Does Music Training Improve Emotion Recognition Abilities? A Critical Review

Marta Martins[a], Ana P. Pinheiro[b], and César F. Lima[a, c, *]

[a] Instituto Universitário de Lisboa (ISCTE-IUL), Lisbon, Portugal

[b] Faculty of Psychology, University of Lisbon, Lisbon, Portugal

[c] University College London, London, UK

Author Note

Marta Martins, Centro de Investigação e Intervenção Social (CIS-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal; Ana P. Pinheiro, Faculdade de Psicologia, Universidade de Lisboa, Lisboa, Portugal; César F. Lima, CIS-IUL, Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal, and Institute of Cognitive Neuroscience, University College London, London, UK.

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

Abstract

There is widespread interest in the possibility that music training enhances non-musical abilities. This possibility has been examined primarily for speech perception and domain-general abilities such as IQ. Although social and emotional processes are central to many musical activities, transfer from music training to socioemotional skills remains underexplored. Here we synthesize results from studies examining associations between music training and emotion recognition in voices and faces. Enhancements are typically observed for vocal emotions but not for faces, although most evidence is cross-sectional. These findings are discussed considering the design features of the studies. Future research could explore further the neurocognitive mechanisms underlying musician-related differences in emotion recognition, the role of predispositions, and the implications for broader aspects of socio-emotional functioning.

Keywords: music training; emotion recognition; socioemotional processes; transfer

Introduction

Music training has been used widely as a model for studying brain plasticity (Habib & Besson, 2009; Herholz & Zatorre, 2012; Kraus & Chandrasekaran, 2010; Moreno & Bidelman, 2014; Pantev & Herholz, 2011), which refers to changes in brain structure and/or function that occur as a consequence of learning and experience. Learning to play a musical instrument is a demanding multimodal task that usually starts in childhood and involves large amounts of practice, often for several years. It requires high-precision motor control, perceptual tuning to the fine-grained acoustics of sounds, integrating information from several sensory modalities, and attention, memory, and emotional processes.

Due to potential theoretical and practical implications, there is a widespread interest is the possibility that music training has consequences that extend to non-musical abilities. Such generalization of knowledge is called 'transfer'. A distinction is often made between *near* transfer, which occurs between domains that are closely related (e.g., between two different pitch discrimination tasks), and *far* transfer, which occurs between domains that have less in common (e.g., between music and mathematics; Barnett & Ceci, 2002). If music training promotes transfer of skills, this would inform debates on learning and plasticity (Herholz & Zatorre, 2012), the organization of cognitive functions (e.g., modularity; Besson et al., 2011), the biological basis of musicality (e.g., Clark et al., 2015), and the use of music as a tool in clinical and educational contexts (Dumont et al., 2017; Grau-Sánchez et al., 2020).

Many studies have examined far transfer from music to speech perception and domain-general abilities such as IQ and executive functions. Music training has been associated with enhancements in speech-in-noise perception, for instance (Coffey et al., 2017a; Coffey et al., 2017b; but see Boebinger et al., 2015), and in prosody perception (e.g., Marques et al., 2007). A theoretical account has been put forward for understanding these effects, based on plasticity mechanisms and the overlap between sensory and cognitive

processes in both music and speech (OPERA hypothesis; Patel, 2011; 2014). Enhancements have also been documented for working memory (Roden et al., 2014), executive functions (Bugos & DeMarie, 2017; Moreno et al., 2011), and IQ (Schellenberg, 2004, 2011). Although it is often claimed that music training *causes* advantages in non-musical tasks, most evidence comes from cross-sectional comparisons between musicians and non-musicians, which do not establish causality (Schellenberg, 2019). Sala and Gobet (2017a, 2017b; 2020) reviewed longitudinal training studies conducted with children, and found that far transfer from music to cognitive and academic skills is in fact weak, and most likely to be found when study designs are not optimal (e.g., no random allocation of participants). Claims about music training and transfer therefore remain a matter of debate.

In this review, we address the association between music training and a central aspect of socioemotional processing—the ability to recognize emotions expressed by others. Music is linked fundamentally to emotional and social processes, but potential transfer of skills to these domains is typically overlooked, in contrast to the large number of studies focusing on speech perception and cognitive abilities. Our evaluation of the available literature on music training and the recognition of vocal and facial emotions focuses on three issues: (1) design features of the studies, (2) specificity and scope of the associations, and (3) underlying cognitive and neural mechanisms.

Music and socioemotional processes

Music's prevalence across human societies relates to its power to express, induce, and regulate emotions (for a review, Swaminathan & Schellenberg, 2015). Such connection between music and emotion is part of our social interactions from early life, for instance as a communicative channel between caregivers and infants. Caregivers' speech to infants is music-like (Fernald et al., 1989), and singing to infants is ubiquitous, with effects on sustained attention, arousal, social regulation, and emotional synchrony (Nakata & Trehub,

2004; Trehub, 2003). Moreover, adolescents and adults report listening to music primarily to regulate their moods (Lonsdale & North, 2011; North et al., 2000). Music also triggers synchronization behaviors, such as dancing or foot tapping, which induce positive emotions (Trost et al., 2017) and promote bonding and prosocial behavior (Tarr et al., 2014).

Similarly to other socioemotional signals, such as faces and vocalizations, music expresses emotions that can be recognized quickly and consistently across listeners (e.g., Bigand et al., 2005; Thompson, 2009). There is also evidence of above-chance recognition of musical emotions across cultures, suggesting that there are universal cues to emotion in music (Balkwill & Thompson, 1999; Fritz et al., 2009). Emotion recognition in music can additionally be systematically linked to factors such as age and music training. For example, the ability to identify negative emotions, but not positive ones, starts to decline in middle age (Castro & Lima, 2014; Lima & Castro, 2011), and music training is associated with enhanced sensitivity to musically expressed emotions (Bhatara et al., 2011; Castro & Lima, 2014; Lima & Castro, 2011).

Consistent with the fundamental link between music and emotion, a meta-analysis of neuroimaging studies shows that music listening recruits core brain systems underlying emotional processes, including the amygdala, nucleus accumbens, hypothalamus, hippocampus, insula, cingulate cortex, and orbitofrontal cortex (Koelsch, 2014). According to Koelsch (2013; 2014), the social functions of music are at the core of these emotional effects. Actively participating in musical activities, especially in a group, would engage social functions that are essential for the survival of the individual and the species. Playing music promotes contact with others, social cognition processes including mental state attribution, co-pathy (emotional states become more homogenous across individuals), communication, coordination of actions, and cooperation, leading to increased social cohesion. By involving these functions, music would fulfil several basic human needs in an effortless manner,

possibly explaining its role in evolution and development, and its power to produce pleasure and reward value.

In line with this argument, Clark and colleagues (2014, 2015) highlight the potential role of musical activities in promoting the ability to analyze others' mental states and motivations. They argue that music has evolved from the call signals of our hominid ancestors as a vehicle to model and express emotional mental states, allowing socially relevant routines to be abstracted, rehearsed, and transformed in a low-cost manner (i.e., without the costs of enacting the corresponding scenarios). The proposal of a connection between music and social cognition is supported by findings indicating that the decoding of mental states in music is impaired in frontotemporal dementia, a paradigmatic acquired disorder of social behavior (Downey et al., 2013). This musical impairment correlates with measures of social inference and empathic capacity and, notably, has a neuroanatomical substrate in medial prefrontal and temporal areas that are implicated in mentalizing and social behavior. Furthermore, adults with a developmental deficit for music processing (congenital amusia) show impairments in socioemotional cognition, including a decreased ability to infer the emotional authenticity of laughter and to decode emotions from vocal and facial expressions (Lima et al., 2016a). These results suggest that intact musical behavior and capacities might be needed for the normal development of socioemotional processing.

