BF_{10}	**>100**	**20.6**	1.59

Note. Bold font indicates BF > 3. BF = Bayes factor.

with melody, $r_p = .060$, $p = .485$, $\mathrm{BF}_{10} = 0.181$, and with rhythm, $r_p = .075$, $p = .381$, $\mathrm{BF}_{10} = 0.207$.

Table 3 provides pairwise correlations among psychoacoustic thresholds. There was substantial or stronger evidence for 23 of the 36 correlations tested, all of which were positive. Each threshold was associated positively with at least two other thresholds. Duration was associated with all other thresholds, pitch with seven, contour speed and backward masking with six, intensity, complex duration, and timbre with five, and gap detection and scale mistuning with two. As in Surprenant and Watson (2001), the magnitude of the correlations was not particularly large, and the overlapping variance never exceeded 21% between any pair of variables. Moreover, in contrast to Johnson et al. (1987), associations were not notably stronger among the various measures of pitch or spectral perception, or among tasks that relied on temporal cues. Thus, there was no empirical justification to form aggregate variables.

Before we asked whether psychoacoustic variables predicted musical expertise, we tested whether musical expertise was associated with demographics, cognition, or personality, which could potentially explain any observed associations. Correlations are provided in Table 4. Although music training was not associated with any control variable, the other three measures of musical expertise were. The observed data provided substantial evidence that (a) melody scores correlated positively with age and education, (b) melody and rhythm scores correlated with WM, and (c) informal musicality was associated positively with extraversion. There was also decisive evidence for a positive correlation between informal musicality and openness. In subsequent analyses, age, education, WM, openness, and extraversion (hereafter control variables) were held constant.

The main results—partial correlations between musical expertise and psychoacoustic thresholds—are reported in Table 5 and illustrated in Figure 1. For each expertise variable, alpha-levels were corrected for multiple (nine) psychoacoustic tests (Holm-Bonferroni method).

Music Training

As shown in Table 5, there was decisive evidence that music training predicted lower thresholds for pitch and scale mistuning, and strong evidence for associations with contour speed and backward masking. Except for backward masking, these associations involved measures of pitch perception.

Because the relevant literature on music training (for review, see Schellenberg & Lima, 2024) typically classifies participants as musicians or nonmusicians, or measures duration of music training, we repeated the analyses reported in Tables 4 and 5 (first column), replacing the music training subscale with years of formal training (Item 36 from the Gold-MSI). Detailed results are provided in Tables S1 and S2 in the online supplemental materials. We first confirmed that the two measures were positively correlated, $r = .809$, $p < .001$, $\mathrm{BF}_{10} > 100$. As with music training, years of training was not correlated with demographics, cognition, or

Table 3

Associations Among Psychoacoustic Variables

Variable	Intensity	Duration	Complex duration	Contour speed	Backward masking	Gap detection	Timbre	Scale mistuning
Pitch								
r	**.244**	**.346**	**.333**	**.456**	**.399**	.143	**.436**	**.366**
p	**.004**	**<.001**	**<.001**	**<.001**	**<.001**	.098	**<.001**	**<.001**
BF_{10}	**6.12**	**>100**	**>100**	**>100**	**>100**	0.417	**>100**	**>100**
Intensity								
r		**.301**	.219	**.337**	**.430**	**.270**	.218	.087
p		**<.001**	.011	**<.001**	**<.001**	**.002**	.011	.326
BF_{10}		**57.0**	2.69	**>100**	**>100**	**15.1**	2.59	0.176
Duration								
r			**.249**	**.274**	**.355**	**.298**	**.386**	**.232**
p			**.004**	**.002**	**<.001**	**<.001**	**<.001**	**.008**
BF_{10}			**7.07**	**14.4**	**>100**	**47.3**	**>100**	**3.70**
Complex duration								
r				**.276**	**.259**	.030	**.418**	.069
p				**.002**	**.002**	.730	**<.001**	.434
BF_{10}				**15.1**	**10.6**	0.115	**>100**	0.148
Contour speed								
r					**.413**	.144	**.409**	.185
p					**<.001**	.104	**<.001**	.038
BF_{10}					**>100**	0.408	**>100**	0.939
Backward masking								
r						.132	**.399**	.196
p						.127	**<.001**	.025
BF_{10}						0.340	**>100**	1.33
Gap detection								
r							.169	.092
p							.051	.296
BF_{10}							0.708	0.188
Timbre								
r								.164
p								.063
BF_{10}								0.605

Note. Bold font indicates $\mathrm{BF} > 3$. BF = Bayes factor.

Table 4
Associations Between Measures of Musical-Expertise Variables and Demographic, Cognitive, and Personality Variables

Variable	Music training	Melody	Rhythm	Informal musicality
Age				
r	.012	**.224**	.194	−.051
p	.889	**.008**	.023	.550
BF_{10}	0.108	**3.35**	1.39	0.127
Gender				
r	−.041	.021	.039	.057
p	.635	.805	.647	.507
BF_{10}	0.119	0.110	0.119	0.133
Education				
r	.093	**.244**	.218	−.006
p	.276	**.004**	.010	.948
BF_{10}	0.192	**6.57**	2.82	0.107
Abstract reasoning				
r	.055	.115	.073	.058
p	.525	.181	.393	.496
BF_{10}	0.130	0.258	0.153	0.134
Short-term memory				
r	.112	.130	.214	.071
p	.193	.129	.012	.408
BF_{10}	0.247	0.334	2.49	0.149
Working memory				
r	.065	**.255**	**.253**	.044
p	.449	**.003**	**.003**	.606
BF_{10}	0.141	**9.77**	**9.02**	0.121
Extraversion				
r	.138	−.036	−.004	**.241**
p	.106	.677	.967	**.004**
BF_{10}	0.387	0.116	0.107	**5.87**
Agreeableness				
r	.072	.109	.031	.149
p	.400	.204	.715	.082
BF_{10}	0.151	0.237	0.114	0.477
Conscientiousness				
r	.061	.085	−.054	.087
p	.476	.322	.531	.311
BF_{10}	0.137	0.173	0.129	0.177
Neuroticism				
r	−.163	−.204	−.169	−.070
p	.056	.016	.048	.418
BF_{10}	0.652	1.85	0.742	0.147
Openness				
r	.160	.070	.041	**.441**
p	.060	.417	.631	**<.001**
BF_{10}	0.612	0.147	0.119	**>100**

Note. Gender is dummy coded (0 = *men*, 1 = *women*). Bold font indicates BF > 3. BF = Bayes factor.

personality. Moreover, years of training correlated with scale mistuning, pitch, and backward masking, but not with contour speed. We did not consider years of training further.

As shown in Table 3, the four psychoacoustic thresholds that were associated with the music training subscale were intercorrelated. We therefore used multiple regression to ask whether any had unique predictive value when the other three, as well as the five control variables, were held constant. Detailed statistical results are provided in Table S3 in the online supplemental materials. The data provided decisive evidence for the fit of the overall model, $R^2 = .321$, $F(9, 115) = 6.03$, $p < .001$, $BF_{10} > 100$. Scale mistuning had a strong independent association with music training, $r_p = -.294$, $p = .001$, $BF_{10} = 34.1$, as did pitch, $r_p = -.284$, $p = .002$, $BF_{10} = 25.1$. For contour speed, $r_p = -.057$, $p = .542$, $BF_{10} = 0.404$, and backward masking, $r_p = -.074$, $p = .425$, $BF_{10} = 0.455$, there was no substantial evidence for partial associations, or for null associations. For control variables, the data provided substantial evidence that extraversion made an independent contribution to the model, $r_p = .218$, $p = .018$ $BF_{10} = 4.18$. Thus, participants with more music training tended to have lower thresholds for pitch and scale mistuning, and to be more extraverted.

Melody

As shown in Table 5, there was decisive evidence that better melody performance was associated with lower thresholds for pitch, contour speed, scale mistuning, strong evidence for timbre, and substantial evidence for backward masking and complex duration. As expected, the three strongest associations ($BF_{10} > 100$) were with pitch-perception measures (pitch, contour speed, scale mistuning), such that psychoacoustic correlates of melody and music training were similar.

