Does Music Training Improve Emotion Recognition?

Longitudinal and Correlational Evidence from Children

Leonor Neves[1, a], Marta Martins[1, a], Ana Isabel Correia[1, a], São Luís Castro[2],

E. Glenn Schellenberg[1, 3], & César F. Lima[1]

[1] Centro de Investigação e Intervenção Social (CIS-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal

[2] Centro de Psicologia da Universidade do Porto (CPUP), Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto (FPCEUP), Porto, Portugal

[3] Department of Psychology, University of Toronto Mississauga, Mississauga, Canada

[a] These authors contributed equally

THIS MANUSCRIPT IS CURRENTLY UNDER REVIEW

Correspondence concerning this article should be addressed to César F. Lima, Instituto Universitário de Lisboa (ISCTE-IUL), Av.ª das Forças Armadas, 1649-026 Lisboa Portugal. E-mail: cesar.lima@iscte-iul.pt

Abstract

Despite widespread claims that music training enhances nonmusical abilities, causal evidence remains inconclusive. Moreover, much research has focused on potential benefits for perception and cognition, but socioemotional skills are often overlooked. Here, we used longitudinal and correlational approaches to ask whether music training improves emotion recognition in voices and faces among school-aged children. We also assessed musical abilities, fine- and gross-motor skills, broader socioemotional functioning, and cognitive abilities including nonverbal reasoning, executive functions, and auditory short-term and working memory. Study 1 ($N = 110$) was a two-year longitudinal intervention conducted in a naturalistic school environment with three groups: music training, basketball training (active control), and no training (passive control). Compared to both control groups, music training improved fine-motor skills and auditory short-term and working memory, but it had no effect on emotion recognition or other cognitive and socioemotional abilities. Study 2 ($N = 192$) compared musically untrained children with those attending a music school. Music training correlated with improved emotion recognition in speech prosody (tone of voice), but this association disappeared after accounting for musical abilities, short-term memory, or socioeconomic status. In contrast, musical abilities correlated with improved emotion recognition in prosody and faces independently of music training and other confounding variables. These findings suggest that while music training may enhance fine-motor skills and auditory memory, observed advantages in emotion recognition likely stem from preexisting musical abilities.

Keywords: music, ability, training, emotion recognition, plasticity

Does Music Training Improve Emotion Recognition?

Longitudinal and Correlational Evidence from Children

The possibility that music training enhances nonmusical abilities has generated much excitement among researchers, the media, and the public. In recent years, many studies have reported that music training enhances auditory abilities (e.g., Dubinsky et al., 2019), speech perception (e.g., Kraus et al., 2014), prosody perception (Moreno et al., 2009), reading and pre-reading skills (e.g., Linnavalli et al., 2018; Moreno et al., 2009), short-term memory (STM; e.g., Zanto et al., 2022), working memory (WM; e.g., Bugos et al., 2022), executive functions (e.g., Frischen et al., 2021; Moreno et al., 2011), and intelligence (e.g., Okely et al., 2022). Putative far-transfer effects from music training have influenced perspectives on behavioral and brain plasticity (Herholz & Zatorre, 2012; Moreno & Bidelman, 2014; Wan & Schlaug, 2010; Schlaug, 2015), domain specificity (e.g., Besson et al., 2011), the biology of music (e.g., Clark et al., 2015), and the use of music in clinical and educational settings (e.g., Jespersen et al., 2022). For example, cognitive benefits of playing music are used to justify interventions for disadvantaged youth (Harmony Project, 2019) and the inclusion of music in school curricula (Barbaroux et al., 2019; Kraus & White-Schwoch, 2020).

But if training in other domains, including chess, working memory, video games, exergames, executive functions, and physical exercise, rarely produces far-transfer effects (Ciria et al., 2023; Gobet & Sala, 2023; Kassai et al., 2019; Sala & Gobet, 2017a; but see, e.g., Pahor et al., 2022), why would music training would be any different? Arguments for nonmusical benefits of music training are also based largely on correlational data, which preclude causal inferences (Schellenberg, 2020). Although associations between music training and enhanced performance on nonmusical tasks abound (e.g., Coffey et al., 2017; Talamini et al., 2017; Schellenberg & Lima, 2024; but see, e.g., Boebinger et al., 2015; Schellenberg et al., 2023), predispositions could play a role. Musically trained and untrained individuals differ in genetic predispositions for music (Wesseldijk et al., 2023), as well as in personality, cognitive abilities, and socioeconomic status (SES; Corrigall et al., 2013), differences that could explain observed associations.

Indeed, evidence that music training causes nonmusical benefits is weak. Some meta-analyses of longitudinal studies conclude that music training produces small cognitive gains (Bigand & Tillmann, 2022; Neves et al., 2022; Román-Caballero et al., 2022), whereas others report null effects after excluding suboptimal studies, such as those without random assignment or active control groups (i.e., comparisons with stimulating but nonmusical activities; Sala & Gobet, 2017b, 2020). Publication bias is another concern (Neves et al., 2022), and studies considered uninformative in critical reviews (Schellenberg & Lima, 2024) contribute to meta-analyses with positive findings, leading to interpretative problems. In short, causal evidence is limited and correlations with music training could reflect predispositions or other confounding factors such as SES.

Because music is linked closely to social and emotional processes (e.g., Clark et al., 2015; Koelsch, 2014; Swaminathan & Schellenberg, 2015), it is surprising that music training has been examined predominantly in relation to cognitive rather than socioemotional benefits. Here, we too examined associations between music training and cognitive abilities, but our primary focus was on a central aspect of socioemotional functioning—the ability to recognize emotions in vocal and facial expressions. Links between music and emotion recognition could stem from overlapping processing mechanisms (Martins et al., 2021; Nussbaum & Schweinberger, 2021; Thompson et al., 2012). Specifically, the same acoustic cues express emotions in music and speech prosody (tone of voice; Coutinho & Dibben, 2013; Curtis & Bharucha, 2010; Juslin & Laukka, 2003), and auditory skills that are relevant for music, such as detecting small differences in pitch, are also important for emotion recognition in voices (e.g., to determine whether someone sounds happy or sad; Globerson et al., 2013). Sensitivity to music could therefore correlate with sensitivity to voices, with higher levels of musical expertise improving vocal-emotion recognition.

Improvements could generalize further to recognizing facial emotions. For example, some individuals with prosopagnosia, a disorder of face recognition, have pitch-discrimination impairments (Barton et al., 2023; Corrow et al., 2019), whereas individuals with congenital amusia, a disorder of music processing, have impaired recognition of vocal *and* facial expressions (Lima et al., 2016). Moreover, music perception and social cognition share neurobiological circuits (van't Hooft et al., 2021), and both music and prosody engage medial prefrontal and anterior cingulate sites (Escoffier et al., 2013; Park et al., 2015) that support supramodal socioemotional processing (Schirmer & Adolphs, 2017; Peelen et al., 2010). Proposals that music plays a central role in social functions (Clark et al., 2015; Koelsch, 2013, 2014) provided additional motivation to ask whether music training improves emotion recognition across auditory and visual modalities.

Adult musicians typically exhibit advantages in vocal-emotion recognition (Martins et al., 2021; Nussbaum & Schweinberger, 2021), including in prosodic stimuli such as sentences with neutral semantics (Lima & Castro, 2011) or pseudowords (Nussbaum et al., 2024). It remains unclear, however, whether the advantage generalizes to faces, and whether it is evident for children, who demonstrate greater neurobehavioral plasticity and are more likely to be taking music lessons at the time of testing (Martins et al., 2021). Furthermore, although training could be the causal agent, individuals with better preexisting musical abilities could also be more likely to take music lessons (Kragness et al., 2021) and to have better auditory skills that facilitate emotion recognition.

In one study of young and middle-aged adults (Lima & Castro, 2011), music training predicted emotion recognition in prosody even after accounting for cognitive abilities. In another study that examined emotion recognition in prosody and nonverbal vocalizations (e.g., laughter, crying), effects of music training and musical abilities were considered separately, with musical abilities assessed with music-perception tests and self-reports (Correia et al., 2022). Musical abilities predicted improved emotion recognition regardless of music training, but music training did not predict emotion recognition when musical abilities were held constant. In fact, musically untrained adults with high levels of musical ability were as good as musicians at recognizing emotions. Thus, preexisting musical abilities may explain why adult musicians show improved emotion recognition.

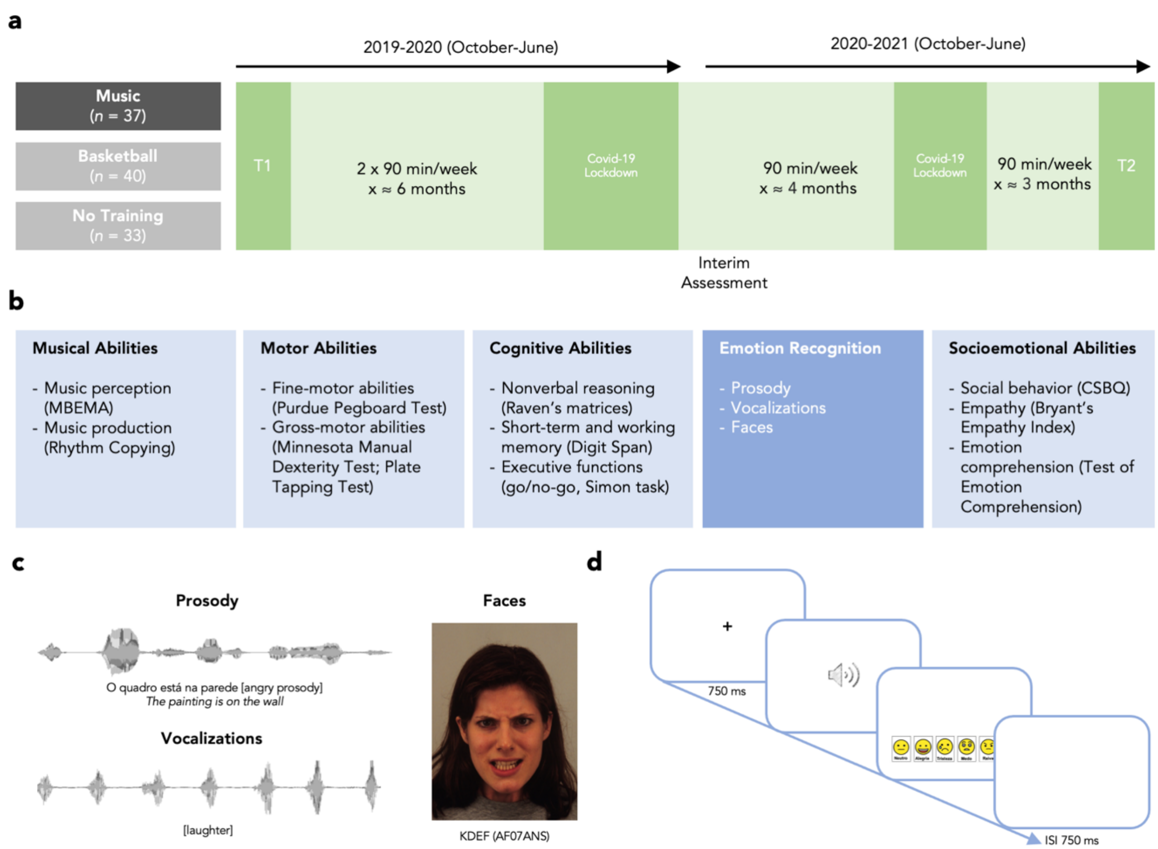
To date, longitudinal studies on music training and emotion recognition are infrequent and inconclusive. Three studies of individuals with cochlear implants reported null effects (Chari et al., 2020; Fuller et al., 2018; Good et al., 2017), but the small samples (7 or fewer participants per group) precluded clear conclusions. In one study of typically developing children, 1 year of group keyboard lessons improved the ability to discriminate fear from anger in prosody (Thompson et al., 2004), but so did drama lessons, which raises the possibility that the effect stems from enjoyable group activities, not from music specifically. Moreover, singing lessons did not have the same effect, the advantage for fear and anger did not generalize to happiness and sadness, and only 43 of 144 children who started the training were tested.