It is therefore plausible to hypothesize that systematic engagement with musical activities could lead to improvements in socioemotional skills and corresponding brain networks. Music could additionally improve socioemotional skills as a consequence of low-level sensory enhancements. For example, music training can fine-tune aspects of auditory processing, including pitch (Habibi et al., 2016; Moreno et al., 2009), timing (Chobert et al., 2014; Frey et al., 2019), and timbre (Putkinen et al., 2019). Processing these acoustic cues is critical for sensory stages of socioemotional perception in the auditory domain, particularly

for vocal emotions. The acoustic profile of musical and vocal emotions is in fact partly

shared, with notable commonalities in the use of acoustic cues such as pitch, intensity, and

high-frequency energy (Curtis & Bharucha, 2010; Ilie & Thompson, 2006; Juslin & Laukka,

2003). For instance, in both music and speech, anger is communicated by fast tempo, rising

pitch, high intensity, and high-frequency energy, whereas sadness and tenderness are

typically associated with reductions in these features (Juslin & Laukka, 2003). The locus for

putative transfer from music to socioemotional processing could thus be at an auditory-

perceptual level of processing, at a higher-order social cognition level, or both.

Emotion recognition abilities

Emotion recognition is a crucial socioemotional skill. In everyday interactions, we

receive information about the emotional states of others through a multitude of nonverbal

cues expressed by the face, body, and voice. The voice is a particularly rich communicative

tool: emotions can be conveyed through prosody in actual speech (*emotional prosody*), and

through nonverbal vocalizations. Both prosody and nonverbal vocalizations are vocal signals,

but their underlying production and perception mechanisms are partly distinct (Pell et al.,

2015; Scott et al., 2010). Emotional prosody corresponds to variations in aspects of speech,

including fundamental frequency, amplitude, timing, and timbre (Banse & Scherer, 1996;

Scherer, 1995). Nonverbal vocalizations are nonspeech vocal sounds, such as laughs,

screams, or sobs.

Forced-choice classification and rating studies indicate that listeners can recognize a

wide range of emotions in facial, prosodic, and nonverbal vocal cues, with performance

levels well above chance (e.g., Bänziger et al., 2012; Langner et al., 2010; Lima, Castro, &

Scott, 2013). This recognition process is fast, automatic (Lima et al., 2019; Tracy & Robins,

2008), and partly governed by universal principles (e.g., Chronaki et al., 2018; Laukka &

Elfenbein, 2020; Sauter et al., 2010, 2015; but see Gendron et al., 2018). Emotion-

recognition abilities also vary widely across individuals. For instance, older adults perform worse than younger adults on facial and vocal emotion-recognition tasks, (e.g., Amorim et al., 2019; Lima et al., 2014; Ruffman et al., 2008), and some studies show associations between poor empathy and poor emotion recognition (e.g., Dawel et al., 2012; Gery et al., 2009; but see Vilaverde et al., 2020). Understanding variability in emotion recognition is important because of its potential implications for everyday behavior and communication. Emotion-recognition abilities relate to personal and social adjustment in children (e.g., Blair & Coles, 2000; Bowen & Nowicki, 2007; Stevens et al., 2001) and adults (e.g., Carton et al., 1999; Dawel et al., 2012; Hall et al., 2009), and are central to the notions of emotional competence and emotional intelligence (Mayer et al., 2003; Scherer & Scherer, 2011). Moreover, intervention studies focused on emotion recognition have shown that these abilities are amenable to training (e.g., Golan et al., 2010; Schlegel et al., 2017).

Several neural systems are involved in emotion recognition, including the amygdala, fusiform gyrus, superior temporal sulcus and gyrus, and medial and lateral prefrontal cortices (e.g., Frühholz et al., 2016; Schirmer, 2018; Schirmer & Adolphs, 2017). Some of these systems are engaged regardless of modality, namely the medial prefrontal cortex and the posterior part of the superior temporal sulcus (e.g., Peelen et al., 2010), and play a more general role in socioemotional processing. Nevertheless, other systems show modality-specific effects. Vocal emotions seem to engage more readily the superior temporal sulcus and gyrus, whereas the amygdala is engaged more readily by facial emotions (Schirmer, 2018; Grandjean, 2020; Young et al., 2020). Recent studies have additionally suggested that the motor system also plays a role in socioemotional processing, both for vocal (Correia et al., 2019; Lima et al., 2016b; McGettigan et al., 2015; O'Nions et al., 2017) and facial expressions (Ethofer et al., 2013; Johnston et al., 2013; Kreifelts et al., 2013; Rochas et al.,

2013). For instance, a stronger recruitment of the motor system predicts better performance in vocal emotion recognition tasks (Correia et al., 2019; McGettigan et al., 2015).

Music training and emotion recognition abilities

Several studies have examined whether music training is associated with enhanced performance in emotion-recognition tasks. The results of these studies are summarized in Figure 1 and Supplementary Tables S1 and S2. Most of them are based on cross-sectional comparisons between trained and untrained participants ($n = 17$), but a few have used longitudinal training designs ($n = 4$).

[Insert Figure 1]

Cross-sectional evidence

Most cross-sectional studies ask how trained and untrained listeners recognize vocal emotions, using prosodic stimuli in the majority of cases (e.g., Correia et al., 2020; Dmitrieva et al., 2006; Farmer et al., 2020; Fuller et al., 2014; Lima & Castro, 2011; Park et al., 2015; Pinheiro et al., 2015), but also melodic analogues of emotional prosody (Thompson et al., 2004; Trimmer & Cuddy, 2008) and purely nonverbal vocalizations (Correia et al., 2020; Parsons et al., 2014; Young et al., 2012). Only a few studies examined emotion recognition for other modalities, including faces (Correia et al., 2020; Farmer et al., 2020; Weijkamp & Sadakata, 2016) and audio-visual stimuli (Farmer et al., 2020; Weijkamp & Sadakata, 2016). The focus is typically on the recognition of specific emotions (e.g., happiness, sadness), evaluated via forced-choice tasks, in which participants select the emotion being expressed by each stimulus from a list of alternatives. Performance is then measured in terms of the accuracy with which emotions are recognized. There are also studies evaluating other aspects of emotional processing, however, including valence and arousal perception (Dibben et al., 2018), and inferences of distress (Parsons et al., 2014; Young et al., 2012) and depression in voices (Nilsonne & Sundberg, 1985). Trained participants are usually adults with at least five

years of classical music training, which is in line with the broader literature on music training

(Zhang et al., 2018), who are compared to individuals with little or no training.

A musician advantage emerges in the majority of studies on vocal emotion

recognition (11 out of 17, 65%). For instance, Lima and Castro (2011) compared forty

musicians and forty untrained listeners, from a wide age range (18-60 years), on their ability

to recognize seven prosodic emotions in sentences with emotionally neutral semantic content.