Nevertheless, associations of melody with psychoacoustic thresholds could have been driven by the musically trained participants, who scored better on the psychoacoustic tasks and the melody subtest. Table 5 also provides correlations with music training additionally held constant. Decisive evidence remained for associations of melody with pitch and contour speed, strong evidence for scale mistuning, and substantial evidence for timbre. Complex duration and backward masking were no longer significant, but the data favored neither the null nor alternative hypothesis. As expected, the strongest associations continued to be with pitch-perception measures.

We then asked if any of these four psychoacoustic thresholds (pitch, contour speed, scale mistuning, timbre) had unique associations with melody, using multiple regression to hold constant the other three as well as music training and the control variables (see Table S4 in the online supplemental materials). The data provided decisive evidence for the fit of the overall model, $R^2 = .434$, $F(10, 113) = 8.67$, $p < .001$, $BF_{10} > 100$. Although there was strong evidence for an independent contribution of contour speed, $r_p = -.257$, $p = .006$, $BF_{10} = 10.2$, there was no longer evidence for a partial association with scale mistuning, $r_p = -.203$, $p = .030$, $BF_{10} = 2.64$, or with pitch, $r_p = -.166$, $p = .076$, $BF_{10} = 1.29$, yet evidence for null associations was not substantial either. For timbre, $r_p = -.025$, $p = .788$, $BF_{10} = 0.313$, however, there was substantial evidence for the null. There was also substantial evidence for partial associations with age, $r_p = .251$, $p = .007$, $BF_{10} = 8.64$, and music training, $r_p = .229$, $p = .014$, $BF_{10} = 4.95$. In short, performance on the melody subtest tended to be better among participants with low thresholds for contour speed, older participants, and those with more music training.

Rhythm

As shown in Table 5, associations with rhythm were evident for all psychoacoustic tasks except gap detection. There was decisive evidence for associations with pitch and timbre, very strong evidence for associations with contour speed, complex duration, and backward masking, and substantial evidence for associations with intensity, duration, and scale mistuning. Contrary to predictions, associations with tasks that measured temporal perception were not particularly strong; in absolute terms, the strongest associations were with pitch and timbre. Very strong associations were also evident, however, for backward masking, which measured fine-grained

Table 5

Partial Correlations Between Measures of Musical Expertise and Psychoacoustic Thresholds

					Music training also held constant		
Psychoacoustic threshold	Music training	Melody	Rhythm	Informal musicality	Melody	Rhythm	Informal musicality
Pitch							
r_p	**−.460**	**−.476**	**−.376**	**−.363**	**−.347**	**−.309**	−.148
p	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	.093
BF_{10}	**>100**	**>100**	**>100**	**>100**	**>100**	**>100**	0.845
Intensity							
r_p	−.118	−.133	**−.232**	−.141	−.093	**−.211**	−.092
p	.179	.127	**.007**	.106	.292	.016	.296
BF_{10}	0.879	1.01	**8.26**	0.976	0.507	**4.71**	0.371
Duration							
r_p	−.086	−.058	**−.202**	**−.274**	−.024	−.187	**−.273**
p	.329	.510	**.020**	**.001**	.787	.032	**.002**
BF_{10}	0.616	0.454	**3.87**	**29.6**	0.321	2.72	**24.8**
Complex duration							
r_p	−.159	**−.199**	**−.292**	−.168	−.146	**−.265**	−.098
p	.070	.023	**<.001**	.055	.097	**.002**	.269
BF_{10}	1.64	**3.52**	**54.0**	1.59	1.06	**21.0**	0.398
Contour speed							
r_p	**−.295**	**−.425**	**−.286**	−.200	**−.344**	**−.237**	−.052
p	**<.001**	**<.001**	**.001**	.025	**<.001**	**.008**	.546
BF_{10}	**45.9**	**>100**	**34.1**	2.93	**>100**	**7.86**	0.271
Backward masking							
r_p	**−.245**	**−.233**	**−.296**	−.178	−.148	**−.252**	−.053
p	**.005**	**.007**	**<.001**	.048	.091	**.004**	.546
BF_{10}	**12.3**	**8.74**	**68.3**	2.02	1.12	**14.9**	0.263
Gap detection							
r_p	.010	−.041	−.079	−.086	−.049	−.083	−.110
p	.913	.644	.373	.328	.578	.347	.212
BF_{10}	0.418	0.414	0.546	0.464	0.362	0.561	0.461
Timbre							
r_p	−.176	**−.280**	**−.317**	−.088	**−.229**	**−.287**	.012
p	.045	**.001**	**<.001**	.321	**.009**	**<.001**	.895
BF_{10}	2.24	**33.8**	**>100**	0.472	**6.69**	**41.8**	0.226
Scale mistuning							
r_p	**−.402**	**−.388**	**−.216**	**−.285**	**−.257**	−.124	−.090
p	**<.001**	**<.001**	**.015**	**.001**	**.004**	.168	.315
BF_{10}	**>100**	**>100**	**4.96**	**36.8**	**14.2**	0.873	0.388

Note. Age, education, working memory, extraversion, and openness-to-experience are held constant. Bold font indicates significant *p* value (Holm-Bonferroni corrected) or BF > 3. BF = Bayes factor.

temporal resolution, and for complex duration, which measured temporal pattern perception. In short, performance on rhythm correlated with performance on temporal tasks, but not more so than it did on nontemporal tasks.

When we asked whether observed associations remained evident when music training was additionally held constant, substantial or stronger evidence was apparent for six of eight partial correlations, with decisive evidence for pitch, very strong evidence for timbre, strong evidence for complex duration and backward masking, and substantial evidence for intensity and contour speed (Table 5). In this instance, frequentist evidence for intensity was not significant after correcting for multiple tests. The partial association with duration was significant with standard statistics but not after correcting for multiple tests, and anecdotal according to the BF. The association with scale mistuning also disappeared, but evidence for the null hypothesis was not substantial.