In the present investigation, we used both longitudinal and correlational approaches to ask whether music training improves emotion recognition in childhood. We also sought to isolate training effects from natural musical abilities and other potential confounding variables, particularly general cognitive abilities and SES. Study 1 was longitudinal, conducted with 6- to 8-year-olds. The training programs were integrated into the school curricula to promote ecological validity (Tervaniemi, 2023). Two classes of children were assigned randomly to music training, two to basketball training (active control), and two to no training (passive control). Figure 1 depicts the design and measures. If music training improves emotion recognition, changes over time should be larger for the music compared to the control groups.

Study 2 was correlational, conducted with 8- to 11-year-olds who were either musically trained (attending a music school) or untrained. They completed the emotion-recognition tasks from Study 1 and measures of musical and cognitive abilities. Based on evidence from adults (Correia et al., 2022; Martins et al., 2021; Nussbaum & Schweinberger, 2021), we expected emotion recognition to correlate with music training *and* musical abilities. We also expected musically trained children to have higher SES and better musical and cognitive abilities, which are selecting factors for music lessons outside of the laboratory (Corrigall et al., 2013; Kragness et al., 2021). If musical predispositions rather than training explain advantages in emotion recognition, (1) musical abilities should predict emotion recognition regardless of music training and other factors (SES, cognitive abilities), but (2) music training should no longer predict emotion recognition when musical abilities or other factors are held constant.

Figure 1

Overview of the Design and Measures of Study 1.



Note. (a) Participants were assessed before (T1) and after (T2) 2 school years of music training, basketball training, or no training. The training consisted of two 90-min sessions per week during the first school year, and one 90-min session per week during the second school year. The programs were interrupted twice due to the Covid-19 lockdowns. (b) The measures were the same at T1 and T2, and covered musical abilities, motor abilities, general cognitive abilities, emotion recognition, and broader socioemotional functioning. (c) Emotion recognition was measured in the auditory modality, from speech prosody and purely nonverbal vocalizations, and in the visual modality from facial expressions. (d) Each trial of the emotion-recognition tasks ended when the children responded, and no feedback was provided. MBEMA = Montreal Battery of Evaluation of Musical Abilities; CSBQ = Child Self-Regulation and Behavior Questionnaire; KDEF = Karolinska Directed Emotional Faces.

Study 1: Longitudinal

All children were tested at Time 1 (T1) before training and approximately 2 years later at Time 2 (T2). The main dependent variables measured emotion recognition in prosody, nonverbal vocalizations, and faces. Additional variables measured music perception and production, fine- and gross-motor abilities, general cognition (nonverbal reasoning, STM, WM, executive functions), and broader socioemotional functioning (social behavior, empathy, emotion comprehension). The test battery allowed us to determine whether any positive effects of training on emotion recognition (1) influenced everyday socioemotional functioning (Neves et al., 2021) and (2) were music-specific or the consequence of auditory (e.g., Patel, 2014) or cognitive improvements (Degé, 2021; Schellenberg & Peretz, 2008). The design also allowed us to examine whether musical ability predicted emotion recognition *before* training.

For the other variables we measured, we expected music training to improve fine-motor abilities, based on the results from Martins et al. (2018), who used a similar design and tests. We were uncertain about musical abilities, however, because we used standardized music perception and production tasks—not tailored to the exact skills practiced during the training—and previous evidence that music training improves performance on these tasks is mixed (Kragness et al., 2021; Martins et al., 2023). Although a meta-analysis suggested that music training improves auditory processing in general, publication bias could not be excluded (Neves et al., 2022). We were also uncertain about other potential far-transfer effects (gross-motor skills, general cognitive abilities, broader socioemotional functioning), because evidence from recent reviews and meta-analyses is mixed or inconclusive (e.g., Bigand & Tillmann, 2022; Martins et al., 2021; Neves et al., 2022; Román-Caballero et al., 2022; Sala & Gobet, 2017b, 2020; Schellenberg & Lima, 2024).

Method

Participants

The study was approved by the ethics committee at Iscte University Institute of Lisbon (reference 28/2019) and by the school boards. Written informed consent was obtained from a parent or legal guardian. All children provided oral assent.

We recruited 128 second graders from three public schools in the metropolitan area of Porto (northern Portugal). A background questionnaire asked parents about their child's age and sex, neurological and/or psychiatric diagnoses, extracurricular activities, and parents' years of education. Four children were excluded because of a neurological disorder (n = 2) or below-average cognitive ability (Raven's score < 25^{th} percentile; n = 2). Others (n = 14) transferred to another school. The final sample included 110 children from six different classes, 54 girls and 56 boys, who were 7.01 years old on average at T1 (SD = 0.46). All were native Portuguese speakers. Mother's and father's education were correlated, r = .707, N = 105, p < .001, and averaged to index SES. On average, parents had 11.1 years of education (SD = 3.58).

Two classes of children were assigned randomly to music training (n = 37, M_{age} = 7.08, 20 girls), two to basketball training (n = 40, M_{age} = 6.96, 17 girls), and two received music or basketball training after the study ended (no-training group, n = 33, M_{age} = 6.99, 17 girls). At T1, the three groups did not differ in age, SES, or sex, ps > .5. Sixty of 110 children had a history of involvement in extracurricular activities other than music lessons ($M_{duration}$ = 23.7 months, SD = 20.9), with swimming being the most common (for 33 of 60 children). Neither the proportion of children with a history of extracurricular activities, p = .246, nor their duration of involvement, p = .817, differed across groups.

We recruited as many children as possible, ensuring that we were above 20 per group (Simmons et al., 2011). Post hoc power analysis conducted with G* Power, version 3.1.9.6 (Faul et al., 2009) indicated that with 110 children, we had substantial power (80%) to detect medium-sized differences in improvement between the music and the other two groups (r = .26, alpha = .05, 3 predictors)[1]. Our sample exceeded that of most prior studies of music training (i.e., larger than >85% and >75% of the samples included in the meta-analyses by Neves et al., 2022, and Román-Caballero et al., 2022, respectively).

[1] Effects of training (three levels) were analyzed with two orthogonal contrasts: a comparison of the music and the control groups, and a comparison of the basketball and the no-training groups. Scores at T1, a third predictor, were held constant. Statistical power for the first contrast was calculated with Test family = t test, Statistical test = linear multiple regression: Fixed model, single regression coefficient, and Type of power analysis = Post hoc: Compute achieved power – given α, sample size, and effect size. Input parameters were two tails, effect size f^2 = 0.0725 (Partial R^2 = 0.06764. partial R = 0.26), α err prob = .05, Total sample size = 110, and number of predictors = 3.

Training Programs

Training was integrated into the school curriculum, with sessions for entire classes conducted by teachers specialized in music or basketball (adapted from Martins et al., 2018). Music instruction was based on the Orff approach and covered four domains: musical awareness (e.g., familiarization with musical instruments and genres), basic musical concepts (e.g., rhythm figures, notes, time signatures), rhythm and pitch skills (e.g., ear training and production of rhythmic and melodic patterns), and performance (e.g., improvisation and imitation using the voice or instruments). Playing was based on instruments such as xylophones, metallophones, and drums, but it also included singing and body percussion. Children learned about musical concepts through playing. As they progressed, more complex melodic and rhythmic patterns were incorporated.

Basketball training focused on technical knowledge and skills. It also covered four domains: physical fitness (e.g., strength and flexibility), motor coordination (e.g., eye-hand, upper and lower limbs), concepts and practice (e.g., rules of basketball, ball-handling skills), and tactical planning (e.g., occupation of space, cooperation). Activities of physical fitness and coordination progressed from general exercises to basketball-specific ones, with an emphasis on visuospatial coordination at the individual and team levels.

Both programs lasted for 2 school years and consisted of 90-min sessions that took place twice a week during the first year, and once a week during the second year. Each school year lasted approximately 9 months (October to June), but the programs were implemented for approximately 13 months instead of 18 because of interruptions due to the Covid-19 lockdowns (April to June in the first year, and February to March in the second year). The music and basketball groups completed the same number of sessions. The teachers were hired specifically for this project, and both had experience teaching elementary school children. A questionnaire completed by the children’s primary teacher in October 2020 suggested that the academic and socioemotional impact of Covid-19 was similar across groups (see Supporting Information).

Materials and Tasks

Emotion Recognition. We measured emotion recognition in prosody, nonverbal vocalizations, and faces. Each of three tasks had 60 trials, with 10 different stimuli expressing each of six emotions (happiness, sadness, anger, disgust, fear, and neutrality). The stimuli were drawn from validated corpora (prosody, Castro & Lima, 2010; vocalizations, Lima et al., 2013; faces, Karolinska Directed Emotional Faces database, Goeleven et al., 2008). Prosodic stimuli were short sentences (*M* = 1473 ms, *SD* = 255) with emotionally neutral semantics (e.g., 'O futebol é um desporto', *Football is a sport*), recorded by female speakers to convey emotions through prosody alone. Nonverbal vocalizations were brief vocal sounds (*M* = 966 ms, *SD* = 259) without verbal content, such as screams or laughs, recorded by female and male speakers. Facial expressions were photographs of male and female actors with no eyeglasses, beards, mustaches, earrings, or visible make-up. Because validation data from adults indicates that recognition accuracy is high across tasks (prosody, 78.4%; vocalizations, 82.2%; faces, 83.0%), we deemed them suitable for children.