Musicians showed improved recognition accuracy, an effect that was similar across the age

range and across emotions (for a recent replication, see Correia et al., 2020). Similar

advantages have been found for recognizing emotions from tone sequences that mimic the

prosody of spoken sentences (Thompson et al., 2004, Experiment 1), for pseudowords (Fuller

et al., 2014), and for inferring whether a vocal expression was produced during a depressive

state (Nilsonne & Sundberg, 1985). In six of the studies reporting an advantage of music

training, however, group differences were restricted to a subgroup of participants (Dmitrieva

et al., 2006; Parsons et al., 2014; Young et al., 2012) or a subset of emotions (Thompson et

al., 2004; Pinheiro et al., 2015; Correia et al., 2020). Thompson and colleagues, for example

(2004, Experiment 2), compared 28 trained listeners and 26 untrained ones on their ability to

recognize prosodic emotions in tone analogues of spoken sentences, and in intact English and

Tagalog sentences. The training advantage was observed across stimulus types, but only for

sadness, fear and neutrality, and not for happiness and anger. More recently, Correia and

colleagues (2020) showed that musician advantages extend to the recognition of purely

nonverbal vocalizations, but again not for all emotions (e.g., trained and untrained

participants had a similar performance in the recognition of anger). Importantly, there are six

studies that failed to find training effects. Specifically, null findings were observed for

emotional-prosody recognition in linguistic stimuli including words, pseudowords, and

sentences (Başkent et al., 2018; Dibben et al., 2018; Mualem & Lavidor, 2015; Park et al.,

2015; Trimmer & Cuddy, 2008), for tone analogues of spoken sentences (Trimmer & Cuddy, 2008), and for humming voices (Weijkamp & Sadakata, 2016). When Trimmer and Cuddy (2008) tested 100 participants with varying levels of music training, they found that emotion recognition in sentences and tone analogues correlated with emotional intelligence but not with music training.

Therefore, there is evidence for enhanced vocal-emotion recognition in musicians, but the advantage is not always apparent. One possibility is that the effect is small (e.g., Correia et al., 2020), such that relatively large samples are required for it to emerge. Consistent with this possibility, four of the studies reporting null results tested 10 to 16 trained participants only (Başkent et al., 2018; Trimmer & Cuddy, 2008; Weijkamp & Sadakata, 2016), or did not specify the number of trained participants (Trimmer & Cuddy, 2008). It is also possible that in some cases the samples included participants who were no longer practicing their instruments. Only three of the studies reporting null results specified that participants were active musicians (Başkent et al., 2018; Dibben et al., 2018; Mualem & Lavidor, 2015). This issue is important because behavioral and brain effects of learning-induced plasticity might be lost if not practiced, in line with evidence from taxi drivers (Woollett et al., 2009) and jugglers (Draganski et al., 2004).

The few studies that examined associations between music training and emotion recognition for modalities other than the voice found *null* effects. In a task involving emotion recognition in static facial expressions, Correia and colleagues (2020) found that trained musicians performed as well as untrained ones. Moreover, Bayesian analyses provided substantial evidence for the null hypothesis. Farmer and colleagues (2020) explored emotion recognition in audio-visual clips including biological motion and prosodic cues, and observed that the training advantage was restricted to the auditory condition. Musicians were more accurate than untrained participants when they assessed prosodic cues alone, but not when

they assessed biological motion or the two types of cues combined. Finally, using a Stroop-like task involving facial and vocal stimuli combined, Weijkamp and Sadakata (2016) found that musicians' emotional evaluations were less susceptible to interference from the to-be-ignored modality, which could reflect improved attention or more efficient audio-visual integration. Performance was similar across groups, however, when the focus was limited to facial expressions.

Although more research on the visual modality and on multimodal integration is warranted, the available evidence suggests that music training is *selectively* related to vocal emotions. What remains poorly explored are the mechanisms underlying this advantage. Because music training correlates with higher domain-general cognitive abilities (e.g., Swaminathan et al., 2017; Swaminathan & Schellenberg, 2018), emotion-recognition advantages could arise because trained individuals tend to be high-functioning, and not because of a meaningful association between music and emotion processes. Cognitive abilities seem like an obvious confounding variable, yet only a few studies have measured them (Figure 1). Among the studies that measured cognitive abilities and documented musician advantages in emotion recognition, all found positive associations between music training and general cognitive abilities (Correia et al., 2020; Lima & Castro, 2011; Thompson et al., 2004). Thompson and colleagues (2004) showed, however, that group differences in emotional prosody recognition could not be explained by cognitive performance as measured by Raven's Advanced Progressive Matrices. Similarly, Lima and Castro (2011) found that the musician advantage in emotional-prosody recognition remained significant after accounting for performance on the Montreal Cognitive Assessment (MoCA), Raven's Advanced Progressive Matrices, and the Stroop test. In the study by Correia and colleagues (2020), digit span did not explain training advantages in the case of nonverbal vocalizations, but it partly did in the case of emotional prosody. When digit span was held constant, the

correlation between music training and prosody recognition only approached significance. Because the digit-span test has a strong auditory component, it remains unclear if it represents a proper confounding variable, or if it reflects more specific aspects of the training. In sum, there is initial support for the idea that the musician advantage in emotion recognition is not an artefact of general cognitive advantages. Nevertheless, the inclusion of cognitive measures needs to become standard practice, and the precise role of different domain-general processes will need to be delineated in future research.

Studies on the neural correlates of vocal-emotional processing in musicians are rare, and they will be crucial to inform the debate on the mechanisms underlying the association between musical experience and vocal emotions. Initial evidence points to an important role of auditory-perceptual enhancements. For example, Strait et al. (2009) recorded brainstem potentials of musicians and non-musicians in response to an infant's crying sound. Musicians exhibited enhanced subcortical timing, as well as enhanced representations of frequencies important for pitch and timbre perception. In an event-related potential study, Pinheiro and colleagues (2015) found that music training relates to distinct responses to emotional prosody in two components that are markers of early sensory processing. Specifically, musicians showed reduced P50 amplitude compared to untrained listeners in response to prosodic emotions in emotionally neutral semantic sentences. For both P50 and N100, musicians showed similar amplitudes regardless of stimulus type (prosodic stimuli with or without semantic content), whereas untrained listeners showed more negative amplitudes for stimuli without semantic content. These two studies indicate that neurocognitive pathways for processing music and vocal emotions may overlap at an early sensory level, and that this overlap may contribute to musician advantages in emotion recognition. A similar view has been used to account for transfer effects from music to speech processing (Patel, 2011; 2014). The role of auditory-perceptual enhancements in emotion-recognition tasks remains poorly

understood. In one recent study, music-perception abilities, as assessed through a range of

tasks (e.g., musical beat perception, pitch discrimination), correlated positively with emotion

recognition in prosody and nonverbal vocalizations (Correia et al., 2020). Importantly, when

these perceptual abilities were held constant, the association between music training and

vocal emotion recognition was no longer significant.

Whether music training also relates to emotion recognition at higher-order levels of

socioemotional processing is another avenue for future studies. On the one hand, the lack of

associations between training and visual-emotion recognition suggests that this might not be

the case. On the other hand, in an fMRI study, Park and colleagues (2015) found that

musicians show enhanced responses to sad prosody, though not for happy or fear prosody, in

regions involved in general socioemotional processing, including the medial prefrontal and

anterior cingulate cortices.

Longitudinal evidence

Central to our understanding of music-training effects is the issue of causality, which

can only be established with longitudinal data, random assignment, and appropriate (active)

control groups (Schellenberg, 2020). We identified only four longitudinal studies that

included emotion-recognition measures (Figure 1 and Table S2). One was conducted with

typically developing children, and the children were tested only once on the emotion

recognition task, such that the design was not truly longitudinal in this regard (Thompson et

al., 2004). The remaining three were conducted with deaf children or adults with cochlear

implants (CIs; Chari et al., 2019; Fuller et al., 2018; Good et al., 2017). All of these studies

focused on emotional-prosody recognition, although Good and colleagues (2017) also

included an audio-visual condition that combined prosody with facial expressions.