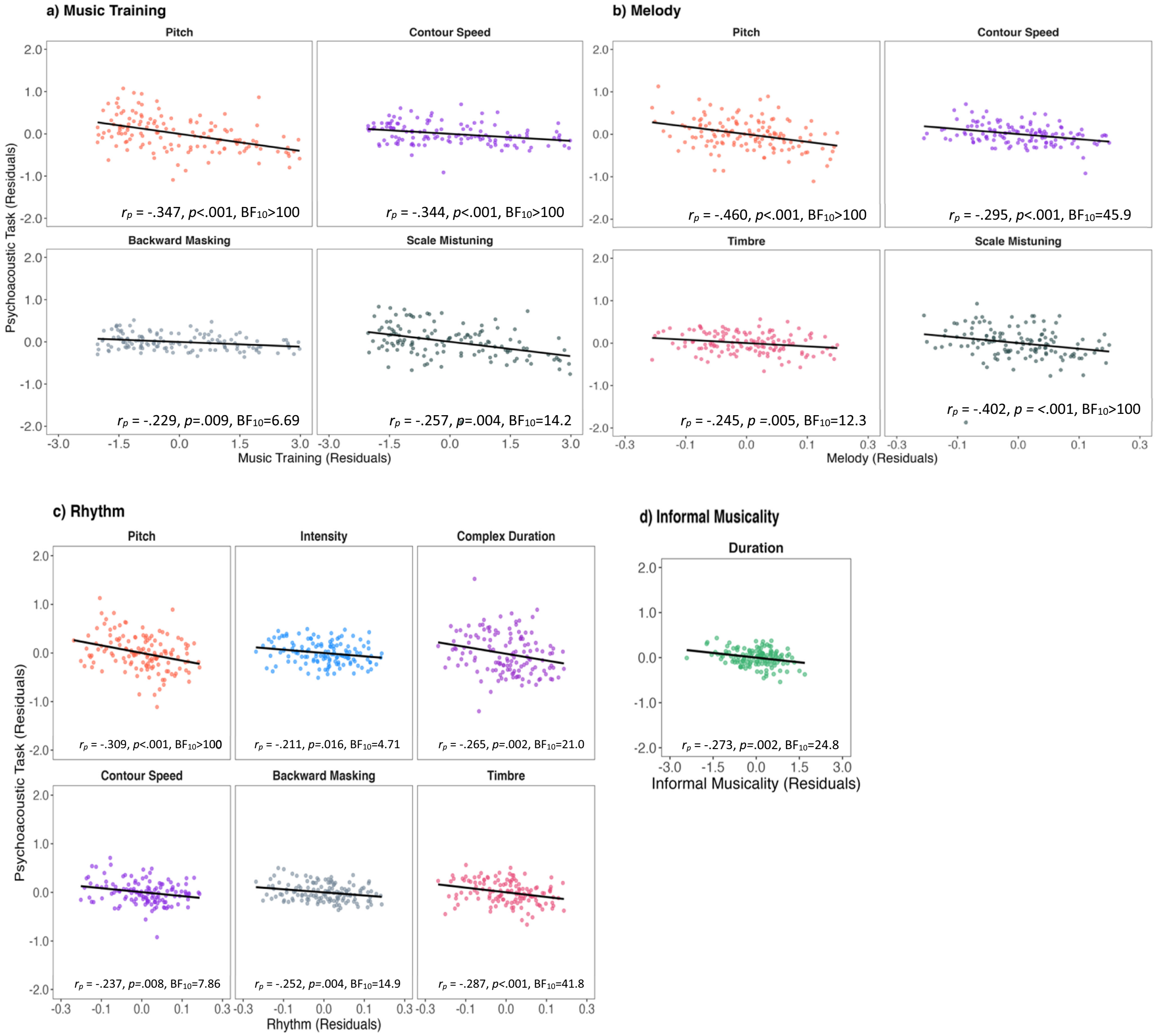
Did any of these six psychoacoustic thresholds (pitch, intensity, complex duration, contour speed, backward masking, and timbre) continue to predict rhythm with the other five held constant as well as music training and the control variables? Although the data provided substantial evidence for overall model fit, $R^2 = .243$, $F(12, 112) = 2.99$, $p = .001$, $\text{BF}_{10} = 5.46$, there was no evidence that any of the six psychoacoustic thresholds had a significant partial correlation with rhythm, $ps \geq .143$, all $\text{BF}_{10} \leq 1.08$ (Table S5 in the online supplemental materials). Nevertheless, there was no substantial evidence for any null hypothesis, either. Similar results were evident for music training and all control variables, $ps \geq .322$, all $\text{BF}_{10} \leq .661$. Thus, although individual differences in rhythm performance were associated with multiple psychoacoustic thresholds, none stood out as particularly important or unimportant when the others were held constant.

Informal Musicality

As shown in Table 5, informal musicality predicted lower thresholds for pitch, duration, and scale mistuning. The observed data provided evidence that was decisive for pitch, very strong for scale mistuning, and strong for duration. When we additionally held music training constant, however, strong evidence remained only for duration. The data did not favor the null or alternative hypothesis

Figure 1

Plots Illustrating Associations Between Musical-Expertise Measures and Psychoacoustic Thresholds



Note. Panels depict significant associations for (a) music training, (b) melody, (c) rhythm, and (d) informal musicality. The data are residuals: Age, education, working memory, extraversion, and openness-to-experience are held constant. Music training is additionally held constant in Panels b, c, and d. BF = Bayes factor. See the online article for the color version of this figure.

for pitch or scale mistuning. In addition to the partial association with duration, the multiple-regression model provided decisive evidence for overall model fit, $R^2 = .513$, $F(7, 129) = 19.39$, $p < .001$, $\mathrm{BF}_{10} > 100$, and for independent contributions of music training, $r_p = .553$, $p < .001$, $\mathrm{BF}_{10} > 100$, and openness, $r_p = .422$, $p < .001$, $\mathrm{BF}_{10} > 100$ (Table S6 in the online supplemental materials). In short, informal musicality tended to be higher among participants with more music training, higher levels of openness, and lower thresholds for duration.

Because the association between with duration was unexpected and difficult to explain, we conducted an exploratory analysis separately for each Gold-MSI subscale that we used to form informal musicality. For example, duration thresholds could be a marker of self-reports of perceptual abilities, which would be relatively interpretable. After accounting for music training and the control variables, there was indeed substantial evidence for an association with the perceptual abilities subscale, $r_p = -.205$, $p = .019$, $\mathrm{BF}_{10} = 3.65$, but there was also strong evidence for an association with the emotions subscale, $r_p = -.267$, $p = .002$, $\mathrm{BF}_{10} = 23.5$. For singing abilities, $r_p = -.148$, $p = .092$, $\mathrm{BF}_{10} = 0.857$, and active engagement, $r_p = -.101$, $p = .253$, $\mathrm{BF}_{10} = 0.552$, the data did not provide evidence for either an association or no association.

Discussion

We examined associations between basic auditory abilities and musical expertise. Auditory abilities were measured as thresholds on psychoacoustic tasks. Musical expertise was defined as music training, scores on objective tests of melody and rhythm discrimination, and self-reports of musical abilities and behaviors other than training, including emotional responses to music, active engagement with music, and perceptual and singing abilities. Although the different measures of musical expertise were intercorrelated, the associations were not particularly strong, which allowed different psychoacoustic thresholds to emerge as relevant depending on how expertise was defined, even after accounting for demographics, cognitive abilities, and personality.

Music Training

The results confirmed our hypothesis that music training (measured as scores on the music training subscale from the Gold-MSI) would be associated with low-level pitch processing. Higher levels of music training predicted better ability to discriminate pure tones varying in frequency (pitch), to compare two short-duration pure tones in terms of which tone was higher (contour speed), and to determine whether the fifth tone (sol) of the major scale was mistuned from equal temperament (scale mistuning). Music training was additionally associated with backward masking, a measure of fine-grained temporal resolution that is considered to index auditory attention (Yoo & Bidelman, 2019). When the four psychoacoustic variables were considered simultaneously, two of the pitch tasks—pitch and scale mistuning—continued to have independent associations with music training.

Previous research has identified that music training, or being a musician, is associated with personality traits, enhanced music perception, speech and language skills, and general cognitive abilities (for review, see Schellenberg & Lima, 2024). Other research has linked musicianship with enhanced perception of pitch (e.g., Bianchi et al., 2016, 2019; Boebinger et al., 2015; Micheyl et al., 2006; Schellenberg & Moreno, 2010). The present study provides large-sample evidence for a link between music training and low-level pitch perception in the general population. How can we interpret this correlation? Perhaps extensive music practice hones listening skills, a possibility consistent with other evidence of learning and exposure. For example, pitch-perception thresholds improve with training (i.e., repeated testing) in the laboratory, even though individual differences in pitch perception are stable (e.g., Christopherson & Humes, 1992; Globerson et al., 2013; Johnson et al., 1987). In some instances, nonmusicians are no longer worse than musicians on a pitch-discrimination task after 4–8 hr of training (Micheyl et al., 2006) or after having their thresholds tested 3 times (Kishon-Rabin et al., 2001).

In one study, individuals with amusia improved after four sessions of training in pure-tone pitch discrimination (Whiteford & Oxenham, 2018). Nevertheless, the magnitude of their deficit compared to controls did not change, and training on a different auditory task unrelated to pitch produced similar improvements for another group of amusic individuals. In other words, such improvements appear to have stemmed from repeated testing. Increases in scores on the MBEA—the standard test for diagnosing amusia—were also similar for both groups, further suggesting a retesting effect. In a different study of amusic participants (Liu et al., 2017), participants were trained over 10 sessions to identify the pitch contour of two piano tones (high–low or low–high). Their performance improved relative to controls on the task, but not on the MBEA. There was also some evidence that only the trained group improved on the pitch-contour subtest of the MBEA, but the timbral changes that occur with pitch changes in complex (e.g., piano) tones (McAdams, 2013) could have provided an additional cue for listeners.