The tasks required six-alternative forced-choice judgments; children selected the emotion that each stimulus expressed. Each task started with 6 practice trials, followed by two blocks of 30 trials. Stimulus order was randomized separately for each child. On each trial, a fixation cross appeared on the screen for 750 ms, the stimulus was presented, and the trial ended when participants responded. Auditory stimuli were played once. Faces remained visible until the child responded (the interstimulus interval, ISI, was a blank screen lasting 750 ms). Emojis on the screen helped to illustrate response options (Figure 1d), a strategy that proved to be effective in previous studies with children (Correia et al., 2019). No feedback was given except on practice trials. Scores were the percentage of correct responses, averaged across all six categories for each task. We had no hypotheses about specific emotions.

Musical Abilities. Music-perception abilities were tested with three subtests from the Montreal Battery of Evaluation of Musical Abilities (MBEMA; Peretz et al., 2013). On the Melody and Rhythm subtests, children heard two melodies on each of 20 trials, and judged whether the second was identical to the first. On half of the trials of the Melody subtest, one of the notes of the second melody was displaced to produce a scale, contour, or interval violation. On half of the trials of the Rhythm subtest, durations of two adjacent tones were swapped. On the Memory subtest, children heard a single melody on each trial and judged whether they heard it previously in the Melody or Rhythm subtests. The melodies were new on 10 of the 20 trials. For each subtest, scores were the number of correct responses.

Music-production abilities were tested with Moore's (2018) revised version of the rhythm-copying subtest of the Music Aptitude Tests (Overy et al., 2003). On each of 20 trials, children heard a wood-block rhythm and repeated it by pressing a key on a keyboard. The trials became progressively more difficult. Scores were the number of correct responses.

General Cognitive Abilities. Nonverbal reasoning was tested with Raven's Colored Progressive Matrices (Raven et al., 1998). The test had 36 items and the score was the number of items answered correctly. Auditory STM and WM were tested with the Digit Span subtest of the Wechsler Intelligence Scale for Children 3rd Edition (Wechsler, 2003). The experimenter read aloud series of single-digit numbers, which the child had to repeat in the same (forward portion, STM) or reverse (backward portion, WM) order. Total raw scores were used.

Executive functions were tested with go/no-go and Simon tasks. The go/no-go task, adapted from Moreno et al. (2011), had four stimuli (red or yellow butterflies, red or yellow birds). Children were asked to press a key for butterfly stimuli (*go* trials), irrespective of color, but to withhold responding for bird stimuli (*no-go* trials). The task had 100 trials (80% go trials, 20% no-go trials). On each trial, a fixation cross lasted for a variable duration (from 500 to 750 ms), the stimulus was presented until a response was provided or for a maximum of 500 ms, and the ISI was a blank screen lasting 500 ms. We calculated *d'* scores from responses on go (hits) and no-go trials (false alarms).

The Simon task was adapted from Bialystok (2006). A cartoon fish facing right or left was presented on either the left or right side of the computer screen. The task was to press a key indicating the direction that the fish was facing, ignoring its position on the screen. For half of the trials, the fish was facing right (or left) on the right (or left) side of the screen (congruent trials). For the other half, the fish was facing right (or left) on the left (or right) side of the screen (incongruent trials). The task had 80 trials, 20 per condition. Each trial began with a fixation cross lasting 500 ms, the stimulus was presented until a response was provided or for a maximum of 3 s, and the ISI lasted 500 ms. The Simon effect corresponded to the difference in accuracy between congruent and incongruent trials. Larger scores reflected more interference from irrelevant information (i.e., poorer performance).

Motor Abilities. Fine-motor abilities, particularly fingertip dexterity and hand-eye coordination, were tested with the Purdue Pegboard test (Tiffin, 1968). Children had 30 s to place, as fast as possible, small pegs into small holes arranged in columns on a board. Three separate scores represented the number of pins placed properly into columns with the child's preferred hand, nonpreferred hand, and both hands; for the task performed with both hands, the score was the number of pairs of pins placed properly.

Gross-motor abilities were tested with the Minnesota Manual Dexterity Test (Desrosiers et al., 1997). The test was similar to the Purdue test but with large round cylinders (3.7 cm in diameter), larger holes, and it was timed. Scores were the time taken to place 60 cylinders in the holes. Children were tested separately with their preferred and nonpreferred hand. The Plate Tapping test from Eurofit (Adam et al., 1993) provided a second measure of gross-motor abilities. It measured arm movement and coordination. Children used one hand to tap two plates in succession (both 20 cm in diameter), one on their left and the other on their right, while the other hand was placed stationary in a rectangle between plates; at the beginning, the tapping hand was positioned in the contralateral plate. Plate centers were separated by 80 cm. Children performed the task with the preferred and nonpreferred hands at the two different positions (tapping and stationary). Preferred and nonpreferred hands were scored separately as the time (in seconds) taken for 50 taps (25 on both plates). For both the Minnesota and the Plate Tapping test, lower scores indicated better (faster) performance.

For all the motor tasks, the hand order (preferred v. nonpreferred) varied across children. For the Purdue Pegboard test, the task with both hands was always performed at the end.

Socioemotional Functioning. Social behavior was measured with the Child Self-Regulation and Behavior Questionnaire (CSBQ), a 33-item questionnaire completed by the teacher (Howard & Melhuish, 2017). Scale items covered seven domains (sociability, externalizing problems, internalizing problems, prosocial behavior, behavioral self-regulation, cognitive self-regulation, and emotional self-regulation) and each item was rated from 1 (*not true*) to 5 (*certainly true*). Total scores were used in the analyses, with higher scores indicating more adaptive social behavior. Maladaptive subscales (e.g., externalizing behaviors) were reversed-coded.

Empathy was measured with the Index of Empathy for Children (Bryant, 1982), a questionnaire completed by the children. For each of 22 items, children agreed or disagreed about whether they would have an empathic response to another individual's emotional situation (e.g., *I get upset when I see a girl being hurt*; *Kids who have no friends probably don't want any*). Scores were the number of responses that indicated empathy (for the examples, agree and disagree respectively) such that the maximum score was 22.

For the Test of Emotion Comprehension (TEC), children were presented with illustrations accompanied by brief stories read aloud by the experimenter (Pons & Harris, 2000; Rocha et al., 2013). The TEC had nine sections, each of which measured different aspects of emotion comprehension, such as understanding the emotional impact of situations, hidden emotions, or mixed emotions. On each trial, children were asked to choose from a set of four facial expressions, the one that corresponded to the emotion conveyed in the story. Children received a score of 0 or 1 for each section, such that the maximum possible score was 9.

Procedure

T1 and T2 assessments were conducted by three researchers. At both timepoints, each child was tested individually in three sessions in a quiet room of their school. The sessions lasted approximately 2 h. The order of the sessions, and of tasks within sessions, varied across children. In one session, children completed two emotion-recognition tasks (prosody and vocalizations), the MBEMA, and the go/no-go task. A second session comprised the Digit Span, rhythm-copying, and motor tasks. A third session included the emotion-recognition task for faces, Raven's matrices, Simon task, TEC, and the empathy questionnaire.

Stimuli were presented via headphones (Sennheiser HD 201) for the auditory tasks, with the volume adjusted to a comfortable level for each child. The emotion-recognition, go/no-go, and Simon tasks were implemented in SuperLab X6 (Cedrus Corporation, San Pedro, CA), running on Apple Macbook Pro laptops. Responses were collected using a seven-button response pad (Cedrus RB-740).

T1 occurred at the beginning of the 2019-2020 school year, and T2 at the end of 2020-2021. Because of Covid-19, an interim assessment was conducted in October 2020, when teachers completed a questionnaire about the impact of the lockdown, and children completed the musical and motor tests. Changes from T1 observed at interim testing were consistent with those at T2 (Tables S1 and S2).

The children also completed a task of authenticity recognition in laughs and cries (adapted from Pinheiro et al., 2021; Neves et al., 2018), and a magnetic resonance imaging session at T1, during which resting-state and structural scans were acquired. These data will be reported elsewhere.

Data Analysis

We used standard frequentist as well as Bayesian statistics (JASP 0.18.3, default priors; JASP Team, 2024). In each analysis, a Bayes factor (BF_{10}, reported with three-digit accuracy) quantified the evidence that the observed data provided for the null and alternative hypotheses. $BF_{10} > 1$ indicated evidence for the alternative hypothesis, and $BF_{10} < 1$ for the null hypothesis. According to Jeffreys' guidelines (Jarosz & Wiley, 2014), values between 1 and 3 provide weak or anecdotal evidence for the alternative hypothesis, whereas values between 3 and 10 provide substantial evidence, and > 10 and > 100 strong and decisive evidence, respectively. Reciprocal values (i.e., 1-.33, .33-.10, < .10, and < .01) correspond, respectively, to weak/anecdotal, substantial, strong, and decisive evidence for the null hypothesis. To illustrate, $BF_{10} = 15$ indicates that the data are 15 times more likely under the alternative than the null hypothesis (i.e., strong evidence for an association/effect), whereas a $BF_{10} = .067$ (1/15) indicates that the data are 15 times more likely under the null hypothesis (i.e., strong evidence for no association/effect). Instead of correcting p values for multiple comparisons, we considered positive results to be reliable only when both $p < .05$ and $BF_{10} > 3$.

Transparency and Openness

The data used for the analyses are available at https://osf.io/u96fa/?view_only=d3d751ebbc2c4d3a84734b0916536b17 (Lima et al., 2024). We report how we determined sample size, all data exclusions, all manipulations, and all measures in the study. The study's design and analysis were not pre-registered.

Results and Discussion

We first examined whether groups differed at T1. As shown in Table 1, one-way between-subject Analyses of Variance (ANOVAs) revealed a group effect for 2 of 21 tests, but BF_{10} was ≤ 1.57 in both instances. We also confirmed that scores improved from T1 to T2 for the whole sample, *p*s < .001 (Table S3), with BF_{10} indicating strong evidence for the Simon Effect (BF_{10} = 15.3) and decisive evidence for all other measures (BF_{10} > 100). In each instance, correlations between scores at T1 and at T2 were significant and strong with $r \geq .323$, $p < .001$, $BF_{10} \geq 41.6$, except for three measures (Simon Effect, r = .233, Go/no-go, r = .192, empathy, r = .233), for which the evidence was weak, $p < .05$, $BF_{10} \leq 2.30$. In general, however, there were no initial group differences, performance improved over time, and individual differences were relatively stable.