Thompson and colleagues (2004) recruited 43 6-year-old children, who had

previously been assigned to one year of keyboard, vocal, drama or no lessons groups ($n \approx 10$

per group), and examined their emotion-recognition abilities. In other words, the children were not assessed on the experimental task before training. In a two-alternative forced choice task, children decided whether prosodic stimuli were happy or sad in one condition, and angry or fearful in a second condition. The groups did not differ in IQ before training, but there were differences after training for the recognition of angry/fearful stimuli, with the music groups performing better than the no lessons group. This advantage raises the possibility of a training benefit, but the music groups did not differ from the drama group, and no effects were observed for happy/sad trials. Moreover, in separate analyses per group, only the keyboard group, not the singing one, significantly outperformed children who had no lessons. These findings thus provide only limited evidence for an effect of music training on emotional prosody recognition, and they do not establish causality because no pre-training emotion-recognition data were available.

The three studies conducted with CI users included random assignment and active control groups, but they all reported null results. Good and colleagues (2017) compared two groups of nine children with CIs, who completed either six months of music or visual-arts training. The two groups were tested for music-perception skills, and for emotion recognition in prosody alone and in audio-visual clips. While music perception skills improved significantly more in the music compared to the visual arts group, emotion recognition did not. Emotion-recognition performance was generally better after training, but the interaction between time (pre/post) and group was not significant. In a different study, Fuller and colleagues (2018) randomly assigned 19 adult CI users to one of three groups that completed six 2-hour training sessions: a pitch/timbre group ($n = 6$), a music-therapy group ($n = 7$), and an active control group that involved activities such as writing, cooking, and woodworking ($n = 6$). The three groups were tested for several aspects of speech perception, melodic-contour identification, quality of life, and emotional-prosody recognition in pseudowords. For

melodic-contour identification, the pitch/timbre group improved significantly more than the music-therapy and control groups. For emotion recognition and for the remaining measures, however, no benefits of training were found. There was a small improvement for the music-therapy group only, but there was no interaction between time and group. Finally, Chari and colleagues (2019) randomly allocated 18 participants to one of three groups that completed four weeks of training: an auditory-motor group ($n = 7$), an auditory-only group ($n = 7$), and a passive control group ($n = 4$). Participants were tested for speech perception, pitch discrimination, melodic-contour identification, and emotional-prosody recognition in sentences. For melodic-contour identification, the auditory-motor group improved significantly more than the other groups, but no benefits were observed for the remaining measures, including emotion recognition.

In short, evidence for a causal role for music training in emotion recognition is currently non-existent. Nevertheless, the null results cannot be taken as evidence for the absence of effects either, because of the limited number of studies and the small sample sizes.

Music training vs. pre-existing factors

Because associations between music training and vocal-emotion recognition are reported primarily from cross-sectional comparisons, it is possible that they stem from pre-existing factors. Cross-sectional studies involving highly trained musicians are valuable because they offer a model for examining the correlates of training when expertise and experience are highest. But they do not allow us to tease apart environmental from genetic contributions, and often they do not account for other relevant confounding variables. Cognitive abilities, but also personality and socio-economic variables, might determine who takes music lessons, and this possibility is rarely considered in cross-sectional comparisons (Swaminathan & Schellenberg, 2018; for a review, Schellenberg, 2020). Importantly, musical abilities, the propensity for music training, and associations between music training and

musical abilities, have a genetic component (for a review, Mosing & Ullén, 2016). Twin

studies indicate that individual differences in practicing music are heritable to a substantial

degree (40%-70%), and when genetic predispositions are held constant, practice is no longer

associated with better musical abilities (Mosing et al., 2014). In the same vein, associations

between music training and IQ, often interpreted as transfer, appear to stem from shared

genetic influences that determine practice behavior *and* cognitive performance (Mosing et al.,

2016).

In short, samples of trained musicians differ from untrained individuals in music

training, but they are also likely to differ in genetic and other environmental factors that

affect our measures of interest. A promising line of research is starting to delineate the

potential role of pre-existing factors in associations between musical and non-musical

domains by asking whether musician-like advantages can be found in untrained individuals.

These studies can be done by identifying individuals who have 'naturally' good musical

abilities. Such good abilities, in the absence of formal training, must be a consequence of

genetics and/or informal engagement with music. For example, Mankel and Bidelman (2018)

found that untrained adults with higher music-perception abilities show a more efficient

neural encoding of noise-degraded speech sounds, mirroring the benefits previously observed

in musicians. Similarly, Swaminathan and Schellenberg (2017) found that better music-

perception abilities are associated with enhanced phoneme perception in a foreign language,

regardless of music training. In a study of 6- to 9-year-old children, Swaminathan and

Schellenberg (2019) documented an association between natural musical abilities and two

aspects of language ability: phoneme perception and grammar. As for emotion-recognition

abilities, Correia and colleagues (2020) found that both self-report and performance-based

measures of music-perception skills correlate positively with emotion recognition in prosody

and nonverbal vocalizations, even with music training held constant. Crucially, untrained

participants with high musical abilities were as good as trained musicians at recognizing both types of vocal expressions.

These response patterns indicate that training might not be a *necessary* condition for advantages in vocal-emotion recognition to emerge. Although cross-sectional differences are often attributed to training (Schellenberg, 2019), similar advantages can be seen in listeners without any training. In addition to highlighting the need to interpret cross-sectional evidence with caution, these findings also point to the multifaceted nature of musical expertise. While the musical expertise literature typically equates musical abilities with formal music training, an exciting avenue of research will be to consider the potential effects of other forms of engagement with music (Müllensiefen et al., 2014), as well as interactions with genetic influences. Such an approach will lead to a more complete understanding of associations between music and emotion recognition.

From emotion recognition to broader aspects of socioemotional functioning

An interesting question is whether music training relates to emotion recognition only, or whether it also relates to broader aspects of socioemotional competence, such as empathy, emotion regulation, mentalizing, or social adjustment. Most studies focusing on these aspects are longitudinal studies with children, and a diversity of measures have been considered, such as general adaptive and maladaptive behavior (Gerry et al., 2012; Schellenberg, 2004, 2006), self-esteem (Costa-Giomi, 2004; Rickart et al., 2013; Rickart et al., 2012), empathy (Rabinowitch et al., 2012), motivation (Degé & Schwarzer, 2017; Rickard et al., 2012), and socioemotional wellbeing (Kim & Kim, 2017; Rose et al., 2017; Welch et al., 2014).

Evidence for positive effects of music training is currently weak and mixed. Schellenberg and colleagues (2015), for instance, tested whether 8- and 9-year-old children improved in prosocial skills after 10 months of music lessons, in comparison with children who took no lessons. Music training led to improvements in sympathy and prosocial

behaviors, but only in children who had poor prosocial skills before training. In a prior study with 6-year-olds, Schellenberg (2004) found that drama lessons improved adaptive social behavior, but music training (including keyboard and vocal training) did not. Similarly, music lessons were not related to social behavior in a later cross-sectional study with 6- to 11-year-old children (Schellenberg, 2006).

Mixed evidence for training effects has also been found for measures of self-esteem. In a study with 9-year-olds, Costa-Giomi (2004) found that individual piano lessons for three years did not improve self-esteem more than no lessons, even though there was a trend-level advantage. By contrast, Rickard and colleagues (2013) found that school-based music classes prevented the decline in self-esteem measures that was observed in the control group, in children from two age cohorts ($\approx$6-year-olds and $\approx$9-year-olds). The different response patterns may stem from the fact that music training was group-based in Rickard and colleagues (2013), and individual in Costa-Giomi (2004). In fact, most studies reporting positive effects of music training on socioemotional skills relied on group-based programs (e.g., Gerry et al., 2012; Schellenberg et al., 2015). Group-based programs arguably emphasize the engagement of the social functions of music (Koeslch, 2014), which could allow for socioemotional effects to emerge. It is crucial to examine systematically whether differential results emerge depending on the specific type of music training (group-based vs. individual), and whether benefits observed in emotion-recognition abilities relate to other socioemotional skills and to everyday outcomes of social functioning.