In any event, does the trainability of pitch perception in the laboratory mean that music training would lead to similar improvements? Repeated testing in the laboratory at one frequency (pitch level) can lead to lasting improvements that generalize to other frequencies and to other tests (Whiteford & Oxenham, 2018). By contrast, "music training" refers to a variety of pedagogies that differ markedly but have one thing in common, teaching the student how to perform music. Current evidence that all of them lead to the same advantage in low-level pitch processing is weak. One meta-analysis reached a positive conclusion for auditory processing (Neves et al., 2022), but unexplained heterogeneity across studies was high, the focus was not specifically on pitch, and publication bias could not be ruled out. Moreover, after reviewing the methods and results of the relevant literature, Schellenberg and Lima (2024) found no evidence that music training has a positive causal effect on pitch perception. An illustrative example comes from a study of children who received 6 months of piano, reading, or no training (Nan et al., 2018). Pitch discrimination as measured by electroencephalogram responses was enhanced for the piano group, but there were no group differences in improvements on a behavioral test of pitch discrimination.

The association between pitch perception and music training could also be driven by a third variable—likely genetic—that influences both traits independently. This hypothesis implies the existence of separate genetic mechanisms for pitch perception and the inclination to take music lessons, such that their association is coincidental. Perhaps a simpler explanation is that individuals with better pitch perception are more likely to take music lessons later in life, a perspective consistent with the available correlational and quasi-experimental evidence, and with the lack of evidence that music lessons are the causal agent. Experimental evidence for this hypothesis requires random assignment and manipulation of pitch-perception abilities in the laboratory, followed by long-term examination of music training and playing. A correlational but longitudinal study would be more practicable yet still informative: If low-level pitch-perception abilities tested early in life predict the amount of music lessons taken subsequently, the timeline would rule out an effect of music training on pitch acuity before the training began. We are not proposing that music training is inconsequential for basic listening skills, but rather that a stronger starting point (or genotype) may contribute to the development of musical expertise. The ultimate manifestation (or phenotype) would be the consequence of the starting point combined with exposure to music and other environmental and genetic factors.

In the present study, music training was independent of demographic, cognitive, and personality variables, null fundings that differ from others (e.g., Corrigall et al., 2013). Although our use of the music training subscale could be implicated, the results did not change when we defined training as years of music lessons. Perhaps such associations are relatively unlikely when samples of undergraduates comprise few individuals with extensive training. In our sample, a sizeable proportion (40.1%, 56 of 138) scored between 1 and 2 on the music training subscale (possible range: 1–7), and only 13% (18 of 138) would be classified as musicians (≥6 years of lessons; J. D. Zhang et al., 2020). A previous sample from the same population had a similar proportion (i.e., 16%) of musicians (Schellenberg et al., 2023); our reanalyses of these older data confirmed that neither the music training subscale nor years of music lessons was associated with any extraneous variables.

Melody Discrimination

In previous studies, performance on the melody subtest of the MET correlated positively with music training, other musical behaviors, age, education, cognitive ability, openness, and by the ability to speak a tone language (Correia, Vincenzi, et al., 2022; Swaminathan et al., 2021). Results from the present study add four psychoacoustic thresholds to this list of correlates. Importantly, melody discrimination correlated with pitch perception even after accounting for music training, which confirms that the association was not driven by musically trained participants.

As expected, psychoacoustic correlates of melody discrimination and music training were similar. Two apparent discrepancies were that melody performance was additionally associated with timbre thresholds but not with backward masking, whereas music training was associated with backward masking but not with timbre. Because melody discrimination had a strong association with backward masking before but not after music training was held constant, auditory attention and temporal resolution (as measured by backward masking) may be associated specifically with music training. For timbre, the magnitude of the correlations ($r = -.212$ v. $r = -.226$) differed minimally, yet Bayesian evidence surpassed the threshold for "substantial" only for melody discrimination (see Table 5).

When significant psychoacoustic and control variables were considered simultaneously, older age predicted better melody performance. Older participants may have been more attentive to the task, but we hesitate to speculate further because the vast majority was under 25 years of age. More germane was the finding that melody discrimination continued to have independent associations with contour speed and scale mistuning, both of which measured low-level aspects of relative pitch. Task similarity could be implicated in these associations. Half of the "different" trials in the melody subtest incorporated a contour change (as in contour speed); the other half had the same contour but different intervals between successive tones (as in scale mistuning). The important difference was that the melody subtest relied on relative pitch processing in a quasi-musical context, with longer sequences of piano tones and pitch changes of at least one semitone. In other words, the present findings suggest that individual differences in relative-pitch processing generalize across stimuli that vary in ecological validity and resemblance to music.

Although our results are correlational, it is unlikely that melody-discrimination abilities caused low-level auditory skills. Moreover, the most likely extraneous variables that would influence both melody discrimination and psychoacoustic thresholds were music training and general cognitive ability, yet the association between melody and psychoacoustic thresholds remained evident even when we considered these potential confounding variables. We speculate, therefore, that ability to perceive and remember pitch relations is one of the building blocks upon which melody perception and discrimination are based. Although this hypothesis is difficult to verify experimentally, our findings clarify that music training is not required for the association to emerge. Future research could also explore the possibility of moderating effects. For example, training could have stronger influence on the development of musical expertise among those who have initially higher (or lower) levels of auditory abilities.

Rhythm Discrimination

Although we hypothesized that rhythm discrimination would be associated most strongly with psychoacoustic tasks that directly measured temporal perception (i.e., duration, complex duration, gap detection), associations with rhythm discrimination extended across eight of nine psychoacoustic thresholds (all but gap detection), with none standing out as particularly important. In fact, when psychoacoustic variables were considered jointly along with music training and WM, the variables as a group predicted rhythm perception but no individual variable made an independent contribution. In absolute terms, moreover, rhythm performance had the strongest association with pitch thresholds, which did not measure temporal processing.

Our finding of a general association between rhythm discrimination and auditory abilities has parallels with previous results indicating that rhythm scores—compared to melody scores—have a stronger association with general cognitive ability (Correia, Vincenzi, et al., 2022). The present results suggest that rhythm discrimination is supported by general auditory skills rather than by temporal perception specifically. If auditory skills are relatively stable over the lifespan, as general cognitive ability is (Plomin et al., 1997), rhythm abilities may appear to be more hard-wired and universal compared to melody abilities, and, consequently, more weakly associated with music training (Correia, Vincenzi, et al., 2022; Swaminathan et al., 2021). For example, in one large sample ($N = 523$), the association between rhythm and years of music training was small in magnitude ($r = .183$) and smaller than the association between melody and training (Swaminathan et al., 2021). In another large multinational sample ($N = 608$) that comprised many professional musicians (Brazil, Italy, Portugal, United States/Canada; Correia, Vincenzi, et al., 2022), the correlation between music training subscale and rhythm was larger in magnitude, $r = .296$, but still smaller than the association between music training and melody.

Other evidence for small or null environmental effects on rhythm perception comes from a previous study that examined 7-year-olds who chose to register in either a music-training or sports program (Villanueva et al., 2024). After 4 years, improvement in the ability to drum in time with an isochronous beat (presented at different tempi) was similar between groups. In a longitudinal investigation (Kragness et al., 2021), rhythm and melody abilities were stable over time and relatively unaffected by music training, yet rhythm skills at study onset were the best predictor of melody and rhythm abilities five years later, which suggests that rhythm has a privileged role in the development of musical ability. Because it is implausible that rhythm-discrimination abilities improved a wide range of psychoacoustic thresholds, the present findings motivated a hypothesis that could be tested in future developmental research: Rhythm abilities may stem from general cognitive ability and auditory skills that remain relatively stable over the lifespan.