Data Reduction

We used principal components analysis to form latent variables from measures that were related conceptually and empirically. This manipulation increased construct validity and reduced the likelihood of measurement-specific or Type I errors. For *musical ability*, a latent variable was extracted from scores on the three MBEMA subtests and the rhythm-copying test. At T1, it accounted for 55.5% of the variance and correlated highly with each original variable (*r*s ≥ .545). At T2, it accounted for 60.0% of the variance and loadings were at least .739. Because latent variables at T1 and T2 both had $M = 0$ and $SD = 1$, main effects of time (reported above) were precluded in subsequent analyses, but we retained the ability to detect interactions between group and time. For example, if the music group improved more than the other groups, more children in the music group would go up in rank order from T1 to T2, such that average improvement (T2 - T1) would be greater.

Table 1

Descriptive Statistics at T1 for the Three Groups of Children in Study 1, with p-*Values and Bayes Factors (BF_{10}) from Group Comparisons.*

	Music		Basketball		No Training			
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*p*-value	BF_{10}
Prosody	60.00	13.93	61.67	16.65	52.73	16.36	.044	1.14
Vocalizations	84.95	7.49	82.67	8.93	82.17	6.06	.263	.256
Faces	82.30	7.41	77.67	9.28	79.55	9.50	.072	.773
MBEMA Melody	12.49	2.41	12.55	2.25	12.00	1.85	.523	.145
MBEMA Rhythm	13.73	3.11	13.37	2.63	13.09	2.44	.622	.125
MBEMA Memory	13.97	3.13	13.77	2.75	13.67	2.39	.896	.093
Rhythm Copying	5.81	3.06	5.05	3.27	5.24	3.30	.566	.136
Raven’s Matrices	22.73	4.45	23.28	4.91	23.03	3.95	.868	.096
Digit Span Forward	5.65	1.18	5.55	1.04	5.76	1.15	.734	.110
Digit Span Backward	3.03	0.96	2.83	0.96	3.00	0.94	.601	.130
Simon Effect	15.54	12.94	15.88	18.10	12.42	16.64	.615	.126
Go/No-go *d’*	1.64	0.76	1.85	0.60	1.89	0.53	.202	.319
Purdue PH	11.68	1.83	11.28	1.78	11.45	1.50	.594	.131
Purdue NH	10.62	1.57	10.23	1.56	10.12	1.58	.364	.195
Purdue BH	8.49	1.79	7.88	1.62	7.88	1.52	.192	.334
Minnesota PH	101.05	12.64	107.18	15.09	106.61	13.37	.112	.526
Minnesota NH	108.76	15.36	114.23	19.70	112.48	11.82	.327	.215
Plate Tapping PH	27.78	5.03	27.45	5.18	26.45	4.08	.494	.151
Plate Tapping NH	30.14	5.37	30.28	6.23	28.55	4.54	.347	.202
CSBQ	3.73	0.56	3.99	0.49	3.67	0.60	.031	1.57
Empathy	11.81	2.94	11.53	2.73	12.39	3.53	.477	.156
TEC	6.19	1.22	6.25	1.37	6.36	1.19	.846	.098

Note. MBEMA = Montreal Battery of Evaluation of Musical Abilities; Purdue = Purdue Pegboard test; Minnesota = Minnesota Manual Dexterity Test; PH = preferred hand; NH = nonpreferred hand; BH = both hands; CSBQ = Child Self-Regulation and Behavior Questionnaire; TEC = Test of Emotion Comprehension.

For *fine-motor skills*, a latent variable was formed from three Purdue subtests (preferred hand, nonpreferred hand, both hands). It accounted for 69.1% and 78.6% of the variance in the original data at T1 and T2, respectively (loadings ≥ .788 and .862). For *gross-motor skills*, a latent variable was extracted from the preferred and nonpreferred subtests of the Minnesota and plate-tapping tests. At T1 and T2, respectively, it accounted for 71.1% and 68.7% of the variance and loadings were greater than .812 and .805. Scores were inverted at both timepoints so that higher scores indicated better performance. Finally, for *executive functions*, the latent variable was formed from the Simon and go/no-go tasks, accounting for 54.5% and 56.5% of the variance at T1 and T2, respectively (loadings = .738 and .751).

Improvement Over Time

We calculated improvement scores (T2 – T1) for each variable. Using planned orthogonal contrasts to maximize power, we tested whether improvements differed (1) between the music and the two control groups, and (2) between the basketball and no-training groups. T1 scores were held constant throughout, which further increased power by removing variance due to regression to the mean. Results are summarized in Table 2, Figure2, and Figure S1. Although our focus was on group differences, effect sizes were estimated with *r* (correlation coefficient) to be consistent with Study 2, which is correlational.

The music group did not improve more than controls on the emotion-recognition tasks. In fact, BF_{10} values favored the null hypothesis, with the evidence being substantial for vocalizations and weak for prosody and faces. For the prosody and faces tasks, *p*-values suggested larger improvements for the no-training over basketball groups, but Bayesian evidence was weak. These null findings did not change when we analyzed data from individual trials with mixed-effects models (Supporting Information), or when specific emotions were considered separately (Table S4). In fact, for the 18 comparisons (3 tasks X 6 emotions) of the music and control groups, all Bayes factors were < 1.

Table 2

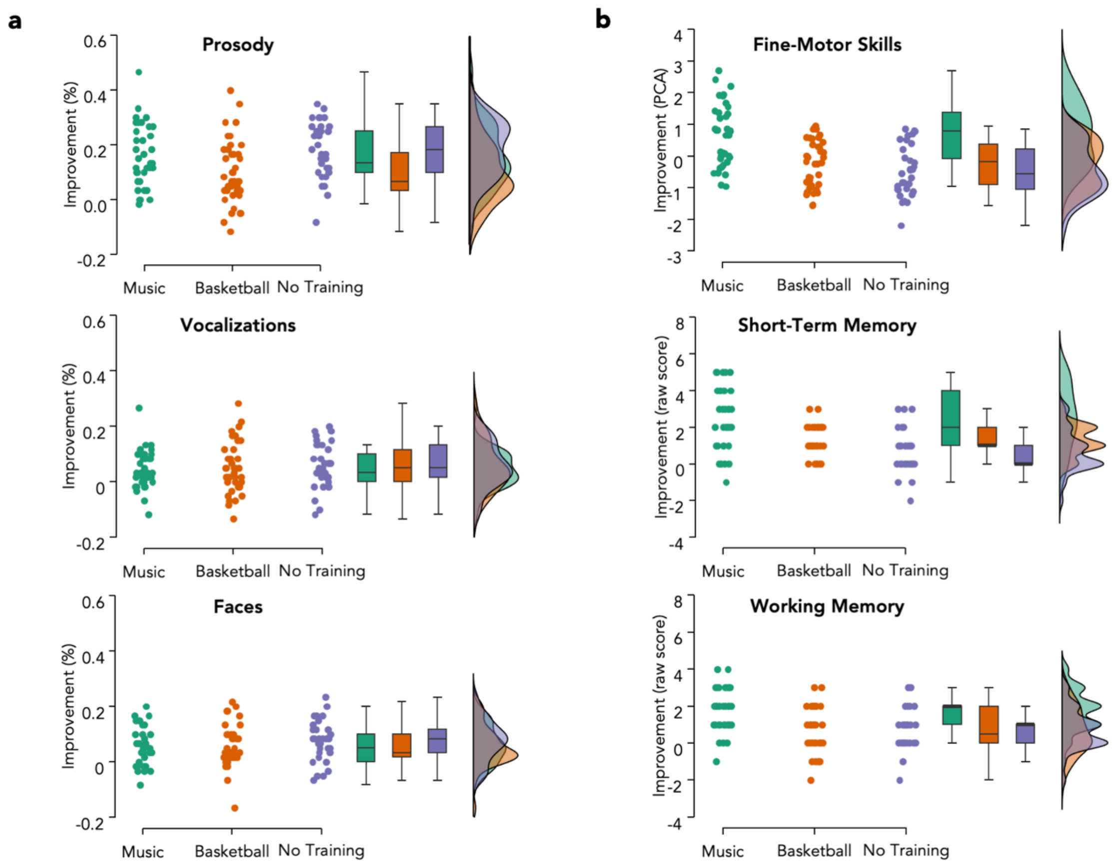
Comparisons of Improvement Over Time in Study 1 Between the Music and the Two Control Groups, and the Basketball and No-Training Groups. T1 Scores Were Held Constant. r *indicates effect size.*

	Music v. Controls			Basketball v. No Training		
Variable	*p*	*r*	BF_{10}	*p*	*r*	BF_{10}
Prosody	.057	.183	.828	.048	-.190	.953
Vocalizations	.539	.060	.191	.902	-.012	.160
Faces	.208	.123	.369	.041	-.198	1.30
Musical Ability	.041	.197	1.73	.078	-.170	1.05
Nonverbal Reasoning	.111	.154	.827	.819	.022	.257
Short-Term Memory	< .001	.498	> 100	.058	.183	1.20
Working Memory	< .001	.450	> 100	.555	-.057	.239
Executive Functions	.600	.051	.210	.729	-.034	.194
Fine-Motor Skills	.667	< .001	> 100	.143	.140	.375
Gross-Motor Skills	.306	.001	28.7	.312	.001	34.7
Social Behavior	.267	.108	.538	.438	-.075	.403
Empathy	.363	-.088	.269	.686	-.039	.195
Emotion Comprehension	.983	.002	.164	.055	.186	.980

Note. Negative values for *r* indicate that the music group had a *lower* mean improvement compared to the other children (comparison Music v. Controls), or that the basketball group had a lower mean improvement than the no-training group (comparison Basketball v. No Training).

Figure 2

Improvement Over Time for the Three Groups of Children in Study 1 in (a) Emotion Recognition and (b) Fine-Motor Skills, Short-Term Memory, and Working Memory.



Note. Improvement corresponds to difference scores (T2 – T1). PCA = latent variable extracted using principal components analysis

For STM, WM, and fine-motor skills, however, there was decisive evidence that the music group improved more than controls, and that the two control groups did not differ. For gross-motor skills, there was again strong evidence that the music group improved more than controls, but the basketball and no-training groups also differed. Follow-up pairwise (Tukey) comparisons revealed that both the music group, $p < .001$, and the basketball group, $p = .003$, improved more than the no-training group, but the music and basketball groups improved similarly, $p = .813$. There were no group differences in improvements for musical ability, nonverbal reasoning, executive functions, social behavior, and empathy. Although the p-value suggested a positive effect of music training on musical ability, Bayesian evidence was weak.