Concluding remarks

The reviewed studies suggest an association between music training and an enhanced ability to recognize emotions in vocal expressions. This association has been shown for a variety of vocal stimuli, including emotional prosody and purely nonverbal vocalizations. Initial evidence further indicates that the advantage might not simply result from better

general cognitive abilities. Nevertheless, it appears to be restricted to audition because it does not extend to the visual domain. This conclusion remains tentative, however, because only a few studies have examined emotional stimuli other than voices. Finding associations between music training and socioemotional processing highlights the importance of including socioemotional processes as part of the quest to identify links between musical and non-musical domains, which have focused primarily on speech-perception and general-cognitive abilities.

A crucial goal of future research is to delineate the precise mechanisms of the musician advantage in vocal-emotion recognition. There is some evidence that auditory-perceptual enhancements might play a role, which is consistent with the idea of overlapping sensory pathways for processing music and vocal-emotional information. Whether this overlap extends to higher-order stages of socioemotional processing remains unclear. EEG and fMRI studies will be useful to address these issues, by helping to determine *when* and *where* in the brain these cross-domain interactions occur. Importantly, the causal direction of the effect also remains to be determined. It has been assumed that music training *causes* emotion-recognition enhancements, but the available longitudinal evidence for this is non-existent. The issue of causality is particularly relevant in light of evidence that musical abilities and training have a genetic component, and that musician-like advantages in emotion recognition can be seen in untrained individuals with naturally good musical abilities. Well-powered and well-designed longitudinal studies, which include random assignment, an active control group, and careful assessment of confounding variables (e.g., socio-economic status, personality), are needed to address the issue of causality. Such studies will improve our understanding of the associations between music and socioemotional processing, and inform debates on music-based interventions in clinical, community, and educational contexts.

References

Amorim, M., Anikin, A., Mendes, A. J., Lima, C. F., Kotz, S. A., & Pinheiro, A. P. (2019).

Changes in vocal emotion recognition across the life span. *Emotion*. Advance online

publication. https://doi.org/10.1037/emo0000692

Balkwill, L. L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception

of emotion in music: Psychophysical and cultural cues. *Music Perception, 17*(1), 43-64.

https://doi.org/10.2307/40285811

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of

Personality and Social Psychology, 70*(3), 614-636. https://doi.org/10.1037/0022-

3514.70.3.614

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2012). Introducing the Geneva Multimodal

expression corpus for experimental research on emotion perception. *Emotion, 12*(5),

1161–1179. https://doi.org/10.1037/a0025827

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A

taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637.

https://doi.org/10.1037/0033-2909.128.4.612

Başkent, D., Fuller, C. D., Galvin III, J. J., Schepel, L., Gaudrain, E., & Free, R. H. (2018).

Musician effect on perception of spectro-temporally degraded speech, vocal emotion,

and music in young adolescents. *The Journal of the Acoustical Society of America,

143*(5), EL311-EL316. https://doi.org/10.1121/1.5034489

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech:

Common processing, attention, and memory. *Frontiers in Psychology, 2*, 94.

https://doi.org/10.3389/fpsyg.2011.00094

Bhatara, A., Tirovolas, A. K., Duan, L. M., Levy, B., & Levitin, D. J. (2011). Perception of

emotional expression in musical performance. *Journal of Experimental Psychology:*

Human Perception and Performance, 37(3), 921–934.

https://doi.org/10.1037/a0021922

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional

scaling of emotional responses to music: The effect of musical expertise and of the

duration of the excerpts. *Cognition & Emotion, 19*(8), 1113-1139.

https://doi.org/10.1080/02699930500204250

Blair, R. J. R., & Coles, M. (2000). Expression recognition and behavioural problems in early

adolescence. *Cognitive Development, 15*(4), 421-434. https://doi.org/10.1016/S0885-

2014(01)00039-9

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

Bowen, E., & Nowicki, S. (2007). The nonverbal decoding ability of children exposed to

family violence or maltreatment: Prospective evidence from a British cohort. *Journal of

Nonverbal Behavior, 31*(3), 169-184. https://doi.org/10.1007/s10919-007-0030-x

Bugos, J. A., & DeMarie, D. (2017). The effects of a short-term music program on preschool

children's executive functions. *Psychology of Music, 45*(6), 855-867.

https://doi.org/10.1177/0305735617692666

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship

well-being in adults. *Journal of Nonverbal Behavior, 23*(1), 91-100.

https://doi.org/10.1023/A:1021339410262

Castro, S. L., & Lima, C. F. (2014). Age and musical expertise influence emotion recognition

in music. *Music Perception, 32*(2), 125-142. https://doi.org/10.1525/mp.2014.32.2.125

Chari, D. A., Barrett, K. C., Patel, A. D., Colgrove, T. R., Jiradejvong, P., Jacobs, L. Y., &

Limb, C. J. (2019). Impact of auditory-motor musical training on melodic pattern

recognition in cochlear implant users. *Otology & Neurotology, 41*(4), 422-431.

https://doi.org/10.1097/MAO.0000000000002525

Chobert, J., François, C., Velay, J., & Besson, M. (2014). Twelve months of active musical

training in 8- to 10-year-old children enhances the preattentive processing of syllabic

duration and voice onset time. *Cerebral Cortex, 24*(4), 956-967.

http://doi.org/10.1093/cercor/bhs377

Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-

cultural recognition of vocal emotion during childhood and adolescence. *Scientific

Reports, 8*, 8659. https://doi.org/10.1038/s41598-018-26889-1

Clark, C. N., Downey, L. E., & Warren, J. D. (2014). Music biology: all this useful beauty.

Current Biology, 24(6), R234-R237. https://doi.org/10.1016/j.cub.2014.02.013

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. J., Chepesiuk, A. M. P., Herholz, S. C., Baillet, S., & Zatorre, R. J. (2017a).

Neural correlates of early sound encoding and their relationship to speech-in-noise

perception. *Frontiers in Neuroscience, 11,* 479.

https://doi.org/10.3389/fnins.2017.00479

Coffey, E. B., Mogilever, N. B., & Zatorre, R. J. (2017b). 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*, 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. (2020). Enhanced recognition of vocal emotions in individuals with naturally

good musical abilities. *Emotion*. Advance online publication.

https://doi.org/10.1037/emo0000770

Costa-Giomi, E. (2004). Effects of three years of piano instruction on children's academic

achievement, school performance and self-esteem. *Psychology of Music, 32*(2), 139-

152. https://doi.org/10.1177/0305735604041491

Curtis, M. E., & Bharucha, J. J. (2010). The minor third communicates sadness in speech,

mirroring its use in music. *Emotion, 10*(3), 335. https://doi.org/10.1037/a0017928

Dawel, A., O'Kearney, R., McKone, E., & Palermo, R. (2012). Not just fear and sadness:

Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal

expressions in psychopathy. *Neuroscience & Biobehavioral Reviews, 36*(10), 2288-

2304. https://doi.org/10.1016/j.neubiorev.2012.08.006

Degé, F., & Schwarzer, G. (2017). The influence of an extended music curriculum at school

on academic self-concept in 9-to 11-year-old children. *Musicae Scientiae, 22*(3), 305-

321. https://doi.org/10.1177/1029864916688508

Dibben, N., Coutinho, E., Vilar, J. A., & Estévez-Pérez, G. (2018). Do individual differences

influence moment-by-moment reports of emotion perceived in music and speech

prosody?. *Frontiers in Behavioral Neuroscience, 12*, 184.

https://doi.org/10.3389/fnbeh.2018.00184

Dmitrieva, E. S., Gel'man, V. Y., Zaitseva, K. A., & Orlov, A. M. (2006). Ontogenetic

features of the psychophysiological mechanisms of perception of the emotional

component of speech in musically gifted children. *Neuroscience and Behavioral

Physiology, 36*(1), 53-62. https://doi.org/10.1007/s11055-005-0162-6

Downey, L. E., Blezat, A., Nicholas, J., Omar, R., Golden, H. L., Mahoney, C. J., Crutch, S.