Informal Musicality

Our final measure of musical expertise—informal musicality—was formed by extracting the common variance (principal component) from four subscales of the Gold-MSI: active engagement, perceptual abilities, singing abilities, and emotions. In contrast to the other expertise measures, we did not expect associations with psychoacoustic thresholds. In line with this view, the strongest associations were with music training and openness, which suggests that individual differences stemmed primarily from experiential and personality variables. After holding music training and personality constant, however, there was still substantial but unexpected evidence that the ability to perceive and discriminate small differences in the duration of pure tones was associated with informal musicality. Exploratory analyses suggested that the association between duration and informal musicality was not driven solely by self-reports on the perceptual abilities subscale. It is therefore difficult to explain why duration was a predictor. Perhaps this uninterpretable finding is also unreplicable.

Broader Implications

Our findings show that music training relates to better auditory thresholds, as do musical abilities even after controlling for training. Although the study design precludes causal inferences, in the absence of clear evidence that music training improves auditory abilities (Schellenberg & Lima, 2024), a preexisting link between musical and auditory abilities offers a simple explanation why musicians perform better on musical and psychoacoustic tests. Auditory skills could provide the bases for musical abilities, which in turn may influence whether an individual is likely to take music lessons (Kragness et al., 2021).

A broader implication concerns the role of musical expertise in associations between musical and nonmusical domains. Musical expertise is often associated with enhancements in other auditory domains, such as prosody (Jansen et al., 2023), vocal emotion recognition (Martins et al., 2021), and speech-in-noise perception (Coffey et al., 2017), and these enhancements are typically assumed to stem from a music-specific mechanism. It is also possible that a common general auditory link, rather than a musical one, plays a role. For speech perception, individual differences in L2 pronunciation are predicted better by general auditory processing than by musical aptitude (Zheng et al., 2022). For music perception, associations between musical ability and vocal-emotion recognition are mediated by lower-level prosodic discrimination skills, which are also associated with pure-tone pitch discrimination (Vigl et al., 2024). Correia, Castro, et al. (2022) found, moreover, that an aggregate measure of music-perception skills—formed from general auditory-processing tasks that included pitch and duration discrimination—mediated associations between music training and vocal-emotion recognition. Thus, auditory skills, rather than musical ability, may represent basic mechanisms that explain why the link between music training and vocal-emotion recognition disappeared when the aggregate ability measure was held constant.

To reiterate, we suggest that low-level listening skills are not merely a consequence of music training, but rather contribute to individual differences in musical ability. Over development, these differences would interact with personality, cognitive variables, and SES to influence who takes music lessons and for how long, and who becomes a working musician.

Limitations

One limitation of the present study is the relatively low number of musicians, which could have under- or overestimated the contribution of music training to the observed associations. In the future, researchers could recruit a larger sample of participants with extensive training (i.e., active musicians with more than six years of formal training; J. D. Zhang et al., 2020). The generalizability of the findings could also be tested by recruiting samples that are more diverse and balanced in terms of men and women.

Psychoacoustic tests measure the precision of different acoustic channels, but they also have disadvantages. For example, performance on these tests recruits domain-general cognitive abilities, such as memory and attention (Litovsky, 2015; Snowling et al., 2018; Y.-X. Zhang et al., 2016). Although we measured and controlled for general cognitive abilities, an alternative approach would be to use implicit measures such as electroencephalogram, which indexes auditory processing during early perceptual stages (e.g., Kujala et al., 2007; H. Sun et al., 2021). The combined use of implicit and explicit measures of auditory acuity could help to delineate the contribution of different levels of auditory processing to musical expertise.

Finally, our results preclude causal inferences because they are correlational. Although longitudinal designs that carefully control the development of musical and auditory abilities over time are necessary to establish causality unequivocally, long-term random assignment to music training or control conditions is impracticable because of differential attrition. One longitudinal but correlational study documented that musical ability improves markedly in childhood, yet individual differences remain stable, with performance at age 8 predicting performance at age 13 (Kragness et al., 2021). Similar stability was evident in a study with over 18,000 children, which found that the best predictor of teachers' ratings of musical ability at age 11 was teacher ratings of artistic ability at age 7 (Müllensiefen et al., 2022). In some instances, however, music training predicts greater improvement in musical abilities over time (Ilari et al., 2016; Müllensiefen et al., 2022). Such inconsistent findings could stem from differences in music training and how musical ability is measured. To the best of our knowledge, a similar exploration of the stability or malleability of general auditory processing, rather than musical ability, has yet to be conducted.

Conclusion

Our findings confirm that basic auditory skills are related to different aspects of musical expertise, including music training and the ability to discriminate melodies or rhythms. They additionally provide evidence that (a) the specific auditory skills that are relevant depend on the aspect of musical expertise being considered and (b) the pitch-processing skills that are associated with music training are also associated with melody discrimination, independently of training. Thus, music training is not necessary for associations between auditory and musical abilities to emerge. Collectively, these findings have implications for the interpretation of associations with music training that are reported in cross-sectional studies. Our results also inform future research on central components of human musicality, and links between nonmusical abilities and musical expertise construed broadly.

References

Alain, C., Zendel, B. R., Hutka, S., & Bidelman, G. M. (2014). Turning down the noise: The benefit of musical training on the aging auditory brain. *Hearing Research*, *308*, 162–173. https://doi.org/10.1016/j.heares.2013.06.008

Albouy, P., Cousineau, M., Caclin, A., Tillmann, B., & Peretz, I. (2016). Impaired encoding of rapid pitch information underlies perception and memory deficits in congenital amusia. *Scientific Reports*, *6*(1), Article 18861. https://doi.org/10.1038/srep18861

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Banai, K., Fisher, S., & Ganot, R. (2012). The effects of context and musical training on auditory temporal-interval discrimination. *Hearing Research*, *284*(1–2), 59–66. https://doi.org/10.1016/j.heares.2011.12.002

Bavin, E. L., Grayden, D. B., Scott, K., & Stefanakis, T. (2010). Testing auditory processing skills and their associations with language in 4–5-year-olds. *Language and Speech*, *53*(1), 31–47. https://doi.org/10.1177/0023830909349151

Benítez-Barrera, C. R., Skoe, E., Huang, J., & Tharpe, A. M. (2022). Evidence for a musician speech-perception-in-noise advantage in school-age children. *Journal of Speech, Language, and Hearing Research*, *65*(10), 3996–4008. https://doi.org/10.1044/2022_JSLHR-22-00134

Bianchi, F., Carney, L. H., Dau, T., & Santurette, S. (2019). Effects of musical training and hearing loss on fundamental frequency discrimination and temporal fine structure processing: Psychophysics and modeling. *Journal of the Association for Research in Otolaryngology*, *20*(3), 263–277. https://doi.org/10.1007/s10162-018-00710-2

Bianchi, F., Santurette, S., Wendt, D., & Dau, T. (2016). Pitch discrimination in musicians and non-musicians: Effects of harmonic resolvability and processing effort. *Journal of the Association for Research in Otolaryngology*, *17*(1), 69–79. https://doi.org/10.1007/s10162-015-0548-2

Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *The Journal of the Acoustical Society of America*, *137*(1), 378–387. https://doi.org/10.1121/1.4904537

Boets, B., Wouters, J., Van Wieringen, A., De Smedt, B., & Ghesquière, P. (2008). Modelling relations between sensory processing, speech perception, orthographic and phonological ability, and literacy achievement. *Brain and Language*, *106*(1), 29–40. https://doi.org/10.1016/j.bandl.2007.12.004

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, *62*(8), 1572–1595. https://doi.org/10.3758/BF03212157

Brito Costa, S., Bem-Haja, P., Moisao, A., Alberty, A., Vicente Castro, F., & De Almeida, H. (2016). Psychometric properties of Portuguese version of Big Five Inventory (BFI). *International Journal of Developmental and Educational Psychology*, *1*(2), 83–94. https://doi.org/10.17060/ijodaep.2015.n2.v1.325

Chierchia, G., Fuhrmann, D., Knoll, L. J., Pi-Sunyer, B. P., Sakhardande, A. L., & Blakemore, S.-J. (2019). The matrix reasoning item bank (MaRs-IB): Novel, open-access abstract reasoning items for adolescents and adults. *Royal Society Open Science*, *6*(10), Article 190232. https://doi.org/10.1098/rsos.190232