In contrast to the null results for music training, musical *ability* had strong positive correlations at T1 with emotion recognition in prosody, $r = .443$, $p < .001$, $BF_{10} > 100$, and in faces, $r = .313$, $p < .001$, $BF_{10} = 28.2$, but not in vocalizations, $r = .163$, $p = .089$, $BF_{10} = .498$. At T2, musical ability predicted emotion recognition across modalities: prosody, $r = .480$, $p < .001$, $BF_{10} > 100$; vocalizations, $r = .289$, $p = .002$, $BF_{10} = 12.3$; faces, $r = .388$, $p < .001$, $BF_{10} > 100$. Individual differences in our aggregate measure of musical ability (i.e., the latent variable) were also stable from T1 to T2, $r = .684$, $p < .001$, $BF_{10} > 100$.

Tests of Moderation

Although music training did not improve emotion recognition at the group level, effects on STM and WM motivated us to ask whether children with larger (or smaller) improvements in auditory memory had also larger improvements in emotion recognition. A two-way interaction between music (v. controls) and increases in STM was added to a model that included both component main effects and the comparison between the basketball and no-training groups. T1 scores for the outcome variable and STM were also included. Increases in STM did not interact with music training for improvements in emotion recognition in prosody, $p = .096$, $BF_{10} = .769$, vocalizations, $p = .171$, $BF_{10} = .564$, or faces, $p = .322$, $BF_{10} = .396$. A similar analysis for WM also indicated that increases in WM did not interact with music training in predicting improvements in prosody, $p = .080$, $BF_{10} = .868$, vocalizations, $p = .068$, $BF_{10} = .914$, or faces, $p = .747$, $BF_{10} = .263$. For both STM and WM, evidence for the null hypothesis was substantial only for faces.

Finally, we asked whether music training improved emotion recognition for children who began the study with lower emotion-recognition abilities. In an earlier study, improvement in social skills among children who took music lessons for a year was evident only among those who had below-average social skills at the beginning of the year (Schellenberg et al., 2015). Accordingly, we included an interaction term between music training and T1 scores in our tests of improvements in emotion recognition. The interaction was not significant (prosody: p = .191, BF_{10} = .377; vocalizations: p = .413, BF_{10} = .258; faces: p = .108, BF_{10} = .681), although evidence for the null hypothesis was substantial only for vocalizations.

In sum, music training improved fine-motor skills, STM, and WM, but it had no causal effect on our main dependent variables—emotion recognition in voices and faces—or on other cognitive and socioemotional abilities. For gross-motor skills, improvements were similar for the music and basketball groups, but greater than those for the no-training group. At the beginning and end of the study, musical ability was correlated positively with emotion recognition.

Study 2: Correlational

In the second study, we compared children with and without music training on tests of emotion recognition in voices and faces. Thus, training varied naturally rather than being experimentally manipulated. If music training improves emotion recognition, a real-world positive association should be evident even after accounting for confounding factors such as SES. The null results from Study 1 made this seem unlikely, although aspects of the training (e.g., pedagogy, school, classroom dynamics) could have played a role. Alternatively, if preexisting musical abilities are an important factor, they should predict emotion recognition independently of training and other confounding variables.

Method

Participants

Participants were 192 children, 99 girls and 93 boys, who were 7.58 years old on average (SD = 1.04). Children with no music training (n = 156) included 110 who were tested in Study 1, and 46 newly recruited from three other schools. Children with music training (n = 36) were 6- to 11-year-olds recruited from a music school, who had 27.7 months of lessons on average (SD = 19.0). Older children were included in the trained group to maximize sample size and duration of training. Musically trained children were significantly older (M = 9.00 years, SD = 1.35) than the untrained ones (M = 7.25 years, SD = 0.58), $p < .001$, but the sex balance did not differ, p = .203. Age was held constant throughout the statistical analyses. Musically trained children were also more likely to have participated in other extracurricular activities, $\chi^2(1, N = 192) = 7.87$, $p = .005$, $\phi = .202$, with 80.6% and 55.1% of musically trained (29/36) and untrained (86/156) children having an average of 40.2 ($SD = 23.8$) and 23.1 ($SD = 21.0$) months of activity, respectively. Because of positive skew, music training and involvement in other extracurricular activities (hereafter *other activities*) were coded as dummy variables (1 = yes, 0 = no).

As in Study 1, mothers' and fathers' years of education were correlated, $r = .858$, $p < .001$, and averaged to index SES ($M = 11.8$ years, $SD = 3.65$). SES data were missing for three children and replaced with the group mean. As in Corrigall et al. (2013), musically trained children came from higher-SES families ($M = 15.2$, $SD = 1.93$) than the untrained children ($M = 11.0$, $SD = 3.50$), $p < .001$.

We sought to recruit as many children as possible. Post hoc power analysis confirmed that we had sufficient power (80%) to detect partial correlations of .2 or greater between music training and emotion recognition, with four other variables held constant.[2]

Procedure

For the children from Study 1, we considered T1 scores. The 82 newly recruited children completed the three emotion-recognition tasks and the tests of musical ability (MBEMA), nonverbal reasoning (Raven's), STM (Digit Span Forward), and WM (Digit Span Backward). We omitted tests of rhythm copying, executive functions, motor ability, and social behavior to shorten the testing session. A Raven's score was missing for one child.

Results and Discussion

We extracted a principal component for musical ability from the three MBEMA subtests. The latent variable accounted for 67.6% of the variance, and it correlated highly with the original variables (rs $\geq .784$).

Descriptive statistics and results from comparisons between musically trained and untrained children on the emotion-recognition tasks are provided in Table 3 and Figure 3 (see Table S5 for simple

---

[2] G*Power settings: Test family = *t* test; Statistical test = linear multiple regression: Fixed model, single regression coefficient; and Type of power analysis = Post hoc: Compute achieved power – given α, sample size, and effect size. Input parameters: two tails, effect size $f^2 = 0.0417$ (Partial $R^2 = 0.04$, partial $R = 0.2$), α err prob = .05, Total sample size = 192, and number of predictors = 5.

associations among all variables). The data provided strong evidence that musically trained children performed better on the prosody task (Figure 3a). For vocalizations and faces, however, the data provided substantial support for the null hypotheses—there was no association between music training and emotion recognition. When we controlled additionally for SES, other activities, musical ability, nonverbal reasoning, STM, and WM, music training no longer predicted the ability to recognize emotions from prosody, $r = .045$, $p = .544$, $BF_{10} = .321$ (Figure 3b). In fact, the Bayes factor indicated substantial support for the null hypothesis. Similar null results were evident when we examined individual emotions separately (Table S6). By contrast, after accounting for music training and all other variables, musical *ability* predicted children's ability to recognize emotions from prosody $r = .299$, $p < .001$, $BF_{10} > 100$, and from faces, $r = .211$, $p = .004$, $BF_{10} = 13.4$ (Figure 3c, d), but not from vocalizations, $r = .117$, $p = .113$, $BF_{10} = 1.18$. For prosody, associations with musical ability were evident for five of six emotions (all but fear, Table S6). For faces, the association with musical ability was evident only for fear (Table S6).

We next examined which variable or variables eliminated the association between music training and prosody. For each predictor variable, we ran a separate model predicting performance on the prosody task as a function of music training, the predictor, and age. The advantage for musically trained children disappeared when SES, $r = .126$, $p = .082$, $BF_{10} = .918$; musical ability, $r = .067$, $p = .360$, $BF_{10} = .259$; or STM, $r = .116$, $p = .112$, $BF_{10} = .699$, was held constant, although it remained evident after controlling for nonverbal reasoning, $r = .238$, $p < .001$, $BF_{10} = 36.4$; or other activities, $r = .179$, $p = .014$, $BF_{10} = 4.01$. When WM was held constant, $r = .167$, $p = .021$, $BF_{10} = 2.78$, the partial association with music training was weak. Note that after accounting for individual differences in musical ability alone, the observed data provided substantial support for the null hypothesis, specifically that there was no partial association between music training and prosody, presumably because of shared variance between musical ability and music training, $r = .239$, $p = .001$, $BF_{10} = 34.1$ (age, SES, other activities, nonverbal ability, STM, and WM held constant).

Table 3

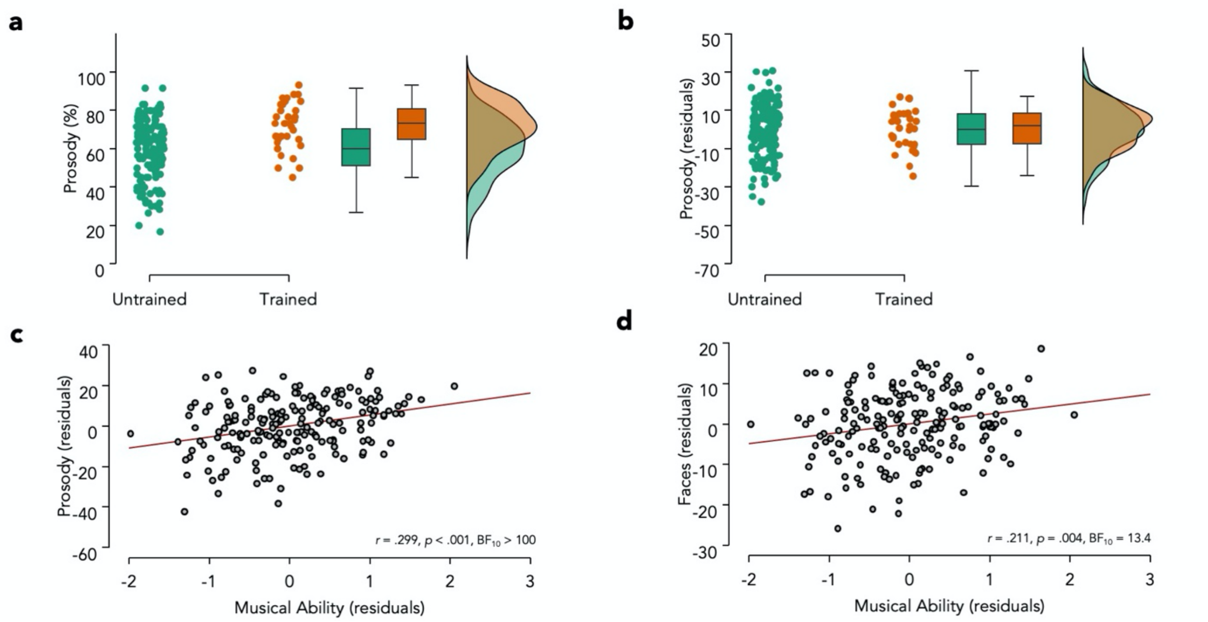
Descriptive Statistics for Musically Trained and Untrained Children in Study 2 (Unadjusted M*s and* SD*s) and* p*-Values, Effect Sizes (*r*), and Bayes Factors (BF_{10}) for Group Comparisons (Age Held Constant).*

Variable	Trained		Untrained				
	M	*SD*	*M*	*SD*	*p*	*r*	BF_{10}
Prosody	71.90	11.94	59.50	14.95	.003	.212	11.9
Vocalizations	85.93	7.15	83.40	7.79	.912	.008	.218
Faces	83.47	8.02	79.46	8.93	.429	.058	.284
Musical Ability	1.14	0.73	-0.26	0.86	< .001	.353	>100
Nonverbal Reasoning	21.17	3.24	22.68	4.69	.064	-.134	1.14
Short-Term Memory	7.33	1.43	5.68	1.15	< .001	.379	>100
Working Memory	4.56	1.61	2.96	1.05	< .001	.294	>100

Note. Negative values for *r* indicate that the trained group had a *lower* mean compared to the untrained group.