   J., & Warren, J. D. (2013). Mentalising music in frontotemporal dementia. *Cortex,*

   49(7), 1844-1855. https://doi.org/10.1016/j.cortex.2012.09.011

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Changes

   in grey matter induced by training. *Nature, 427*(6972), 311-312.

   https://doi.org/10.1038/427311a

Dumont, E., Syurina, E. V., Feron, F. J., & van Hooren, S. (2017). Music interventions and

   child development: A critical review and further directions. *Frontiers in Psychology, 8*,

   1694. https://doi.org/10.3389/fpsyg.2017.01694

Ethofer, T., Bretscher, J., Wiethoff, S., Bisch, J., Schlipf, S., Wildgruber, D., & Kreifelts, B.

   (2013). Functional responses and structural connections of cortical areas for processing

   faces and voices in the superior temporal sulcus. *Neuroimage, 76*, 45-56.

   https://doi.org/10.1016/j.neuroimage.2013.02.064

Farmer, E., Jicol, C., & Petrini, K. (2020). Musicianship Enhances Perception But Not

   Feeling of Emotion From Others' Social Interaction Through Speech Prosody. *Music*

   Perception, 37(4), 323-338. https://doi.org/10.1525/mp.2020.37.4.323

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I.

   (1989). A cross-language study of prosodic modifications in mothers' and fathers'

   speech to preverbal infants. *Journal of Child Language, 16*(3), 477-501.

   https://doi.org/10.1017/S0305000900010679

Frey, A., François, C., Chobert, J., Velay, J., Habib, M., & Besson, M. (2019). Music training

   positively influences the preattentive perception of voice onset time in children with

   dyslexia: A longitudinal study. *Brain Sciences, 9*(91), 1-24.

   https://doi.org/10.3390/brainsci9040091

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A. D., &
Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current
Biology, 19*(7), 573-576. https://doi.org/10.1016/j.cub.2009.02.058

Frühholz, S., Trost, W., & Kotz, S. A. (2016). The sound of emotions—Towards a unifying
neural network perspective of affective sound processing. *Neuroscience &
Biobehavioral Reviews, 68*, 96-110. https://doi.org/10.1016/j.neubiorev.2016.05.002

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

Fuller, C. D., Galvin III, J. J., Maat, B., Free, R. H., & Başkent, D. (2014). The musician
effect: does it persist under degraded pitch conditions of cochlear implant simulations?.
Frontiers in Neuroscience, 8, 179. https://doi.org/10.3389/fnins.2014.00179

Gendron, M., Crivelli, C., & Barrett, L. F. (2018). Universality reconsidered: Diversity in
making meaning of facial expressions. *Current Directions in Psychological Science,
27*(4), 211-219. https://doi.org/10.1177/0963721417746794

Gerry, D., Unrau, A., & Trainor, L. J. (2012). Active music classes in infancy enhance
musical, communicative and social development. *Developmental Science, 15*(3), 398-
407. https://doi.org/10.1111/j.1467-7687.2012.01142.x

Gery, I., Miljkovitch, R., Berthoz, S., & Soussignan, R. (2009). Empathy and recognition of
facial expressions of emotion in sex offenders, non-sex offenders and normal controls.
Psychiatry Research, 165(3), 252-262. https://doi.org/10.1016/j.psychres.2007.11.006

Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., & Baron-Cohen,
S. (2010). Enhancing emotion recognition in children with autism spectrum conditions:
An intervention using animated vehicles with real emotional faces. *Journal of Autism*

and Developmental Disorders, 40(3), 269-279. https://doi.org/10.1007/s10803-009-0862-9

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. https://doi.org/10.1097/AUD.0000000000000402

Grandjean, D. (2020). Brain networks of emotional prosody processing. *Emotion Review*. https://doi.org/10.1177/1754073919898522

Grau-Sánchez, J., Münte, T. F., Altenmüller, E., Duarte, E., & Rodríguez-Fornells, A. (2020). Potential benefits of music playing in stroke upper limb motor rehabilitation. *Neuroscience & Biobehavioral Reviews, 112*, 585-599. https://doi.org/10.1016/j.neubiorev.2020.02.027

Habib, M., & Besson, M. (2009). What do music training and musical experience teach us about brain plasticity?. *Music Perception, 26*(3), 279-285. https://doi.org/10.1525/mp.2009.26.3.279

Habibi, A., Cahn, B. R., Damasio, A., & Damasio, H. (2016). Neural correlates of accelerated auditory processing in children engaged in music training. *Developmental Cognitive Neuroscience, 21*, 1-14. https://doi.org/10.1016/j.dcn.2016.04.003

Hall, J. A., Andrzejewski, S. A., & Yopchick, J. E. (2009). Psychosocial correlates of interpersonal sensitivity: A meta-analysis. *Journal of Nonverbal Behavior, 33*(3), 149-180. https://doi.org/10.1007/s10919-009-0070-5

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

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception, 23*(4), 319-330. https://doi.org/10.1525/mp.2006.23.4.319

Johnston, P., Mayes, A., Hughes, M., & Young, A. W. (2013). Brain networks subserving the evaluation of static and dynamic facial expressions. *Cortex, 49*(9), 2462-2472. https://doi.org/10.1016/j.cortex.2013.01.002

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

Kim, H. S., & Kim, H. S. (2017). Effect of a musical instrument performance program on emotional intelligence, anxiety, and aggression in Korean elementary school children. *Psychology of Music, 46*(3), 440-453. https://doi.org/10.1177/0305735617729028

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

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

Kreifelts, B., Jacob, H., Brück, C., Erb, M., Ethofer, T., & Wildgruber, D. (2013). Non-verbal emotion communication training induces specific changes in brain function and structure. *Frontiers in Human Neuroscience, 7*, 648. https://doi.org/10.3389/fnhum.2013.00648

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24*(8), 1377-1388. https://doi.org/10.1080/02699930903485076

Laukka, P., & Elfenbein, H. A. (2020). Cross-Cultural Emotion Recognition and In-Group

    Advantage in Vocal Expression: A Meta-Analysis. *Emotion Review*.

    https://doi.org/10.1177/1754073919897295

Lima, C. F., Alves, T., Scott, S. K., & Castro, S. L. (2014). In the ear of the beholder: how

    age shapes emotion processing in nonverbal vocalizations. *Emotion, 14*(1), 145.

    https://doi.org/10.1037/a0034287

Lima, C. F., Anikin, A., Monteiro, A. C., Scott, S. K., & Castro, S. L. (2019). Automaticity in

    the recognition of nonverbal emotional vocalizations. *Emotion, 19*(2), 219.

    https://doi.org/10.1037/emo0000429

Lima, C. F., Brancatisano, O., Fancourt, A., Müllensiefen, D., Scott, S. K., Warren, J. D., &

    Stewart, L. (2016a). Impaired socio-emotional processing in a developmental music

    disorder. *Scientific Reports, 6*, 34911. https://doi.org/10.1038/srep34911

Lima, C. F., & Castro, S. L. (2011). Emotion recognition in music changes across the adult

    life span. *Cognition and Emotion, 25*(4), 585-598.

    https://doi.org/10.1080/02699931.2010.502449

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., Krishnan, S., & Scott, S. K. (2016b). Roles of supplementary motor areas in

    auditory processing and auditory imagery. *Trends in Neurosciences, 39*(8), 527-542.

    https://doi.org/10.1016/j.tins.2016.06.003

Lonsdale, A. J., & North, A. C. (2011). Why do we listen to music? A uses and gratifications

    analysis. *British Journal of Psychology, 102*(1), 108-134.

    https://doi.org/10.1348/000712610X506831

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

Marques, C., Moreno, S., Castro, S.L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience, 19*(9), 1453-1463. https://doi.org/10.1162/jocn.2007.19.9.1453

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion, 3*(9), 7–105. https://doi.org/10.1037/1528-3542.3.1.97

McGettigan, C., Walsh, E., Jessop, R., Agnew, Z. K., Sauter, D. A., Warren, J. E., & Scott, S. K. (2015). Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cerebral Cortex, 25*(1), 246-257. https://doi.org/10.1093/cercor/bht227

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

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

Mosing, M. A., Madison, G., Pedersen, N. L., Kuja-Halkola, R., & Ullén, F. (2014). Practice

does not make perfect: no causal effect of music practice on music ability.