Christopherson, L. A., & Humes, L. E. (1992). Some psychometric properties of the Test of Basic Auditory Capabilities (TBAC). *Journal of Speech, Language, and Hearing Research*, *35*(4), 929–935. https://doi.org/10.1044/jshr.3504.929

Coffey, E. B. J., Mogilever, N. B., & Zatorre, R. J. (2017). Speech-in-noise perception in musicians: A review. *Hearing Research*, *352*, 49–69. https://doi.org/10.1016/j.heares.2017.02.006

Conzelmann, K., & Süß, H.-M. (2015). Auditory intelligence: Theoretical considerations and empirical findings. *Learning and Individual Differences*, *40*, 27–40. https://doi.org/10.1016/j.lindif.2015.03.029

Correia, A. I., Castro, S. L., 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

Correia, A. I., Vincenzi, M., Vanzella, P., Pinheiro, A. P., Lima, C. F., & Schellenberg, E. G. (2022). Can musical ability be tested online? *Behavior Research Methods*, *54*(2), 955–969. https://doi.org/10.3758/s13428-021-01641-2

Correia, A. I., Vincenzi, M., Vanzella, P., Pinheiro, A. P., Schellenberg, E. G., & Lima, C. F. (2023). Individual differences in musical ability among adults with no music training. *Quarterly Journal of Experimental Psychology*, *76*(7), 1585–1598. https://doi.org/10.1177/17470218221128557

Corrigall, K. A., Schellenberg, E. G., & Misura, N. M. (2013). Music training, cognition, and personality. *Frontiers in Psychology*, *4*, Article 222. https://doi.org/10.3389/fpsyg.2013.00222

Criscuolo, A., Bonetti, L., Särkämö, T., Kliuchko, M., & Brattico, E. (2019). On the association between musical training, intelligence and executive functions in adulthood. *Frontiers in Psychology*, *10*, Article 1704. https://doi.org/10.3389/fpsyg.2019.01704

Criscuolo, A., Pando-Naude, V., Bonetti, L., Vuust, P., & Brattico, E. (2022). An ALE meta-analytic review of musical expertise. *Scientific Reports*, *12*(1), Article 11726. https://doi.org/10.1038/s41598-022-14959-4

Donai, J. J., & Jennings, M. B. (2016). Gaps-in-noise detection and gender identification from noise-vocoded vowel segments: Comparing performance of active musicians to non-musicians. *The Journal of the Acoustical Society of America*, *139*(5), EL128–EL134. https://doi.org/10.1121/1.4947070

Elliott, L. L. (1971). Backward and forward masking. *International Journal of Audiology*, *10*(2), 65–76. https://doi.org/10.3109/00206097109072544

Endo, N., Ito, T., Watanabe, K., & Nakazawa, K. (2021). Enhancement of loudness discrimination acuity for self-generated sound is independent of musical experience. *PLoS ONE*, *16*(12), Article e0260859. https://doi.org/10.1371/journal.pone.0260859

Engel, A. C., Bueno, C. D., & Sleifer, P. (2019). Treinamento musical e habilidades do processamento auditivo em crianças: Revisão sistemática [Music training and auditory processing skills in children: A systematic review]. *Audiology—Communication Research*, *24*, Article e2116. https://doi.org/10.1590/2317-6431-2018-2116

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fechner, G. T. (1889). *Elemente der psychophysik* (2nd ed.). Breitkopf and Härtel.

Foxton, J. M. (2004). Characterization of deficits in pitch perception underlying ‘tone deafness’. *Brain*, *127*(4), 801–810. https://doi.org/10.1093/brain/awh105

Geringer, J. M. (1993). Loudness estimations of noise, synthesizer, and music excerpts by musicians and nonmusicians. *Psychomusicology: A Journal of Research in Music Cognition*, *12*(1), 22–30. https://doi.org/10.1037/h0094120

Geringer, J. M. (1995). Continuous loudness judgments of dynamics in recorded music excerpts. *Journal of Research in Music Education*, *43*(1), 22–35. https://doi.org/10.2307/3345789

Globerson, E., Amir, N., Golan, O., Kishon-Rabin, L., & Lavidor, M. (2013). Psychoacoustic abilities as predictors of vocal emotion recognition. *Attention, Perception, & Psychophysics*, *75*(8), 1799–1810. https://doi.org/10.3758/s13414-013-0518-x

Gordon, E. (1967). The musical aptitude profile. *Music Educators Journal*, *53*(6), 52–54. https://doi.org/10.2307/3390915

Grassi, M., Meneghetti, C., Toffalini, E., & Borella, E. (2017). Auditory and cognitive performance in elderly musicians and nonmusicians. *PLoS ONE*, *12*(11), Article e0187881. https://doi.org/10.1371/journal.pone.0187881

Graves, J. E., Pralus, A., Fornoni, L., Oxenham, A. J., Tillmann, B., & Caclin, A. (2023). Consonance perception in congenital amusia: Behavioral and brain responses to harmonicity and beating cues. *Journal of Cognitive Neuroscience*, *35*(5), 765–780. https://doi.org/10.1162/jocn_a_01973

Harrison, P. M. C., Collins, T., & Müllensiefen, D. (2017). Applying modern psychometric techniques to melodic discrimination testing: Item response theory, computerised adaptive testing, and automatic item generation. *Scientific Reports*, *7*(1), Article 3618. https://doi.org/10.1038/s41598-017-03586-z

Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron*, *76*(3), 486–502. https://doi.org/10.1016/j.neuron.2012.10.011

Honing, H. (2018). On the biological basis of musicality. *Annals of the New York Academy of Sciences*, *1423*(1), 51–56. https://doi.org/10.1111/nyas.13638

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science*, *15*(5), 356–360. https://doi.org/10.1111/j.0956-7976.2004.00683.x

Ilari, B. S., Keller, P., Damasio, H., & Habibi, A. (2016). The development of musical skills of underprivileged children over the course of 1 year: A study in the context of an El Sistema-Inspired program. *Frontiers in Psychology*, *7*, Article 62. https://doi.org/10.3389/fpsyg.2016.00062

Jakobson, L. S., Cuddy, L. L., & Kilgour, A. R. (2003). Time tagging: A key to musicians' superior memory. *Music Perception*, *20*(3), 307–313. https://doi.org/10.1525/mp.2003.20.3.307

Jansen, N., Harding, E. E., Loerts, H., Başkent, D., & Lowie, W. (2023). The relation between musical abilities and speech prosody perception: A meta-analysis. *Journal of Phonetics*, *101*, Article 101278. https://doi.org/10.1016/j.wocn.2023.101278

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), Article 2. https://doi.org/10.7771/1932-6246.1167

JASP Team. (2024). *JASP* (Version 0.19.1) [Computer software]. https://jasp-stats.org/

Jeffreys, H. (1998). *Theory of probability* (3rd ed.). Oxford University Press.

Jeon, J. Y., & Fricke, F. R. (1997). Duration of perceived and performed sounds. *Psychology of Music*, *25*(1), 70–83. https://doi.org/10.1177/0305735697251006

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). The Guilford Press.