Figure 3

Performance of Musically Untrained and Trained Children on the Prosody Task in Study 2 (a) Before and (b) After Adjusting for Confounding Variables. Associations Between Musical Ability and Emotion Recognition in (c) Prosody and (d) Faces After Adjusting for Confounding Variables.



Note. In (b), emotion recognition was adjusted for age, SES, other activities, musical ability, nonverbal reasoning, STM, and WM. In (c) and (d), musical ability and emotion recognition were adjusted for age, music training, SES, other activities, nonverbal reasoning, STM, and WM.

As shown in Table 3, groups comparisons on our cognitive measures revealed decisive evidence for superior STM and WM among trained compared to untrained children, although the two groups performed similarly for nonverbal reasoning. The partial association between music training and STM remained evident when we controlled additionally for SES, other activities, and musical ability, $r = .233$, $p = .001$, $BF_{10} = 30.2$, as it did for WM, $r = .174$, $p = .017$, $BF_{10} = 3.35$, with the data providing very strong and substantial evidence, respectively.

General Discussion

Does music training enhance children's ability to recognize emotions? In Study 1, we found no evidence that 2 years of music training leads to greater improvements in emotion recognition compared to basketball or no training. In Study 2, although music training correlated positively with improved recognition of emotional prosody, the association disappeared after adjusting for confounding variables. By contrast, musical *ability* predicted emotion recognition in both prosody and faces independently of music training and other variables.

Associations between music training and emotion recognition in voices have been reported for adults (Martins et al., 2021; Nussbaum & Schweinberger, 2021). Our results extended these findings to children, but they also raised doubts about the assumption that training is causally implicated. Although Study 1 was the first well-powered longitudinal study on this topic, with a larger sample and longer duration of training than typical music-training studies (e.g., Neves et al., 2022; Román-Caballero et al., 2022; Sala & Gobet, 2020), null effects were observed across modalities for emotion recognition. Evidence for the null hypothesis was not strong, but it was even weaker for the alternative hypothesis, including when we focused specifically on participants with low emotion-recognition at baseline or on those with larger improvements in STM and WM. Moreover, after adjusting for confounding variables in Study 2, the data provided substantial evidence for the null hypothesis: there is no direct association between music training and emotion recognition in prosody.

Our null findings are commensurate with outcomes from earlier small-scale longitudinal studies on emotion recognition, which tested individuals with cochlear implants (Chari et al., 2020; Fuller et al., 2018; Good et al., 2017). In one study of typically developing young adults, however, Mualem and Lavidor (2015) found larger improvements in recognizing emotions from prosody after 4 weeks of music compared to art training. The small sample (n = 12 per group) and a music-training program that specifically targeted emotion recognition raise doubts about whether this finding is reliable and/or generalizable.

In Thompson et al.'s (2004) study, children who had 1 year of keyboard lessons were better at discriminating anger from fear in prosody than children who had singing lessons or no lessons. The sample was again small ($n \leq 13$ per group), and it is unclear why the advantage did not extend to comparisons between happiness and sadness, or why keyboard lessons showed an effect but singing lessons did not. Additionally, the effect was not specific to music—drama lessons showed similar benefits—and sampling bias may have played a role (as noted above). In any event, if links between music training and emotion recognition are epiphenomenal, as our findings indicate, underpowered studies would be particularly likely to yield inconsistent or uninterpretable results.

In Study 1, the large sample and relatively long training make it unlikely that the null results were Type II errors. We also maximized statistical power by using orthogonal contrasts to compare groups of children. Furthermore, although the training was interrupted twice because of Covid-19, we confirmed that there were no major differences between groups in how they were affected, and the interruptions did not seem to impede developmental gains. Over time, children improved consistently across all measures, with group-specific improvements on several of them.

In Study 2, the association between music training and recognizing emotions from prosody disappeared after adjusting for SES, musical ability, or STM, but it remained evident after adjusting for nonverbal reasoning, WM, or other activities. These findings from children replicate those observed previously with adults, when an association between music training and emotion recognition in prosody persisted after accounting for nonverbal reasoning (Lima & Castro, 2011), but disappeared after adjusting for STM and musical abilities (Correia et al., 2022). Because measures of STM and music perception both rely on auditory cognition, individual differences in auditory skills may explain musicians' advantages for recognizing prosodic emotions. This indirect association could also help to explain previous failures to find a link between music training and emotion recognition (e.g., Park et al., 2015; Trimmer & Cuddy, 2008).

In contrast to music training, musical ability was a strong predictor of emotion recognition. Children with higher levels of musical ability were better at recognizing emotions in both prosody and faces, with a particularly robust association observed for prosody, whether the sample included children with (Study 2) or without (Study 1) music training. At the opposite end of the spectrum, individuals with congenital amusia, who have atypically low musical abilities, also have difficulty recognizing emotions from prosody and faces (Lima et al., 2016; Thompson et al., 2012). The positive association between musicality and emotion recognition is consistent with the proposal that shared processing mechanisms across domains explain the close link between music and social cognition (e.g., Clark et al., 2015). These shared mechanisms may extend beyond the auditory domain, as indicated by neuroimaging findings from adults (Escoffier et al., 2013; Park et al., 2015; van't Hooft et al., 2021). At present, however, it remains uncertain whether the link between musical ability and emotion recognition is general or auditory-specific, particularly in light of evidence that adults' musical abilities are associated with recognizing emotions from voices but not from faces (Correia et al., 2022). Moreover, the association with faces observed here was significant only for fear. Perhaps age plays a moderating role.

Considered jointly, our longitudinal and correlational findings suggest that predispositions explain observed associations between music training and emotion recognition. Indeed, musical abilities have an established genetic component (Wesseldijk et al., 2023), and children who are naturally more musical are more likely to take music lessons (Kragness et al., 2021). Because such children also perform better on emotion-recognition tasks, artifactual correlations between music training and emotion recognition would be likely to emerge. In principle, music training could enhance this association, as Mankel and Bidelman (2018) proposed for speech-in-noise perception. Nevertheless, the absence of evidence for causality in Study 1 and the lack of a direct link with training in Study 2 are inconsistent with this suggestion. Indeed, a recent meta-analysis reported that musical abilities are better than music training at predicting prosody perception, whether prosody's function is linguistic or emotional (Jansen et al., 2023).

Concerning music-training effects on other measures, we found only weak causal evidence (Study 1) for improvements on our objective tests of musical ability, which seems counterintuitive. Nevertheless, the tasks were not designed to assess the skills trained directly by the music intervention, such that task improvement would represent a case of transfer. In a related study, the magnitude of improvement in rhythm abilities was greater for children with 6 months of music training compared to controls, but the two groups did not differ at post-test (Martins et al., 2023). Moreover, in a 5-year longitudinal study of children who self-selected into training, Kragness et al. (2021) found no evidence that training caused improvement on same/different music-discrimination tasks. If self-selected training, when motivation is substantial, does not improve musical ability, such an effect would be even more unlikely when motivation is lower because of random assignment. In short, the strong association between music training and musical ability observed in Study 2 is consistent with the proposal that musical ability predicts training.

For gross-motor skills, improvements in Study 1 were higher for the music compared to the no-training group, yet the music and basketball groups were similarly advantaged, presumably a consequence of participating in organized physical activity. A music-specific advantage was evident, however, for fine-motor skills. This effect replicated results from an earlier study, which showed a similar fine-motor advantage for music training over basketball or no training (Martins et al., 2018). The Orff-based teaching method likely played a role in the effect reported here and earlier (Martins et al., 2018), because of its emphasis on motor control and performance, which required children to practice targeted movements involved in playing percussion instruments. This hypothesis could be tested in the future by examining whether other types of music training (e.g., singing) lead to similar results.

For nonverbal reasoning, executive functions, and broader socioemotional functioning, there was no evidence that music training improved performance, in line with the view that far transfer is rare (Gobet & Sala, 2023; Sala & Gobet, 2020; Schellenberg & Lima, 2024). These null findings will contribute to future summaries of the literature, which are needed to resolve discrepancies across recent meta-analyses (for meta-analysis reporting positive effects of music training, see e.g., Bigand & Tillmann, 2022; Román-Caballero et al., 2022). In contrast, we found robust but unexpected improvements for STM and WM. Although music training tends to be associated positively with memory in general (Talamini et al., 2017), associations do not guarantee causation, and previous longitudinal studies with children reported null effects of music training compared to active control groups on auditory STM or WM (D'Souza & Wiseheart, 2018; Kosokabe et al., 2021; Nan et al., 2018). In some instances, moreover, musical ability correlates more strongly than music training does with STM and WM (Swaminathan & Schellenberg, 2018, 2020), perhaps because musical ability and STM/WM are measured with similar (same/different) tasks.

On the one hand, the lack of a strong prediction combined with multiple variables increases the probability of Type I errors. On the other hand, the robust results for STM and WM seem unlikely to be due to random error. Schellenberg and Lima (2024) noted that positive results from music-training studies tend to be accompanied by methodological and statistical problems. In our battery of tests, Digit Span, which we used to measure both STM and WM, was the only one that involved direct face-to-face, one-on-one contact between the experimenter and the child. Specifically, stimuli were presented by the experimenters, who read aloud numbers for children to repeat. Because the experimenters were aware of group assignment, we cannot exclude the possibility that expectancies influenced the results. This aspect of the procedure could also explain why the music group showed larger increases in memory for digits but not for tunes from the MBEMA, which children heard over headphones. In any event, links between music training and STM or WM remain the focus of much research from different labs (e.g., Grassi et al., 2023), which could clarify the nature of the associations.

The present studies have limitations. Although the impact of Covid-19 lockdowns appeared similar across groups according to questionnaires completed by the teachers, we cannot rule out the possibility that the interruptions affected the results. For example, findings that were close to being significant, such as those of music training effects on musical abilities, could have been stronger had the training proceeded under more stable conditions for the full planned duration. Additionally, as in most research on transfer, we did not measure skills targeted directly by the interventions. Although group-specific improvements on several measures (STM, WM, fine- and gross-motor skills) indicate that the training was effective, future longitudinal studies could identify the key skills trained by the programs and assess how they develop over time. Such assessments would validate the intervention and specify the components responsible for any observed transfer effects.