Psychological Science, 25(9), 1795-1803. https://doi.org/10.1177/0956797614541990

Mosing, M. A., Madison, G., Pedersen, N. L., & Ullén, F. (2016). Investigating cognitive

transfer within the framework of music practice: Genetic pleiotropy rather than

causality. *Developmental Science, 19*(3), 504-512. https://doi.org/10.1111/desc.12306

Mosing, M. A., & Ullén, F. (2016). Genetic influences on musical giftedness, talent, and

practice. In Gary E. McPherson (Ed.), *Musical Prodigies: Interpretations from*

Psychology, Education, Musicology, and Ethnomusicology (pp. 156-167). Oxford

University Press.

Mualem, O., & Lavidor, M. (2015). Music education intervention improves vocal emotion

recognition. *International Journal of Music Education, 33*(4), 413-425.

https://doi.org/10.1177/0255761415584292

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of

nonmusicians: An index for assessing musical sophistication in the general population.

PLOS ONE, 9, e89642. https://doi.org/10.1371/journal.pone.0089642

Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing.

Infant Behavior and Development, 27(4), 455-464.

https://doi.org/10.1016/j.infbeh.2004.03.002

Nilsonne, Å., & Sundberg, J. (1985). Differences in ability of musicians and nonmusicians to

judge emotional state from the fundamental frequency of voice samples. *Music*

Perception, 2(4), 507-516. https://doi.org/10.2307/40285316

North, A. C., & Hargreaves, D. J. O Neill, SA (2000). The importance of music to

adolescents. *British Journal of Educational Psychology, 70*, 255-272.

https://doi.org/10.1348/000709900158083

O'Nions, E., Lima, C. F., Scott, S. K., Roberts, R., McCrory, E. J., & Viding, E. (2017).

Reduced laughter contagion in boys at risk for psychopathy. *Current Biology, 27*(19),

3049-3055. https://doi.org/10.1016/j.cub.2017.08.062

Pantev, C., & Herholz, S. C. (2011). Plasticity of the human auditory cortex related to

musical training. *Neuroscience & Biobehavioral Reviews, 35*(10), 2140-2154.

https://doi.org/10.1016/j.neubiorev.2011.06.010

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*,

1049. https://doi.org/10.3389/fnhum.2014.01049

Parsons, C. E., Young, K. S., Jegindø, E. M. E., Vuust, P., Stein, A., & Kringelbach, M. L.

(2014). Music training and empathy positively impact adults' sensitivity to infant

distress. *Frontiers in Psychology, 5*, 1440. https://doi.org/10.3389/fpsyg.2014.01440

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The

OPERA hypothesis. *Frontiers in psychology, 2*, 142.

https://doi.org/10.3389/fpsyg.2011.00142

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

Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., & Rigoulot, S. (2015).

Preferential decoding of emotion from human non-linguistic vocalizations versus

speech prosody. *Biological Psychology, 111*, 14-25.

https://doi.org/10.1016/j.biopsycho.2015.08.008

Pinheiro, A. P., Vasconcelos, M., Dias, M., Arrais, N., & Gonçalves, Ó. F. (2015). The music

of language: an ERP investigation of the effects of musical training on emotional

prosody processing. *Brain and Language, 140*, 24-34.

https://doi.org/10.1016/j.bandl.2014.10.009

Putkinen, V., Tervaniemi, M., & Huotilainen, M. (2019). Musical playschool activities are

linked to faster auditory development during preschool-age: a longitudinal ERP study.

Scientific Reports, 9, 11310. https://doi.org/10.1038/s41598-019-47467-z

Rabinowitch, T. C., Cross, I., & Burnard, P. (2012). Long-term musical group interaction has

a positive influence on empathy in children. *Psychology of Music, 41*(4), 484-498.

https://doi.org/10.1177/0305735612440609

Rickard, N. S., Appelman, P., James, R., Murphy, F., Gill, A., & Bambrick, C. (2013).

Orchestrating life skills: The effect of increased school-based music classes on

children's social competence and self-esteem. *International Journal of Music

Education, 31*(3), 292-309. https://doi.org/10.1177/0255761411434824

Rickard, N. S., Bambrick, C. J., & Gill, A. (2012). Absence of widespread psychosocial and

cognitive effects of school-based music instruction in 10–13-year-old students.

International Journal of Music Education, 30(1), 57-78.

https://doi.org/10.1177/0255761411431399

Rochas, V., Gelmini, L., Krolak-Salmon, P., Poulet, E., Saoud, M., Brunelin, J., & Bediou, B.

(2013). Disrupting pre-SMA activity impairs facial happiness recognition: an event-

related TMS study. *Cerebral Cortex, 23*(7), 1517-1525.

https://doi.org/10.1093/cercor/bhs133

Roden, I., Grube, D., Bongard, S., & Kreutz, G. (2014). Does music training enhance

    working memory performance? Findings from a quasi-experimental longitudinal study.

    Psychology of Music, 42(2), 284-298. https://doi.org/10.1177/0305735612471239

Rose, D., Jones Bartoli, A., & Heaton, P. (2017). Measuring the impact of musical learning

    on cognitive, behavioural and socio-emotional wellbeing development in children.

    Psychology of Music, 47(2), 284-303. https://doi.org/10.1177/0305735617744887

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review

    of emotion recognition and aging: Implications for neuropsychological models of

    aging. *Neuroscience & Biobehavioral Reviews, 32*(4), 863-881.

    https://doi.org/10.1016/j.neubiorev.2008.01.001

Sala, G., & Gobet, F. (2017a). Does far transfer exist? Negative evidence from chess, music,

    and working memory training. *Current Directions in Psychological Science, 26*, 515-

    520. https://doi.org/10.1177/0963721417712760

Sala, G., & Gobet, F. (2017b). 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. (2020). Cognitive and academic benefits of music training with

    children: A multilevel meta-analysis. Preprint. https://doi.org/10.31234/osf.io/7s8wr

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of

    basic emotions through nonverbal emotional vocalizations. *Proceedings of the National*

    Academy of Sciences, 107(6), 2408-2412. https://doi.org/10.1073/pnas.0908239106

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2015). Emotional vocalizations are

    recognized across cultures regardless of the valence of distractors. *Psychological*

    Science, 26(3), 354-356. https://doi.org/10.1177/0956797614560771

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15*(8), 511-514. https://doi.org/10.1111/j.0956-7976.2004.00711.x