Johnson, D. M., Watson, C. S., & Jensen, J. K. (1987). Individual differences in auditory capabilities. I. *The Journal of the Acoustical Society of America*, *81*(2), 427–438. https://doi.org/10.1121/1.394907

Kidd, G. R., Watson, C. S., & Gygi, B. (2007). Individual differences in auditory abilities. *The Journal of the Acoustical Society of America*, *122*(1), 418–435. https://doi.org/10.1121/1.2743154

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic and Clinical Physiology and Pharmacology*, *12*(2), 125–143. https://doi.org/10.1515/JBCPP.2001.12.2.125

Koelsch, S. (2011). Toward a neural basis of music perception—A review and updated model. *Frontiers in Psychology*, *2*, Article 110. https://doi.org/10.3389/fpsyg.2011.00110

Kragness, H. E., Swaminathan, S., Cirelli, L. K., & Schellenberg, E. G. (2021). Individual differences in musical ability are stable over time in childhood. *Developmental Science*, *24*(4), Article e1308. https://doi.org/10.1111/desc.13081

Kraus, N., & Banai, K. (2007). Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science*, *16*(2), 105–110. https://doi.org/10.1111/j.1467-8721.2007.00485.x

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, *126*(1), 159–179. https://doi.org/10.1037/0033-2909.126.1.159

Kuckelkorn, K. L., De Manzano, Ö., & Ullén, F. (2021). Musical expertise and personality—Differences related to occupational choice and instrument categories. *Personality and Individual Differences*, *173*, Article 110573. https://doi.org/10.1016/j.paid.2020.110573

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology*, *74*(1), 1–19. https://doi.org/10.1016/j.biopsycho.2006.06.001

Larrouy-Maestri, P. (2018). "I know it when I hear it": On listeners' perception of mistuning. *Music & Science*, *1*, 1–17. https://doi.org/10.1177/2059204318784582

Larrouy-Maestri, P., Harrison, P. M. C., & Müllensiefen, D. (2019). The mistuning perception test: A new measurement instrument. *Behavior Research Methods*, *51*(2), 663–675. https://doi.org/10.3758/s13428-019-01225-1

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), Article e52508. https://doi.org/10.1371/journal.pone.0052508

Lee, H., & Müllensiefen, D. (2020). The Timbre Perception Test (TPT): A new interactive musical assessment tool to measure timbre perception ability. *Attention, Perception, & Psychophysics*, *82*(7), 3658–3675. https://doi.org/10.3758/s13414-020-02058-3

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2B), 467–477. https://doi.org/10.1121/1.1912375

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, *4*(5), 187–196. https://doi.org/10.1016/S1364-6613(00)01471-6

Lima, C. F., Anikin, A., Monteiro, A. C., Scott, S. K., & Castro, S. L. (2019). Automaticity in the recognition of nonverbal emotional vocalizations. *Emotion*, *19*(2), 219–233. https://doi.org/10.1037/emo0000429

Lima, C. F., Brancatisano, O., Fancourt, A., Müllensiefen, D., Scott, S. K., Warren, J. D., & Stewart, L. (2016). Impaired socio-emotional processing in a developmental music disorder. *Scientific Reports*, *6*(1), Article 34911. https://doi.org/10.1038/srep34911

Lima, C. F., Correia, A. I., Müllensiefen, D., & Castro, S. L. (2020). Goldsmiths Musical Sophistication Index (Gold-MSI): Portuguese version and associations with socio-demographic factors, personality and music preferences. *Psychology of Music*, *48*(3), 376–388. https://doi.org/10.1177/0305735618801997

Litovsky, R. (2015). Development of the auditory system. In M. J. Aminoff, F. Boller, & D. F. Swabb (Eds.), *Handbook of clinical neurology* (Vol. 129, pp. 55–72). Elsevier. https://doi.org/10.1016/B978-0-444-62630-1.00003-2

Liu, F., Jiang, C., Francart, T., Chan, A. H. D., & Wong, P. C. M. (2017). Perceptual learning of pitch direction in congenital amusia. *Music Perception*, *34*(3), 335–351. https://doi.org/10.1525/mp.2017.34.3.335

Mankel, K., & Bidelman, G. M. (2018). Inherent auditory skills rather than formal music training shape the neural encoding of speech. *Proceedings of the National Academy of Sciences*, *115*(51), 13129–13134. https://doi.org/10.1073/pnas.1811793115

Martin, C. D., Niziolek, C. A., Duñabeitia, J. A., Perez, A., Hernandez, D., Carreiras, M., & Houde, J. F. (2018). Online adaptation to altered auditory feedback is predicted by auditory acuity and not by domain-general executive control resources. *Frontiers in Human Neuroscience*, *12*, Article 91. https://doi.org/10.3389/fnhum.2018.00091

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

McAdams, S. (2013). Musical timbre perception. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 35–67). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-381460-9.00002-X

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81–90. https://doi.org/10.1037/0022-3514.52.1.81

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McPherson, G. E., & Thompson, W. F. (1998). Assessing music performance: Issues and influences. *Research Studies in Music Education*, *10*(1), 12–24. https://doi.org/10.1177/1321103X9801000102

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*(1–2), 36–47. https://doi.org/10.1016/j.heares.2006.05.004

Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). Academic Press.

Mueller, J. L., Friederici, A. D., & Männel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences*, *109*(39), 15953–15958. https://doi.org/10.1073/pnas.1204319109

Müllensiefen, D., Elvers, P., & Frieler, K. (2022). Musical development during adolescence: Perceptual skills, cognitive resources, and musical training. *Annals of the New York Academy of Sciences*, *1518*(1), 264–281. https://doi.org/10.1111/nyas.14911

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), Article e89642. https://doi.org/10.1371/journal.pone.0089642

Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., Gabrieli, J. D. E., & Desimone, R. (2018). Piano training enhances the neural processing of pitch and improves speech perception in mandarin-speaking children. *Proceedings of the National Academy of Sciences*, *115*(28), E6630–E6639. https://doi.org/10.1073/pnas.1808412115

Neves, L., Correia, A. I., Castro, S. L., Martins, D., & Lima, C. F. (2022). Does music training enhance auditory and linguistic processing? A systematic review and meta-analysis of behavioral and brain evidence. *Neuroscience & Biobehavioral Reviews*, *140*, Article 104777. https://doi.org/10.1016/j.neubiorev.2022.104777

Palomar-García, M.-Á., Hernández, M., Olcina, G., Adrián-Ventura, J., Costumero, V., Miró-Padilla, A., Villar-Rodríguez, E., & Ávila, C. (2020). Auditory and frontal anatomic correlates of pitch discrimination in musicians, non-musicians, and children without musical training. *Brain Structure and Function*, *225*(9), 2735–2744. https://doi.org/10.1007/s00429-020-02151-1

Parbery-Clark, A., Strait, D. L., & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, *49*(12), 3338–3345. https://doi.org/10.1016/j.neuropsychologia.2011.08.007

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, *999*(1), 58–75. https://doi.org/10.1196/annals.1284.006

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*(7), 688–691. https://doi.org/10.1038/nn1083

Phillips, D. P. (1999). Auditory gap detection, perceptual channels, and temporal resolution in speech perception. *Journal of the American Academy of Audiology*, *10*(06), 343–354. https://doi.org/10.1055/s-0042-1748505

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 976–986. https://doi.org/10.1037/0096-1523.20.5.976

Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 years: A parent-offspring adoption study. *Psychological Science*, *8*(6), 442–447. https://doi.org/10.1111/j.1467-9280.1997.tb00458.x

Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception*, *24*(1), 37–48. https://doi.org/10.1525/mp.2006.24.1.37

Raven, J. C. (1965). *Advanced progressive matrices sets I and II*. Oxford Psychologists Press.

Saito, K. (2023). How does having a good ear promote successful second language speech acquisition in adulthood? Introducing Auditory Precision Hypothesis-L2. *Language Teaching*, *56*(4), 522–538. https://doi.org/10.1017/S0261444822000453

Schellenberg, E. G. (2006). Long-term positive associations between music lessons and IQ. *Journal of Educational Psychology*, *98*(2), 457–468. https://doi.org/10.1037/0022-0663.98.2.457

Schellenberg, E. G. (2020). Music training, individual differences, and plasticity. In M. S. C. Thomas, D. Mareschal, & I. Dumontheil (Eds.), *Educational neuroscience* (1st ed., pp. 415–441). Routledge. https://doi.org/10.4324/9781003016830-21

Schellenberg, E. G., Correia, A. I., & Lima, C. F. (2023). Is musical expertise associated with self-reported foreign-language ability? *Journal of Experimental Psychology: Human Perception and Performance*, *49*(7), 1083–1089. https://doi.org/10.1037/xhp0001116

Schellenberg, E. G., & Lima, C. F. (2024). Music training and nonmusical abilities. *Annual Review of Psychology*, *75*(1), 87–128. https://doi.org/10.1146/annurev-psych-032323-051354

Schellenberg, E. G., & Moreno, S. (2010). Music lessons, pitch processing, and g. *Psychology of Music*, *38*(2), 209–221. https://doi.org/10.1177/0305735609339473

Schneider, P., Groß, C., Bernhofs, V., Christiner, M., Benner, J., Turker, S., Zeidler, B. M., & Seither-Preisler, A. (2022). Short-term plasticity of neuro-auditory processing induced by musical active listening training. *Annals of the New York Academy of Sciences*, *1517*(1), 176–190. https://doi.org/10.1111/nyas.14899

Seashore, C. E., Lewis, D., & Saetveit, J. G. (1956). *Seashore measures of musical talents*. Psychological Corp.