Future studies could also compare different types of music training to examine generalizability. We used Orff-based training in Study 1. In Study 2 and virtually all cross-sectional comparisons, music training differed across individuals, such that direct comparisons across pedagogies were not possible. In fact, such comparisons are rarely addressed in the literature, in which the concept of music training tends to be broad and undefined (Schellenberg & Lima, 2024). Finally, we tested children's ability to recognize emotions in stimuli produced by adults, and only by female adults in the case of prosody. Future research could use more diverse (e.g., child produced) stimuli. Nevertheless, the available data show that school-age children can recognize adult expressions with above-chance accuracy (e.g., Amorim et al., 2021; Correia et al., 2019), and that the actor's age or sex does not play a major role in recognition accuracy (Amorim et al., 2021).

In summary, our findings suggest that observed associations between music training and emotion recognition stem from preexisting musical abilities, rather than from plasticity and far transfer. This conclusion, alongside the lack of training effects for most of our measures, aligns with the overall rarity of far transfer (Gobet & Sala, 2023) and the established genetic influences on musical abilities and behaviors (Wesseldijk et al., 2023). Nevertheless, improvements for fine-motor skills and auditory memory that we observed raise the possibility that transfer occurs in specific domains and contexts. To resolve ongoing controversies about potential training effects, we suggest further consideration of longitudinal evidence, combined with a systematic examination of the nonmusical correlates of musical abilities. In the absence of causal evidence, preexisting associations between music and nonmusical abilities offer the simplest explanation for the many advantages seen in musicians.

Author Note

This work was funded by the Portuguese Foundation for Science and Technology (FCT) through a PhD studentship awarded to L.N. (SFRH/BD/135604/2018), a Scientific Employment Stimulus grant awarded to E.G.S. (CEECIND/03266/2018), Centro de Psicologia da Universidade do Porto (CPUP UIBD/00050/2020), and a project grant awarded to C.F.L. (PTDC/PSI-GER/28274/2017), co-funded by the European Regional Development Fund (ERDF) through the Lisbon Regional Operational Program (LISBOA-01-0145-FEDER-028274) and the Operational Program for Competitiveness and Internationalization (POCI-01-0145-FEDER-028274).

We thank the school administrators, teachers, parents, and all the children who took part in the studies.

The data used for the analyses are available at https://osf.io/u96fa/?view_only=d3d751ebbc2c4d3a84734b0916536b17 (Lima et al., 2024).

References

Adam, C., Klissouras, V., Ravazzolo, M., Renson, R., Tuxworth, W., Kemper, HCG., van Mechelen, W., Hlobil, H., Beunen, G., & Levarlet-Joye, H. (1993). *EUROFIT - European test of physical fitness (2nd ed.).* Council of Europe.

Amorim, M., Anikin, A., Mendes, A. J., Lima, C. F., Kotz, S. A., & Pinheiro, A. P. (2021). Changes in vocal emotion recognition across the life span. *Emotion, 21,* 315-325. https://doi.org/10.1371/emo0000692

Barbaroux, M., Dittinger, E., & Besson, M. (2019). Music training with Démos program positively influences cognitive functions in children from low socio-economic backgrounds. *PLoS One*, *14*(5), Article e0216874. https://doi.org/10.1371/journal.pone.0216874

Barton, J. J., Stubbs, J. L., Paquette, S., Duchaine, B., Schlaug, G., & Corrow, S. L. (2023). Music perception in acquired prosopagnosia. *Neuropsychologia*, *183*, Article 108540. https://doi.org/10.1016/j.neuropsychologia.2023.108540

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*, Article 94. https://doi.org/10.3389/fpsyg.2011.00094

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology, 60*(1), 68-79. htpps://doi.org/10.1037/cjep2006008

Bigand, E., & Tillmann, B. (2022). Near and far transfer: Is music special? *Memory & Cognition*, *50*(2), 339-347. https://doi.org/10.3758/s13421-021-01226-6

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

Bryant, B. K. (1982). An index of empathy for children and adolescents. *Child Development*, *53*(2), 413-425. https://doi.org/10.2307/1128984

Bugos, J. A., DeMarie, D., Stokes, C., & Power, L. P. (2022). Multimodal music training enhances executive functions in children: Results of a randomized controlled trial. *Annals of the New York Academy of Sciences*, *1516*(1), 95-105. https://doi.org/10.1111/nyas.14857

Castro, S. L., & Lima, C. F. (2010). Recognizing emotions in spoken language: A validated set of Portuguese sentences and pseudo-sentences for research on emotional prosody. *Behavior Research Methods*, *42*(1)*,*74-81. https://doi.org/10.3758/BRM.42.1.74

Chari, D. A., Barrett, K. C., Patel, A. D., Colgrove, T. R., Jiradejvong, P., Jacobs, L. Y., & Limb, C. J. (2020). Impact of auditory-motor musical training on melodic pattern recognition in cochlear implant users. *Otology & Neurotology*, *41*(4), e422-e431. https://doi.org/10.1097/MAO.0000000000002525

Ciria, L. F., Román-Caballero, R., Vadillo, M. A., Holgado, D., Luque-Casado, A., Perakakis, P., & Sanabria, D. (2023). An umbrella review of randomized control trials on the effects of physical exercise on cognition. *Nature Human Behaviour*, *7,* 928-941. https://doi.org/10.1038/s41562-023-01554-4

Clark, C. N., Downey, L. E., & Warren, J. D. (2015). Brain disorders and the biological role of music. *Social Cognitive and Affective Neuroscience*, *10*(3), 444-452. https://doi.org/10.1093/scan/nsu079

Coffey, E. B., 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

Correia, A. I., Branco, P., Martins, M., Reis, A. M., Martins, N., Castro, S. L., & Lima, C. F. (2019). Resting-state connectivity reveals a role for sensorimotor systems in vocal emotional processing in children. *NeuroImage, 201,* Article 116052. https://doi.org/10.1016/j.neuroimage.2019.116052

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

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

Corrow, S. L., Stubbs, J. L., Schlaug, G., Buss, S., Paquette, S., Duchaine, B., & Barton, J. J. (2019). Perception of musical pitch in developmental prosopagnosia. *Neuropsychologia*, *124*, 87-97. https://doi.org/10.1016/j.neuropsychologia.2018.12.022

Coutinho, E., & Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition & Emotion*, *27*(4), 658-684. https://doi.org/10.1080/02699931.2012.732559

Curtis, M. E., & Bharucha, J. J. (2010). The minor third communicates sadness in speech, mirroring its use in music. *Emotion*, *10*(3), 335-348. https://doi.org/10.1037/a0017928

Degé, F. (2021). Music lessons and cognitive abilities in children: How far transfer could be possible. *Frontiers in Psychology*, *11*, Article 557807. https://doi.org/10.3389/fpsyg.2020.557807

Desrosiers, J., Rochette, A., Hebert, R., & Bravo, G. (1997). The Minnesota Manual Dexterity Test: Reliability, validity and reference values studies with healthy elderly people. *Canadian Journal of Occupational Therapy*, *64*(5), 270-276. https://doi.org/10.1177/000841749706400504

D'Souza, A. A., & Wiseheart, M. (2018). Cognitive effects of music and dance training in children. *Archives of Scientific Psychology, 6*(1), 178–192. https://doi.org/10.1037/arc0000048

Dubinsky, E., Wood, E. A., Nespoli, G., & Russo, F. A. (2019). Short-term choir singing supports speech-in-noise perception and neural pitch strength in older adults with age-related hearing loss. *Frontiers in Neuroscience*, *13*, Article 1153. https://doi.org/10.3389/fnins.2019.01153

Escoffier, N., Zhong, J., Schirmer, A., & Qiu, A. (2013). Emotional expressions in voice and music: Same code, same effect? *Human Brain Mapping*, *34*(8), 1796-1810. https://doi.org/10.1002/hbm.22029

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*, 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Frischen, U., Schwarzer, G., & Degé, F. (2021). Music lessons enhance executive functions in 6-to 7-year-old children. *Learning and Instruction*, *74*, Article 101442. https://doi.org/10.1016/j.learninstruc.2021.101442

Fuller, C. D., Galvin III, J. J., Maat, B., Başkent, D., & Free, R. H. (2018). Comparison of two music training approaches on music and speech perception in cochlear implant users. *Trends in Hearing*, *22*, 1-22. https://doi.org/10.1177/2331216518765379

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

Gobet, F., & Sala, G. (2023). Cognitive training: A field in search of a phenomenon. *Perspectives on Psychological Science*, *18*(1), 125-141. https://doi.org/10.1177/17456916221091830

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition and Emotion*, *22*(6), 1094-1118. https://doi.org/10.1080/02699930701626582

Good, A., Gordon, K. A., Papsin, B. C., Nespoli, G., Hopyan, T., Peretz, I., & Russo, F. A. (2017). Benefits of music training for perception of emotional speech prosody in deaf children with cochlear implants. *Ear and Hearing*, *38*(4), 455-464. https://doi.org/10.1097/AUD.0000000000000402

Grassi, M., Talamini, F., Ferreri, L., Swaminathan, S., Tillmann, B., Caclin, A., … Wilbiks, J. (2023, February 6). Do musicians have better short-term memory than nonmusicians? A multi-lab study. https://doi.org/10.17605/OSF.IO/69GZF

Harmony Project (2019). Harmony Project: Our Story - Research. *Harmony Project*. https://www.harmony-project.org/research

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

Howard, S. J., & Melhuish, E. (2017). An early years toolbox for assessing early executive function, language, self-regulation, and social development: Validity, reliability, and preliminary norms. *Journal of Psychoeducational Assessment*, *35*(3), 255-275. https://doi.org/10.1177/073428291663300

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), 2-9. http://doi.org/10.7771/1932-6246.1167

JASP Team (2024). *JASP* (Version 0.18.3) [Computer software].