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

Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *British Journal of Psychology, 102*, 283-302. https://doi.org/10.1111/j.2044- 8295.2010.02000.x
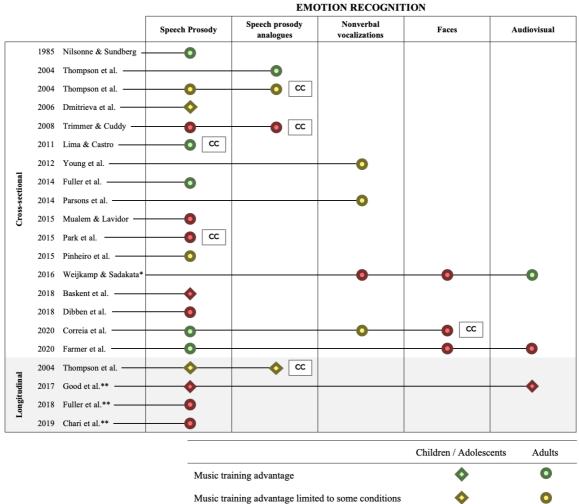
Schellenberg, E. G. (2019). Correlation= causation? Music training, psychology, and neuroscience. *Psychology of Aesthetics, Creativity, and the Arts. Advance online publication.* https://doi.org/10.1037/aca0000263

Schellenberg, E. G. (2020). Music Training, Individual Differences, and Plasticity. In M. S. C. Thomas, D. Mareschal, & I. Dumontheil (Eds.), *Educational Neuroscience: Development Across the Life Span* (1st ed. pp. 413-438). Routledge. https://doi.org/10.4324/9781003016830

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

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*(3), 235-248. https://doi.org/10.1016/s0892-1997(05)80231-0

Scherer, K. R., & Scherer, U. (2011). Assessing the ability to recognize facial and vocal expressions of emotion: Construction and validation of the Emotion Recognition Index. *Journal of Nonverbal Behavior, 35*(4), 305. https://doi.org/10.1007/s10919-011-0115-4

Schirmer, A. (2018). Is the voice an auditory face? An ALE meta-analysis comparing vocal

and facial emotion processing. *Social Cognitive and Affective Neuroscience, 13*(1), 1-

13. https://doi.org/10.1093/scan/nsx142

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

Schlegel, K., Vicaria, I. M., Isaacowitz, D. M., & Hall, J. A. (2017). Effectiveness of a short

audiovisual emotion recognition training program in adults. *Motivation and Emotion,

41*(5), 646-660. https://doi.org/10.1007/s11031-017-9631-9

Scott, S. K., Sauter, D. & McGettigan, C. (2010). Brain mechanisms for processing perceived

emotional vocalizations in humans In *Handbook of mammalian vocalization: an

integrative neuroscience approach* (ed. Brudzynski, S. M.) 187–198. London:

Academic Press.

Stevens, D., Charman, T., & Blair, R. J. R. (2001). Recognition of emotion in facial

expressions and vocal tones in children with psychopathic tendencies. *The Journal of

Genetic Psychology, 162*(2), 201-211. https://doi.org/10.1080/00221320109597961

Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural

efficiency–effects of training on subcortical processing of vocal expressions of

emotion. *European Journal of Neuroscience, 29*(3), 661-668.

https://doi.org/10.1111/j.1460-9568.2009.06617.x

Swaminathan, S., & Schellenberg, E. G. (2015). Current emotion research in music

psychology. *Emotion Review, 7*(2), 189-197.

https://doi.org/10.1177/1754073914558282

Swaminathan, S., & Schellenberg, E. G. (2017). Musical competence and phoneme

perception in a foreign language. *Psychonomic Bulletin & Review, 24*(6), 1929-1934.

https://doi.org/10.3758/s13423-017-1244-5

Swaminathan, S., & Schellenberg, E. G. (2018). Musical competence is predicted by music

training, cognitive abilities, and personality. *Scientific Reports, 8*, 9223.

https://doi.org/10.1038/s41598-018-27571-2

Swaminathan, S., & Schellenberg, E. G. (2019). Musical ability, music training, and language

ability in childhood. *Journal of Experimental Psychology: Learning, Memory, and

Cognition*. Advance online publication. https://doi.org/10.1037/xlm0000798

Swaminathan, S., Schellenberg, E. G., & Khalil, S. (2017). Revisiting the association

between music lessons and intelligence: Training effects or music aptitude?.

Intelligence, 62, 119-124. https://doi.org/10.1016/j.intell.2017.03.005

Tarr, B., Launay, J., & Dunbar, R. I. (2014). Music and social bonding: "self-other" merging

and neurohormonal mechanisms. *Frontiers in Psychology, 5*, 1096.

https://doi.org/10.3389/fpsyg.2014.01096

Thompson, W. F. 2009. *Music, thought, and feeling: Understanding the psychology of music*.

New York: Oxford University Press.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do

music lessons help?. *Emotion, 4*(1), 46. https://doi.org/10.1037/1528-3542.4.1.46

Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion,

8*(1), 81. https://doi.org/10.1037/1528-3542.8.1.81

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience, 6*(7),

669-673. https://doi.org/10.1038/nn1084

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

Trost, W. J., Labbé, C., & Grandjean, D. (2017). Rhythmic entrainment as a musical affect induction mechanism. *Neuropsychologia, 96*, 96-110. https://doi.org/10.1016/j.neuropsychologia.2017.01.004

Vilaverde, R. F., Correia, A. I., & Lima, C. F. (2020). Higher trait mindfulness is associated with empathy but not with emotion recognition abilities. *Royal Society Open Science, 7,* 192077. https://doi.org/10.1098/rsos.192077

Weijkamp, J., & Sadakata, M. (2016). Attention to affective audio-visual information: Comparison between musicians and non-musicians. *Psychology of Music, 45*(2), 204-215. https://doi.org/10.1177/0305735616654216

Welch, G. F., Himonides, E., Saunders, J., Papageorgi, I., & Sarazin, M. (2014). Singing and social inclusion. *Frontiers in Psychology, 5*, 803. https://doi.org/10.3389/fpsyg.2014.00803

Woollett, K., Spiers, H.J., and Maguire, E.A. (2009). Talent in the taxi: a model system for exploring expertise. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364* (1522), 1407–1416. https://doi.org/10.1098/rstb.2008.0288

Young, A. W., Frühholz, S., & Schweinberger, S. R. (2020). Face and voice perception: Understanding commonalities and differences. *Trends in Cognitive Sciences, 24*(5), 398-410. https://doi.org/10.1016/j.tics.2020.02.001

Young, K. S., Parsons, C. E., Stein, A., & Kringelbach, M. L. (2012). Interpreting infant vocal distress: The ameliorative effect of musical training in depression. *Emotion, 12*(6), 1200. https://doi.org/10.1037/a0028705

Zhang, J. D., Susino, M., McPherson, G. E., & Schubert, E. (2018). The definition of a

    musician in music psychology: A literature review and the six-year rule. *Psychology of*

    Music, 48(3), 389-409. https://doi.org/10.1177/0305735618804038

Figure 1. Studies on emotion recognition in musically trained individuals, organized according to design (cross-sectional or longitudinal) and stimulus type (speech prosody, speech prosody analogues, nonverbal vocalizations, faces, and audio-visual). Studies with adults are indicated by circles, whereas those with children are represented by rhombuses. Green symbols indicate studies in which musically trained individuals presented a performance advantage relative to those with no music training; yellow symbols indicate studies in which the music training advantage was limited to some conditions (e.g., one out of several emotions); and red symbols indicate studies in which no benefits of music training were found. Studies that confirmed that emotion recognition advantages could not be accounted for by domain-general cognitive differences show a "CC" signal. (* nonverbal vocalizations consisted of humming voices; ** participants were deaf individuals with cochlear implant).

EMOTION RECOGNITION