Shorey, A. E., King, C. J., Whiteford, K. L., & Stilp, C. E. (2024). Musical training is not associated with spectral context effects in instrument sound categorization. *Attention, Perception, & Psychophysics*, *86*(3), 991–1007. https://doi.org/10.3758/s13414-023-02839-6

Slevc, L. R., Davey, N. S., Buschkuehl, M., & Jaeggi, S. M. (2016). Tuning the mind: Exploring the connections between musical ability and executive functions. *Cognition*, *152*, 199–211. https://doi.org/10.1016/j.cognition.2016.03.017

Snowling, M. J., Gooch, D., McArthur, G., & Hulme, C. (2018). Language skills, but not frequency discrimination, predict reading skills in children at risk of dyslexia. *Psychological Science*, *29*(8), 1270–1282. https://doi.org/10.1177/0956797618763090

Soranzo, A., & Grassi, M. (2014). Psychoacoustics: A comprehensive Matlab toolbox for auditory testing. *Frontiers in Psychology*, *5*, Article 712. https://doi.org/10.3389/fpsyg.2014.00712

Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*(1–2), 22–29. https://doi.org/10.1016/j.heares.2009.12.021

Sun, H., Saito, K., & Tierney, A. (2021). A longitudinal investigation of explicit and implicit auditory processing in l2 segmental and suprasegmental acquisition. *Studies in Second Language Acquisition*, *43*(3), 551–573. https://doi.org/10.1017/S0272263120000649

Sun, Y., Lu, X., Ho, H. T., & Thompson, W. F. (2017). Pitch discrimination associated with phonological awareness: Evidence from congenital amusia. *Scientific Reports*, *7*(1), Article 44285. https://doi.org/10.1038/srep44285

Surprenant, A. M., & Watson, C. S. (2001). Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners. *The Journal of the Acoustical Society of America*, *110*(4), 2085–2095. https://doi.org/10.1121/1.1404973

Swaminathan, S., Kragness, H. E., & Schellenberg, E. G. (2021). The Musical Ear Test: Norms and correlates from a large sample of Canadian undergraduates. *Behavior Research Methods*, *53*(5), 2007–2024. https://doi.org/10.3758/s13428-020-01528-8

Swaminathan, S., Schellenberg, E. G., & Khalil, S. (2017). Revisiting the association between music lessons and intelligence: Training effects or music aptitude? *Intelligence*, *62*, 119–124. https://doi.org/10.1016/j.intell.2017.03.005

Swaminathan, S., Schellenberg, E. G., & Venkatesan, K. (2018). Explaining the association between music training and reading in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(6), 992–999. https://doi.org/10.1037/xlm0000493

Tan, Y. T., McPherson, G. E., Peretz, I., Berkovic, S. F., & Wilson, S. J. (2014). The genetic basis of music ability. *Frontiers in Psychology*, *5*, Article 658. https://doi.org/10.3389/fpsyg.2014.00658

Thompson, W. F., Marin, M. M., & Stewart, L. (2012). Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proceedings of the National Academy of Sciences (USA)*, *109*(46), 19027–19032. https://doi.org/10.1073/pnas.1210344109

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

Ullén, F., Hambrick, D. Z., & Mosing, M. A. (2016). Rethinking expertise: A multifactorial gene–environment interaction model of expert performance. *Psychological Bulletin*, *142*(4), 427–446. https://doi.org/10.1037/bul0000033

Ullén, F., Mosing, M. A., Holm, L., Eriksson, H., & Madison, G. (2014). Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test. *Personality and Individual Differences*, *63*, 87–93. https://doi.org/10.1016/j.paid.2014.01.057

Vigl, J., Talamini, F., Strauss, H., & Zentner, M. (2024). Prosodic discrimination skills mediate the association between musical aptitude and vocal emotion recognition ability. *Scientific Reports*, *14*(1), Article 16462. https://doi.org/10.1038/s41598-024-66889-y

Villanueva, J., Ilari, B., & Habibi, A. (2024). Long-term music instruction is partially associated with the development of socioemotional skills. *PLOS ONE*, *19*(7), Article e0307373. https://doi.org/10.1371/journal.pone.0307373

Vincenzi, M., Correia, A. I., Vanzella, P., Pinheiro, A. P., Lima, C. F., & Schellenberg, E. G. (2022). Associations between music training and cognitive abilities: The special case of professional musicians. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. https://doi.org/10.1037/aca0000481

Vuvan, D. T., Nunes-Silva, M., & Peretz, I. (2015). Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia. *Cortex*, *69*, 186–200. https://doi.org/10.1016/j.cortex.2015.05.002

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, *20*(3), 188–196. https://doi.org/10.1016/j.lindif.2010.02.004

Wechsler, D. (2008). *WAIS-III: Manual da Escala de Inteligência de Wechsler para Adultos* [WAIS-III: Wechsler Adult Intelligence Scale] (3rd Edition). Cegoc-Tea [Hogrefe].

Werker, J. F. (2018). Perceptual beginnings to language acquisition. *Applied Psycholinguistics*, *39*(4), 703–728. https://doi.org/10.1017/S0142716418000152

Wesseldijk, L. W., Ullén, F., & Mosing, M. A. (2023). Music and genetics. *Neuroscience & Biobehavioral Reviews*, *152*, Article 105302. https://doi.org/10.1016/j.neubiorev.2023.105302

Whiteford, K. L., & Oxenham, A. J. (2017). Auditory deficits in amusia extend beyond poor pitch perception. *Neuropsychologia*, *99*, 213–224. https://doi.org/10.1016/j.neuropsychologia.2017.03.018

Whiteford, K. L., & Oxenham, A. J. (2018). Learning for pitch and melody discrimination in congenital amusia. *Cortex*, *103*, 164–178. https://doi.org/10.1016/j.cortex.2018.03.012

Won, J. H., Drennan, W. R., Kang, R. S., & Rubinstein, J. T. (2010). Psychoacoustic abilities associated with music perception in cochlear implant users. *Ear & Hearing*, *31*(6), 796–805. https://doi.org/10.1097/AUD.0b013e3181e8b7bd

Yoo, J., & Bidelman, G. M. (2019). Linguistic, perceptual, and cognitive factors underlying musicians' benefits in noise-degraded speech perception. *Hearing Research*, *377*, 189–195. https://doi.org/10.1016/j.heares.2019.03.021

Zhang, J. D., Susino, M., McPherson, G. E., & Schubert, E. (2020). The definition of a musician in music psychology: A literature review and the six-year rule. *Psychology of Music*, *48*(3), 389–409. https://doi.org/10.1177/0305735618804038

Zhang, Y.-X., Moore, D. R., Guiraud, J., Molloy, K., Yan, T.-T., & Amitay, S. (2016). Auditory discrimination learning: Role of working memory. *PLOS ONE*, *11*(1), Article e0147320. https://doi.org/10.1371/journal.pone.0147320

Zheng, C., Saito, K., & Tierney, A. (2022). Successful second language pronunciation learning is linked to domain-general auditory processing rather than music aptitude. *Second Language Research*, *38*(3), 477–497. https://doi.org/10.1177/0267658320978493

Received August 28, 2024
Revision received January 15, 2025
Accepted January 17, 2025 ■