Jespersen, K. V., Gebauer, L., & Vuust, P. (2022). White paper: Music interventions in health care. *Center for the Music in the Brain, Aarhus University.* https://danishsoundcluster.dk/white-paper-music-interventions-in-health-care/

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

Kassai, R., Futo, J., Demetrovics, Z., & Takacs, Z. K. (2019). A meta-analysis of the experimental evidence on the near-and far-transfer effects among children's executive function skills. *Psychological Bulletin*, *145*(2), 165-188. https://doi.org/10.1037/bul0000180

Koelsch, S. (2013). From social contact to social cohesion – The 7 Cs. *Music and Medicine, 5*(4), 204-209. https://doi.org/10.1177/1943862113508588

Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience*, *15*(3), 170-180. https://doi.org/10.1038/nrn3666

Kosokabe, T., Mizusaki, M., Nagaoka, W., Honda, M., Suzuki, N., Naoi, R., & Moriguchi, Y. (2021). Self-directed dramatic and music play programs enhance executive function in Japanese children. *Trends in Neuroscience and Education*, *24*, Article 100158. https://doi.org/10.1016/j.tine.2021.100158

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 e13081. https://doi.org/10.1111/desc.13081

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *Journal of Neuroscience*, *34*(36), 11913-11918. https://doi.org/10.1523/JNEUROSCI.1881-14.2014

Kraus, N., & White-Schwoch, T. (2020). The argument for music education. *American Scientist*, *108*(4), 210-214. https://doi.org/10.1511/2020.108.4.210

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,* Article 34911. https://doi.org/10.1038/srep34911

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion, 11*(5), 1021-1031. https://doi.org/10.1037/a0024521

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods*, *45*(4), 1234-1245. https://doi.org/10.3758/s13428-013-0324-3

Lima, C. F., Schellenberg, E. G., Neves, L., Martins, M., Correia, A. I., & Castro, S. L. (2024, March 25). Music training and emotion recognition. Retrieved from https://osf.io/u96fa/?view_only=d3d751ebbc2c4d3a84734b0916536b17

Linnavalli, T., Putkinen, V., Lipsanen, J., Huotilainen, M., & Tervaniemi, M. (2018). Music playschool enhances children's linguistic skills. *Scientific Reports*, *8*, Article 8767. https://doi.org/10.1038/s41598-018-27126-5

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

Martins, M., Neves, L., Rodrigues, P., Vasconcelos, O., & Castro, S. L. (2018). Orff-based music training enhances children's manual dexterity and bimanual coordination. *Frontiers in Psychology, 9,* Article 2616. https://doi.org/10.3389/fpsyg.2018.02616

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

Martins, M., Reis, A. M., Gaser, C., & Castro, S. L. (2023). Individual differences in rhythm perception modulate music-related motor learning: A neurobehavioral training study with children. *Scientific Reports, 13*, Article 21552. https://doi.org/10.1038/s41598-023-48132-2

Moore, E. (2018). *Music and dyslexia: Investigating the role of auditory-motor timing skills in the transfer from musical learning to language skills* [Unpublished doctoral dissertation, University of Edinburgh].

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, *22*(11), 1425-1433. https://doi.org/10.1177/0956797611416999

Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing Research*, *308*, 84-97. https://doi.org/10.1016/j.heares.2013.09.012

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712-723. https://doi.org/10.1093/cercor/bhn120

Mualem, O., & Lavidor, M. (2015). Music education intervention improves vocal emotion recognition. *International Journal of Music Education*, *33*(4), 413-425. https://doi.org/10.1177/02557614155842

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., Cordeiro, C., Scott, S. K., Castro, S. L., & Lima, C. F. (2018). High emotional contagion and empathy are associated with enhanced detection of emotional authenticity in laughter. *Quarterly Journal of Experimental Psychology, 71*(11), 2355-2363. https://doi.org/10.1177/1747021817741800

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

Neves, L., Martins, M., Correia, A. I., Castro, S. L., & Lima, C. F. (2021). Associations between vocal emotion recognition and socio-emotional adjustment in children. *Royal Society Open Science*, *8*(11), Article 211412. https://doi.org/10.1098/rsos.211412

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

Nussbaum, C., & Schweinberger, S. R. (2021). Links between musicality and vocal emotion perception. *Emotion Review*, *13*(3), 211-224. https://doi.org/10.1177/17540739211022803

Okely, J. A., Overy, K., & Deary, I. J. (2022). Experience of playing a musical instrument and lifetime change in general cognitive ability: Evidence from the Lothian Birth Cohort 1936. *Psychological Science*, *33*(9), 1495-1508. https://doi.org/10.1177/09567976221092726

Overy, K. (2003). Dyslexia and music: From timing deficits to musical intervention. *Annals of the New York Academy of Sciences*, *999*(1), 497-505. https://doi.org/10.1196/annals.1284.060

Pahor, A., Seitz, A. R., & Jaeggi, S. M. (2022). Near transfer to an unrelated N-back task mediates the effect of N-back working memory training on matrix reasoning. *Nature Human Behaviour*, *6*, 1243-1256. https://doi.org/10.1038/s41562-022-01384-w

Park, M., Gutyrchik, E., Welker, L., Carl, P., Pöppel, E., Zaytseva, Y., Meindl, T., Blautzik, J., Reiser, M., & Bao, Y. (2015). Sadness is unique: Neural processing of emotions in speech prosody in musicians and non-musicians. *Frontiers in Human Neuroscience*, *8*, Article 1049. https://doi.org/10.3389/fnhum.2014.01049

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing research*, *308*, 98-108. https://doi.org/10.1016/j.heares.2013.08.011

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *Journal of Neuroscience*, *30*(30), 10127-10134. https://doi.org/10.1523/JNEUROSCI.2161-10.2010

Peretz, I., Gosselin, N., Nan, Y., Caron-Caplette, E., Trehub, S. E., & Béland, R. (2013). A novel tool for evaluating children's musical abilities across age and culture. *Frontiers in Systems Neuroscience*, *7*, Article 30. https://doi.org/10.3389/fnsys.2013.00030

Pinheiro, A. P., Anikin, A., Conde, T., Sarzedas, J., Chen, S., Scott, S. K., & Lima, C. F. (2021). Emotional authenticity modulates affective and social trait inferences from voices. *Philosophical Transactions of the Royal Society B*, *376*, Article 20200402. https://doi.org/10.1098/rstb.2020.0402

Pons, F., & Harris, P. (2000). *Test of Emotion Comprehension – TEC*. University of Oxford.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales.* Oxford Psychologists Press.

Rocha, A. Roazzi, A. Silva, A. Candeias, A. Minervino, C., Roazzi, M., & Pons, F. (2013). Test of Emotional Comprehension: Exploring the underlying structure through Confirmatory Factor Analysis and Similarity Structure Analysis. In A. Roazzi, B. C. de Souza, & W. Bilsky (Eds.), *Searching for structure in complex social, cultural and psychological phenomen*a (pp. 74-95). Recife, PE: FTA.

Román-Caballero, R., Vadillo, M. A., Trainor, L. J., & Lupiáñez, J. (2022). Please don't stop the music: A meta-analysis of the cognitive and academic benefits of instrumental musical training in childhood and adolescence. *Educational Research Review, 35,* Article 100436. https://doi.org/10.1016/j.edurev.2022.100436

Sala, G., & Gobet, F. (2017a). When the music's over. Does music skill transfer to children's and young adolescents' cognitive and academic skills? A meta-analysis. *Educational Research Review*, *20*, 55-67. https://doi.org/10.1016/j.edurev.2016.11.005

Sala, G., & Gobet, F. (2017b). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science*, *26*(6), 515-520. https://doi.org/10.1177/0963721417712760

Sala, G., & Gobet, F. (2020). Cognitive and academic benefits of music training with children: A multilevel meta-analysis. *Memory & Cognition*, *48*(8), 1429-1441. https://doi.org/10.3758/s13421-020-01060-2

Schellenberg, E. G. (2020). Correlation= causation? Music training, psychology, and neuroscience. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(4), 475-480. https://doi.org/10.1037/aca0000263

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., Corrigall, K. A., Dys, S. P., & Malti, T. (2015). Group music training and children's prosocial skills. *PLoS One*, *10*(10), Article e0141449. https://doi.org/10.1371/journal.pone.0141449

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

Schellenberg, E. G., & Peretz, I. (2008). Music, language and cognition: Unresolved issues. *Trends in Cognitive Sciences, 12*(2), 45.46. https://doi.org/10.1016/j.tics.2007.11.005

Schirmer, A., & Adolphs, R. (2017). Emotion perception from face, voice, and touch: Comparisons and convergence. *Trends in Cognitive Sciences*, *21*(3), 216-228. https://doi.org/10.1016/j.tics.2017.01.001

Schlaug, G. (2015). Musicians and music making as a model for the study of brain plasticity. *Progress in Brain Research*, *217*, 37-55. https://doi.org/ 10.1016/bs.pbr.2014.11.020

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359-1366. https://doi.org/10.1177/0956797611417632

Swaminathan, S., & Schellenberg, E. G. (2015). Current emotion research in music psychology. *Emotion Review*, *7*(2), 189-197. https://doi.org/10.1177/175407391455828

Swaminathan, S., & Schellenberg, E. G. (2018). Musical competence is predicted by music training, cognitive abilities, and personality. *Scientific Reports*, *8*(1), Article 9223. https://doi.org/10.1038/s41598-018-27571-2

Swaminathan, S., & Schellenberg, E. G. (2020). Musical ability, music training, and language ability in childhood. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(12), 2340-2348. https://doi.org/10.1037/xlm0000798

Talamini, F., Altoè, G., Carretti, B., & Grassi, M. (2017). Musicians have better memory than nonmusicians: A meta-analysis. *PloS One*, *12*(10), Article e0186773. https://doi.org/10.1371/journal.pone.0186773

Tervaniemi, M. (2023). The neuroscience of music–towards ecological validity. *Trends in Neurosciences*, *46*(5), 355-364. https://doi.org/ https://doi.org/10.1016/j.tins.2023.03.00

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*, *109*(46), 19027-19032. https://doi.org/10.1073/pnas.1210344109

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

Tiffin, J. (1968). *Purdue Pegboard: Examiner Manual*. Science Research Associates.

Trimmer, C. G., & Cuddy, L. L. (2008). Emotional intelligence, not music training, predicts recognition of emotional speech prosody. *Emotion, 8*(6), 838-849. https://doi.org/10.1037/a0014080

Van't Hooft, J. J., Pijnenburg, Y. A., Sikkes, S. A., Scheltens, P., Spikman, J. M., Jaschke, A. C., Warren, J. D., & Tijms, B. M. (2021). Frontotemporal dementia, music perception and social cognition share neurobiological circuits: A meta-analysis. *Brain and Cognition*, *148*, Article 105660. https://doi.org/10.1016/j.bandc.2020.105660

Wan, C. Y., & Schlaug, G. (2010). Music making as a tool for promoting brain plasticity across the life span. *The Neuroscientist*, *16*(5), 566-577. https://doi.org/ 10.1177/1073858410377805

Wechsler, D. (2003). *Escala de Inteligência de Wechsler para Crianças – III* [Wechsler Intelligence Scale for Children–3rd ed.]. Hogrefe.

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

Zanto, T. P., Johnson, V., Ostrand, A., & Gazzaley, A. (2022). How musical rhythm training improves short-term memory for faces. *Proceedings of the National Academy of Sciences*, *119*(41), Article e2201655119. https://doi.org/10.1073/pnas.2201655119